UNIVERSITÉ DE MONTRÉAL

FACULTÉ DE MÉDECINE

DÉPARTEMENT DE MÉDECINE SOCIALE ET PRÉVENTIVE

Université de Montréal
C.P. 6128, Succ. A
Montreal H3C 3J7

WA
18.2
A677
1983

INSPQ - Montréal
3 5567 00002 6911

SANTÉCOM

Institut national de santé publique du Québec
4835, avenue Christophe-Colomb, bureau 200
Montréal (Québec) H2J 3G8
Tél.: (514) 597-0606

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"

COURS INTENSIF ET INTEGRE

30 mai au 17 juin 1983

PROGRAMME ET RECUEIL DE TEXTES DE BASE

APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE

COURS INTENSIF ET INTEGRE

UNIVERSITE DE MONTREAL

6-17 juin 1983

PROGRAMME

PREMIERE SEMAINE:

Lundi le 6 juin 1983:

A.M.	8:30 - 9:30	Inscription des participants 2375 Côte Ste-Catherine Salle 2041
	9:30 - 9:45	Mot de bienvenue du doyen de la Faculté de médecine Dr Gauthier
	9:45 - 10:00	Organisation du cours et objectifs généraux Dr Jenicek
	10:00 - 11:15	Le contexte social dans lequel s'insère le système de santé Dr Desrosiers
	11:15 - 11:30	Pause
	11:30 - 12:30	Le Système de santé au Québec Groupe I: Atelier de révision et d'intégration des connaissances acquises durant la première semaine Dr Desrosiers Groupe II: Analyse de l'évolution récente du système de santé du Québec Dr Blain
P.M.	14:00 - 16:30	La planification des programmes de santé. Approche générale Drs Pineault et Contandriopoulos
	17:00 - 18:00	Cocktail de bienvenue Salle 2041

Mardi le 7 juin 1983:

A.M.	9:00 -	12:30	Module I. Identification des problèmes de santé et détermination des priorités Drs Pineault et Contandriopoulos
P.M.	14:00 -	16:30	Ateliers (au choix) - Système de santé - Société Dr Desrosiers - Analyse Coût - Avantage Dr Lance

Mercredi le 8 juin 1983:

A.M.	9:00 -	10:30	Validité et validation des méthodes du diagnostic Les indicateurs de santé Dr Jenicek
	10:30 -	10:45	Pause
	10:45 -	12:30	Identification des problèmes de santé dans une communauté: l'apport de l'anthropologie Dr Bibeau
P.M.	14:00 -	16:30	Ateliers (au choix) - Méthodes visant au consensus Drs Pineault et Contandriopoulos - Problèmes de santé des adolescents Dr Lescop

Jeudi le 9 juin 1983:

A.M.	9:00 -	10:00	La population et ses caractéristiques épidémiologiques. Etudes de prévale e (descriptives) à des fins d'identification des groupes prioritaires pour l'intervention selon les caractéristiques de personnes, temps et endroit Dr Jenicek
	10:00 -	10:15	Pause

Jeudi le 9 juin 1983:

A.M.	10:15 -	12:30	Etudes descriptives et leurs aspects statistiques Dr Lambert
P.M.	14:00 -	16:30	Ateliers (au choix) - Indicateurs de santé: exercices Dr Jenicek - Problèmes de santé des jeunes enfants Dr Seguin

Vendredi le 10 juin 1983:

A.M.	9:00 -	10:30	Les notions de risques Raisonnement causal et les études analytiques. Etudes par observation et par essai Dr Ducic
	10:30 -	10:45	Pause
	10:45 -	12:15	Etudes analytiques et leurs aspects statistiques Dr Lambert
	12:15 -	12:30	Synthèse du Module I Dr Pineault Evaluation de la première semaine Tous les professeurs
P.M.			Temps libre: travail individuel

DEUXIEME SEMAINE:

Lundi le 13 juin 1983:

A.M.	9:00 -	12:30	Module II. Intervention: La programmation Drs Pineault et Contandriopoulos
P.M.	14:00 -	16:30	Ateliers (au choix) - Etudes épidémiologiques: exemples illustratifs Dr Ducic - P.E.R.T. Dr Champagne - Programmes et programmation en médecine vétérinaire préventive Dr Simard

Mardi le 14 juin 1983:

A.M.	9:00 - 12:30	Etat du champ de l'éducation sanitaire Dr O'Neill
P.M.	14:00 - 16:30	Esquisse d'une approche alternative à la programmation d'activités de promotion de la santé Dr O'Neill

Mercredi le 15 juin 1983:

A.M.	9:00 - 12:30	Analyse de cas concrets d'intervention en santé communautaire dans une perspective anthropologique Dr Bibeau
P.M.	14:00 - 16:30	Ateliers (au choix) - Atelier ouvert sur l'épidémiologie Dr Ducic - Tout ce que vous avez toujours voulu savoir sur les statistiques et sans jamais avoir osé le demander Dr Lambert - Systèmes d'information socio-sanitaire Mme Valois Drs Jenicek et Contandriopoulos

Jeudi le 16 juin 1983:

A.M.	9:00 - 12:00	Evaluation des programmes et services de santé Table ronde Modérateur: Dr Pineault
P.M.	14:00 - 16:30	Ateliers (au choix) - Enquêtes épidémiologiques Drs Jenicek et Ducic - Validité des enquêtes épidémiologiques Dr Infante-Rivard

Vendredi le 17 juin 1983:

A.M.	9:30 -	11:30	Séance plénière Revue synthèse et évaluation du cours Tous les professeurs
	12:00 -	13:00	Cocktail de clôture

FIN DU PROGRAMME

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 6 juin 1983
Introduction du cours. Sujet 1: LE CONTEXTE SOCIAL DANS LEQUEL S'INSERE LE SYSTEME DE SANTE ET QUI LE DETERMINE

Responsable: Georges Desrosiers

Objectifs: A la fin de ce cours, les participants devraient:
- Pouvoir définir la santé communautaire.
- Pouvoir définir le concept de système de santé.
- Pouvoir identifier les principales caractéristiques de la société qui ont un impact sur le système de santé et montrer la façon dont leur influence se fait sentir.

Méthode pédagogique: Exposé par le professeur et discussion.

Lecture de base: Georges Desrosiers. "Les méthodes d'analyse des systèmes de santé." Avril 1983.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 6 juin 1983
Introduction du cours. Sujet 2: ANALYSE DE L'EVOLUTION RECENTE DU SYSTEME DE SANTE DU QUEBEC

Responsables: Gilbert Blain et Georges Desrosiers

Groupe I: Georges Desrosiers

Objectifs: - Pour ceux qui ont participé à la pré-session: atelier de révision et d'intégration des connaissances acquises durant la première semaine

Méthode pédagogique: Epreuve d'évaluation des connaissances. Correction de l'épreuve en groupe et commentaires

Lecture de base: Le système de santé du Québec dans le cadre canadien. Avril 1983

Groupe II: Gilbert Blain

Objectifs: A la fin de ce cours, les participants devraient:
- Pouvoir décrire sommairement et commenter l'évolution du système de santé du Québec au cours des vingt dernières années

Méthode pédagogique: Exposé du professeur et discussion

Lecture de base: E. Maquet et L. Mignot, "La réforme sanitaire au Québec, Révolution ou illusion". La Revue du praticien, Paris, France, Tome XXVIII, No 34, 1978.

G. Blain, "Ivan Illich et le rapport Castonguay: deux visions de la santé et de la médecine". L'Union médicale du Canada, Tome 104, Novembre 1975, pp. 1690-93

G.Blain, "La réforme doit prendre son "second souffle". L'Union médicale du Canada, Tome 104, Janvier 1975, pp. 38-46

G. Blain, "D'une réforme à l'autre". Le médecin du Québec. Vol. 13, No 10, Oct. 1978, pp. 43-55.

G. Blain, "Le déclin de l'Etat-providence". Administration hospitalière et sociale. Septembre-Octobre 1982.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 6 juin 1983
Introduction du cours. Sujet 3 - LA PLANIFICATION DES PROGRAMMES DE SANTE: APPROCHE GENERALE

Responsables: André-Pierre Contandriopoulos et Raynald Pineault.

Objectifs: A la fin de ce cours les participants devraient:

- Avoir une vue d'ensemble du fonctionnement dynamique du système de santé, c'est-à-dire des divers phénomènes qui constituent ce système et de leurs relations réciproques.
- Comprendre comment la démarche suivie dans le cours intègre les apports des disciplines telles que l'épidémiologie, l'économique, la biostatistique, etc...
- Etre familiers avec les différentes étapes de la planification et de la programmation en santé communautaire.
- Pouvoir situer le processus de planification par rapport aux éléments socio-politiques inhérents à une telle démarche.

Méthode pédagogique: Exposé par André-Pierre Contandriopoulos et Raynald Pineault et discussion avec les participants.

Lecture de base: Pineault, R., "Eléments et étapes d'élaboration d'un programme de santé communautaire", L'Union médicale du Canada, Tome 105, août 1976, pp. 1208-1214.

Pineault, R., "La planification des services de santé: une perspective épidémiologique", Gestions hospitalières, No 200, décembre 1980, pp. 951-959.

Contandriopoulos, A.P., "Le processus d'utilisation des services de santé", Mimoe, Montréal, 1983.

Contandriopoulos, A.P., "Le système de santé et la prévision de l'offre des services de santé", dans Modélisation et Maîtrise des Systèmes, Hommes et techniques, Paris 1977.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 7 juin 1983
Module 1, sujet 1 - IDENTIFICATION DES PROBLEMES DE SANTE ET DETERMINATION DES PRIORITES

Responsables: André-Pierre Contandriopoulos et Raynald Pineault

Objectifs: A la fin de ce cours, les participants devraient:

- Etre sensibilisés aux différentes approches utilisées pour identifier les problèmes de santé d'une communauté ainsi que les niveaux d'analyse de ces problèmes (santé-services-ressources).

- Connaître le concept de santé, les questions reliées au codage de la santé, les indicateurs et les indices pour mesurer l'état de santé.

- Pouvoir identifier et appliquer des critères pour déterminer des priorités d'intervention en matière de santé communautaire.

Méthode pédagogique: Exposé par André-Pierre Contandriopoulos et Raynald Pineault et discussion avec les participants.

Lecture de base: Goldberg M. et al, "Indicateurs de santé et sanométrie: les aspects conceptuels des recherches récentes sur la mesure de l'état de santé d'une population. Rev. Epid. et santé publique, 1979, 27, 51-68 et 133-152.

Clemenhagen C. et Champagne F., "Program planning in a small community health care setting". Health Care Management Review, vol 7, no 1, winter 1982, 47-55

Chambers L.W. and Woodward C.A., Guide to health needs assessment: A critique of available sources of health and health care information. The Canadian Public Health Assoc.,1980.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 8 juin 1983
Module I, sujet 2 (lère partie): VALIDITE ET VALIDATION DES METHODES DU DIAGNOSTIC INDIVIDUEL

Reponsable: Milos Jenicek

Objectifs: A la fin de ce cours les participants devraient:
- Etre en mesure de juger l'importance de la validité des instruments et des méthodes du diagnostic individuel, comment cette validité influe sur l'établissement des indicateurs de santé et sur l'établissement du tableau sanitaire de la collectivité, dans les étapes initiales de la programmation sanitaire.

Méthode pédagogique: Exposé, exemples pratiques, discussion.

Sujets:
Diagnostic définitif et diagnostic présomptif.
Mesure de la validité interne d'un test.
Mesure de la validité externe d'un test.
Critères du choix d'un programme de dépistage en fonction des caractéristiques de la maladie, de la population-cible et des tests disponibles.

Lecture de base: Jenicek, M. et Cléroux, R. "Normalité et diagnostic clinique et enquêtes de dépistage." pp. 23-41 in: Epidémiologie. Principes, techniques, applications. St-Hyacinthe et Paris: EDISEM et Maloine, 1982.

Schoenberg B.S. "Descriptive epidemiology." Adv. Neurol., 19: 17-42, 1978. (Lire: Defining the disease under investigation, pp. 18-25)

Wagner, G., Tautu, P. and Wolber, U. "Problems of medical diagnosis." A Bibliography. Meth. Inform. Med., 17: No 1 entier, 1978.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 8 juin 1983
Module I, sujet 2 (2ième partie): INDICATEURS DE SANTE
DIAGNOSTIC AU NIVEAU DE LA COLLECTIVITE

Responsable: Milos Jenicek

Objectifs: A la fin de ce cours les participants devraient:
- Connaître l'éventail des indicateurs de santé de la collectivité et leur valeur relative en fonction des différents problèmes sanitaires abordés et des interventions sanitaires envisagées.

Méthode pédagogique: Exposé, exemples et problèmes pratiques.

Sujets: Concept de la normalité chez l'individu.
Collectivité: normalité de la situation épidémiologique.
Taux et ratios.
Morbidité, mortalité, létalité.
Prévalence, incidence, durée de la maladie.
Sources de données. Biais de Berkson.
Interprétation et choix des indicateurs à des fins de programmes de santé.

Lecture de base: Jenicek, M. et Cléroux, R. "Mesure de la santé de la collectivité. Indicateurs de santé." pp. 45-77 in: Epidémiologie. Principes, techniques, applications. St-Hyacinthe et Paris: EDISEM et Maloine, 1982.

Schoenberg, B.S. "Descriptive epidemiology." Adv. Neurol., 19: 17-42, 1978. (Lire: Calculating epidemiologic indices)

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET MOTHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 8 juin 1983
Module 1, sujet 3 - L'IDENTIFICATION DES PROBLEMES DE SANTE DANS UNE COMMUNAUTE: L'APPORT DE L'ANTHROPOLOGIE

Responsable: Gilles Bibeau

Objectifs: A la fin de ce cours les participants:

- Seront plus familiers avec certains concepts anthropologiques de base: culture, ethnoscience, savoirs populaire et scientifique, représentations en matière de santé, croyances, mouvements sociaux (Touraine), "health belief model" et sa critique, etc.

- Saisiront les différences entre une éducation sanitaire qui se fonde sur l'épidémiologie (une place importante y est faite aux habitudes de vie) et celle qui se fonde sur une approche ethnologique et sémiologique (cette dernière fait plutôt une large place aux aspects cognitifs et phénoménologiques); nécessité de conjuguer les deux approches.

- Seront sensibilisés au fait que les changements au niveau individuel s'insèrent dans le contexte plus large de changements culturels.

Méthode pédagogique: - Présentation magistrale
- Discussion

Lecture de base: Bibeau, G., Pelletier, L., Le discours sur la santé et la maladie dans deux populations de Québec, à paraître dans la Collection Mercure, Musée de l'Homme, Ottawa.

Herzlich, C., Santé et Maladie - Analyse d'une représentation populaire, Paris: Mouton, 1969, p. 13-29.

Bibeau, G., Tremblay, J., Résistance et soumission face à l'émergence d'une autre conception de la santé, in Union Médicale du Canada, vol. 109 (mai 1980) 5, 789-90.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 9 juin 1983
Module I, sujet 4: LA POPULATION ET SES CARACTERISTIQUES EPIDEMIOLOGIQUES

Responsable: Milos Jenicek

Objectifs: A la fin de ce cours les participants devraient:
- Comprendre la démarche épidémiologique générale d'une enquête dans la collectivité.
- Savoir interpréter à des fins pratiques les caractéristiques de personnes, de temps et de l'endroit.
- Saisir les qualités et limitations des études descriptives à des fins de la programmation sanitaire.

Méthode pédagogique: Exposé, exercices pratiques, discussion.

Sujets: Démarche épidémiologique générale d'une enquête dans la collectivité.
Objectifs des études descriptives.
Définition et interprétation des caractéristiques de personnes, du temps et de l'endroit.

Lecture de base: Jenicek, M. et Cléroux R. "Réalisation et interprétation des études descriptives." pp. 95-118 in: Epidémiologie. Principes, techniques, applications. St-Hyacinthe et Paris: EDISEM et Maloine, 1982.

Picheral H. "Géographie médicale. Géographie des maladies. Géographie de la santé." L'espace géographique. 11: 161-175, 1982.

Bernard, P.-M., Tremblay, J. et Begin, C. "Prévalence de l'usage de la cigarette en milieu scolaire: Québec, 1978." Un. Méd. Canada. 109: 1724-1733, 1980.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 9 juin 1983
Module I, sujet 5 - ETUDES DESCRIPTIVES ET LEURS ASPECTS STATISTIQUES

Responsable: Jean Lambert

Objectifs: A la fin de ce cours, les participants devraient pouvoir:
- Saisir les notions d'études populationnelles et échantillonnales.
- Comprendre les concepts de représentativité, d'estimation et de précision dans les études échantillonnales.
- Se familiariser avec les principaux types d'échantillons.

Méthode pédagogique: Exposé, exemples et problèmes pratiques.

Sujets: Population. Echantillon. Erreurs aléatoires et non aléatoires. Paramètres et estimateurs ponctuels et par intervalle. Types d'échantillons.

Lecture de base: Lambert, J. "Concepts de base en biostatistique", Montréal, Librairie de l'Université de Montréal, 1981.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 10 juin 1983
Module I, sujet 6 - LES NOTIONS DU RISQUE

Responsable : Slobodan Ducic

Objectifs: A la fin de ce cours, les participants devraient:
- Comprendre les principes du raisonnement causal.
- Connaître la structure des études analytiques.
- Etre capables d'identifier les biais, les facteurs distracteurs et porter les jugements qui s'imposent.

Méthode pédagogique: Exposé, exercices, exemples pratiques.

Sujets: Principes du raisonnement causal. Mesure d'une association causale. Principe, avantages et désavantages des études prospectives et rétrospectives. Recherche des causes d'une éclosion. Applications de la démarche analytique aux études de l'étiologie des maladies et de l'efficacité des interventions.

Lecture de base: Chapitre 7 de l'Introduction à l'épidémiologie.

Références complémentaires: Evans, A.S. Causation and disease: A chronological journey. The Thomas Parran Lecture. Am. J. Epidemiol., 108:249-258, 1978.

Fournier, P. et Bocquet, J.P. Les études rétrospectives en épidémiologie clinique. Le Concours Médical, 105-09:889-899, 1983.

Feinstein, A.R. et Horwitz, R.I. Double standards scientific methods and epidemiologic research. The New England Journal of Medicine. Vol. 37, no 26, pp. 1611-1617.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 10 juin 1983
Module I, sujet 7 - ETUDES ANALYTIQUES ET LEURS ASPECTS STATISTIQUES

Responsable: Jean Lambert

Objectifs: A la fin de ce cours, les participants devraient pouvoir:
- Saisir le concept d'une décision statistique.
- Prendre connaissance des principaux devis expérimentaux rencontrés en recherche en santé communautaire.
- Se familiariser avec les principaux tests d'hypothèse.

Méthode pédagogique: Exposé, exemples et problèmes pratiques.

Sujets: Structure d'un test d'hypothèse. Devis pré-expérimentaux, quasi-expérimentaux et expérimentaux. Comparaisons de moyennes et de proportions. Etudes de relations: régression et corrélation.

Lecture de base: Lambert, J. "Concepts de base en biostatistique", Montréal, Librairie de l'Université de Montréal, 1981.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 13 juin 1983
Module II, sujet 1 - INTERVENTION: LA PROGRAMMATION

Responsables: André-Pierre Contandriopoulos et Raynald Pineault

Objectifs: A la fin de ce cours, les participants:

- Seront familiers avec les concepts relatifs à l'établissement des objectifs.
- Seront capables d'analyser les différents types d'activités rencontrées dans un programme de santé.
- Comprendront et seront capables d'appliquer les méthodes nécessaires pour traduire les besoins de santé en besoins de services puis en besoins de ressources.
- Seront sensibilisés à certaines méthodes utilisées dans l'implantation et la réalisation de programmes notamment le PERT.

Méthode pédagogique: Exposé par André-Pierre Contandriopoulos et Raynald Pineault et discussion avec les participants.

Lecture de base: Dever, G.E., Alan, "Health Resource Analysis" pp. 261-304 du Volume "Community Health Analysis", 1980, Aspen publication.

Spiegel, A.D., et Hyman, H.H., "Promoting Implementation", Chapter 6, Tiré de: Basic Health Planning Methods, Aspen, 1978, pp. 239-316.

Pineault, R., et Lessard, R., "Le système de santé du Québec: objectifs de système ou objectifs de santé?" Miméo 1983.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 14 juin 1983
Module II. Sujet 2 - LES POSSIBILITES ET LIMITES DES ACTIVITES DE PROMOTION DE LA SANTE:

(1) Etat du champ de l'éducation sanitaire

Responsable: Michel O'Neill, professeur
Ecole des sciences infirmières,
Université Laval, Québec

Objectifs: A la fin de ce cours, les participants devraient:

- Pouvoir apprécier quel genre de succès ont eu les programmes d'éducation sanitaire mis sur pied depuis 1950 tant dans les pays développés que dans les pays en développement.

Sujets:
- Les biais épistémologiques de la santé communautaire
- Une définition de l'éducation sanitaire
- La récente popularité de l'éducation sanitaire et ses causes.
- L'échec des programmes "traditionnels" d'éducation sanitaire.

Méthodes pédagogiques:
- Exposé magistral suivi de discussion
- Lecture (préalable, si possible) des textes suivants:

Lecture de base: -O'Neill, Michel, "La modification des comportements reliés à la santé". Revue de la littérature théorique et suggestions pour l'organisation de programmes de santé communautaire"; 1'Union médicale du Canada, 109(s): 733-739 et 109(6): 921-28, 1980.

- O'Neill, Michel, "Vers une problématique de l'éducation sanitaire au Québec; Montréal, Librairie de l'Université de Montréal, chapitre 2, pp. 18-40, 1979.

- Richards, David N., "Methods and effectiveness of health education: The past, present and future of social scientific involvment"; Social Science and Medicine, 9:141-156, 1975.

- Green, Lawrence W., "Determining the impact and effectiveness of health education as it relates federal policy", Health Education Monographs, 6 (supplement 1): 28-66, 1978.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 14 juin 1983
Module II. Sujet 3 - LES POSSIBILITES ET LIMITES DES ACTIVITES DE PROMOTION DE LA SANTE:

(2) Esquisse d'une approche alternative à la programmation d'activités de promotion de la santé

Responsable: Michel O'Neill

Objectifs: Au terme de ce cours les participants devraient:

- Etre capable de saisir l'émergence d'un nouveau paradigme en éducation sanitaire et de son application éventuelle.

Sujets:
- Emergence de "l'approche critique"
- Quelques-unes de ses caractéristiques principales
- Une reformulation de l'approche critique: prémisses
- Conséquences de ces prémisses pour le développement de programmes.

Méthodes pédagogiques:
- Exposé magistral suivi de discussion
- Lecture (préalable, si possible) des textes suivants:

Lecture de base:
- Labonté, Ronald et Susan Peufold, "Analyse critique des perspectives canadiennes ou promotion de la santé", Education sanitaire, 19(3-4): 4-10 1981.

- Brown, Richard E. et Glenn E. Margo, "Health educators: can the reformers be reformed", International Journal of Health Services, 8(1): 3-25, 1978.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 15 juin 1983
Module II. Sujet 4 - ANALYSE DE CAS CONCRETS D'INTERVENTION EN SANTE COMMUNAUTAIRE DANS UNE PERSPECTIVE ANTHROPOLOGIQUE

Responsable: Gilles Bibeau

Objectifs: Au terme de ce cours les participants pourront sur la base des programmes présentement utilisés dans les sociétés occidentales pour changer les comportements reliés au tabac, à l'alcool et aux drogues, saisir les limites de ces programmes dans une perspective anthropologique.

Sujet: Sur la base de programmes concrets d'éducation sanitaire mis en place depuis les années 50, on tentera de découvrir les raisons des échecs ou des demi-succès de ces interventions: on examinera plus spécialement les programmes dans le domaine nutritionnel, dans la lutte contre les parasites intestinaux, dans la promotion de l'eau potable, etc. qui furent mis en oeuvre pour la plupart dans les sociétés non-occidentales.

Méthodes pédagogiques:
- Présentation magistrale
- Discussion en groupe

Lecture de base:
- Foster, G.M., Anderson, B.G., Medical Anthropology, New York: John Wiley and Sons, 1978, p. 223-42.

- Landy, D., Culture, Disease and Healing, New York: Mac Millan, 1977, p. 231-50.

- Tremblay, J., Bibeau, G., Hébert, F., Rousseau, N., Les Mécanismes de reproduction du tabagisme chez les pré-adolescents québécois, rapport ronéotypé. Août 1981, p. 92-103.

UNIVERSITE DE MONTREAL
Faculté de médecine
Département de médecine sociale et préventive

"APPROCHES ET METHODES EN SANTE COMMUNAUTAIRE"
Cours intensif et intégré

Rencontre du 16 juin 1983
Module III. Sujet 1 - EVALUATION DES PROGRAMMES ET SERVICES DE SANTE (Table ronde)

Modérateur: Raynald Pineault

Participants: A.P. Contandriopoulos, S. Ducic, J. Lambert

Objectifs: A la fin de cet avant-midi les participants seront aptes à:

- Préciser les notions de qualité
- Situer leurs actions par rapport à un schéma général d'évaluation.
- Situer l'évaluation dans le processus de planification en santé communautaire.
- Distinguer la recherche évaluative de l'évaluation administrative.
- Etre sensibilisés aux devis de recherche utilisés en évaluation et des outils statistiques appropriés pour ce genre de recherche.
- Etre conscients de l'effet possible des facteurs confondants dans l'évaluation des effets d'une manoeuvre ou d'une intervention.

Méthodes pédagogiques: - Bref exposé par chacun des professeurs participants
- Discussion et synthèse

Lecture de base: - R. Pineault, "Rationalisation de l'évaluation et du contrôle dans les organismes de santé". Adm. hosp. et soc., Mars avril 1977, pp. 14-26.

- Cook, T.D. et al., "Modèles expérimentaux et quasi-expérimentaux en recherche évaluative: une introduction", dans Lecomte L. et Rutman L. Introduction aux méthodes de recherche évaluative, Les Presses de l'Université Laval, 1982, pp. 105-141.

- Beaudoin, A., "La réalisation de l'évaluation dans la pratique: critères et limites" dans Introduction aux méthodes de recherche évaluative, op. cit., pp. 155-176.

LECTURE POUR LUNDI LE 6 JUIN 1983

LE SYSTEME DE SANTE DU QUEBEC
DANS LE CADRE CANADIEN

Le Québec est une des provinces de la Fédération canadienne. Dans le partage des pouvoirs entre les deux niveaux de gouvernement, l'administration des services de santé et des services sociaux est de compétence exclusive des provinces, à quelques exceptions près. C'est à cause de cette particularité qu'on peut parler d'un système de santé propre au Québec ou à chacune des autres provinces. Par ailleurs, on ne peut réellement comprendre le système de santé du Québec sans le situer à l'intérieur du cadre canadien.

De tout temps les sociétés ont cherché à se protéger contre le fléau de la maladie en créant des institutions qui aujourd'hui se regroupent à l'intérieur d'un ensemble qu'on appelle le système de santé. Comme toutes les autres institutions sociales, le système de santé a sa finalité propre qui pour lui est la protection contre la maladie et la promotion de la santé. Mais il est aussi une finalité commune à toutes les institutions sociales qu'on appelle la reproduction sociale, qui consiste à chercher à maintenir la société telle qu'elle existe dans sa forme actuelle, avec ses relations de pouvoir entre des groupes d'intérêts différents et les inégalités sociales qui le plus souvent en résultent et qui dans le domaine de la santé conduisent aux inégalités persistantes devant la maladie et la mort. La société fournit au système de santé les ressources matérielles et humaines dont il a besoin pour accomplir sa mission. Les transformations qui sont apportées à ces ressources, leurs relations fonctionnelles, les fins auxquelles elles servent, sont toutes conditionnées par le contexte social. Les caractéristiques les plus siginficatives de ce contexte sont géognaphiques, démographiques, politico-économiques, culturelles, sociales et épidémiologiques.

Dans une première partie, nous allons considérer le cadre canadien en analysant ses principales caractéristiques qui affectent le système de santé et ensuite l'évolution du système de santé lui-même. Dans une deuxième partie, nous allons considérer de la même façon le système de santé québécois.

Le cadre canadien

Géographie:

Le Canada est le deuxième pays du monde en étendue après l'U.R.S.S. avec une superficie de 9,900,000 kilomètres carrés. Il est borné tout le long de sa longue frontière sud qui divise le continent en deux d'est en ouest, par les Etats-Unis. A l'est par l'océan Atlantique, à l'ouest par l'océan Pacifique et au nord par l'océan Artique qui le sépare de l'U.R.S.S. Au nord-ouest il partage une frontière commune avec l'état américain de l'Alaska. Le climat est froid. Il se pose d'importants problèmes de communications, surtout durant l'hiver, particulièrement pour les régions éloignées des grands centres

Démographie:

C'est un pays peu peuplé qui en 1981 comptait 24,343,000 habitants dont le plus grand nombre sont concentrés le long de la frontière américaine.

Système économique:

Le Canada est un pays à système capitaliste c'est-à-dire que la propriété de la plus grande partie des moyens de production est privée. Cependant, il y a deux différences importantes avec les Etats-Unis. La première est qu'une partie importante du capital investi provient de sociétés multinationales surtout américaines. La deuxième est que l'état central et les états provinciaux possèdent un important secteur nationalisé: une partie des chemins de fer, une partie des sociétés de transport aérien, le transport municipal, l'électricité, une partie de l'exploitation et de la distribution du pétrole et du gaz naturel, des sociétés minières diverses, etc.

Politique:

Le système politique est celui d'une démocratie parlementaire de type britannique. C'est un état fédéral à deux niveaux de gouvernements autonomes, responsables envers la population, le gouvernement fédéral et les dix gouvernements provinciaux. Il y a trois grands partis politiques fédéraux: le parti social démocrate, le parti libéral et le parti conservateur. Le parti social démocrate est né dans les Prairies au début du siècle. Il a toujours préconisé un régime national d'assurance-santé. Il a été traditionnellement appuyé par les coopératives agricoles de l'Ouest et par les syndicats ouvriers. Le parti libéral, devant la montée du parti social-démocrate dans l'entre-deux guerres et pour d'autres raisons, a inséré dans son programme l'assurance-maladie dès cette époque. Le parti conservateur a plus ou moins suivi le parti libéral dans cette orientation. Ce qui fait qu'au Canada, contrairement aux Etats-Unis, il y a eu un consensus politique pour la création d'un régime public d'assurance-hospitalisation et maladie dès les années 50.

Les valeurs de la société canadienne:

Le Canada moderne tire son origine de trois groupes différents, les descendants des français installés en Nouvelle France, les Loyalistes américains refusant la révolution, les immigrants des iles Britanniques auxquels se sont ajoutés par la suite des vagues d'immigrants d'autres pays d'Europe et du monde.

Le rejet de l'expérience américaine amena les Canadiens à s'attacher aux institutions sociales existantes. Ils n'avaient pas d'autres choix. Ces institutions avaient les caractéristiques historiques suivantes: le gradualisme, c'est-à-dire l'évolution plutôt que la révolution, la monarchie parlementaire, la paix, l'ordre et le bon gouvernement, les conventions et la tradition et, enfin, la diversité.

Une des caractéristiques de l'héritage culturel canadien a été la dépendance qui a engendré la croyance dans l'action collective. Par exemple, le développement économique a toujours été dépendant des exportations extérieures et des investissements étrangers. Cette dépendance et les problèmes posés par l'environnement physique commandaient une action collective pour soutenir le développement économique. C'est ainsi que se sont développés dès la fin du 19e siècle les sociétés d'état tel que mentionné précédemment.

Face au "melting pot" culturel américain, le Canada est une société pluraliste favorisant ou acceptant des groupes indépendants, différents et souvent ayant des intérêts opposés, intérêts économiques, politiques, religieux, ethniques, scientifiques, professionnels et autres. Les partis politiques manoeuvrent pour respecter ces différences par un jeu de balance donnant un peu à tous. Ces valeurs dominantes ont eu une influence très forte sur l'évolution du système de santé. C'est ainsi que la croyance en la responsabilité collective a fait accepter très tôt le rôle dominant de l'Etat dans le financement et l'organisation des services de santé. Le besoin de sécurité a fait accepter un système de sécurité sociale à caractère universel le plus complet possible incluant la garantie contre les risques maladies et accidents, contrairement aux Etats-Unis où traditionnellement on s'est efforcé de limiter la sécurité sociale à ceux qui n'ont pas réussi. Le sens de l'ordre a concouru à la création d'un système coordonné, assez intégré, les éléments ayant une certaine dépendance les uns par rapport aux autres; tandis qu'aux Etats-Unis, l'individualisme et l'esprit de compétition ont favorisé l'éclatement du réseau. La croyance au gradualisme explique la succession des phases dans l'établissement d'un système public complet et universel. La croyance en la responsabilité individuelle, héritée à la fois des Américains et des Anglais, a favorisé l'implication des citoyens dans l'administration locale des institutions. La croyance en la liberté héritée aussi des Anglais et des Américains a concouru à maintenir l'autonomie des établissements, la liberté du choix du médecin et de l'hôpital par le malade et la liberté du médecin de choisir ses malades.

Caractéristiques épidémiologiques:

L'évolution de la mortalité et de la morbidité a subit la même évolution au Canada que dans les autres pays industrialisés. Au début du siècle les principales causes de mortalité étaient les maladies infectieuses telles que la grippe, la pneumonie, la tuberculose, la gastro-entérite, la néphrite chronique et la diphtérie. Dans les années 70, ce sont les maladies cardio-vasculaires, les cancers de toutes natures et les accidents qui constituent les principales causes de mortalité au Canada.

Evolution du système de santé:

Bien que l'administration des services de santé et des services sociaux à quelques exceptions près, relève exclusivement des provinces comme nous l'avons mentionné plus haut, le gouvernement fédéral a joué un rôle prépondérant dans l'évolution du système de santé, en offrant aux provinces des subventions conditionnelles leur permettant de développer les systèmes de soins à l'intérieur d'un cadre déterminé en grande partie par lui.

C'est à la suite de la grande crise économique de 1929 et aussi sous l'influence des réformes dans le domaine social entreprises en Grande-Bretagne que le gouvernement fédéral s'intéressa aux moyens à mettre en oeuvre pour garantir à tous les citoyens la protection contre les risques de la maladie, du chômage, de la vieillesse et de l'invalidité. On songea dès le début des années quarante à un régime public universel de protection contre la maladie mais devant l'opposition des provinces à céder leurs pouvoirs, le gouvernement fédéral décida d'utiliser une stratégie différente pour assurer le développement des services de santé à la mesure des possibilités de la nation et accessibles à tous les citoyens indépendamment de leur lieu de résidence et de leur capacité de payer. Cette stratégie a consisté à offrir aux provinces de leur accorder des subventions conditonnelles pour des mesures partielles. C'est la méthode étapiste. La première étape fut celle des subventions nationales et à la santé en 1948 pour encourager les provinces à augmenter leur équipement hospitalier. Cette mesure fut suivie en 1958 de la loi sur l'Assurance-hospitalisation par laquelle le gouvernement fédéral payait, à certaines conditions dont la gratuité complète, en moyenne 50% du coût des frais d'exploitation des hôpitaux aux provinces qui acceptaient d'adhérer au régime avec un système de péréquation pour les provinces plus pauvres. En 1965, fut institué le programme d'Aide à la santé qui prévoyait des subventions aux provinces pour le financement des équipements servant à la formation des personnels de santé, médecins, infirmières et para-médicaux. Ce programme fut suivi en 1968 de la loi sur l'Assurance-maladie, prévoyant une contribution fédérale aux provinces de 50% des coûts des frais des médecins à condition que les grands principes de base, gratuité totale, couverture universelle, transférabilité d'une province à l'autre et administration par un organisme public, soient respectés. Dans ce système, les administrations hospitalières gardaient leur autonomie de gestion et la profession médicale son caractère libéral, par le maintien de la liberté d'accès à la spécialisation, la liberté d'installation, la liberté thérapeutique, le système à l'acte même pour les médecins hospitaliers, mais dans le cadre du tiers payant. Les médecins ont même conservé la liberté de pouvoir exercer en dehors du régime à condition toutefois que le nombre de non-participants d'une province ne soit pas suffisamment élevé pour remettre en question l'accès aux soins gratuits. Le fait de procéder en deux étapes, les frais hospitaliers d'abord, puis les frais des médecins dix ans plus tard, a conduit à ce qu'on peut appeler l'hospitalocentrisme, c'est-à-dire la polarisation sur l'hôpital. L'hôpital canadien est devenu, au cours des années 60, non seulement le centre de distribution des soins spécialisés et ultra-spécialisés, mais aussi d'une bonne partie des soins de première ligne par l'expansion des services d'urgence et de consultations externes. Le voisinage de ces services de première ligne avec les services médico-techniques de plus en plus raffinés a favorisé la surutilisation.

Le système de financement par lequel le gouvernement fédéral payait 50% des coûts engagés par les provinces, appelé en anglais "open-ended system" a eu comme conséquence d'encourager les provinces à dépenser, le fédéral réglant automatiquement la moitié de la note. De part et d'autre, il n'y voit là aucun incitatif à limiter l'accroissement des coûts. C'est ainsi que la proportion du produit national brut consacrée aux dépenses de santé est passée de 4,35% en 1955 à 7% en 1970. On peut donc dire qu'après plus de 25 ans de débats et de mesures législatives successives l'objectif d'accessibilité dans le respect de l'autonomie des établissements et des professions avait été atteint.

Cette première phase du développement du système de santé canadien a correspondu à une phase expansionniste de l'économie. L'Etat a créé l'infrastructure et a garanti la consommation, l'industrie privée pour sa part, a fourni l'équipement médical et les produits pharmaceutiques. On a ainsi créé une véritable industrie de la santé qu'on a justifiée par l'idéologie vantant les mérites de la science et de la technologie et son application par la spécialisation médicale au traitement des maladies.

A partir du début des années 70, le gouvernement fédéral s'intéresse à la rationalisation, c'est-à-dire à la limitation de l'accroissement des coûts et à l'efficacité. On met en lumière les nouveaux problèmes de santé qu'on nomme les maladies de la civilisation, comme les maladies cardio-vasculaires, les cancers, les accidents et les maladies mentales survenant chez une population de plus en plus vieillissante. Des commissions d'enquête se sont penchées sur les coûts de la santé, sur les services de première ligne et de type communautaire, sur les facteurs de risque. Tout cela a été bien résumé dans un petit ouvrage publié en 1974 par le ministère de la Santé et du Bien-être social qui a fait son tour du monde: "Nouvelles perspectives de la santé des Canadiens". Dans le but d'accélérer ce changement d'orientation, le gouvernement fédéral et les provinces signèrent une nouvelle entente en 1977 sur le partage des coûts. Cette entente prévoyait que la contribution fédérale ne serait plus versée au titre de régimes différents assujettis de conditions particulières, mais sous forme de masse globale ou de remise de points d'impôt que les provinces sont libres d'affecter à leur guise pourvu que la gratuité et l'universalité de la couverture des soins hospitaliers et médicaux soient respectées. Elle prévoyait en outre que l'augmentation annuelle de la part fédérale ne se fera plus en fonction des coûts engagés par les provinces mais en fonction de l'augmentation du produit national brut, ce qui est de nature à mettre un frein à l'escalade des coûts qui, de toute façon n'ont pas dépassé depuis 1970 les 7,1% du produit national brut. Cette phase de rationalisation du système de santé correspondait à une période de récession économique qui exigeait, dans le système économique actuel, une restriction des dépenses publiques dans le but de favoriser la relance économique par l'entreprise privée. Au discours expansionniste des années 60 s'en substitue un autre. On fait état des limites de la technologie, de l'impuissance de la médecine à vaincre les maladies dites de la civilisation, de l'application souvent intempestive des innovations techniques en médecine, de l'importance de la modification des comportements individuels dans la prévention des maladies, de la nécessité de passer d'un système de distribution des soins centré sur la haute technique et la spécialisation et sur la dominance des professionnels à un système orienté en fonction des soins primaires, globaux à l'échelon local,

avec la participation des citoyens. A mesure que la grande entreprise augmentait sa pression auprès des gouvernements pour les inciter à réduire les dépenses publiques, ce discours gagnait en intensité. On va même depuis quelques années jusqu'à remettre en cause la gratuité des soins on parle dans plusieurs provinces d'introduire un ticket modérateur. Mais, jusqu'à maintenant, le gouvernement fédéral n'a pas manifesté l'intention de changer les règles du jeu à cet égard.

Le système de santé du Québec

Le Québec, avec ses 1,540,500 kilomètre carrés constitue la plus grande des provinces canadiennes. La population était de 6,438,400 habitants en 1981 dont environ 82% de francophones.

Avant 1960, le système de santé du Québec, comme celui des autres provinces, reposait sur l'entreprise privée. D'une part, les communautés religieuses et, à un moindre degré, les associations de bienfaisance, administraient les hôpitaux, les hospices et les établissements d'enseignement secondaire et universitaire, d'autre part, les médecins exerçaient dans un cadre complètement libéral, l'Etat se contentant d'intervenir pour subventionner l'entreprise privée ou pour occuper les champs qu'elles se refusaient d'occuper comme l'hygiène du milieu et la médecine préventive. Les régimes d'aide mutuelle, d'assurances originant des producteurs de soins comme les hôpitaux et les médecins, ou les assurances commerciales étaient moins développés au Québec qu'en Ontario et aux Etats-Unis.

L'année 1960 a marqué un tournant dans l'évolution politique et sociale du Québec par l'arrivée au pouvoir d'un parti réformiste, après plus de 15 ans d'un régime ultra-conservateur. Durant les années qui suivirent, fut mis en oeuvre un train de réformes politiques, économiques et sociales qui ont permis au Québec d'opérer un important rattrapage et, dans plusieurs domaines, de se hisser à l'avant-garde des réformes sociales au Canada. La décennie des années 60 commença par l'Assurance-hospitalisation et se termina par l'Assurance-maladie, le Québec ayant accepté de participer à ces deux régimes proposés par le gouvernement fédéral. L'Assurance-hospitalisation a permis une extension considérable du secteur hospitalier. L'augmentation du personnel, la syndicalisation massive de tout le secteur para-public et la demande de gestionnaires rompus aux méthodes modernes d'administration alliée au tarissement des vocations religieuses, ont amené les communautés religieuses à abandonner complètement la direction des établissements de santé sans provoquer de secousses dans la société. Contrairement à ce qui s'est produit dans le reste du Canada, la profession médicale commença à s'organiser en syndicats professionnels autonomes dès le début des années 60. Les omnipraticiens formèrent la première fédération autonome suivie, quelques années plus tard, de la fédération des médecins spécialistes et de la fédération des médecins internes et résidents.

Avant 1960, l'Ordre des médecins, vieille institution para-publique fondée en 1848, en plus de réglementer les conditions d'accès aux études médicales, de délivrer le permis d'exercice de la médecine, de réprimer l'exercice illégal de la médecine, de faire respecter le code de déontologie et d'établir des tarifs d'honoraires indicatifs, représentait les médecins pour la défense du caractère libéral de la médecine, à ce titre, jouait le rôle d'organisme de pression. Vers 1965, il abandonna volontairement ce dernier rôle aux syndicats et laissa la fixation des tarifs au libre jeu de la négociation.

En 1966, fut établi le régime de l'assistance médicale en vertu duquel les allocataires et les personnes à charge reçoivent gratuitement les soins médicaux et chirurgicaux dont ils ont besoin. Les deux fédérations d'omnipraticiens et de spécialistes obtinrent le régime de la rémunération à l'acte selon un tarif fixé par négociation collective dans le cadre du tiers payant. Avant cette entente, les médecins avaient l'habitude de fournir gratuitement leurs services aux assistés sociaux. Durant cette même année, les radiologistes firent la grève et optinrent de passer du régime du salariat à celui du paiment à l'acte dans les hôpitaux. Comme on peut le constater, ces premières mesures consacrèrent l'autonomie de la profession médicale dans un régime public d'assurance-maladie. C'est le système de la médecine libérale subventionnée.

Le droit de grève dans le secteur public et para-public fut accordé par l'Etat en 1966. La première grève générale des employés d'hôpitaux fut déclenchée en 1967. Par la suite, rares furent les renouvellements d'entente qui ne furent pas précédés d'une grève générale. C'est aussi en 1966 que le gouvernement créa une commission d'enquête chargée d'étudier les divers aspects de la santé, du bien-être social et de la sécurité du revenu. Cette commission présidée par un actuaire s'adjoignit une équipe formée d'universitaires et de technocrates. Constatant le taux relativement bas de certains indicateurs de santé par rapport à l'Ontario et aux provinces de l'Ouest, découvrant un système qui s'était développé de façon incohérente au gré des intérêts locaux et professionnels, s'inspirant largement des études et des expériences faites ailleurs, notamment aux Etats-Unis, ils proposèrent une réforme globale de tout le système de santé fondée sur une conception de la médecine orientée sur la santé plutôt que la maladie. Pour ce faire, ils fixèrent comme objectifs généraux d'améliorer l'état de santé de la population et l'état du milieu physique et social; améliorer l'état de santé par une médecine globale axée sur la personne incluant des soins complets, préventifs, curatifs et de réadaptation, continus et personnels. Pour atteindre ces objectifs, ils proposent un système public cohérent et intégré de santé et de services sociaux par les moyens suivants: la création d'un Ministère des Affaires sociales, responsable de la santé et des services sociaux; l'abandon par ce Ministère de son rôle de pourvoyeur de services, c'est-à-dire tout le réseau des services de médecine préventive au profit de celui de planificateur et responsable de l'allocation et de l'utilisation des ressources, de l'établissement et de la surveillance

des normes; la décentralisation de la gestion par la création de conseils régionaux; la participation des professionnels et des usagers à l'identification des besoins et à la gestion des établissements; l'introduction du concept de gestion par programmes, d'évaluation permanente de l'utilisation des ressources et de l'efficacité des programmes; l'intégration des soins préventifs et curatifs dans les établissements du réseau; la création d'un réseau d'établissements publics de première ligne; l'orientation de la recherche en santé vers des objectifs précis plutôt que de maintenir le système de subventions selon les seuls intérêts des chercheurs; l'adaptation du système d'enseignement des sciences de la santé aux besoins nouveaux du système de santé; la réorganisation des structures des professions de la santé par une meilleure définition du rôle des ordres, par une délimitation plus précise de leurs champs d'activité, par l'ouverture de leurs conseils d'administration à des non-professionnels et par la création d'un Conseil des professions chargé de la surveillance des ordres et de la surveillance de l'application d'un code des professions. La Commission dépose son rapport en 1970.

Au début de 1970, un nouveau gouvernement prend le pouvoir avec la promesse d'établir immédiatement l'assurance-maladie aux conditions posées par le gouvernement fédéral. L'ex-président de la Commission devient ministre de la santé. Son équipe entre avec lui au ministère et occupe tous les postes-clés. Le premier défi du nouveau ministre est la résistance de la Fédération des médecins spécialistes qui exige le droit de tout médecin d'exercer en dehors du régime, de demander directement ses honoraires aux patients et à ces derniers d'obtenir un remboursement d'au moins 75%, et que l'échelle de rémunération proposée soit plus élevée. Les négociations sont rompues, la grève générale de tous les médecins spécialistes éclate affectant surtout les hôpitaux parce que la Fédération des médecins omnipraticiens de son côté accepte le régime proposé par le gouvernement. Comme cette grève survient en octobre, au moment même d'une grave crise politique qui secoue le Québec et le Canada tout entier par l'action des terroristes du Front de libération du Québec, les spécialistes, n'ayant aucun support de la population, sont obligés de retourner au travail sous le coup d'une loi spéciale votée par le parlement du Québec.

Ce régime d'assurance-maladie consacrait l'existence d'un réseau privé subventionné intégralement parallèle au réseau public et pouvant éventuellement lui faire concurrence.

En 1971 et 1972, plusieurs mesures législatives et réglementaires inspirées largement du rapport de la Commission furent adoptées qui fixèrent les objectifs et le cadre de développement du nouveau système de santé. Dans les lignes qui vont suivre, je vais décrire brièvement les caractéristiques des différents éléments de ce système et l'évolution qu'ils ont connue jusqu'à maintenant.

Il y a création d'un nouveau Ministère des Affaires sociales regroupant la santé, les services sociaux et la sécurité du revenu selon le schéma décrit plus haut. Entre le ministère et les établissements on ajoute un nouveau palier de décision, les conseils régionaux de la santé et des services sociaux. Au début, ils n'ont que des pouvoirs de recommandation sur la planification et l'allocation des ressources mais, par la suite, ils reçoivent des mandats de plus en plus étendus qui constituent une véritable délégation de pouvoir, une véritable décentralisation de la part du Ministère des Affaires sociales. Ces organismes au nombre de douze, sont régis par des conseils d'administration formés de représentants des établissements du réseau, des municipalités, des groupes socio-économiques et des universités, le cas échéant. Ils n'ont pas de pouvoirs de taxation propres. Leur mandat et leur financement venant directement du Ministère des Affaires sociales, c'est à cet organisme qu'ils ont à rendre compte.

La distribution des soins et des services sociaux est assurée par un réseau d'établissements publics autonomes mais liés les uns aux autres par des contrats de services et par une représentation au sein des conseils d'administration. Les hôpitaux sont classifiés en centres hospitaliers pour soins de courte et de longue durée et les premiers en centre ultra-spécialisés, spécialisés et généraux. Leur conseil d'administration sont formés de représentants des usagers du centre, des professionnels et des employés généraux, des groupes socio-économiques du territoire desservi, d'autres établissements avec qui ils ont des contrats de services et des universités pour ceux qui sont affiliés pour fin d'enseignement. Ils jouissent de l'autonomie de gestion pour nommer les médecins, engager le personnel et administrer les services à l'intérieur d'un budget global qu'ils négocient chaque année avec le Ministère des Affaires sociales. Les salaires cependant sont fixés par les conventions collectives négociées à l'échelle provinciale. Pour les achats de nouveau matériel technique, les transformations, les agrandissements et les développements de services ils sont soumis à la double tutelle des conseils régionaux et du Ministère.

Les services de prévention, administrés antérieurement par les services du Ministère de la Santé et les grandes municipalités sont confiés à trente-deux centres hospitaliers qui se voient ainsi investis d'une mission nouvelle, inédite. Ces nouveaux services portent le nom de Département de santé communautaire. Ils desservent des territoires pouvant varier de 100,000 à 350,000 de population. D'après la loi, ils ont comme principale responsabilité de faire l'étude des besoins de la population du territoire desservi dans les domaines de la santé physique et mentale et d'élaborer, de mettre en oeuvre et d'évaluer des programmes en collaboration avec les établissements du territoire. Pour ce faire, ils jouissent d'un budget dit protégé en sus du budget global de l'établissement dont ils font partie. Dans les premières années qui ont suivi leur création, ils se sont consacrés à la tâche difficile d'intégrer les parties éparses de l'ancien réseau de santé publique, unités de médecine préventive, services de santé municipaux, service de santé scolaire, dispensaires antituberculeux, services de soins à domicile. Ces services ont été regroupés par programmes, rendant ainsi leur fonctionnement plus souple. Ils

ont aussi créé de nouveaux programmes notamment dans les domaines des soins aux personnes âgées et de la santé au travail. Ils ont pu bénéficier de budgets sensiblement augmentés par rapport aux services qu'ils remplaçaient. Ils se sont pourvus d'un personnel d'encadrement et de consultation, nombreux, mieux formé et plus diversifié. Ils ont pu aussi établir des liens plus faciles avec les médecins, les infirmières et les administrateurs hospitaliers. Ils ont eu le concours des départements de Médecine sociale et préventive des universités autant pour la formation de leurs personnels que pour le support à la planification et à la programmation. Le transfert d'un personnel d'exécution relativement nombreux surtout formé d'infirmières et la responsabilité de fournir eux-mêmes la plus grande partie des services de prévention n'ont pas été sans avoir une influence profonde sur l'orientation de la plupart d'entre eux jusqu'à aujourd'hui. C'est-à-dire que la mission de planification et d'évaluation a dû céder le pas à la mission d'exécution de programme, rendant d'autant plus difficile leur mission de coordination des activités de prévention des établissements du territoire avec lesquels, dans bien des cas, ils entraient en compétition. Il faut mentionner que cette fonction de dispensation de services devait être assumée dans le projet original par les nouveaux centres locaux de services communautaires (C.L.S.C.) dont nous parlerons plus loin. Ce fut soit pour desservir des territoires non encore pourvus de C.L.S.C., soit parce que plusieurs C.L.S.C. ne semblaient pas prêts ou disposés à assurer ces services que les Départements de santé communautaire sont entrés dans ce champ à la demande expresse du Ministère des Affaires sociales.

Une des grandes lacunes dans les structures du système de santé du Québec avait été l'absence d'établissements publics de soins de première ligne face aux cabinets et aux polycliniques privées. Pour combler cette lacune, on a créé des centres locaux de services communautaires (C.L.S.C.), établissements qui "en rendant accessibles les services de santé et les services sociaux, les services de prévention et les services d'action communautaire, visent dans une approche globale et multidisciplinaire, à renover l'état de santé d'une population donnée, à améliorer le milieu dans une perspective de développement"(1) Selon le plan original, ces établissements devaient être la porte d'entrée du système. On en prévoyait originalement 200 pour l'ensemble du territoire, soit 1 pour 10 à 40,000 habitants environ. Ce nombre a été réduit à 160 récemment. A ce jour il y en a 100 en fonctionnement. Les services qu'ils fournissent comprennent: la médecine générale, la médecine du travail, les immunisations, les soins pré et post natals, le planning familial, l'éducation sanitaire, les services sociaux généraux comme l'assistance économico-sociale, l'assistance psycho-sociale et les pratiques d'actions préventives. S'ajoutent des programmes plus spécifiques qui font le plus appel à l'approche globale et communautaire et à la pluridisciplinarité comme le maintien à domicile des personnes âgées, l'aide aux handicapés physiques et mentaux, la garde des enfants, l'aide aux jeunes enfants et aux adolescents sur le plan santé et réadaptation sociale. Tous ne fournissent pas tous ces services. Le dosage varie selon qu'ils se situent en milieu rural, urbain, résidentiel ou défavorisé et selon le nombre et le type de professionnels dont ils disposent. En général, les services sont organisés sous

(1) La Fédération des C.L.S.C. du Québec, Avril 1977.

forme de programmes avec participation pluridisciplinaire. On cherche à aller au devant de certaines catégories de population. Le personnel professionnel est varié. Il se compose d'infirmières, de travailleurs sociaux individuels et de travailleurs sociaux communautaires, de médecins et d'autres professionnels. Les médecins qui y exercent sont relativement peu nombreux par rapport à l'ensemble du corps médical omnipraticien. Ils ne forment encore aujourd'hui, pas plus 6 à 7% de l'ensemble des effectifs. Ils sont, en général, salariés contrairement aux médecins libéraux. Compte tenu que le réseau est encore inachevé, du petit nombre de médecins qui exercent par rapport à l'ensemble du corps médical, du fait que les soins préventifs sont encore partagés avec les départements de santé communautaire, malgré une directive récente du ministère à l'effet de transférer les programmes et les personnels des départements de santé communautaire et des centres de services sociaux dans les C.L.S.C., ils ne constituent encore qu'une des portes d'entrée du système. En ce qui concerne les soins médicaux curatifs, ils ne jouent encore qu'un rôle très secondaire. Leurs conseils d'administration sont composés de 5 membres élus parmi les usagers du territoire desservi sur 11 membres au total. Les usagers sont aussi impliqués dans la conception et la réalisation des programmes. Ces établissements constituent avec les départements de santé communautaire, les éléments les plus originaux et les plus innovateurs du système. Ils ont attiré des administrateurs, des médecins, des infirmières et des travailleurs sociaux plus ouverts à de nouveaux modes de pratique de nature à réaliser les grands objectifs établis. Ce sont des institutions qui dérangent dans la mesure où leur approche est souvent l'antithèse de l'approche traditionnelle, bien qu'au cours des années, ils aient été amenés à aligner la plus grande partie de leurs pratiques sur les orientations décidées par le Ministère des Affaires sociales dont ils dépendent entièrement pour leur financement.

Les travailleurs sociaux, sauf ceux qui oeuvrent dans les centres locaux de services communautaires, sont regroupés dans des établissements régionaux qu'on appelle centres de services sociaux. Ces centres ont pour mission de fournir des services sociaux spécialisés, services sociaux hospitaliers, scolaires, conseillers matrimoniaux, placement des malades chroniques dans les établissements, etc. Ils fournissent aussi des services sociaux individuels de première ligne dans les territoires dépourvus de C.L.S.C. et même dans les territoires qui en sont pourvus, ce qui a constitué un sujet de conflit avec ces derniers.

Enfin, il existe un dernier type d'établissements connu sous le nom de centres d'accueil qui regroupent les garderies ou crèches, les centres de transition, les centres de réadaptation et les centres d'hébergement.

Tout système de santé s'appuie sur une structure d'enseignement qui permet de former les personnels spécialisés nécessaires et une structure de recherche. Au milieu des années 60, le Québec a procédé à une réforme en profondeur de son système d'éducation. Tout l'enseignement professionnel y compris les stages a été placé complètement sous la responsabilité des

établissements publics d'enseignement relevant du Ministère de l'éducation. Les établissements de soins qui répondent aux critères peuvent être reconnus comme milieux de stages, la surveillance de ces stages incombant cependant aux établissements d'enseignement. Même la formation post-doctorale dans les spécialités médicales relève des universités. On constate que depuis une douzaine d'années, il s'est fait un effort considérable dans les facultés de médecine et les autres facultés des sciences de la santé pour adapter l'enseignement aux réalités nouvelles. Par exemple, toutes les facultés de médecine ont créé des départements de médecine sociale et préventive et de médecine familiale. Les programmes des études médicales contiennent des enseignements et des stages en sciences du comportement et dans les différentes branches de la médecine sociale comme la biostatistique, l'épidémiologie, l'analyse des systèmes de santé, l'économie de la santé, la sociologie, l'anthropologie, l'hygiène du milieu et la santé au travail. On expose tôt les étudiants à des modes d'exercice autres que les stages hospitaliers traditionnels. Les facultés de médecine et de sciences infirmières offrent des programmes de formation de 2e cycle et même de 3e cycle en santé communautaire ou en épidémiologie. Au moment de la création des départements de santé communautaire, les 4 facultés de médecine du Québec se sont regroupées pour offrir un programme de formation en cours d'emploi aux cadres de ces nouveaux départements. Cette intervention des universités a concouru non seulement à augmenter les connaissances mais aussi à créer une unité de pensée et d'action.

La recherche fondamentale dans les sciences de la santé se fait presque entièrement dans les universités ou les hôpitaux et les instituts affiliés aux universités. La recherche épidémiologique et opérationnelle se fait dans les universités et les centres affiliés mais aussi au Ministère des Affaires sociales. Les principales sources de financement sont les subventions du gouvernement fédéral et à un moindre degré du Conseil de la recherche en santé du Québec. De plus en plus, les organismes subventionnaires fixent des thèmes de recherche en fonction de leurs priorités.

Il est important de mentionner qu'en 1971, la Direction générale de l'hygiène du milieu est transférée du Ministère de la Santé au Ministère des Affaires municipales et de l'environnement. C'est dire que les responsables de la santé publique, principalement le Ministère des Affaires sociales et les départements de santé communautaire doivent dorénavant passer par le canal formaliste et assez lourd des relations inter-ministérielles pour toutes les questions relatives à l'hygiène du milieu. A notre connaissance, il n'y a pas eu à ce jour d'analyse des conséquences de ce transfert au plan fonctionnel.

Les grandes réformes des années 60 et du début des années 70 n'ont pas touché au secteur névralgique de la santé au travail. Ce n'est qu'en 1979 que le gouvernement adopta une législation modifiant profondément les structures sanitaires du milieu du travail. Un nouvel organisme est créé, la Commission

de la Santé et de la Sécurité au Travail (CSST) chargé de la mise en oeuvre de la politique de santé et de sécurité au travail. Son conseil d'administration se compose de onze membres nommés par le Ministère du Travail et de la Main d'oeuvre, Ministère de tutelle, dont cinq parmi une liste fournie par les syndicats, cinq parmi une liste fournie par les associations d'employeurs et d'un président directeur général. Le financement est assuré en très grande partie par des cotisations aux entreprises. Le mandat de la Commission couvre l'inspection, la réparation et la prévention. La Commission élabore des programmes cadres de santé en milieu de travail, mais il revient aux départements de santé communautaire d'en assurer l'application dans leurs territoires après avoir passé contrats avec la CSST pour le financement de ces programmes. Les soins individuels aux travailleurs peuvent se donner dans les hôpitaux, les C.L.S.C., les entreprises et les cabinets de médecins. Le médecin du travail de l'entreprise est choisi et nommé par le comité paritaire de l'entreprise à même une liste fournie par le Département de santé communautaire. Il est rémunéré par la Régie d'Assurance-Maladie du Québec et non par l'employeur et fait partie du personnel scientifique du Département de santé communautaire du territoire. Il doit élaborer le programme de prévention de l'entreprise lequel doit être approuvé par le Comité paritaire.

Nous venons de donner un aperçu des structures du système de santé québécois. Qu'en est-il de sa dynamique, de son évolution depuis 1970? La politique de rationalisation et de contrôle des dépenses entreprises au niveau fédéral dont nous avons parlé plus haut a eu son écho au niveau provincial. Nous avons mentionné que le pourcentage du produit national brut affecté aux dépenses de santé n'a à peu près pas changé de 1970 à 1981. Le même phénomène a été observé au niveau provincial. Deux mesures, l'une au niveau fédéral, l'autre au niveau provincial expliquent en grande partie cette performance économique. Les provinces, notamment le Québec, ont procédé durant cette période à une réduction du taux des lits de soins de courte durée au profit des lits pour convalescents et chroniques et des ressources communautaires. Au Québec ce taux est passé de 5.24 pour 1000 en 1974 à 4.34 en 1977. Etant donné l'importance des dépenses au titre des lits de soins de courte durée dans le système de santé, on peut juger de l'impact d'une telle mesure. Parallèlement à cela, le gouvernement fédéral a imposé au milieu des annés 70 un gel des salaires qui s'est ajouté à l'impact de la réduction des lits.

Une politique de rationalisation conduit logiquement à un resserrement des contrôles sur les établissements et les professions autonomes. C'est ce qui s'est produit au Québec depuis 1970. Les hôpitaux ont perdu graduellement une partie de leur autonomie. Les médecins se sont fait imposer des conventions collectives de plus en plus contraignantes au plan financier et au plan du contrôle de l'acte professionnel. Avec le glissement du centre des décisions importantes des conseils d'administration des hôpitaux vers le Ministère des Affaires sociales et les conseils régionaux, les médecins hospitaliers tout en réussissant à maintenir assez bien leur position comme groupe de pression à l'intérieur des établissements, se sont ressentis de ce déplacement.

Dans un autre ordre d'idée, la participation des citoyens usagers aux différents paliers de décision a été un échec de l'avis de la plupart des observateurs. Les professionnels et les bureaucrates demeurent les deux groupes les plus puissants à l'intérieur du système, ne laissant à peu près pas de place à une troisième force.

Nous venons de passer rapidement en revue l'impact des réformes sur les coûts, l'utilisation de certaines ressources et les rapports de force. Quelle est l'influence que ces changements ont pu avoir sur la santé de la population? Il n'entre pas dans le cadre de ce travail d'introduction au système de santé de répondre à une telle question, si tant est qu'on puisse réellement y répondre compte tenu de l'influence des facteurs socioéconomiques et des conditons du milieu sur la santé.

Georges Desrosiers

Avril 1983

BIBLIOGRAPHIE

** CRISE DE L'ETAT-PROVIDENCE:

1. L'ETAT PROTECTEUR EN CRISE
Rapport de la conférence sur les politiques sociales dans les années '80. Paris, OCDE, 1981.

2. LA CRISE DE L'ETAT-PROVIDENCE
Pierre Rosanvallon. Paris, Seuil, 1981

3. DEMAIN LE CAPITALISME (Ch. 4: L'Etat-providence démystifié)
Henri Lepage. Paris, Librairie Générale Française, 1978.

4. L'ECONOMISTE ET LA CHOSE PUBLIQUE
Jean-Luc Migué. Québec, Les Presses de l'Univ. du Québec, 1979.

5. LA SANTE RATIONNEE
J. de Kervasdoué et coll.. Paris, Ed. Economica, 1981.

** REFORME DES SERVICES DE SANTE AU QUEBEC (1970-1980)

6. RAPPORT DE LA COMMISSION D'ENQUETE SUR LA SANTE ET LE BIEN ETRE SOCIAL
Québec, Editeur officiel du Québec, 1970.

7. ON NE CHANGE PAS LA SOCIETE PAR DECRET
Michel Crozier. Paris, Grasset (Coll. Pluriel), 1979.

8. DU PAIN ET DES SERVICES: LA REFORME DE LA SANTE ET DES SERVICES SOCIAUX AU QUEBEC
Frédéric Lesemann. Montréal, Editions Coop. Albert Saint-Martin, 1981.

** VERS UN NOUVEL ART DE VIVRE?

9. NOUVELLE PERSPECTIVE DE LA SANTE DES CANADIENS
Marc Lalonde. Ottawa, Ministère de la santé nationale et du bien-être social, 1974.

10. NEMESIS MEDICALE
Ivan Illich. Paris, Seuil, 1975.

11. LA LUMIERE MEDICALE: LES ILLUSIONS DE LA PREVENTION
Norbert Bensaid. Paris, Editions du Seuil, 1981.

Département d'Administration de la Santé
Université de Montréal, 1983

Extrait de: La Revue du praticien
Paris, France, 19 rue Hautefeuille
Tome XXVIII, No 34, 15 juin 1978.

Tribune libre

La réforme sanitaire au Québec

Révolution ou illusion

par Eric Maquet et Laurent Mignot

Toute juvénile qu'elle soit, la réforme du système de santé au Québec a déjà fait couler beaucoup d'encre. Pour les uns, c'est une révolution, pour les autres, c'est une illusion. En tout cas, elle existe : il suffit, en France, d'écouter les espérances de la jeune génération des professionnels de santé pour trouver, au sein des nombreux projets théoriques, deux exemples concrets : Grenoble et le Québec. C'est sûrement la force actuelle du système québécois que d'offrir un réel champ d'expérience aux concepts « fumeux » de santé communautaire, d'équipe pluridisciplinaire, de contrôle des usagers, de salariat. Cette réforme sanitaire est la dernière née de ce que l'on a appelé la « révolution tranquille ». En effet, jusqu'en 1960, la société québécoise est dominée par l'élite anglophone et a un caractère rural, clérical et paternaliste. A cette époque, le parti libéral est porté au pouvoir sur le thème « il est temps que ça change ». Le Québec va alors devenir le pays de l'imagination et du changement. Dès les années 1960, ce sont les réformes de l'éducation et de l'économie, à partir de 1970 celle de la santé.

Cette réforme sanitaire sera annoncée et inaugurée par des groupes de travailleurs organisés en centres de santé autogérés, les « cliniques populaires » ; elles resteront très marginales, mais leur impact idéologique sera considérable à une époque où la commission provinciale menée par Castongay est chargée de définir les nouvelles règles du jeu en matière sanitaire ; celle-ci propose une rationnalisation visant surtout à baisser les coûts de santé en s'appuyant sur trois principes novateurs. En premier lieu, elle insiste sur l'urgence d'une médecine globale ; la maladie étant souvent tributaire de facteurs sociaux tels la pauvreté ou l'environnement, une intervention globale à la fois palliative et préventive paraît indispensable pour juguler les facteurs de morbidité. L'action sera donc menée par une équipe pluridisciplinaire de professionnels responsables de la santé d'une population. Elle est centrée sur la personne dans sa totalité plutôt qu'uniquement sur sa maladie spécifique et cherche à agir sur les conditions de vie et de travail qui conduisent à la maladie. La priorité est donc donnée aux interventions de premières lignes regroupant services préventifs, curatifs et de réhabilitation, tant au niveau social et mental que physique, seul moyen de régler les problèmes au moindre coût et de façon cohérente. Le deuxième objectif est de susciter la participation des usagers dans la mise en place des programmes sanitaires les concernant, en ayant recours à l'information et à la constitution d'association de quartier. Enfin, le rapport affirme la nécessité d'un rééquilibrage des responsabilités professionnelles face à la communauté, tenant compte des compétences et non plus de statuts hiérarchiques. De tels principes visant à long terme à réduire la consommation médicale et le champ d'intervention hospitalier, faisant appel à la partici-

(suite page 2686)

La réforme sanitaire au Québec. Révolution ou illusion

(suite de la page 2663)

pation et à la mise en place des pouvoirs locaux, pouvaient paraître anachroniques dans le contexte libéral capitaliste nord-américain. Quoi qu'il en soit, Castongay est nommé ministre de la Santé en 1970 et proclame la gratuité des soins, la réduction des budgets hospitaliers, la création d'un ministère des Affaires sociales, la formation, au sein des hôpitaux, de départements de santé communautaire chargés d'y faire naître des préoccupations sociales et préventives, la mise en place d'unités de soins de première ligne (les centres locaux de service communautaire : C.L.S.C.) où médecins, infirmiers, travailleurs sociaux et organisateurs communautaires doivent ensemble prendre en charge la totalité des besoins sociaux et sanitaires d'une population définie et dont la gestion est basée sur le principe de la participation.

Tout semblait en place pour un changement radical et pourtant les insuccès sont fréquents, en particulier en ce qui concerne les C.L.S.C. qui, a priori, devaient être le fer de lance de la réforme, capable de modifier les comportements. 200 C.L.S.C. étaient prévus en 1972, 80 ont vu le jour, perdus au milieu des 400 nouvelles polycliniques libérales constituées dans le même temps, d'autant plus facilement que la gratuité des soins aboutissait à une augmentation considérable de la consommation médicale. De plus, du fait de leur autonomie, ces centres prirent une coloration originale, parfois opposée à l'esprit de la réforme et qu'il est possible de regrouper en quatre types. Les premiers, implantés en milieu urbain, pris en main par des groupes très politisés, sont devenus des outils de lutte politique presqu'immédiatement en grève et ayant finalement peu d'impact sur la population. Les seconds, ayant mal assimilé les vues de la réforme, se sont transformés en distributeur de service de soin, ne se distingant des polycliniques que par le salariat et l'absence de motivation. Les troisièmes, englués dans une idéologie trop rigide, n'en finissent pas de sortir de la phase d'implantation devenant de véritables centres de recherche théorique sur la santé. Un certain nombre, heureusement, parviennent à être réellement des centres de santé : les trois modules, médical, social et communautaire travaillent de concert, le partage du pouvoir est redéfini et malgré des imperfections, on peut voir enfin le fonctionnement d'une équipe pluridisciplinaire, la définition d'un pouvoir autre que médical, la notion de prise en charge globale des problèmes de santé, la signification et l'intérêt du contrôle par les usagers. Ces expériences sont d'un intérêt considérable, car elles ouvrent des portes à d'autres pratiques jusqu'ici imaginées mais jamais concrétisées. Mais pourquoi donc restent-elles minoritaires alors qu'elles incarnent parfaitement l'esprit de la réforme ? Ceci tient probablement à plusieurs faits. D'une part, l'hostilité médicale s'est manifestée de suite et s'est concrétisée dans la mise en place de polycliniques concurrentielles, d'autre part, en cherchant à politiser au maximum les débats, en discutant de façon excessive la redistribution des pouvoirs, en s'attachant à respecter aveuglément l'idéologie de la réforme, ou au contraire, en s'installant dans un statut de fonctionnaire, les professionnels ont souvent perturbé le fonctionnement des centres et contribué à leur discrédit. De plus, le rôle des usagers est resté assez modeste au niveau des C.L.S.C. : il ne suffit pas de mettre quelques concitoyens dans un conseil d'administration pour parler de participation ; en effet, préoccupés des besoins curratifs, les québécois n'étaient pas prêts pour une réflexion élaborée sur la santé et ont mal saisi toutes les ouvertures que contenait la réforme. Enfin et surtout, le pouvoir politique n'a jamais affiché une attitude cohérente vis-à-vis d'une réforme qu'il a pourtant proposée, mais dont il semble ne plus souhaiter le plein succès : l'implantation des centres a été limitée, aucune directive n'est venue jalonnée leur croissance et la médecine libérale fut plutôt favorisée. Pourquoi ces hésitations actuelles ? N'y avait-il pas une certaine naïveté à proposer une telle restructuration dans un tel contexte socio-politique ; il est en effet aisé de décider la gratuité des soins, il est beaucoup plus difficile d'introduire, dans un contexte capitaliste, des orientations visant à sortir les problèmes de santé de l'orbite de la production et de la consommation. Néanmoins, l'expérience québécoise est irremplaçable dans sa recherche de nouvelles pratiques médico-sociales ; elle en montre aussi les difficultés et souligne la nécessité d'une volonté politique et l'intérêt de réalisations militantes, ponctuelles et convaincantes telles que celles déjà observées en France.

E. Maquet et L. Mignot
Chefs de clinique *

(*) Chargés de mission au Québec en septembre 1977 par l'intermédiaire de l'Office Franco-Québecois, 5, rue de Logelbach, 75017 Paris.

41A

dossier D

LE SYSTÈME DES HONORAIRES MODULÉS

RÉSUMÉ-SYNTHÈSE ET RECOMMANDATIONS

L'ass. maladie a 100 ans
(Bismarck 1883)

MARS 1980

La méthodologie adoptée pour évaluer le système de soins a consisté, en utilisant les résultats des études disponibles et les statistiques existantes, à apprécier l'accessibilité, la qualité, la productivité et la satisfaction des professionnels pour chaque secteur, ainsi que l'administrabilité et la satisfaction de la population pour l'ensemble des secteurs.

1. Accessibilité

Le degré d'accessibilité au système de soins a été évalué en analysant successivement l'accessibilité provinciale, l'accessibilité régionale, l'accessibilité sociale et l'accessibilité organisationnelle.

L'accessibilité provinciale concerne la relation entre les effectifs existants, leur évolution récente et leur prévision, et les effectifs requis pour offrir des services de qualité à toute la population de la province.

Elle est généralement bonne, à l'exception du secteur dentaire où il continuera à exister une certaine pénurie durant les dix prochaines années.

La croissance prévue des effectifs sera rapide dans les quatre professions et la proportion des femmes augmentera. Chez les médecins, la proportion des omnipraticiens continuera à croître.

L'accessibilité régionale concerne la répartition géographique des effectifs et l'évolution de cette répartition en fonction des besoins des populations.

Le nombre de professionnels par rapport à la population a augmenté dans toutes les régions; toutefois cette amélioration de la répartition géographique découle plus de l'abondance des effectifs dans les régions métropolitaines que de la prise en considération des besoins des populations. La répartition des médecins spécialistes et surtout des dentistes reste problématique.

L'accessibilité sociale concerne l'existence et l'importance des divers obstacles (psychologiques, sociaux, économiques, culturels) que la population peut rencontrer dans l'utilisation des services de santé.

D'une façon générale, l'introduction du régime d'assurance-maladie a permis une amélioration de l'accessibilité sociale en réduisant les barrières financières à l'accès aux services. Toutefois, la couverture des services dentaires, optométriques et pharmaceutiques reste incomplète. Et les taux de participation pour les populations couvertes sont trop faibles dans le cas des services dentaires et optométriques.

De plus, l'accès aux services optométriques et pharmaceutiques est limité par l'existence d'une fonction commerciale associée à l'exercice professionnel (prix des prothèses visuelles, variation du prix des médicaments). Enfin, des variations importantes dans l'utilisation des services dentaires existent entre les classes sociales.

L'accessibilité organisationnelle concerne la disponibilité des ressources existantes.

Globalement l'accès aux ressources existantes est apparu acceptable. Toutefois, il existe des problèmes d'accessibilité à certains

services dentaires (les services en établissement pour les populations particulières, l'urgence, etc.). Les services des optométristes ne sont pas disponibles en établissement. L'accès aux services pharmaceutiques dans plusieurs régions est limité en raison du petit nombre de pharmaciens par pharmacie.

2. Qualité

La qualité des services a été appréciée sous trois aspects: la globalité, la continuité, la qualité technique.

Pour évaluer la globalité, il faut voir si le système de soins favorise la dispensation de la gamme complète des services (prévention, cure, réadaptation) à l'endroit le plus approprié (cabinet, domicile, établissement). Il faut examiner aussi dans quelle mesure il existe une intégration des différents intervenants dans le système de soins et une orientation des services vers la prévention auprès d'individus et de populations considérés dans leur environnement économique et social.

Dans le système actuel, la globalité apparaît particulièrement insuffisante. De façon générale, les activités qui ne sont pas reliées directement aux soins aux patients, tels la prévention, les tâches administratives, l'enseignement, ne sont pas favorisées. Dans plusieurs cas, elles ne sont pas rémunérées, ou si elles le sont, c'est selon des modes différents de ceux qui s'appliquent aux soins aux malades. Ce passage d'un mode à l'autre impose des contraintes lors de la combinaison et de l'harmonisation des différentes activités.

Ainsi, dans le cas des services médicaux, la prévention auprès des populations particulières est peu encouragée. Le temps voué au contact humain est souvent déprécié par rapport aux dimensions techniques de l'exercice de la médecine.

La tarification associée aux actes préventifs ayant été abaissée pour les services dentaires, il y aura probablement une diminution du nombre de ces actes. Cette diminution, associée au faible développement des programmes de prévention dans le secteur public et au moratoire sur la fluoruration de l'eau, accentuera la prépondérance que le système accorde aux soins curatifs.

En optométrie, l'absence de couverture des services de rééducation visuelle limite l'accessibilité sociale et, ce faisant, la globalité. En outre, les revenus provenant de la distribution de prothèses visuelles incitent à limiter les activités aux services reliés à la distribution de prothèses.

Une situation semblable existe dans le domaine des services pharmaceutiques. La distribution de médicaments est encouragée alors que la communication d'informations aux patients et aux prescripteurs, l'analyse des profils médicamenteux, la prévention sont négligées.

L'objectif de continuité est atteint lorsqu'il y a une interaction cohérente entre les différents professionnels pour que la prise en charge de l'individu (prévention, cure, réadaptation) soit assurée à tout moment et quel que soit le lieu de l'intervention.

Dans le système actuel, cet objectif n'est pas atteint. En effet, la pratique individualiste qui caractérise la rémunération à

l'acte ne favorise pas la contribution à des équipes pluridisciplinaires, ni l'intégration des activités dans un système cohérent où les différentes catégories de services seraient harmonisées par rapport à des objectifs relatifs à la continuité.

La communication entre professionnels est déficiente, en particulier dans le domaine oculo-visuel. Le même phénomène est observé dans le domaine pharmaceutique, puisque les contacts médecins-pharmaciens sont trop peu nombreux.

La qualité technique des services mesure le dègré de concordance entre les services fournis et les normes de bonne pratique.

Bien qu'il existe au Québec un certain contrôle de l'activité professionnelle, de nombreuses interrogations subsistent.

Le maintien des compétences n'est pas favorisé. L'encadrement de l'activité professionnelle existe principalement dans le milieu hospitalier; la majorité des dentistes, les pharmaciens d'officine, la totalité des optométristes, et de nombreux omnipraticiens sont toutefois absents de ce milieu.

Le conflit d'intérêt entre une activité professionnelle et une activité commerciale réduit la qualité. Le chevauchement des champs de pratique (optométrie, ophtalmologie et optique d'ordonnances, médecine dentaire et hygiène dentaire) éloigne les professionnels de leur domaine spécifique de formation.

Enfin et surtout, le paiement à l'acte, en mettant l'accent sur la quantité, pousse les professionnels à ne pas respecter suffisamment les normes de qualité.

3. Productivité

La productivité est la relation entre les ressources existantes et le volume des services offerts.

L'utilisation accrue des examens complets plus rémunérateurs que les examens sommaires, la facturation d'actes mineurs (par exemple exérèse de cérumen, funduscopie, pansements de plus de vingt centimètres carrés) et l'élargissement des indications de certains actes diagnostiques ou thérapeutiques (par exemple électrocardiogrammes, scopies digestives, radiographies dentaires, étude simple de la vision des couleurs) entraînent un accroissement artificiel du nombre des actes et du coût des services, c'est-à-dire une pseudo-productivité. Celle-ci témoigne de la relation entre le choix des actes facturés et le niveau de revenu visé, et elle s'explique par le conflit d'intérêt entre la liberté thérapeutique et une rémunération liée au nombre et au type de services dispensés dans le cadre d'un régime d'assurance universelle.

Les coûts de cette pseudo-productivité débordent les coûts de l'assurance-maladie lorsque les services médicaux dispensés requièrent des ressources hospitalières. De même le conflit d'intérêt va parfois à l'encontre d'une utilisation rationnelle des ressources en retardant l'introduction de technologies automatisées ou en restreignant le regroupement de certains services très spécialisés. Enfin, ce même conflit d'intérêt ne garantit pas la productivité optimale de certains départements hospitaliers lorsque le médecin a la possibilité d'orienter sélectivement des patients vers sa clinique privée pour des examens ou des traitements.

La faible utilisation du personnel auxiliaire diminue la productivité des médecins et des dentistes, mais elle représente un

comportement économique cohérent avec les plafonnements des revenus bruts, les restrictions au paiement d'actes qui n'ont pas été posés par le dispensateur lui-même, et le fait que le tarif des actes est fixe et qu'il ne peut donc être ajusté selon les coûts de fonctionnement des cabinets.

Le chevauchement des champs de pratique dans les secteurs des services dentaires et des services oculo-visuels a pour effet une utilisation non optimale des compétences et une productivité inférieure à ce qu'elle pourrait être.

4. La satisfaction des professionnels

Les professionnels sont généralement satisfaits du régime d'assurance-maladie et ils jouissent d'une liberté étendue dans leur pratique.

Leur rémunération place les médecins québécois aux premiers rangs parmi les provinces canadiennes, et les situe favorablement par rapport aux professionnels à leur compte. Leurs heures de travail diminuent et la durée de leurs vacances augmente.

Toutefois, si c'est le degré de concordance entre les attentes des professionnels à l'égard du système de soins et les situations vécues qui définit les satisfaction, il y a des aspects de la situation actuelle qui sont source de malaise. Le caractère centralisé du régime conduit à des formes de contrôle des paiements irritantes pour certains. Les optométristes souhaitent une meilleure intégration dans le réseau des établissements. Pour les pharmaciens, le contexte commercial de la pratique et le décalage entre cette dernière et leur formation sont sources de frustation.

Malgré tout, la participation au régime est très forte et les demandes d'admissions dans les facultés et écoles demeurent très nombreuses.

5. Administrabilité

L'administrabilité concerne le niveau de complexité de la gestion du système de santé, et la capacité du système à évaluer ses performances, à planifier son développement et à contrôler la croissance de ses coûts.

Dans le cadre de la rémunération à l'acte, l'application de mécanismes de contrôle de la pratique individuelle par une organisation centrale est relativement inefficace principalement parce qu'il est souvent difficile de distinguer si les différences dans le nombre et le type de services dispensés sont reliées à des variations du jugement clinique ou au conflit d'intérêt inhérent à la rémunération à l'acte. En outre, les modes de paiement à vacation et à salaire ou à honoraires fixes ont été mis en place en l'absence des structures administratives habituellement requises et cette absence n'a pas été compensée ou atténuée par la transmission d'un minimum de renseignements sur l'activité rémunérée.

Le contrôle de la croissance des coûts est plus tributaire du rationnement budgétaire global appliqué aux établissements que d'une réallocation sélective et rationnelle des ressources selon des orientations et des objectifs d'ensemble qui tiendraient compte des populations à desservir.

Le système de rémunération à l'acte ne fonctionne pas de façon intégrée avec l'ensemble du système de soins. La négociation s'est

substituée au processus de planification et de gestion. De toute façon, il est impossible qu'une mosaïque de codes d'actes et de stimulants tarifaires spécifiques puisse traduire adéquatement les orientations et les objectifs du système de soins.

Le morcellement de l'activité professionnelle clinique associé à la rémunération à l'acte fragmente la responsabilité du professionnel en actes spécifiques, et empêche une responsabilité globale des professionnels face aux besoins d'une population. De même, la tarification à l'acte incite à une pratique individuelle indépendante des ressources matérielles et humaines des établissements; par contre, ces mêmes ressources sont nécessaires aux professionnels pour exercer leurs fonctions et sont paradoxalement soumises à des restrictions budgétaires, créant ainsi des tensions.

6. La satisfaction de la population

Elle n'est pas facile à évaluer car le consommateur est peu amené à se servir de son jugement critique lorsqu'il doit utiliser les services de santé. De plus l'insatisfaction générée par cette utilisation n'est pas toujours canalisée vers les organismes ayant pour fonction de recevoir les critiques et les plaintes et susceptibles de fournir des statistiques à ce sujet.

De façon générale, l'accessibilité aux services et la qualité des soins sont perçues comme acceptables. La non-gratuité des soins dentaires et les honoraires demandés pour ces soins sont cependant sources de plaintes. Le coût des prothèses visuelles et des médicaments amène également des commentaires négatifs de la part des consommateurs.

Enfin une préoccupation grandissante de la population envers sa

santé entraîne des interrogations et des contestations collectives face au système de soins.

IV. Causes des principales faiblesses et hypothèses de solution

Les faiblesses observées sont en bonne partie reliées, d'une part, aux modes de rémunération en vigueur, c'est-à-dire principalement au paiement à l'acte dans un régime public d'assurance-maladie et, d'autre part, à certaines caractéristiques de l'organisation des services professionnels.

Faiblesses reliées au paiement à l'acte

Le paiement à l'acte avec tiers payant a des caractéristiques qui entravent l'atteinte de plusieurs objectifs.

Ce mode de rémunération ne permet pas de rémunérer la totalité de l'activité professionnelle. Cette dernière est un ensemble complexe d'interactions entre le professionnel, les patients et l'environnement dans lequel elle s'exerce. Cette activité ne peut être entièrement définie par des actes aussi nombreux et précis soient-ils; ainsi la prévention auprès de populations, l'administration, le suivi des patients, la psychiatrie dans une large mesure et d'une façon générale toutes les activités qui dépendent plus de l'interaction entre le professionnel et son patient que de la nature de l'acte posé, sont difficiles ou impossibles à rémunérer à l'acte.

C'est pour pallier ces faiblesses que le paiement à l'acte a été complété par du paiement au temps: salaire, vacations, actes avec une référence explicite au temps (psychothérapie). Cette juxtaposition de modes de rémunération de nature différente qui complique l'admi-

IVAN ILLICH ET LE RAPPORT CASTONGUAY : DEUX VISIONS DE LA SANTÉ ET DE LA MÉDECINE

Gilbert BLAIN [1]

CINQ ANS APRÈS

Il s'est déjà écoulé cinq ans depuis la parution, en septembre 1970, du rapport Castonguay sur « La Santé » (Volume IV) et près de quatre ans depuis la sanction, en décembre 1971, de la loi sur les services de santé et les services sociaux du Québec. En passant du rapport à la loi, certaines recommandations ont été modifiées pour diverses raisons, mais sans remettre en question les principaux objectifs de la réforme proposée.

Depuis 1970, un nombre impressionnant de groupements ou d'individus ont formulé, dans des mémoires soumis à des commissions parlementaires ou dans des articles publiés dans des revues, des commentaires louangeurs ou critiques sur les conclusions du rapport ou les dispositions de la loi, sur le comportement du gouvernement ou sur celui des professionnels de la santé ou des administrateurs.

Tous ces commentaires avaient un point en commun: ils portaient tous sur les modalités de la réforme et non sur ses objectifs, sur les caractéristiques du nouveau système recommandé et non sur l'idéologie dont il découlait. À notre connaissance, personne n'a contesté la valeur de cette idéologie ou le bienfondé de ces objectifs. Tout le monde a souscrit d'emblée au but fixé, quitte à discuter sur les chemins proposés pour s'y rendre.

Mais voici que vient de paraître un petit livre qui touche à l'essentiel. L'auteur de ce livre n'est pas québécois. On peut même douter que M. Ivan Illich connaisse le rapport Castonguay sur « La Santé ». Néanmoins son récent ouvrage, intitulé « Némésis médicale », se présente comme une véritable contrepartie du rapport sur les plus fondamentaux de tous les points: l'idéologie sous-jacente et les objectifs proposés.

Pour démontrer qu'Ivan Illich s'inscrit en faux contre les principaux objectifs de la réforme des services de santé entreprise au Québec, il est sans doute opportun de dresser une liste sommaire de ces objectifs puis de rechercher, dans le texte de son livre, les opinions d'Illich sur chacun.

[1] Directeur, Département d'administration de la santé, Université de Montréal.

OBJECTIFS DE LA RÉFORME QUÉBÉCOISE

1. — *Médecine globale*

L'objectif premier de la réforme des services de santé proposée dans le rapport Castonguay est la réalisation du concept de la médecine globale [1]. La médecine globale, d'après le rapport, implique la « prise en charge » d'une population par une équipe multidisciplinaire [2].

De cette prise en charge, il découle qu'on doit « chercher à connaître et à satisfaire les besoins de santé non exprimés par la population » [3] et même « prévoir des moyens de rejoindre l'individu qui a besoin de soins mais néglige la consultation » [4].

2. — *Accès universel*

D'après le rapport, « il faut sans retard établir au Québec un régime de la santé accessible à tous » [5]. Bien plus, « on doit concevoir le régime de la santé de façon à rejoindre et prendre en charge tous les Québécois, quels que soient leur lieu de résidence et leurs conditions socio-économiques » [6].

3. — *Continuité des soins*

La médecine globale, selon le rapport, présente trois caractéristiques: soins complets, continus et personnels [7].

Le sens de ces trois mots est précisé dans le rapport. Quant à la continuité des soins, l'explication est la suivante: « L'individu s'attend à ce qu'on s'intéresse à lui de près tant que subsiste un problème de santé. Le personnel donne tous les soins, au domicile du malade ou à l'endroit approprié, selon son état et la nature de ses besoins... [8] ... La continuité des soins tout au long de la vie, peu importe l'endroit où il se trouve... apparaît comme le fondement nécessaire à une médecine globale » [9].

4. — *Accent sur la prévention*

Le rapport Castonguay sur « La Santé » ne parlait pas de médecine préventive. C'est que le concept de médecine globale, dans l'esprit de ses promoteurs, intègre les activités préventives et curatives en un seul système.

Néanmoins l'importance jusqu'ici négligée des mesures préventives est clairement soulignée dans le rapport où l'on trouve, par exemple, la phrase suivante: « L'accent doit maintenant porter sur la prévention de même que sur la surveillance de l'état de santé de la population » [10].

PUIS VINT IVAN ILLICH

En juin 1974, dans un article publié par la revue Esprit, Ivan Illich, ce contestataire de l'industrialisation à outrance de la société, dévoilait ses nouvelles batteries: il allait maintenant s'attaquer à la « médicalisation » de la vie, à « l'exploitation professionnelle du contribuable par l'institution médicale » qu'il considère comme souvent inutile et même dangereuse.

Dans cet article, Ivan Illich dénonce trois sortes d'iatrogénèses dont la dernière, l'iatrogénèse culturelle (qu'il appelle « structurelle »), lui semble la plus pernicieuse. Il écrit: « Déposséder les gens de leur douleur, du pouvoir de se définir comme sains ou malades et du droit de se laisser mourir constitue le troisième niveau de l'iatrogénèse... *Je propose de choisir ce dernier domaine d'impact négatif de la médecine sur le niveau de santé comme pivot de la controverse publique qu'il faut maintenant susciter* » [11].

Cette controverse publique, elle a débuté, à l'automne de 1974, dans les pages de la revue « Le Nouvel Observateur » où deux articles de Michel Bosquet ont déclenché une avalanche de réponses et de contre-réponses publiées pendant plusieurs semaines [12]. Cette controverse a maintenant pris toute son ampleur depuis la parution récente du livre d'Ivan Illich intitulé *Némésis médicale* [13].

Le titre un peu ésotérique de cet ouvrage est ainsi défini par l'auteur: « Némésis médicale est plus que la somme de toutes les fautes professionnelles, des négligences, du cynisme de caste, de l'injuste répartition des soins décrétée par les médecins, de l'invalidité par diktat médical. C'est plus que la dégradation malsaine des structures sociales par la surmédicalisation tentaculaire. C'est plus même que l'encouragement médical de l'impuissance de l'homme face à la douleur, la maladie et la mort. Némésis médicale, c'est l'auto-déréglage institutionnel de l'homme vers le cauchemar. C'est l'expropriation du vouloir-vivre de l'homme par un service d'entretien qui se charge de le maintenir en état de marche au bénéfice du système industriel » [14].

À partir de cette position de base, il sera facile de démontrer, avec quelques citations extraites de son livre, qu'Ivan Illich conteste radicalement les objectifs mêmes de la réforme des services de santé entreprise au Québec.

CONTESTATION DES OBJECTIFS DE LA RÉFORME QUÉBÉCOISE

1. — *Médecine globale*

L'émergence de la notion de traitement global est considérée par Illich comme un événement « sinistre » [15]. Pour lui, la notion de patient s'éclipse alors devant celle « d'homme administré » [16].

« La prise en charge institutionnelle de la population par le système médical, écrit-il, enlève progressivement au citoyen la maîtrise de la salubrité, dans le travail et dans le loisir, la nourriture et le repos, la politique et le milieu, elle constitue un facteur essentiel de l'inadaptation croissante de l'homme à son environnement » [17].

« Il faut avoir une vision profondément déformée de la réalité... pour faire croire au public... que la santé d'une société augmente au fur et à mesure que ses membres en viennent à dépendre de prothèses sous forme de médicaments, thérapeutiques, enfermements divers et contrôles préventifs » [17].

2. — *Accès universel*

On n'est pas surpris de voir Ivan Illich soutenir que l'assistance médicale universelle est un « mauvais rêve » [18] et qu'un moindre accès au système médical tel qu'il est actuellement... serait un bénéfice [19].

« Il semblerait raisonnable, affirme-t-il, de s'opposer à toute affectation de fonds publics ayant pour objectif d'égaliser les facilités d'accès à des entreprises médicales potentiellement dommageables... Cependant, dans une société libérale et médicalisée, le gouvernement ne peut affecter les ressources publiques qu'en accord avec la volonté générale, même si ces ressources n'ont aucune chance de répondre effectivement à ce qui est désiré. Dans ces circonstances, l'égalisation de l'accès aux types de services existants ne peut que renforcer le pouvoir des vendeurs d'organes de rechange ou celui des technocrates qui se fixent pour objectif la régulation universelle de la pression artérielle... L'enjeu est l'accès égal pour tous à la dépendance vis-à-vis des professionnels, aux illusions et aux dommages qu'ils causent » [19].

3. — *Continuité des soins*

La continuité des soins dans un système de médecine globale n'est rien d'autre, pour Ivan Illich, qu'un effort de récupération du consommateur de soins, un « détournement de la clientèle en faveur d'autres institutions hégémoniques » [20].

« On peut même aller, remarque-t-il avec ironie, jusqu'à encourager le consommateur à fixer lui-même le seuil au-delà duquel le rôle de l'institution médicale devrait cesser, mais c'est parce que d'autres services institutionnalisés, également « producteurs de santé », l'attendent impatiemment » [20].

« La nouvelle politique médicale... transforme le monde en hôpital pour des patients à vie » [21].

4. — *Accent sur la prévention*

Illich dénonce avec verdeur ce qu'il appelle le dernier en date des signes de statut social chez les bourgeois: la consommation de soins préventifs [22].

Il écrit: « L'extension du contrôle professionnel aux soins dispensés à des gens en parfaite santé est une nouvelle manifestation de la médicalisation de la vie » [23].

« Au-delà d'un certain niveau d'effort, la somme des actes préventifs, diagnostiques et thérapeutiques, ayant pour cibles les maladies spécifiques d'une population, d'un groupe d'âge ou d'individus, abaisse nécessairement le niveau global de santé de toute la société, en réduisant ce qui précisément constitue la santé de chaque individu: son autonomie personnelle » [24].

Un certain programme de dépistage précoce, nouvellement institué, semble indigner particulièrement Ivan Illich puisqu'il le dénonce pas moins de cinq fois dans son ouvrage: c'est le contrôle préventif de la tension artérielle [25]. « La grande propagande internationale faite au début de 1975, écrit-il, en faveur de la régulation de la pression artérielle de populations entières moyennant l'intervention médicale semble irresponsable » [26]... car alors « la société n'est plus qu'une vaste clinique où tous les citoyens sont des patients dont on surveille et régularise constamment la pression artérielle... » [27].

QUOI FAIRE SELON ILLICH ?

Les citations qui précèdent démontrent clairement qu'Ivan Illich, s'il eut vécu au Québec, eut sans doute été un adversaire acharné de la réforme des services de santé proposée dans le rapport Castonguay, parce que, dans son optique, cette réforme ne peut qu'accentuer encore la dépendance des citoyens envers un système médical en expansion illimitée et ainsi renforcer le processus de médicalisation, alors que c'est le contraire qu'il faudrait rechercher. « Seule une réduction substantielle de l'output global de l'entreprise médicale peut permettre aux hommes de retrouver leur autonomie et par là leur santé » [28].

Illich ne prétend pas « fournir des propositions positives visant à établir des alternatives à la présente façon industrielle de se comporter » [29]. D'ailleurs, croit-il, « la discussion prématurée de telle ou telle organisation nouvelle du système médical ne serait que pure diversion » [30]. Pour lui, il faut donc, pour le moment, se contenter de dénoncer...

Toutefois, il est possible, à la lecture de son livre, de glaner çà et là divers éléments d'une réforme du régime de la santé telle que souhaitée par Ivan Illich. Une liste sommaire comprendrait au moins les points suivants:

a) démédicalisation du concept de maladie [31], *b*) arrêt de la prolifération des professionnels de la santé [32], *c*) soins, dans la mesure du possible, par soi-même, sa famille, ses voisins [33], *d*) apprentissage de « l'art de souffrir » [34], *e*) apprentissage de « l'art de mourir » [35].

On ne trouve pas l'expression « art de vivre » dans l'ouvrage d'Ivan Illich mais c'est bien de cela qu'il s'agit, de la responsabilité que tout homme devrait assumer d'établir délibérément, dans sa vie, un équilibre qui lui convienne entre l'autonomie individuelle et la dépendance envers les autres (y compris les professionnels de la médecine) et de maintenir ensuite ou de faire varier selon les circonstances cet équilibre toujours précaire...

Résumé

Cet article démontre, avec quelques citations extraites de son livre intitulé « Némésis médicale », comment on peut considérer qu'Ivan Illich conteste les principaux objectifs de la réforme des services de santé entreprise au Québec: la médecine globale, l'accès universel aux services de santé, la continuité des soins, l'accent sur la prévention. Une réforme du régime de la santé telle que la souhaiterait Illich est ensuite présentée.

Summary

This article shows, with a few quotations from his book entitled "Medical Nemesis", how it can be argued that Ivan Illich challenges the main objectives of the health services reform undertaken in the province of Quebec: comprehensive medicine, universal access to health services, continuing care, emphasis on prevention. Illich's own outlook on a reform of the health care system is then presented.

BIBLIOGRAPHIE

1. *Rapport de la Commission d'enquête sur la Santé et le Bien-être social*. Québec, Editeur officiel du Québec, 1970. Vol. IV, Tome 2, p. 15.
2. *Id.*, p. 35, par. 568 et p. 42, par. 595-2).
3. *Id.*, p. 41, par. 593.
4. *Id.*, p. 31, par. 543.
5. *Id.*, p. 18, par. 498.
6. *Id.*, p. 19, par. 506.
7. *Id.*, p. 15, par. 489.
8. *Id.*, p. 32, par. 547.

9. *Id.*, p. 15, par. 490.
10. *Id.*, p. 42, par. 593.
11. Illich, Ivan: L'expropriation de la santé. *Esprit*, vol. 42, no 6, juin 74, pp. 931-42.
12. Bosquet, Michel: Quand la médecine rend malade. *Le Nouvel Observateur*, No 519 (21-27 oct. 74) et No 520 (28 oct. — 3 nov. 74).
13. Illich, Ivan: *Némésis médicale*. Paris, Seuil, 1975.
14. *Id.*, p. 206.
15. *Id.*, p. 127.
16. *Id.*, p. 126.
17. *Id.*, p. 48.
18. *Id.*, p. 101.
19. *Id.*, pp. 114-15.
20. *Id.*, p. 110.
21. *Id.*, pp. 126-27.
22. *Id.*, p. 66.
23. *Id.*, p. 67.
24. *Id.*, p. 17.
25. *Id.*, pp. 34-35, 65, 67, 116, 163.
26. *Id.*, p. 34.
27. *Id.*, p. 163.
28. *Id.*, p. 104.
29. *Id.*, p. 97.
30. *Id.*, p. 7.
31. *Id.*, p. 110.
32. *Id.*, p. 55.
33. *Id.*, pp. 88, 89, 153, 167.
34. *Id.*, pp. 145, 150.
35. *Id.*, pp. 178, 181, 201 (note 307).

LA RÉFORME DOIT PRENDRE SON « SECOND SOUFFLE » [1]

Gilbert BLAIN [2]

BILAN IMPOSSIBLE

On a souvent défini le bilan comme une photographie instantanée qui représente la situation d'une entreprise à un moment donné, comme immobilisée en action.

Mais, comme vous le savez tous, il est impossible, à moins d'avoir un appareil très perfectionné, de prendre un instantané d'une scène qui bouge trop vite. On obtient alors une image embrouillée où les contours sont plutôt flous ou bien une image fantaisiste où les éléments mobiles sont répétés plusieurs fois, comme par exemple un cheval à huit ou douze pattes selon la vitesse à laquelle il galope pendant qu'on le photographie.

Le problème se complique encore si non seulement la scène est en mouvement, mais aussi la caméra. Alors il est probable que l'image obtenue sera complètement déformée et qu'il sera impossible d'y reconnaître quoi que ce soit.

Ainsi il me paraît extrêmement difficile de prendre une photographie instantanée de la réforme des services de santé et des services sociaux actuellement en cours au Québec. D'une part, la scène change très vite sous nos yeux. D'autre part, l'œil avec lequel nous considérons cette scène change également car nos attitudes et nos opinions évoluent parfois à une vitesse qui nous surprend. Comment alors réussir à appréhender le phénomène dans son dynamisme et à l'apprécier par rapport à des valeurs sûres ? Quant à moi, je vous déclare tout net que faire présentement un bilan de la réforme des services de santé et des services sociaux entreprise au Québec m'apparaît comme une « mission impossible ».

Cependant, plutôt que de déclarer forfait, je voudrais ré-examiner avec vous quelques-uns des objectifs que la Commission Castonguay-Nepveu avait fixés pour cette réforme et quelques-unes des propositions faites par la Commission quant aux moyens d'atteindre ces objectifs par comparaison avec les dispositions de la loi sur les services de santé et les services sociaux et certaines de ses conséquences déjà perceptibles.

[1] Congrès annuel, Fédération des administrateurs des services de santé et des services sociaux du Québec, Québec, le 13 novembre 1974.

[2] Directeur, Département d'administration de la santé, Université de Montréal.

Je vous invite donc à partager mes réflexions, mes interrogations et même mes appréhensions sur les sujets suivants : la médecine globale et la pratique d'équipe, la médecine préventive et la santé publique, la participation, le découpage régional, l'organisme régional, la porte d'entrée dans le système et enfin les méthodes propices à la réalisation de la réforme.

LA MÉDECINE GLOBALE ET LA PRATIQUE D'ÉQUIPE

L'objectif premier de la réforme proposée par la Commission Castonguay-Nepveu était la réalisation du concept de médecine globale [1].

La médecine globale, selon le rapport de la Commission, implique « la prise en charge » d'une population par une équipe multidisciplinaire qui assure à cette population des soins complets, continus et personnels, comprenant prévention, conseil, éducation, dépistage, diagnostic, traitement et réadaptation [2].

Depuis qu'on a commencé l'application de la loi sur les services de santé et les services sociaux, on ne voit nulle part, au Québec, la réalisation du concept de la médecine globale si ce n'est, jusqu'à un certain point, dans les unités de médecine familiale dont le prototype semble être la clinique familiale du C.H.U. de Sherbrooke. Dans cette clinique, « chaque équipe multidisciplinaire prend charge d'environ 200 familles et distribue tous les soins généraux » [3]. On trouve donc là l'essence même de la médecine globale, c'est-à-dire une population déterminée et une équipe déterminée reliées l'une à l'autre par une sorte de contrat exclusif assurant le service d'une part et la clientèle d'autre part. En effet, si l'on ne fait pas affaire à une population captive d'une façon ou d'une autre, il est inutile d'en mesurer les besoins et de s'organiser pour y répondre d'une manière globale.

D'autre part, la médecine globale requiert la pratique en équipe. Or l'organisation et le fonctionnement des équipes multidisciplinaires font encore l'objet d'études et d'expériences dans plusieurs milieux et nulle part on ne semble avoir trouvé la formule qui assurera à la fois un fonctionnement efficace et la satisfaction des patients et des soignants. Le succès de la pratique d'équipe dépend de nombreux facteurs dont l'interaction est encore mal connue. Je voudrais m'arrêter à trois de ces facteurs : la formation des membres de l'équipe, la direction de l'équipe et le mode de rémunération des soignants.

L'orientation particulière, dans la formation des futurs professionnels, vers le travail en équipes multidisciplinaires semble un prérequis sur lequel tous s'accordent [4]. Cependant, personne ne sait comment on peut former des étudiants en vue des équipes multidisciplinaires lorsque de telles équipes n'existent pas dans les institutions où ils sont formés. Les tentatives de formation des futurs professionnels au travail multidisciplinaire, dans les collèges et les universités, ont été jusqu'à aujourd'hui fort peu nombreuses et l'évolution des milieux d'enseignement est en retard sur ce point par rapport aux milieux d'exercice.

Un autre problème à régler est celui de la direction de l'équipe. Certains entrevoient, pour l'avenir, une espèce d'équipe démocratique où toutes les décisions seraient prises par consensus et dans laquelle se développerait une personnalité collective d'où seraient exclues toute susceptibilité et toute jalousie. Le malade se confierait à cette personne morale que constitue le groupe et dont le porte-parole pourrait varier selon les circonstances ou selon les désirs du malade... D'autres sont convaincus que cet idéal est utopique et que tout effort dans cette voie comporte de grands risques d'inefficacité et même de paralysie de l'équipe. D'aucuns proposent même la formation d'un nouveau « généraliste » qui participerait à la fois du médecin, du psychologue et du travailleur social et qui pourrait coordonner les interventions des multiples spécialistes. Plusieurs expériences ont été tentées dans divers milieux mais nulle part on ne semble avoir déjà trouvé le mode d'intégration idéal des divers professionnels les uns par rapport aux autres dans les équipes multidisciplinaires [5].

Un dernier problème consiste dans le mode de rémunération du médecin par rapport à celui des autres travailleurs de la santé. Il est certain que si l'on veut favoriser la médecine préventive ainsi que la substitution dans l'exécution de certaines tâches entre les professionnels, membres de l'équipe soignante, un régime de rémunération identique pour tous s'impose. En effet, il est bien évident que dans une hypothèse de rémunération à l'acte, le médecin ne saurait être entraîné à poser des « actes préventifs » non rémunérés non plus qu'à déléguer des actes curatifs pour en perdre la rémunération.

Il reste donc de nombreux problèmes à régler avant que des modes de pratique en équipes multidisciplinaires soient assez largement expérimentés, acceptés et généralisés. A mon avis, on ferait mieux de s'attaquer à ces problèmes par des expériences pratiques comme celles des unités de médecine familiale, plutôt que d'ergoter à l'infini sur le sens qu'il faut donner à des textes législatifs ambigus.

LA MÉDECINE PRÉVENTIVE ET LA SANTÉ PUBLIQUE

Les deux expressions « médecine préventive » et « santé publique » sont presque impossibles à trouver dans le Rapport Castonguay-Nepveu ou dans la loi sur les services de santé et les services sociaux. Mais les réalités qu'elles signifient n'en demeurent pas moins essentielles à l'intérieur du réseau des services de santé et des services sociaux.

Ces deux expressions ne sont pas synonymes. Les deux sortes d'action qu'elles représentent poursuivent le même objectif mais la cible visée par cette action n'est pas la même non plus que les méthodes utilisées.

L'expression « médecine préventive » s'applique généralement aux activités de promotion de la santé et de protection contre la maladie exercées par chaque médecin auprès de chacun de ses patients tandis que l'expression « santé publique » désigne certains services communautaires ainsi que l'ensemble des mesures collectives d'éducation et parfois de coercition que prend la société et qui visent la population tout entière ou certains groupes particuliers en ce qui concerne la salubrité de l'environnement et le lien entre les habitudes de vie et la qualité de la vie.

Le Rapport Castonguay-Nepveu ne parlait pas de « médecine préventive ». Ce n'est pas que la Commission négligeait l'importance de la prévention. Bien au contraire, le concept de « médecine globale » qu'elle préconisait devait, dans l'esprit de ses promoteurs, intégrer les activités préventives et curatives, surtout au niveau du centre local. Dans le système idéal proposé par la Commission, tout soignant doit devenir un hygiéniste et tout établissement doit accorder, dans ses objectifs, autant d'intérêt aux mesures préventives qu'aux soins curatifs. Cet idéal n'est peut-être pas utopique mais pour se réaliser, il nécessitera certainement une ré-orientation dans les programmes de formation des professionnels de la santé et des services sociaux.

Quant à la « santé publique », elle était assurée jusqu'à aujourd'hui, dans notre province, par deux sortes d'organismes : d'abord des bureaux de santé municipaux dans une quinzaine d'agglomérations urbaines, puis environ soixante-quinze unités sanitaires dans les comtés ruraux et semi-ruraux, regroupées depuis 1966 en huit régions par le ministère de la Santé.

La Commission Castonguay-Nepveu déplorait en termes assez mordants que l'on ait laissé dépérir ce secteur des services de santé au cours des dernières années et proposait de le revaloriser en l'intégrant carrément dans le nouveau système de médecine globale que doit dispenser le réseau des établissements de santé et de services sociaux [6].

Toutefois les recommandations de la Commission relativement aux diverses activités de santé publique n'étaient pas très précises sauf à propos de la médecine scolaire dont on assignait la responsabilité future aux hôpitaux [7]. D'autre part, on pouvait déduire, bien que ce ne fût pas dit clairement, que les programmes d'immunisation, d'hygiène maternelle, infantile, dentaire, etc. devraient être confiés aux centres locaux.

On a pu croire un moment que le législateur avait complètement oublié la santé publique dans l'élaboration de la loi sur les services de santé et les services sociaux mais la publication des règlements a par la suite calmé les inquiétudes sur ce sujet.

De par ces règlements, une nouvelle structure est constituée, un nouveau nom est trouvé et un nouvel essor semble assuré aux programmes de santé publique.

A long terme, les services de santé municipaux et les unités sanitaires sont censés disparaître après que toutes leurs fonctions auront été convenablement assumées par les trente-et-un nouveaux départements de santé communautaire. Pour le moment, toutefois, nous n'en sommes qu'à la phase de transition.

On peut déplorer le fait que le règlement ait posé un double défi aux départements de santé communautaire. En effet, il est à craindre que dans certains milieux, la fonction interne d'administration hospitalière qui est confiée à ce département en ce qui regarde le fonctionnement du service d'urgence de première ligne et des services externes généraux n'accapare d'une manière excessive le temps et l'énergie des responsables du département et les empêche de se consacrer plus intensivement à leur fonction externe en rapport avec tous les établissements sanitaires et sociaux de la région, les institutions d'enseignement, les industries, les centres de loisirs et tous les milieux où peuvent s'appliquer les programmes de santé communautaire.

LA PARTICIPATION

C'est à l'époque où le Rapport Castonguay-Nepveu a été élaboré, soit de 1966 à 1972, que les manifestations en faveur de la participation ont secoué la société en Amérique et en Europe, et particulièrement en France où la révolte étudiante de mai '68 faillit même faire tomber le gouvernement. Il faut bien constater que cette flambée de « participationnite » semble déjà éteinte, comme si la société unidimensionnelle que dénonçait Marcuse l'avait

récupéré et assimilé lui-même avec ses œuvres et ses disciples, comme des produits de consommation, lui donnant par là raison d'une manière absolue.

Quand on parle de participation, un certain nombre de distinctions viennent d'emblée à l'esprit : d'abord entre la participation interne des employés et la participation externe des usagers ou des citoyens ; puis, en ce qui regarde les citoyens, entre la participation à la définition des besoins et la participation à la gestion ; enfin entre les différents degrés dans la participation : de l'information à la décision en passant par la consultation. Il n'est pas certain que tous ces modes de participation s'imposent partout et toujours.

D'autre part, si l'on croit en la valeur de la participation et qu'on essaie de la réaliser, on fait face à deux problèmes d'ordre complémentaire. On se demande, d'une part, comment secouer l'apathie des masses et, d'autre part, comment réagir devant l'agitation de certaines minorités très actives. La solution de l'un et l'autre de ces deux problèmes est aussi difficile à trouver.

« Toute dépolitisation favorise l'ordre établi, l'immobilité, le conservatisme », écrit le politicologue français Maurice Duverger qui accuse les classes nanties et satisfaites de chercher à démobiliser les classes opprimées et insatisfaites en les persuadant que les luttes politiques sont malpropres [8]. D'autre part, les soi-disant « leaders naturels », apparus lors des phénomènes de contestation populaire, s'arrogent le droit de parler au nom des secteurs de la population dont ils sont issus sans en avoir reçu aucun mandat.

Alors se posent un certain nombre de questions : Qui peut parler au nom du peuple ? Comment susciter la participation d'une population aux décisions qui la concernent ? Que doivent faire les dirigeants de la société s'ils estiment que la population ne voit pas son intérêt à long terme ? Comment, enfin, peut-on réconcilier les objectifs des technocrates qui tendent à rationaliser et à bureaucratiser le système avec ceux des « citoyens de la base » qui veulent adapter le système aux conditions locales et l'humaniser ?

La Commission Castonguay-Nepveu a cru répondre à ces questions en préconisant la participation de toutes les manières: interne et externe, à la définition des besoins comme à la gestion, et, en ce qui regarde les citoyens, selon les trois degrés possibles: information par des réunions périodiques, consultation par des conseils consultatifs externes et rôle décisionnel dans les conseils d'administration des établissements et de l'organisme régional [9].

La Commission précisait comme suit son attitude quant à la participation des citoyens dans les conseils d'administration: « Ces personnes ne représentent pas la communauté mais siègent... à titre personnel; c'est dire que, non mandataires d'un groupe ou de la population, elles peuvent prendre des décisions sans l'autorisation préalable de ceux qui les ont élues. La population conserve intact son pouvoir de contestation » [10].

On peut même dire que la Commission a poussé le souci de la participation très loin en préconisant une compénétration des divers niveaux du régime par un système compliqué de représentation réciproque des établissements dans les conseils d'administration les uns des autres. Peut-être a-t-on cru que l'enchevêtrement administratif des établissements serait une garantie de l'efficacité du réseau. Il y a là, me semble-t-il, un risque de fausse sécurité, car ces interrelations structurales ne sauraient certes remplacer les liens fonctionnels à établir en vue d'assurer la continuité des soins...

Quoi qu'il en soit, c'est peut-être en ce qui concerne la participation que la loi sur les services de santé et les services sociaux a le mieux respecté l'esprit du Rapport Castonguay-Nepveu.

Ainsi, il est indéniable qu'une démocratisation des conseils d'administration s'est produite, que certaines personnes participent maintenant à la définition des besoins et à la gestion des établissements, qui jamais, avant la nouvelle loi sur les services de santé et les services sociaux, n'auraient pu pénétrer dans ces « cénacles » maintenant ouverts à de nouvelles catégories de participants.

Il reste à voir si ces nouveaux participants sauront faire passer le bien commun avant leurs intérêts particuliers, s'ils sont prêts à consacrer le temps et l'énergie nécessaires pour acquérir une compétence suffisante quant aux questions en jeu et s'ils accepteront de partager la responsabilité des décisions prises en commun...

LE DÉCOUPAGE RÉGIONAL

La Commission Castonguay-Nepveu recommandait de diviser la province en trois régions, pour le moment, autour des facultés de médecine de Montréal, Québec et Sherbrooke [11].

Elle était arrivée à ce découpage après avoir considéré un certain nombre de critères de deux ordres: d'abord des critères techniques propres au domaine de la santé, puis des critères dits généraux, valables pour divers services publics.

Les critères techniques fondamentaux, considérés selon l'expérience de plusieurs autres pays, étaient

au nombre de trois: bassin de population de 1 à 2 millions (avec minimum de 600,000 et maximum de 6 millions), présence d'au moins un centre universitaire et capacité d'organiser une gamme complète de soins.

Après que les critères d'ordre technique aient situé le noyau et les limites générales de la région sanitaire, les critères dits généraux devaient permettre d'en arriver à un découpage précis du territoire. Ces critères étaient au nombre de cinq: la distance, la distribution actuelle de la clientèle, le sentiment d'appartenance régionale, les régions administratives et l'équipement actuel [12].

Dans le découpage qui finalement a prévalu au moment de la rédaction de la loi, deux des critères dits généraux, soit les régions administratives actuelles et le sentiment d'appartenance régionale, ont éclipsé tous les autres critères. De l'aveu même d'un porte-parole du ministère, la préférence a été accordée à un « découpage régional basé sur un impératif politique (la participation) plutôt que technocratique (la complémentarité des services) » [13].

Le ministère a donc voulu favoriser avant tout l'émergence d'une « conscience régionale », en créant des régions de dimensions plus restreintes mais sociologiquement plus significatives pour les citoyens appelés à participer au fonctionnement de leurs institutions.

Toutefois, ce découpage s'est fait au mépris des trois critères dits techniques que la Commission Castonguay-Nepveu avait présentés comme les premiers à considérer. Il faut maintenant s'accommoder des conséquences du choix qui a été fait: par exemple, une région qui possède un centre universitaire mais une population trop petite et qui, sous bien des aspects, se trouve sur-équipée; une autre région dont la population respecte le critère technique établi par la Commission Castonguay-Nepveu quant au nombre mais qui, ne possédant pas de centre universitaire sur son territoire, réclame l'aide de toutes les universités de la province... Il reste à voir si les avantages escomptés du découpage en douze plutôt qu'en trois régions se réaliseront et s'ils en contrebalanceront les inconvénients...

L'ORGANISME RÉGIONAL

Le Rapport Castonguay-Nepveu préconisait une réforme dont les tendances régionalistes étaient très fortes. L'organisme régional qu'on y proposait, appelé « Office régional de la santé », pouvait presque être assimilé à un ministère régional de la santé car il aurait exercé des pouvoirs étendus et contrôlé un budget important. Ce contrôle d'une enveloppe budgétaire trop considérable est d'ailleurs l'une des raisons pour laquelle la formule fut écartée, car le gouvernement ne pouvait pas, semble-t-il, se résoudre à confier des budgets globaux d'une telle importance à des organismes régionaux [14].

Le premier projet de loi sur les services de santé et les services sociaux proposait donc de créer des organismes appelés « Offices régionaux des affaires sociales » dépourvus des pouvoirs autonomes que la Commission avait recommandé de leur attribuer. Les mémoires présentés à la Commission des Affaires sociales, avant la rédaction définitive de la loi, ont bien démontré l'absence de consensus dans la population sur le rôle possible d'un organisme régional: en effet, les uns recommandaient d'en réduire encore les pouvoirs; d'autres souhaitaient revenir aux propositions de la Commission Castonguay-Nepveu; d'autres enfin tenaient une position moyenne favorisant l'instauration d'une « force communautaire » extra-gouvernementale au niveau régional [15, 16].

Le législateur a tranché le nœud gordien à la manière d'Alexandre, d'un coup d'épée. Il a décidé de créer des « Conseils régionaux » à qui il a confié simultanément cinq rôles différents, soit ceux d'animateur social, d'agent d'information, d'ombudsman spécialisé, de conseiller en administration et enfin de planificateur régional.

Les fonctions très variées que la loi a attribuées au conseil régional de la santé et des services sociaux en font une espèce de prisme dont une facette est tournée vers la population, une autre vers les établissements et une autre vers le ministère. Il existe un risque que le conseil régional soit perçu par chacun des trois comme un instrument de pression sur les deux autres et que cette perception l'empêche, dans une certaine mesure, de jouer son rôle autonome de planification et de coordination régionales.

Toutefois, il faut ajouter que M. Claude Castonguay, alors ministre des Affaires sociales, déclarait le 14 février 1972: « Ces conseils régionaux se transformeront éventuellement en de véritables organismes régionaux jouissant d'une large autonomie sur le plan de la programmation régionale du développement des services de santé et des services sociaux » [17]. Ces paroles n'étaient pas vaines, semble-t-il, et on peut en voir une confirmation dans une déclaration faite, le 5 septembre dernier, par le ministre des Affaires sociales à la conférence de la Canadian Medical and Biological Engineering Society. Considérant la possibilité que des équipes de spécialistes en génie bio-médical soient formées au niveau régional, M. Forget déclarait : « Une telle

équipe affiliée à une université et fonctionnant sous les auspices d'un Conseil régional de la santé et des services sociaux pourrait recevoir un mandat général de rationalisation des activités techniques bio-médicales. Ce mandat pourrait être rendu efficace envers les établissements par le regroupement sous l'autorité du Conseil régional de la santé et des services sociaux des budgets de renouvellement de l'équipement bio-médical de tous les centres hospitaliers de la région et par la mise en commun du total estimatif des sommes consacrées normalement à l'entretien de ce matériel... » [18].

C'est bientôt que la sincérité du ministère sera mise à l'épreuve. Sera-t-il prêt, ainsi que M. Castonguay le déclarait et que M. Forget semble disposé à le faire, à déléguer aux conseils régionaux une « large autonomie » sur le plan du développement et de la programmation régionales et à lui confier certains budgets regroupés au niveau de la région ? L'avenir le dira.

LA PORTE D'ENTRÉE DANS LE SYSTÈME

Le Rapport Castonguay-Nepveu prévoyait l'existence d'une seule porte d'entrée dans le système de soins : le « Centre local de santé » ou C.L.S. Le Rapport contenait même les phrases suivantes : « Nous ne croyons pas que l'exercice individuel de la médecine disparaisse totalement, du moins à court ou à moyen terme. Cependant, nous estimons qu'un médecin exerçant seul doit s'affilier à un C.L.S., de façon que ses malades trouvent la complémentarité et la continuité des soins généraux qu'il ne peut lui-même offrir » [19].

Pour ce centre local, la Commission Castonguay-Nepveu préconisait trois formules possibles « pour garder au régime le maximum de souplesse et à ses agents le maximum de liberté ». En fait, ces trois formules se ramenaient à deux c'est-à-dire soit une corporation publique constituée sur l'initiative d'un groupe de citoyens ou d'un groupe de professionnels, soit une corporation de type privé. « Nous sommes conscients, ajoutait le Rapport, que l'organisation de la distribution des soins généraux selon ces types de formules doit se faire de façon progressive et, en outre, que ces formules ne sont pas limitatives et qu'on peut sans doute en imaginer d'autres, selon les besoins et les circonstances » [20]

Du Rapport de la Commission à la loi sur les services de santé et les services sociaux, deux changements importants se sont opérés. Tout d'abord, sans aucune recommandation en ce sens de la part de la Commission et par un processus encore inexpliqué, les centres locaux de santé sont devenus les centres locaux de services communautaires. Puis la grande souplesse préconisée relativement à la formule d'organisation des centres locaux s'est rigidifiée en une seule formule très précise, prescrite par la loi et les règlements.

Le centre local de services communautaires, d'après la définition qu'en donne la loi et selon les directives du ministère, doit remplir non pas une mais trois missions, c'est-à-dire celles de fournir des services médicaux et des services sociaux mais aussi celle de constituer un noyau d'action communautaire ou, comme on dit aussi, d'animation sociale.

Certains croient qu'un tel organisme n'est pas viable, que des fonctions de service peuvent très difficilement coexister dans la même structure avec la fonction d'animation sociale, car les premières requièrent un système d'action ordonné et des interventions soucieuses d'efficacité tandis que la seconde implique le bouillonnement des idées, la discussion permanente et souvent même la contestation de l'ordre établi. C'est, en tout cas, ce que semble démontrer la douloureuse expérience de certains centres locaux à Montréal et à Québec. L'opposition presque irréductible entre les fonctions de service et celle d'animation sociale est clairement exposée dans un article récent de la revue MacLean, intitulé : « Docteur, j'ai mal au C.L.S.C. » [21].

D'autre part, la loi a ramené à un modèle unique la grande variété de modèles que la Commission Castonguay-Nepveu préconisait dans l'organisation des centres locaux. En effet, la loi semble vouloir créer un réseau exclusivement public d'établissements de soins généraux de première ligne, auquel devraient se rallier tôt ou tard tous les professionnels de la santé et des services sociaux exerçant à ce niveau.

Dans le domaine de l'éducation, la réforme a laissé coexister, à côté du secteur public, un secteur privé qu'on subventionne généreusement d'ailleurs et qui lui fait concurrence. Dans le domaine des services de santé et des services sociaux, si on voulait abolir le secteur privé au niveau des soins de première ligne, il fallait prendre d'emblée deux mesures radicales : d'abord organiser rapidement un réseau complet de centres locaux publics à travers tout le Québec, puis négocier avec les médecins omnipraticiens les conditions de leur intégration dans ce réseau. Si, d'autre part, on avait voulu conserver, dans le domaine des affaires sociales comme dans celui de l'éducation, un secteur privé qui puisse concurrencer le secteur public, il fallait rédiger la loi autrement et s'en tenir aux propositions de la Commission Castonguay-Nepveu sur ce sujet.

Une fois la loi adoptée telle que rédigée, il semble que le ministère ne s'est pas résolu à suivre la ligne qu'il s'était lui-même tracée. Tout d'abord, il a entrepris bien timidement et bien lentement l'implantation des centres locaux. Puis, invoquant le fait que ces centres locaux sont des établissements juridiquement autonomes, il a adopté une étrange politique de nier la représentativité de la Fédération des médecins omnipraticiens du Québec en ce qui regarde la négociation des traitements et des conditions de travail des médecins généralistes dans ces centres et il a invité ceux-ci à signer des contrats individuels. Il en est évidemment résulté une opposition farouche de la Fédération à la participation de ses membres dans les centres locaux.

Dans un article très pénétrant et s'appuyant sur plusieurs exemples, sous le titre : « Les C.L.S.C. en crise », un journaliste de Québec prétendait récemment qu'en ce qui regarde les centres locaux, le ministère, par certaines de ses manœuvres, « compromet ses propres politiques » [22].

Faut-il croire que l'implantation de nouveaux centres locaux serait maintenant remise en question comme certains se plaisent à le répéter ? Si tel était le cas, la réforme des services de santé et des services sociaux serait menacée d'avortement dans un de ses éléments essentiels, celui de la médecine globale qui ne peut se réaliser que dans et par une nouvelle structure au niveau local à travers toute la province.

D'autre part, si le ministère décide — et plusieurs le souhaitent — que la création de centres locaux doit être intensifiée jusqu'à la constitution d'un réseau complet, une condition me semble impérieuse. Pour que cet objectif soit réalisable, des assouplissements s'imposent quant aux modèles possibles de centres locaux, comme le démontre le docteur Léonard Langlois dans un récent article très pondéré et pragmatique [23]. L'intransigeance pourrait saborder la réforme mais la souplesse pourra la sauver.

LES MÉTHODES PROPICES AU SUCCÈS DE LA RÉFORME

Depuis qu'on a commencé à appliquer les recommandations du Rapport Castonguay-Nepveu, nous avons été témoins d'une véritable corrida législative comme on en trouve peu d'exemples dans l'histoire du Québec.

Le 24 décembre 71, l'Assemblée nationale du Québec adoptait une loi sur les services de santé et les services sociaux qui devait entrer en vigueur le 1er juin 72. Cette loi remplaçait la loi des hôpitaux de 62 et elle abrogeait ou modifiait une dizaine d'autres lois.

Le 8 novembre 72, le lieutenant-gouverneur en conseil décrétait un ensemble de règlements qui sont entrés en vigueur le 1er janvier 73 soit six mois après la loi. Il faut reconnaître ici, entre la mise en vigueur de la loi et celle des règlements, une célérité à laquelle nous n'étions pas habitués car, entre l'adoption de l'ancienne loi des hôpitaux et la publication de ses règlements, il s'était écoulé une période de sept ans, de 62 à 69. Sept ans, ce fut bien long mais six mois, est-ce que ce fut assez ? On peut peut-être en douter...

En effet, depuis le 1er janvier 73, de nombreux amendements ont été apportés à ces règlements par suite des arrêtés en conseil du 4 avril, 1er août, 12 et 15 septembre et 19 décembre 73. Par ces cinq arrêtés en conseil, 75 articles des règlements, sur environ 325, ont été modifiés, soit près du quart.

Enfin, en juillet dernier, le ministre des Affaires sociales déposait les projets de lois numéros 40 et 41 qui transforment environ 60 des 168 articles de la loi, soit plus du tiers.

Parallèlement à ce cheminement de la loi sur les services de santé et de ses règlements, il faut mentionner l'adotion de trois lois très importantes : la loi de la protection du malade mental, sanctionnée le 30 juin 72 ; celle de la protection de la santé publique, sanctionnée le 21 décembre 72 ; enfin le code des professions, sanctionné le 6 juillet 73, qui abrogeait une quinzaine de lois des corporations professionnelles mais était lui-même accompagné de plus de vingt nouvelles lois relatives chacune à une profession différente.

Pour compléter ce tableau, un arrêté en conseil du 13 avril 73 amendait certains règlements de l'assurance-hospitalisation tandis qu'une nouvelle loi modifiant la loi de l'assurance-maladie et la loi de la Régie de l'assurance-maladie était sanctionnée le 6 juillet 73.

Devant cette avalanche de nouvelles lois et de demandes d'interprétation qui fusent de partout, devant les innombrables contrats d'affiliation ou de service qu'il faudra rédiger entre les établissements ou entre les établissements et les institutions d'enseignement, les avocats jubilent, certes, mais les administrateurs sont moins enthousiastes de participer à une partie où les règles du jeu changent si souvent.

En ce qui concerne la loi sur les services de santé et les services sociaux et ses règlements, il est permis de se demander comment il se fait que de nouveaux textes législatifs ont besoin de retouches si majeures après si peu de temps. On est bien forcé de conclure que de deux choses l'une : ou bien ces textes

ont été rédigés à la hâte, sans prévoir certaines difficultés ou contradictions possibles, ou bien la réforme des services de santé et des services sociaux évolue si rapidement que, de mois en mois, les textes législatifs ne conviennent plus et doivent être adaptés comme s'il fallait dorénavant refaire la loi à chaque occasion de l'appliquer. Quelle que soit l'explication, la conséquence est la même : l'application de la réforme s'empêtre dans des querelles juridiques et des affrontements sur les interprétations variées et parfois contraires que l'on donne à divers textes selon qu'on appartient à tel ou tel groupe dans la société.

D'ailleurs, il n'est pas sûr que la méthode législative employée comme premier instrument de la réforme des services de santé et des services sociaux était bien compatible avec l'esprit d'initiative et d'aventure qui inspirait chaque page du Rapport Castonguay-Nepveu. En effet, à l'heure des innovations et des projets-pilotes, on pourrait croire qu'il était prématuré de pétrifier d'avance des structures expérimentales dans des règles juridiques très rigides.

Certes la voie législative et coercitive semble, à première vue, plus expéditive mais d'autres méthodes, en apparence plus lentes et plus complexes, peuvent à la longue se révéler plus efficaces : la persuasion des intéressés, la formation des jeunes, la concertation avec divers corps intermédiaires, la propagande auprès du grand public, la stratégie politique et administrative, les mesures incitatives d'ordre social ou financier... Quand viendra l'heure du bilan de la réforme, les méthodes qui auront été adoptées seront sans doute jugées sur leurs résultats à long terme et non sur la vitesse de la création par l'Assemblée nationale des nouvelles structures juridiques.

Au point où nous en sommes, la réforme législative, espérons-le, est terminée. La réforme culturelle reste à faire. Non pas à coups de lois mais par l'information et l'éducation. Pour qu'elle se réalise, il faut susciter l'intérêt et la collaboration du public, l'engagement et même l'enthousiasme des professionnels soignants et des administrateurs des services de santé et des services sociaux.

Les cadres juridiques sont en place mais ce ne sont pas les structures du système de distribution des soins qui vont changer la mentalité sociale et les attitudes collectives.

Un leadership administratif s'impose maintenant pour conduire à bien la réforme culturelle. Quant à moi, je suis convaincu qu'on peut trouver, aux divers niveaux du réseau des affaires sociales, la compétence requise pour exercer ce leadership et pour imprimer à la réforme le dynamisme nécessaire à sa réalisation complète.

La réforme des services de santé et des services sociaux doit prendre son second souffle. La force de ce second souffle dépendra des énergies de tous ceux qui accepteront d'y contribuer.

Résumé

Dans cet article, l'auteur réexamine quelques-uns des objectifs que la Commission d'Enquête sur la Santé et le Bien-être social avait fixés pour la réforme des services de santé et des services sociaux dans la province de Québec. Les moyens que la Commission proposait pour atteindre ces objectifs sont comparés avec les dispositions de la loi sur les services de santé et les services sociaux et certaines de ses conséquences déjà perceptibles.

Les sujets suivants sont passés en revue : la médecine globale et la pratique d'équipe, la médecine préventive et la santé publique, la participation, le découpage régional, l'organisme régional, la porte d'entrée dans le système et enfin les méthodes propices au succès de la réforme.

Summary

In this article, the author re-examines some of the goals that had been set by the Commission of Inquiry on Health and Social Welfare for the reform of health and social services in the province of Quebec. The means suggested by the Commission for reaching these goals are compared with the provisions of the new Act respecting health services and social services and with some of its already apparent consequences.

The following subjects are reviewed: comprehensive medicine and team practice, preventive medicine and public health, participation, regionalization and regional structures, the point of entry into the system and, finally, recommended methods for a successful reform.

BIBLIOGRAPHIE

1. Rapport de la Commission d'Enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. IV, Tome 2, p. 15.
2. Ibidem, Vol. IV, Tome 2, p. 15 paragraphe 489, p. 31 paragraphe 542 et p. 35 paragraphe 566.
3. Langlois, Léonard: *Effets d'une clinique multidisciplinaire sur l'utilisation des lits d'hôpitaux. La Vie médicale au Canada français*, 3: No 9, septembre 1974, p. 890.
4. Vinet, Alain: « Formation médicale et travail d'équipe ». *La Vie médicale au Canada français*, 2: No 1, janvier 1974, p. 63.
5. Horwitz, John: *Team Practice and the Specialist*. Springfield, Ill., Charles C. Thomas, 1970.
6. Rapport de la Commission d'Enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. IV, Tome 1, p. 104 paragraphes 257-258 et p. 112 paragraphe 292.
7. Ibidem. Vol. IV, Tome 2, p. 88, paragraphes 40-42.
8. Duverger, Maurice: *Introduction à la politique*. Paris, Gallimard, Collection Idées, No 44, 1964, p. 21.
9. Rapport de la Commission d'enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. IV, Tome 2, pp. 104 à 129, paragraphes 676-677-678-688-740-747.
10. Rapport de la Commission d'Enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. IV, Tome 2, p. 107, paragraphe 687.
11. Ibidem, Vol. IV, Tome 4, p. 176, paragraphes 1695-1696.
12. Ibidem, Vol. IV, Tome 2, pp. 53 à 63.

13. Forget, Claude E.: « Développement et implantation de l'idée de régionalisation des services de santé et des services sociaux au Québec ». *Administration publique du Canada*, Vol. 17, No 1, printemps 1974, pp. 26-37 (voir p. 36).

14. Ibidem, p. 28.

15. Lemieux, Vincent et al.: « La régulation des affaires sociales: une analyse politique ». *Administration publique du Canada*, Vol. 17, No 1, printemps 1974, pp. 37-65 (voir pp. 48 à 50).

16. Ronaud, François et Von Schoenberg, Brigitte: « L'implantation des conseils régionaux de la santé et des services sociaux: analyse d'un processus politique ». *Revue canadienne de science politique*, Vol. VII, No 1, mars 1974, pp. 53-70 (voir pp. 53-55).

17. Castonguay, Claude: Allocution prononcée devant l'Académie des sciences morales et politiques, à Paris, le 14 février 1972. Texte minéographié publié par le ministère des Affaires sociales du Québec, p. 12.

18. Forget, Claude E.: « La contribution du génie bio-médical dans la distribution des services de santé ». Allocution prononcée à l'Hôtel Reine Elizabeth lors de la 5e conférence de la « Canadian Medical and Biological Engineering Society », le 5 septembre 1974. Texte minéographié publié par le ministère des Affaires sociales du Québec, page 7.

19. Rapport de la Commission d'Enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. IV, Tome 2, p. 103, paragraphe 673.

20. Ibidem, Vol. IV, Tome 2, p. 103, paragraphe 672.

21. Ribeyron, Marie-Thérèse: « Docteur, j'ai mal au C.L.S.C. » *Le Maclean*, septembre 1974, p. 28.

22. Bellefeuille, Roger: « Les C.L.S.C. en crise ». *Le Soleil*, 28 septembre 1974.

23. Langlois, Léonard: « Des assouplissements s'imposent ». *Le médecin du Québec*, 9: No 6, juin 1974, p. 7.

55

D'UNE REFORME A L'AUTRE

par Gilbert Blain

Depuis une cinquantaine d'années, trois importantes réformes se sont produites dans le domaine des Affaires sociales au Québec. Ces réformes se sont chevauchées dans le temps. Ainsi, la première n'était pas encore terminée au moment où la deuxième a été entreprise. Ces deux réformes ne sont pas encore complétées qu'une autre vient d'être amorcée. Présentement, la première s'achève et la deuxième se poursuit, pendant qu'on se demande quelles sont les chances de succès de la troisième.

L'objet de ces trois réformes n'est pas le même pour chacune. La première portait sur le transfert de la responsabilité financière de l'individu à la collectivité, face aux risques de la vie; elle a donné naissance à l'*Etat-providence*. La deuxième portait sur l'organisation des services; elle a favorisé le développement de l'*Etat technocratique*. La dernière porte sur l'art de vivre: elle rendra peut-être à

Le Dr Gilbert Blain est professeur titulaire au Département d'administration de la santé, de l'université de Montréal.

*Texte présenté lors d'une session d'étude de l'Association des directrices de services bénévoles hospitaliers du Québec, à l'hôpital Sainte-Justine de Montréal, le 22 mai 1978.

l'individu une certaine autonomie et un certain sens de sa propre responsabilité face à l'*Etat paternaliste et omniprésent* . . .

On pourrait présenter ces trois réformes de la manière suivante:

- *Etablissement de la sécurité sociale:* graduellement depuis le début du vingtième siècle et, d'une manière plus intensive, depuis 1960 . . .
- *Tentative de réorganisation des services et de rationalisation du système socio-sanitaire:* depuis la publication du Rapport Castonguay-Nepveu sur la santé, en 1970 . . .
- *Projet de modifier les habitudes de vie et l'environnement:* depuis la publication de l'ouvrage de Marc Lalonde, *Nouvelle perspective de la santé des Canadiens*, en 1974 . . .

Etablissement de la sécurité sociale

De la fondation de Québec jusqu'au début du vingtième siècle, il s'est écoulé trois cents ans. Pendant cette période, comment nos ancêtres ont-ils fait face aux problèmes usuels comme la maladie ou les accidents, l'infirmité ou l'invalidité, le décès du père de famille, etc., ou aux catastrophes imprévues comme l'incendie, l'inondation, le chômage prolongé, etc.?

Un fait est certain: ils ne recevaient généralement aucune assistance de l'Etat. C'était l'époque de l'*entr'aide mutuelle*, de la *charité*, de la *philantropie*. En cas d'incendie, les voisins organisaient une «corvée» pour reconstruire la maison. En cas de chômage ou de maladie, les parents — ou, à leur défaut, les organisations paroissiales — fournissaient les secours nécessaires. En cas d'accident de travail, un ouvrier estropié ne devait compter sur aucune compensation monétaire; bien au contraire, il pouvait plutôt perdre son emploi et devenir une charge pour sa famille . . .

Chaque citoyen affrontait individuellement les risques de la vie. Les mécanismes de secours mutuel dépendaient de la bonne volonté des proches. A cette époque, certaines expressions n'avaient pas cours: jusque vers la fin du dix-neuvième siècle, personne ne réclamait la *sécurité sociale* ni l'*assurance-chômage* ni le *revenu annuel garanti*. . .

Ces expressions, qui font maintenant partie du vocabulaire courant, démontrent clairement comment le pendule est passé d'une extrême à l'autre en moins d'un demi-siècle. Graduellement, l'Etat s'est substitué aux individus et un transfert de responsabilités s'est effectué, du domaine personnel et privé au domaine collectif et public. En ce qui regarde les fardeaux financiers reliés aux aléas de la vie, on est passé du *risque individuel* à une *sécurité sociale* de plus en plus complète.

La sécurité sociale est un concept d'application très récent dans l'histoire de notre pays et de notre province. Sans préciser si elles sont d'origine fédérale ou provinciale, la liste suivante comprend les principales mesures gouvernementales d'*assistance* (pour le citoyen incapable de pourvoir à ses besoins) ou d'*assurance* (pour protéger le citoyen contre certains risques éventuels), avec leur date d'instauration *au Québec*: indemnités aux accidentés du travail (1909); assistance publique (1921); assurance-chômage (1940); allocations familiales (1945); pensions de vieillesse (1952); assurance-hospitalisation (1961); régime de rentes à la retraite (1966); aide sociale (1969); assurance-maladie (1970).

Notre régime de sécurité sociale est maintenant très étendu. Certes, il reste quelques améliorations à y apporter. Ainsi, par exemple, il faudra encore compléter la gamme des services inclus dans l'assurance-maladie et instaurer le revenu annuel garanti. On pourra alors dire qu'au point de vue financier, le citoyen est parfaitement protégé contre l'adversité, de la naissance au décès ou bien, selon les expressions popularisées par les propagandistes de l'Etat-providence, «du berceau au tombeau», ou encore: «de l'accouchement à l'enterrement», ou mieux encore: «de la conception à la résurrection» . . .

Du côté des citoyens, l'avènement de l'Etat-providence a eu deux conséquences quant à l'utilisation des services socio-sanitaires: d'abord un accroissement notable de la demande et de la consommation, puis la constatation que l'accessibilité financière ne suffit pas pour assurer une distribution équitable des services. Ainsi, on s'est aperçu que d'autres facteurs que le coût pouvaient entraver l'accès aux services socio-sanitaires: éloignement géographique, pénurie ou mauvaise répartition de la main-d'oeuvre, inadéquation entre les ressources et les besoins, obstacles d'ordre psycho-social, etc.[1]

Du côté de l'Etat, le financement des mesures de sécurité sociale a eu deux effets inévitables: d'abord une augmentation rapide des dépenses publiques et, par conséquent, une élévation correspondante du niveau des taxes; puis la prolifération d'une envahissante bureaucratie publique ou para-publique pour régle-

menter, distribuer et contrôler les prestations et les services...

Alors, certains ont cru qu'il fallait restructurer le système des services de santé et des services sociaux, en vue de poursuivre un but humanitaire d'abord : améliorer l'accessibilité; mais aussi un but économique: freiner l'escalade vertigineuse des coûts. Après la réforme du *financement*, on a donc entrepris la réforme de l'*organisation*.

Réorganisation des services et rationalisation du système

Réorganiser, restructurer, rationaliser, ces trois mots impliquent des changements majeurs, une réforme en profondeur. C'est, en effet, un «grand bouleversement» qui s'est produit au Québec après le Rapport de la Commission Castonguay et l'accession de son président au poste de ministre des Affaires sociales, en 1970. Alors que d'autres provinces tentaient d'appliquer un traitement plutôt conservateur pour remédier aux maux de leur système socio-sanitaire, le Québec a eu recours à un traitement qu'on a comparé à la «chirurgie expérimentale»[2].

D'une manière très rapide et très brusque, le contexte juridique des affaires sociales a été chambardé du tout au tout. En moins de quatre ans, on a adopté, coup sur coup, dans un marathon législatif sans précédent, près de quarante nouvelles lois[3]. La plupart de ces nouvelles lois s'articulaient autour de deux pièces maîtresses: un *Code des professions* et une *Loi sur les services de santé et les services sociaux*, qui précisaient les deux éléments essentiels de la réforme proposée que sont la réorganisation des *corporations professionnelles* d'une part, des *établissements* et du *réseau* lui-même, d'autre part.

Du côté des professions, l'objectif visé était double: une plus grande protection du public par de nouveaux mécanismes de contrôle et une plus grande souplesse entre les catégories professionnelles, en vue de faciliter la formation des équipes multidisciplinaires.

Quant aux établissements, l'objectif visé était quadruple: améliorer l'accessibilité, la continuité, la qualité et l'efficacité des services offerts. Pour atteindre cet objectif, le moyen proposé était la constitution d'un réseau cohérent où les établissements et les programmes seraient inter-reliés à l'intérieur d'un territoire donné, à la fois horizontalement, selon *la variété* des besoins à satisfaire, et verticalement, selon *le niveau* de spécialisation nécessaire.

Certains critiques ont parlé de «structurite» à propos de la réforme proposée, car ses promoteurs semblaient croire naïvement qu'un nouveau cadre juridique et de nouvelles structures administratives devaient en assurer le succès. Ainsi, de par la législation, pas moins de *sept* nouvelles structures ont été constituées, dont *cinq* entièrement nouvelles et *deux* par regroupement d'organisations déjà existantes.

Les cinq structures entièrement nouvelles sont les suivantes:

- *Au niveau provincial:*

— la Régie de l'assurance-maladie (R.A.M.Q.)
— l'Office des professions (O.P.Q.)

- *Au niveau régional:*

— les Conseils régionaux de la santé et des services sociaux (C.R.S.S.S.) au nombre de 12, correspondant aux 10 régions socio-économiques du Québec avec une subdivision en 3 pour la région de Montréal;
— les Départements de santé communautaire (D.S.C.), responsables de la santé publique sur le territoire québécois divisé, à cette fin, en 32 «zones»; ces Départements prennent la relève des unités sanitaires et des bureaux de santé municipaux qui ont été abolis.

- *Au niveau local:*

— les Centres locaux de services communautaires (C.L.S.C.), présentement au nombre d'environ 80 et dont on projette un total d'à peu près 200 dans autant de «districts» socio-sanitaires dont les limites demeurent à préciser.

Les deux nouvelles structures établies par regroupement et réorganisation de structures anciennes sont les suivantes:

- *Au niveau provincial:*

— le ministère des Affaires sociales (M.A.S.) constitué par la fusion des deux ministères de la Santé et du Bien-être social.

- *Au niveau régional:*

— formés par le regroupement des diverses agences sociales autrefois autonomes, les Centres de services sociaux (C.S.S.) présentement au nombre de 14, soit un par région socio-sanitaire, sauf dans la région du Montréal-Métropolitain où il y en a trois, la région du Bas-Saint-Laurent et de la Gaspésie où il y en a deux, et la région du Nouveau Québec où il n'y en a pas encore.

Il faut également mentionner certains changements *de noms* introduits par cette réforme. Ainsi, les quelque 250 hôpitaux du Québec sont devenus des «Cen-

très hospitaliers» (C.H.), tandis que les hospices, résidences protégées, garderies, orphelinats et maisons de réadaptation, qui forment un total d'environ 1 000 établissements, ont été réunis sous l'appellation de «Centres d'accueil» (C.A.) [4].

L'heure n'est pas encore venue de faire le bilan définitif de ce «grand dérangement» qui s'est produit dans le monde de la santé et des services sociaux, sous l'impulsion de M. Claude Castonguay. Pour le moment, certains auteurs expriment des opinions plutôt favorables, et deux d'entre eux n'hésitent pas à employer le mot «révolution»[5,6]. D'autres, par contre, se montrent sceptiques. L'un des critiques les plus virulents se demande même si l'on assiste vraiment à une réforme, ou si l'on n'est pas plutôt victime d'une «illusion»[7]. D'une extrême à l'autre, ces jugements ne peuvent être que provisoires, puisque personne ne prétendra évaluer une réforme présentement en plein essor comme si elle était déjà terminée.

Toutefois, il est un point sur lequel on s'accorde généralement. Malgré une certaine rhétorique sur le changement des comportements, qu'on trouvait dans quelques passages du Rapport Castonguay et qu'on retrouve dans quelques articles des lois subséquentes, la réforme aura été plus technocratique que populaire et elle aura modifié davantage les structures que la culture. La désillusion que certains manifestent devant l'évolution de cette réforme provient sans doute des vaines espérances qu'ils entretenaient quant à sa portée psychologique et sociale. D'ailleurs, la méthode juridique, qui a été adoptée pour l'application de cette réforme, n'était pas particulièrement propre à transformer les mentalités. Bien au contraire, elle supposait les mentalités déjà transformées, ce qui n'était pas le cas, sauf dans quelques milieux plutôt restreints. La conversion des esprits reste à faire et elle constitue précisément l'objet de la troisième réforme qui vient d'être initiée.

Modification des habitudes et de l'environnement

Le citoyen est maintenant protégé par l'*Etat payeur* et pris en charge par l'*Etat organisateur*. Grâce aux mesures de sécurité sociale, les services socio-sanitaires sont devenus gratuits (au moment de l'utilisation). Au point de vue géographique, grâce à la rationalisation du système, la main-d'oeuvre et les équipements sont mieux répartis. La continuité des soins est assurée par les liaisons à l'intérieur du réseau; la qualité des services est évaluée par divers systèmes de contrôle appliqués aux professionnels et aux établissements.

Lorsqu'on offre à une population des services gratuits, accessibles, continus et de qualité surveillée, faut-il s'étonner que la consommation de ces services monte en flèche, surtout si le public est de plus en plus informé des développements scientifiques les plus récents et si les professionnels ont une formation orientée vers le recours à la technique? Dans ces conditions, il est difficile de renverser les tendances adoptées, et par les consommateurs et par les fournisseurs, dans leur comportement. Pourtant, c'est à un tel renversement que nous invite M. Marc Lalonde dans un ouvrage intitulé *Nouvelle perspective de la santé des Canadiens*. La nouvelle réforme proposée dans cet ouvrage n'est ni *financière* ni *administrative*. On pourrait plutôt la qualifier de *philosophique* ou de *morale*[8]. Elle constitue non «un pas en avant» mais «un pas dans une autre direction», et peut-être même un retour en arrière vers l'exercice d'une plus grande responsabilité personnelle de la part de chaque citoyen quant à sa propre santé.

Les idées exprimées dans cet ouvrage ne sont pas nouvelles. Bien au contraire, elles sont propagées par les hygiénistes depuis Hippocrate, et on les retrouve d'ailleurs en partie dans le Rapport Castonguay, noyées dans la multitude des sujets qui y sont traités. Ce qu'il y a de nouveau quant à cette «Nouvelle perspective», c'est la position de son auteur, la clarté de son exposé, sa concision, sa force de démonstration.

L'ouvrage du ministre Lalonde semble écrit par un observateur qui réfléchit sur la santé des Canadiens, avec candeur mais aussi avec lucidité. «De quoi meurent les Canadiens?» se demande cet observateur. Un examen rapide des statistiques lui démontre que les principales causes de mortalité et de morbidité sont les maladies cardio-vasculaires, les accidents de la route et autres accidents, le cancer du poumon et autres maladies respiratoires, le suicide, la cirrhose du foie, etc. En rapport avec ces causes de décès ou d'invalidité, il n'est pas difficile à cet observateur de découvrir que les facteurs de risque sont généralement liés au *comportement individuel* et à l'*environnement* non seulement physique mais aussi social et culturel . . .

Il constate que «de puissantes forces adverses sont venues freiner les efforts visant à hausser le niveau de santé des Cana-

diens»: pollution de l'environnement, vie en milieu urbain, manque d'exercice, abus d'alcool, de tabac et de drogues, habitudes alimentaires axées davantage sur la satisfaction que sur les besoins[9]. Il en conclut «que l'action future visant à améliorer l'état de santé des Canadiens devra s'orienter davantage vers l'assainissement du milieu, la réduction des risques auxquels l'individu s'expose délibérément et la connaissance plus approfondie de la biologie humaine[10].

Ainsi prend forme ce que l'auteur appelle, d'après une expression proposée par Hubert Laframboise, la «conception globale de la santé» composée de quatre éléments: la biologie humaine, l'environnement, les habitudes de vie et l'organisation des soins. Le Ministre déclare que le gouvernement du Canada se propose dorénavant d'apporter aux trois premiers éléments autant d'importance qu'au quatrième[11].

L'objectif visé est double, encore une fois humanitaire et économique: «une vie meilleure, plus longue et plus saine, pour un nombre accru de Canadiens»[12], mais aussi «mettre un frein au taux annuel d'augmentation du coût des soins de santé, sinon la société ne pourra bientôt plus les supporter»[13]. L'originalité de ce nouveau projet de réforme réside non pas dans son objectif mais dans la méthode proposée, qui est de *changer les comportements*, d'abord en faisant appel à la *raison*, mais aussi en ayant recours aux techniques modernes d'*influence* et, dans certains cas, à la *coercition* législative. En d'autres mots, il s'agit d'informer, éduquer, persuader et, si nécessaire, réglementer.

Le ministère fédéral de la Santé a manifesté beaucoup d'initiative et de leadership dans la mise en marche de la nouvelle réforme. En 1976, soit deux ans après la parution de son livre, M. Lalonde faisait état de quelques «actions concrètes» déjà accomplies:

- efforts «en bonne partie couronnés de succès» en vue de rendre obligatoire le port de la ceinture de sécurité dans les automobiles, et en vue de réduire la vitesse sur les routes;
- campagne d'information et de discussion sur le problème de l'alcoolisme, sous le titre *Dialogue sur l'alcool;*
- efforts pour promouvoir l'activité et le conditionnement physique, comme le programme «Participaction» ou la mise en circulation de la «Physi-trousse» *(Fit-Kit);*
- dépliants publicitaires sur la santé, adressés à plus de 3,5 millions de foyers canadiens avec les chèques d'allocations familiales;
- enquête nationale sur la santé, qui doit être entreprise en 1978, pour mieux mesurer et évaluer les risques auxquels est soumise la santé des Canadiens[14].

Certaines provinces ont emboîté le pas dans ce projet de réforme des mentalités, en ce qui regarde la préservation de la santé. Outre la *Loi sur la sécurité des véhicules automobiles* incluant le port obligatoire de la ceinture de sécurité, on peut donner quelques autres exemples d'actions concrètes entreprises au Québec :

- concertation, dont toutes les modalités ne sont pas encore définies, entre gouvernement, employeurs et syndicats en vue de promouvoir la santé et la sécurité au travail;
- campagne en vue de l'épuration de l'air, qui a entraîné des interdictions de fumer dans plusieurs hôpitaux, et un projet de loi visant à «régir l'usage du tabac dans les endroits publics»[15];
- programme de publicité doté d'un budget de $800 000 et devant se manifester, sur une période de quatre mois, par plus de 2 500 messages à la télévision et à la radio et environ 500 annonces dans le métro de Montréal. Cette publicité porte surtout sur la nutrition mais aussi sur les dangers du tabac et des médicaments[16];
- campagne de «contre-publicité» pour sensibiliser la population au phénomène de la pollution, et allocation d'un montant de $ 200 000 à des groupes de citoyens pour produire des documents écrits ou audio-visuels sur ce sujet[17].

La réforme est donc bien engagée, du moins au niveau gouvernemental. Reste à savoir si «la base» va s'y intéresser, si les citoyens vont entrer dans la danse, si la transformation du milieu et des modes de vie va devenir une préoccupation commune, non seulement des gouvernants mais aussi des gouvernés . . . Sur ce point, le ministre Lalonde a souligné l'importante contribution qu'on peut attendre des associations sans but lucratif du monde de la santé, comme la Croix-Rouge par exemple, mais aussi, ose-t-il espérer, des sociétés commerciales comme, par exemple, les compagnies d'automobiles, d'alcool ou de médicaments . . .[18]. Avec un brin d'humour noir, le ministre déclare: «*If you can't lick'em, get'em to join you!*» . . .[19]

Outre ses chances de succès, deux autres aspects de cette réforme sont présentement l'objet de spéculations: d'abord l'aspect *éthique* quant aux moyens d'action proposés, puis l'aspect

philosophique quant aux objectifs poursuivis. Le premier à s'interroger sur ces questions est M. Lalonde lui-même. En 1976, il déclarait: «Je désire notamment m'arrêter aux écueils qui nous guettent lorsque nous cherchons à améliorer la condition humaine. Jusqu'où pouvons-nous aller dans le but d'améliorer cette condition humaine, et sur quelle philosophie pouvons-nous nous appuyer lorsque la responsabilité nous incombe d'agir à cet égard soit comme législateur, soit comme éducateur?»[30]

Quant à l'aspect éthique de l'intervention de l'Etat dans la vie des citoyens, il en était brièvement question dans *Nouvelle perspective*[31]. Le ministre Lalonde est, par la suite, revenu sur ce thème au cours de plusieurs conférences. D'une part, il rejette à la fois l'*autoritarisme* (prohibition totale du tabac ou de l'alcool) et la *manipulation* (un dosage approprié de récompenses et de punitions pour inciter les fumeurs ou les alcooliques à suivre une cure de désintoxication)[32]. D'autre part, il n'hésite pas à faire appel, dans son livre, à la notion de marché social dont les dangers, selon lui, ne sont pas plus grands que ceux du marché commercial. Il s'agit alors de «combattre le feu par le feu»[33], c'est-à-dire d'emprunter au monde des affaires ses propres techniques publicitaires pour lui riposter lorsqu'il menace notre santé. Alors on peut se demander qui sortira vainqueur de cette guerre d'influence dont le comportement du citoyen sera l'enjeu. Peut-on espérer que cette méthode d'éducation «par les contraires» sera propre à promouvoir l'autonomie personnelle?

Quant à l'aspect philosophique de cette réforme, il embrasse la conception qu'on peut se faire du sens de la vie, de l'idéal de la santé et des objectifs d'une politique de la santé dans une société démocratique et pluraliste. Sur ce sujet, le ministre Lalonde pose le principe suivant: «A la base de toute amélioration de la condition humaine, y compris la santé, doit se retrouver cette clarté et cette lucidité de l'esprit qui amènent un individu à prendre des décisions et à adopter un comportement favorisant le développement et l'épanouissement personnels.»[34] Comment réussir à atteindre l'*épanouissement personnel?* Les opinions peuvent varier sur ce sujet. Ces opinions sont liées aux aspirations des individus et des sociétés dont découlent également les attitudes quant à la santé et à la maladie.

Dans son ouvrage *Les rêves de la raison*, René Dubos écrit: «Quelque théorie qu'élaborent les médecins, les savants des laboratoires et les sociologues, c'est, bien entendu, la société qui doit décider des dangers auxquels elle désire le plus échapper et de la forme de santé qu'elle ambitionne — si elle attache plus de prix à la sécurité qu'à l'aventure, si elle est prête à compromettre l'avenir en échange du confort présent»[35]... «Santé et maladie sont des concepts... Leur signification est conditionnée par les exigences du milieu social et plus encore par les objectifs que l'individu s'impose à lui-même. Le laboureur, le pilote d'essai, le philosophe ne se fixent pas le même idéal de santé.»[36]

Comme le remarque M. Lalonde dans son livre, il faut «juger si certains risques ou certaines privations en valent la peine»... «Ce genre d'appréciation doit être effectué sur trois plans: d'abord sur le plan *individuel* quant à nos propres habitudes de vie, ensuite sur le plan *social* relativement aux valeurs prônées par la société et, enfin, sur le plan *gouvernemental* du point de vue des crédits affectés à la conservation de la santé et des restrictions que les gouvernements imposent à la population en tant que protecteurs du citoyen.»[37] Il importe de débattre ces sujets d'une manière démocratique. «Nous devons faire reposer nos décisions et nos interventions, déclare le ministre Lalonde, sur l'assentiment le plus large possible de la population... Nous devons nous appuyer sur l'accord de nos concitoyens... L'obligation nous incombe d'obtenir cet accord et d'éviter de l'obtenir par des formes de persuasion autoritaires ou manipulatrices.»[38] Il faudra donc prendre le temps de rechercher un consensus quant à l'art de vivre, à supposer qu'un tel consensus soit possible...

Réussira-t-on dans un avenir prévisible, en Amérique du Nord, par l'appel à la raison et la discussion démocratique appuyés par le marché social et la réglementation gouvernementale, à créer un «homme nouveau» *qui se sente responsable de sa santé*, comme il existe, paraît-il, en Chine ou à Cuba, et cela sans révolution, sinon celle des esprits? Quel défi à relever! □

Bibliographie

1. Rapport de la Commission d'enquête sur la Santé et le Bien-être social. Québec, Editeur officiel du Québec, 1970, Vol. 4, Tome 1, pp. 118-122.
2. BLAIN, Gilbert, «Le dilemme posé par le Rapport Castonguay sur la santé», *L'hôpital d'aujourd'hui*, Vol. XVII, n° 8, août 1971, p. 22.
3. Voir la liste des nouvelles lois dans *Le Médecin du Québec*, Vol. 10, n° 2, fév. 75, p. 21.

4. Pour des chiffres précis quant au nombre d'établissements des diverses catégories au Québec, voir le tableau statistique sur la répartition des établissements dans le *Répertoire des établissements de santé et de services sociaux*, Québec, M.A.S., 1977, p. 590.
5. CANNONE, François, «Au Québec, une révolution?», *Revue Autrement*, n° 9, mai 77, pp. 62-71.
6. PRESSAT, R. et LABERGE-NADEAU, C., «L'organisation des services de santé communautaire au Québec», *Concours Médical*, 21 mai 77, pp. 3531-3540.
7. RENAUD, Marc, «Réforme ou illusion? Une analyse des interventions de l'Etat québécois dans le domaine de la santé», *Sociologie et société*, Vol. 9, n° 1, avril 77, pp. 127-152.
8. LALONDE, Marc, *Nouvelle perspective de la santé des Canadiens*, Ottawa, ministère de la Santé nationale et du Bien-être social, 1974, p. 38.
9. Ibidem, p. 5.
10. Ibidem, p. 19.
11. Ibidem, p. 6.
12. Ibidem, p. 6.
13. Ibidem, p. 29.
14. LALONDE, Marc, «Notre politique de santé: le passé et l'avenir», Allocution au 18e Séminaire canado-américain, à Windsor, le 11 nov. 1976, texte miméographié, publié par le ministère de la Santé nationale et du Bien-être social, p. 12.
15. Anonyme, «Les fumeurs pourraient être la cible de mesures législatives», *Le Devoir*, 29 mars 78, p. 27.
16. Anonyme, «Stop smoking, eat wisely», *The Montreal Star*, 4 mars 78, p. A-12.
17. Anonyme, «Pollution: une seconde campagne de contre-publicité», *La Presse*, 16 mai 78.
18. LALONDE, Marc, Notes for an address by the Hon. Marc Lalonde to the meeting of the Red Cross Society, Toronto, Dec. 6, 1976, texte miméographié, publié par le ministère, pp. 6 et 9.
19. LALONDE, Marc, «Beyond a New Perspective», Notes for an address to the 104th annual meeting of the American Public Health Association, Miami, Oct. 18, 1976, texte miméographié, publié par le ministère, p. 7.
20. LALONDE, Marc, Notes pour une allocution de l'Hon. Marc Lalonde à la IXe Conférence internationale d'éducation pour la santé, Ottawa, 29 août 1976, texte miméographié, publié par le ministère, p. 1.
21. Voir référence n° 8, pp. 38-39.
22. Voir référence n° 14, pp. 13 et 14.
23. Voir référence n° 19, p. 7.
24. Voir référence n° 20, p. 2.
25. DUBOS, René, *Les rêves de la raison*, Paris, Denoël, 1964, p. 112.
26. Ibidem, p. 122.
27. Voir référence n° 8, p. 10.
28. Voir référence n° 14, p. 13.

Centre d'éducation médicale continue Université McGill et Centre hospitalier Douglas

La phychiatrie médico-légale: caractère confidentiel consentement et engagement

Le 3 Novembre 1978

au Pavillon Douglas du Centre hospitalier Douglas

Renseignements:
Centre d'éducation médicale continue

Université McGill
1110, avenue des Pins ouest
Montréal H3A 1A3

Régime d'assurance collective de la F.M.O.Q.

CE RÉGIME EST SOUSCRIT PAR LES PRÉVOYANTS DU CANADA ASSUREUR PRINCIPAL

- assurance-vie
- assurance-revenu en cas d'invalidité
- assurance pour frais de bureau
- mort accidentelle ou mutilation par accident

pour tout renseignement veuillez contacter:

bernard legault
& associés ltée
2860, boul. tracy
duvernay, québec
H7E 1L5
tél.: montréal 514-325-1472
québec 418-842-0854

OU

agence d'assurances cusson
fillion & associés inc.
5675, élie suite 103
st-hubert, québec
J3Y 6B3
tel.: montréal 514-656-6621
québec 418-824-4011

Le "réseau" des affaires sociales au Québec

selon

le rapport Castonguay et la réforme consécutive

Unité territoriale	Rapport Castonguay Vol. IV Sept. 70	Projet de loi Bill 65 (juin 71)	Loi adoptée Chapitre 48 (Déc. 71)	Règlement (en vigueur le 1 janv. 73)	Nombre
- REGION	ORS →	ORAS →	CRSSS		12
Région ou Sub-région	----------	CSS →	CSS		14
- ZONE	----------	----------	----------	DSC Note 1	32
------	CCS / CHU →	CH →	CH →	(CHCD	155
				(CHSP	91
------	----------	CA →	CA →	(-Garderie	377
				(-Transition	32
				(-Réadaptation	131
				(-Hébergement Note 2	402
- DISTRICT	CLS →	CLSC →	CLSC Note 3		108

Note 1: Certains D.S.C. ont des "modules" de service

Note 2: Il y a aussi les "pavillons" (10-30 lits) et les "familles d'accueil" (0-9 lits)

Note 3: Il y a aussi les "points de service" (succursales permanentes) et les "points de chute" (lieux temporaires de distribution de service par une équipe volante)

Département d'Administration de la Santé
Université de Montréal

Le déclin de l'État-providence?

63

par Gilbert Blain, m.d.

Gilbert Blain est professeur titulaire au Département d'Administration de la santé de l'Université de Montréal.

Ce texte a été présenté au Congrès du Conseil canadien de développement social tenu à Montréal le 21 juin dernier.

L'ÉTAT-PROVIDENCE N'A PAS CENT ANS

On célébrera, en 1983, le centième anniversaire de naissance de l'État-providence. En effet, c'est le 15 juin 1883 que fut adoptée la première loi sur l'assurance-maladie au monde. C'était en Allemagne. Cette loi fut suivie de deux autres: sur les accidents de travail en 1884 et sur l'assurance-vieillesse et invalidité en 1889. Ces trois lois devaient valoir au chancelier Bismarck le titre de *père* de l'État-providence.

En France, c'est en 1898 qu'on promulga une loi sur les accidents de travail. Trois autres lois furent ensuite passées: sur les retraites en 1910, sur les assurances sociales en 1930 et sur les allocations familiales en 1932. Le système actuel de sécurité sociale vit le jour en 1946 mais l'assurance-chômage ne débuta qu'en 1958.

En Grande-Bretagne, c'est en 1908 que Lloyd George faisait voter une loi sur l'assistance aux vieillards et, en 1911, une autre créant un système d'assurance-maladie et chômage. Après la guerre, l'application du plan Beveridge de 1942 devait assurer la relance du *Welfare State*[1].

Au Québec, la première mesure de sécurité sociale, adoptée par l'État, portait sur les indemnités aux accidentés du travail en 1909. Depuis cette date, les gouvernements provincial et fédéral (selon les cas, *l'un ou l'autre* agissant seul ou bien *l'un et l'autre* conjointement ou en parallèle) ont sans cesse contribué à l'expansion de l'État-providence.

Après 1909, les principaux programmes de sécurité sociale ont été instaurés au Québec dans l'ordre suivant: assistance-publique en 1921, assurance-chômage en 1940, allocations familiales en 1945, pensions de vieillesse en 1952, assurance-hospitalisation en 1961, régime de rentes à la retraite en 1966, aide sociale en 1969 et assurance-maladie en 1970.

DÉJÀ LE DÉCLIN?

Depuis cent ans, selon des modalités diverses et à une vitesse qui varie d'un pays à l'autre, l'État-providence s'est développé dans presque toutes les nations du monde occidental. Nulle part, pourrait-on croire, il n'a déjà atteint sa forme définitive mais partout, il est toujours allé s'amplifiant depuis un siècle.

Qui aurait pu penser que, dans ces pays, on en viendrait tout à coup à remettre en cause non seulement le rythme de croissance ou les modalités de fonctionnement de l'État-providence mais même certains des principes sur lesquels il repose? Pourtant, depuis quelques années déjà, les publications se multiplient dont les titres sonnent l'alarme quant à la santé de l'État-providence:

— L'État protecteur en crise[2]

— La crise de l'État-providence[3]

— L'État-providence démystifié[4]

— Le social, c'est fini[5]

Que se passe-t-il donc? On dirait que subitement l'État-providence est essoufflé. *Orwell se serait-il trompé?* se demande une journaliste française. *Contraitement à ce qu'imaginait l'auteur de "1984", le Big Brother occidental semblerait plutôt avoir tendance, dans les années qui viennent, à relâcher sa pesante sollicitude et à baisser les bras: Je suis désolé, les gars, mais je n'ai plus les moyens...*[6]

En fait, l'État-providence connaît présentement une double crise: financière et administrative.

CRISE FINANCIÈRE

La crise financière de l'État-providence s'explique de deux façons:

d'abord par le ralentissement (pour ne pas dire l'arrêt) de la croissance économique, ensuite par la révolte du contribuable.

- **Ralentissement de la croissance**

Il est reconnu que la généralisation du mieux-être dépend de la croissance économique et du plein emploi. Certains soutiennent même que la prospérité a fait davantage pour élever le niveau de vie que l'ensemble des mesures de redistribution du revenu adoptées par les États-providences. *Le progrès social*, écrit le politicologue Georges Burdeau, *est tributaire de l'abondance plus que de la justice.*[7] Ainsi, pendant les années de croissance économique, s'est établi ce que Pierre Rosanvallon appelle une *solidarité mécanique* des citoyens, c'est-à-dire que l'État-providence s'est développé sans douleur. Comme le gâteau grossissait, la part de chacun pouvait augmenter même si les inégalités persistaient.

Présentement, l'État n'arrive plus à boucler son budget. Ses recettes diminuent et ses dépenses augmentent. Vu les difficultés de l'industrie et du commerce, l'État perçoit moins de revenus, et en même temps, il doit débourser davantage pour les prestations d'assurance-chômage et d'aide sociale. Dans ce contexte de dépression économique, *la croissance ne joue plus son rôle de lubrifiant social.*[8] Le maintien de l'État-providence exigerait alors une véritable expression de solidarité entre les diverses couches de la société, des mécanismes plus audacieux de redistribution des revenus, un partage différent du gâteau qui ne grossit plus et qui, peut-être même, va devenir plus petit...

- **Révolte du contribuable**

Or, loin de consentir à un *nouveau contrat social*, les citoyens qui payent les impôts et ainsi portent le fardeau de l'État-providence, commencent à se rebiffer. Non seulement regimbent-ils contre les contributions exigées par l'État, mais ils ont recours à toutes sortes de moyens pour y échapper: travail au noir, production autonome, troc de marchandises ou échange de services, sans parler des astuces de comptabilité dans les déclarations au fisc. Bien plus, en 1978, en Californie, un contribuable rebelle, Howard Jarvis, réussit à faire diminuer les taxes par l'adoption de la célèbre Proposition 13. La même chose s'est produite en 1980 au Massachusetts avec la Proposition 2½. Au Québec, on a vu naître récemment un mouvement populaire de *taxaction* comparable à celui de la Californie...

Une question se pose donc maintenant: quelle est la limite de la solidarité sociale? A-t-on atteint, aux États-Unis et au Canada, un niveau intolérable d'imposition? Certes, il n'existe aucune *norme* sur ce sujet mais seulement des comparaisons. Un tableau publié par l'O.C.D.E.* présente *le pourcentage des recettes publiques (i.e. l'ensemble des impôts et des cotisations sociales) par rapport au P.I.B.* dans les pays membres en 1976[9]. Dans la majorité de ces pays, le pourcentage est plus élevé qu'aux États-Unis (29,3) ou qu'au Canada (32,9). Allemagne et Royaume-Uni: 36,7. France: 39,4. Sept pays au-dessus de 40. Deux au-dessus de 50... S'il apprécie les services de l'État-providence, un Canadien ne devrait-il pas accepter sans maugréer les hausses d'impôt requises pour leur maintien et leur expansion?

Pour raffermir l'État-providence dans les pays où il vacille, il faudrait un *mouvement de civilisation* car, comme le remarque Pierre Rosanvallon, *s'il y a des limites, elles ne peuvent être que sociétales et culturelles.*[10] Le problème, c'est que justement l'État-providence fait face, ces temps-ci, non seulement à une crise de financement mais aussi à une crise de confiance envers ses méthodes administratives.

CRISE ADMINISTRATIVE

La crise de confiance qui frappe l'État-providence porte sur deux points: son efficacité (atteint-il son but?) et son efficience (s'il atteint son but, à quel coût?).

- **L'État-providence est-il efficace?**

L'*égalité* de tous, la *sécurité* pour tous, la *solidarité* entre tous sont-elles réalisées dans les États-dits-providences? Certains auteurs se font fort de montrer que, dans ces pays, la *disparité*, l'*anxiété* et la *rivalité* persistent au point que, selon eux, l'État-providence n'est guère plus juste, sinon moins, que l'État libéral.

Tous ces auteurs ne nient pas les défaillances du marché, note Pierre Rosanvallon, *ce qu'ils tentent de démontrer, c'est qu'elles sont moindres que celles de la bureaucratie.*[11] *Le Welfare State*, écrit Henri Lepage, *tel qu'il fonctionne actuellement n'est bien souvent qu'une vaste escroquerie. Il faut montrer aux citoyens que tout ce qu'on leur vend au nom de la solidarité nationale aboutit trop souvent à des résultats très différents des intentions affichées...*[12]

Pour la distribution de certains services, on substitue, dans l'État-providence, le *marché politique* au *marché économique*. Quel que soit le domaine concerné (éducation, santé, logement, urbanisme, écologie, transport, etc...) l'État-providence profite surtout à ceux qui peuvent trouver leur chemin dans son labyrinthe bureaucratique et faire jouer selon leurs intérêts les mécanismes politiques. Ainsi, le *lobbying* remplace le *marketing*. Tant pis pour celui qui n'est pas relié à un groupe de pression ou à un mouvement social, ou qui ne sait pas avoir recours aux mass media ou aux manifestations populaires... Arrive-t-on ainsi à l'égalité entre les citoyens?

Comme l'État-providence se définit non seulement par ses programmes de redistribution sociale mais aussi par son interventionnisme économique (v.g. détaxation ou subventions pour les entreprises dans certains secteurs *mous*), les groupes de producteurs et les compagnies commerciales entrent dans la ronde autour de l'État, avec les groupes de consommateurs et les associations

* L'O.C.D.E. (Organisation de Coopération et de Développement Économiques) est une association qui regroupe environ 25 pays industrialisés (Europe de l'Ouest, Scandinavie, Amérique du Nord, Japon, Australie...)

d'individus. *Le gouvernement devient l'instrument de promotion d'intérêts particuliers*[13] et les transferts de revenus sont faits en faveur de minorités habiles et actives aux dépens de la majorité silencieuse et inerte. Ainsi, contrairement au but poursuivi, le système favorise souvent les producteurs plutôt que les consommateurs...[14]

- **L'État-providence est-il efficient?**

Dans les cas où l'État-providence atteint son but, le coût des services fournis est-il concurrentiel avec celui des mêmes services que l'entreprise privée pourrait fournir? Sur ce sujet, certains critiques ont entrepris récemment un impitoyable procès de l'État-providence dont ils attaquent non pas les bases philosophiques mais l'inefficience bureaucratique. Deux économistes, J.-Luc Migué[15] et Henri Lepage[16], multiplient les exemples dans divers secteurs économiques et sociaux, en vue de démontrer que les biens et services produits par le secteur public sont produits *à des coûts excessifs*[17].

Le cas le plus surprenant est sans aucun doute celui des *pompiers privés*.[18] Pour sa protection contre les incendies, une ville de l'Arizona a conclu un contrat avec une firme commerciale de pompiers. On rapporte que la qualité du service n'a pas été altérée.[19] D'autre part, les calculs d'un analyste ont démontré qu'avec les pompiers privés, *le service qui revenait, au moment de l'étude, à 3,78 $ par habitant, coûterait 7,10 $, soit le double, s'il était confié à une administration publique type*.[20]

L'exemple des pompiers et plusieurs autres similaires entraînent Migué à formuler une terrifiante *règle de deux: en moyenne, il en coûterait environ deux fois plus cher de produire dans le secteur public et parapublic ce que le marché peut offrir*.[21] Il exagère? Peu importe. Il n'en faut pas tant pour provoquer une crise de confiance...

PERSPECTIVES... DE *GAUCHE* OU DE *DROITE*

Paralysé par la conjoncture économique, critiqué à la fois par les radicaux de gauche qui en font la *béquille du capitalisme* et par ceux de droite qui voient en lui un *ersatz du socialisme*, l'État-providence passe certes un mauvais moment.

Serait-il tombé en discrédit? Là-dessus, les opinions exprimées à une Conférence organisée par l'O.C.D.E. en 1980 diffèrent sensiblement. *À l'heure actuelle, l'idée dominante serait plutôt que la progression continue de l'État protecteur n'est ni probable, ni même désirable*, déclare un représentant de la France.[22] Un professeur de l'Université de Californie lui réplique: *L'hypothèse qui a servi de point de départ à nombre de rapports soumis à la Conférence, et suivant laquelle l'État protecteur serait en train de perdre la faveur du public, est erronée*.[23]

DEMANDE D'ADHÉSION À LA FASSQ

Nom Directorat

Institution

Adresse de correspondance Téléphone

Je désire recevoir de la documentation sur la FASSQ ☐

Je désire devenir membre de la Fédération en 1982 ☐

Ci-joint un chèque de $250.00 ☐

Ci-joint 5 chèques de $50.00 ☐

Je désire que ma cotisation soit retenue à la source ☐

VEUILLEZ ADRESSER VOTRE DEMANDE À LA

Fédération des Administrateurs des Services de Santé et des Services Sociaux du Québec

4237, rue de Bordeaux Montréal, H2H 1Z4

Qui croire? Et surtout que faire maintenant? En simplifiant un peu grossièrement, on peut examiner l'avenir selon deux perspectives: l'une qu'on pourrait dire *de gauche* (encore plus d'intervention étatique, de planification, de réglementation...), l'autre, *de droite* (retour partiel à l'entreprise privée, à la concurrence, au laisser faire...).

Le tableau intitulé *ENJEUX ET PERSPECTIVES EN 1982* présente côte-à-côte les solutions qu'on peut proposer aux problèmes qu'affronte présentement l'État-providence selon que l'on adopte la perspective de gauche ou celle de droite. Tous les points de comparaison de ce tableau mériteraient certes quelque débat mais un seul retiendra notre attention pour le moment.

LE RATIONNEMENT

Comme on le voit au bas du tableau, on peut envisager des mécanismes de rationnement dans l'une ou l'autre des perspectives.

Rationnement ne veut pas nécessairement dire privation mais plutôt *répartition de quantités déterminées et parfois limitées*. L'origine étymologique du mot est *ratio* au sens de mesure. Le rationnement idéal assurerait à chacun sa juste mesure, à chacun son compte. La privation découle de la rareté.

Le rationnement a toujours existé. Dans la société libérale, les lois du marché appliquent un rationnement non planifié qui dépend des moyens et de la volonté de payer de chacun selon les prix demandés. Dans l'État-providence, les règles de la bureaucratie appliquent un rationnement délibéré et organisé selon des critères théoriquement rationnels.

LA CRISE DE L'ÉTAT-PROVIDENCE:
ENJEUX ET PERSPECTIVES EN 1982

ENJEUX	PERSPECTIVE *de gauche*	PERSPECTIVE *de droite*
Financement	— augmentation des impôts et des cotisations sociales	— frais modérateur par service — franchise annuelle
Couverture des besoins et des risques	— universelle et complète le plus possible	— sélective quant aux personnes — incomplète quant aux services
Orientation des réformes	— unification des services en un seul système sous administration étatique	— transfert au marché de certains risques — *privatisation* de certains services
Évolution de l'État	— croissance — technocratisation — bureaucratisation — enregimentation des citoyens	— décroissance — encouragement aux groupes privés et bénévoles — autonomisation des citoyens
Rôle des experts en administration	— établir des normes et des moyens de mesurer la production, l'utilisation, la rationalisation...	— trouver des moyens d'améliorer la concurrence, la productivité, le rendement...
Comment augmenter l'efficience?	— planification coercitive — réglementation accrue — contrôles	— concertation volontaire — déréglementation — incitatifs
Moyens de limiter les dépenses	— contingentement de l'offre	— découragement de la demande
Mécanismes de rationnement	— règles politiques — queues, listes d'attente	— règles du marché — choix individuels

On pouvait lire, récemment, dans le magazine *Perspectives* du journal *La Presse*, un appel au public en faveur de la Banque d'yeux du Québec qui n'avait en réserve, à ce moment-là, que 50 cornées alors que plus de 1000 personnes attendaient une greffe.[24] Comme les cornées ne sont pas vendues, l'allocation (i.e. le rationnement) doit certainement se faire selon certaines priorités médicales et administratives. Vu le manque de cornées, plusieurs patients en sont privés, du moins temporairement...

Le rationnement, selon la perspective de gauche, est-il plus juste, plus facile à supporter que le rationnement selon la perspective de droite? Chacun peut donner à cette question une réponse différente selon son idéologie mais aussi selon son pragmatisme. Les uns professent une idéologie jusqu'à être prêts à en tolérer tous les inconvénients. Les autres tempèrent leur idéologie et veulent prendre en considération l'efficacité et l'efficience d'un système avant de porter un jugement plus nuancé.

Nous connaîtrons sans doute, au cours des années qui viennent, un rationnement plus sévère des services de l'État mais non, espérons-le, la privation de ces services lorsqu'ils seront nécessaires. S'il fallait que cette éventualité se produise, le déclin de l'État-providence nous forcerait à inventer de nouvelles formules en vue d'instaurer une société qui soit plus efficace en même temps que plus juste.

Le maintien ou le renouveau des mesures sociales dépendra certes, dans l'avenir, des possibilités de relance économique mais aussi et peut-être surtout de la force du sentiment de solidarité entre les citoyens, de l'imagination des planificateurs sociaux et du courage des hommes politiques. ●●

RÉFÉRENCES

1 Pour un bref historique des origines de l'État-providence, voir: Rosanvallon, Pierre. *La crise de l'État-providence*. Paris, Seuil, 1981, pp. 141-158.

2 En collaboration. *L'État-protecteur en crise*. Rapport de la Conférence sur les politiques sociales dans les années 80. Paris, O.C.D.E., 1981.

3 Voir référence no 1.

4 Lepage, Henri. *Demain, le capitalisme*. Paris, Librairie Générale Française, 1978. Chapitre 4.

5 Vingré, Michel. *Le social, c'est fini*. Paris, Collection Autrement, 1980.

6 De Kervasdoué, J. et coll. *La santé rationnée?* Paris, Éd. Economica, 1981, p. 214.

7 Burdeau, Georges. *L'État*. Paris, Seuil, Collection Politique, 1970, p. 160.

8 Voir référence no 1, p. 31.

9 Ibidem, p. 187.

10 Ibidem, p. 18.

11 Ibidem, p. 61.

12 Voir référence no 4, p. 423.

13 Migué, J.-Luc. *L'économiste et la chose publique*. Québec, Les Presses de l'Université du Québec, 1979, p. 154.

14 Voir référence no 4, p. 234 et no 13, p. 167.

15 Voir référence no 13.

16 Voir référence no 4.

17 Voir référence no 13, p. 233.

18 Voir référence no 4, p. 277 et no 13, p. 181.

19 Voir référence no 4, p. 278.

20 Voir référence no 13, p. 181.

21 Voir référence no 13, p. 184.

22 Voir référence no 2, p. 175.

23 Voir référence no 2, p. 219.

24 *Magazine* Perspectives, Vol. 24, No 5, 30 janv. 82, p. 12.

susceptibles d'influencer le malade à reprendre une place active dans son milieu de travail après avoir subi une chirurgie coronarienne.

Résumé

Parmi les 213 patients opérés à l'Hôpital du Sacré-Cœur pour pontage aorto-coronarien, une estimation conservatrice montre que 72% sont retournés au travail, après un délai médian postopératoire de 4 mois. Le sexe et le type d'angine (stable ou instable) avant l'opération ne sont pas des facteurs déterminants, tandis que l'âge exerce une influence légère mais non significative dans notre série. La proportion relativement élevée de patients au travail résulte possiblement de la sélection des patients ainsi que de l'attitude positive de l'équipe médico-chirurgicale à cet égard.

Summary

Out of 213 patients operated at Hôpital du Sacré-Cœur for aortocoronary bypass, a conservative estimate shows that 72% had returned to work, after a median post-operative period of 4 months. Sex and type of angina (stable or unstable) preceeding the operation do not appear as determinating factors, whereas age exerts a slight but non-significant influence in our series. The relatively high proportion of patients at work may result from patient selection and a positive attitude of the medico-surgical team with regard to a return to active life.

REMERCIEMENTS

Nous remercions Mlle Henriette Tenaille pour nous avoir communiqué ses résultats non publiés; Mlle Francine Désourdy et Mme Colette Roy pour leur aide précieuse; ainsi que la Fondation du Québec des maladies du cœur pour son support financier.

BIBLIOGRAPHIE

1. Pelletier et coll.: Entretiens de cardiologie de langue française, Montréal, 1975.
2. Voir article précédent: David, Tenaille, Blain et Tremblay, et communication personnelle.
3. Communication personnelle.

ÉLÉMENTS ET ÉTAPES D'ÉLABORATION D'UN PROGRAMME DE SANTÉ COMMUNAUTAIRE

Raynald PINEAULT, Ph.D. *

La gestion par programme n'est pas un concept nouveau en administration. Déjà en 1924 General Motors l'utilisait. Mais ce n'est véritablement que depuis la dernière guerre mondiale que cette méthode de gestion a connu un essor considérable aux États-Unis et au Canada. En effet vers les années 1945, le ministère de la Défense nationale américain, à l'instigation du secrétaire Robert McNamara, cherchait à créer un lien entre la planification et l'établissement du budget. Cet effort a conduit au développement d'une méthode de gestion que l'on connaît aujourd'hui sous le nom de PPBS (Planning-Programming-Budgeting System) et qu'on peut traduire en français par rationalisation des choix budgétaires. Cette méthode consiste essentiellement à relier les ressources aux objectifs à atteindre, à mesurer les résultats obtenus et à ajuster en conséquence soit les ressources, soit les objectifs. Adopter la méthode de gestion par programme implique donc deux tâches permanentes: (1) la définition d'objectifs clairs et précis, et (2) une analyse de tous les objectifs, sous-objectifs, activités et ressources ainsi que des relations entre ces divers éléments.

Le champ de la santé communautaire se prête particulièrement bien à ce genre d'analyse. D'une part, les objectifs définis en termes de modification de l'état de santé ou du moins du comportement relatif aux soins de prévention sont le plus souvent identifiables et quantifiables. De plus, bien que les activités impliquées dans la réalisation de ces objectifs soient multiples, elles ne sont pas complexes au point de rendre impraticable une analyse systémique de ces activités et de leurs relations mutuelles.

Voyons donc maintenant quels sont les éléments d'un programme de santé communautaire et quelles sont les étapes de son élaboration.

PROGRAMME ET SERVICES

Précisons tout de suite que nous faisons la distinction entre, d'une part, l'organisation des services de santé pour répondre aux besoins courants d'une

* Professeur adjoint, département d'administration de la santé, Université de Montréal.

Toute demande de tirés à part doit lui être adressée à l'Université de Montréal.

population et, d'autre part, l'exécution d'un ensemble cohérent et intégré d'activités en vue d'atteindre un objectif déterminé. Dans ce dernier cas, on peut vraiment parler de programme, c'est-à-dire d'activités organisées que l'on se propose d'accomplir en vue d'arriver à un résultat défini pour une population déterminée. Cette définition met également en relief les éléments fondamentaux d'un programme, à savoir: (1) les résultats ou objectifs qu'on veut atteindre; (2) les activités à accomplir et les ressources à mobiliser pour réaliser ces objectifs; et (3) la population pour laquelle le programme est conçu.

ÉTAPES D'ÉLABORATION DU PROGRAMME

Les étapes que le planificateur suit dans l'élaboration d'un programme sont représentées à la Fig. 1.

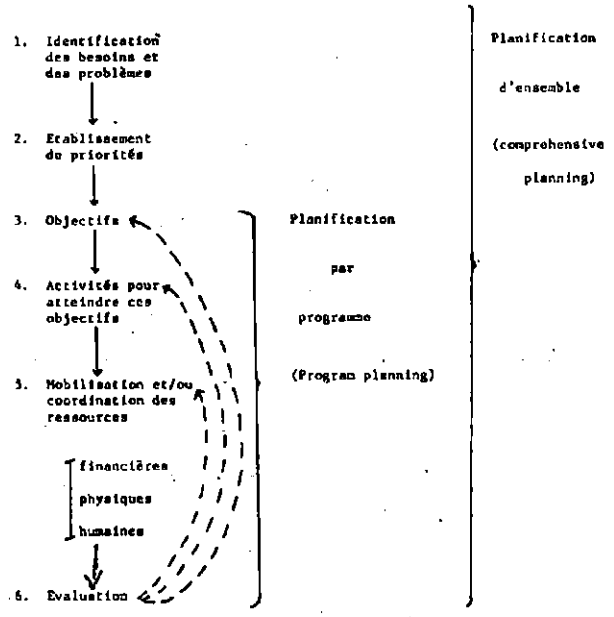


Fig. 1 — Etapes de la planification des programmes de santé.

Il faut noter ici que les deux premières étapes, identification des problèmes et détermination des priorités, constituent à proprement parler, des étapes préliminaires à la planification par programme (program planning) [1]. En d'autres termes, elles constituent le cadre général dans lequel les priorités entre les divers problèmes identifiés et les programmes qu'ils commandent peuvent être déterminées. Ainsi, la planification par programme se situe dans le cadre plus vaste de la planification d'ensemble (comprehensive planning). Ces étapes préliminaires sont néanmoins nécessaires si l'on veut que la justification d'un programme de santé repose sur une base solide. Les autres étapes présentées à la Fig. 1 correspondent aux éléments d'un programme de santé mentionnés plus haut.

Nous allons maintenant discuter en détail chacune de ces étapes.

1 — IDENTIFICATION DES BESOINS ET DES PROBLÈMES

Dès la première étape, le planificateur se voit confronté avec la notion complexe de besoin. Les besoins d'une population peuvent différer considérablement selon qu'ils sont définis et identifiés par les individus eux-mêmes, par les professionnels de la santé ou par un organisme public comme le ministère des Affaires Sociales. Ces trois points de vue peuvent même parfois être incompatibles. Le problème survient surtout lorsque les points de vue des professionnels ou des organismes politiques dominent sur celui des individus. De cette mauvaise identification des besoins réels peuvent résulter des programmes de santé qui sont d'avance voués à l'échec parce qu'ils ne correspondent pas aux problèmes que les individus de cette population considèrent comme prioritaires. Bref, agir en fonction des besoins de la population peut s'avérer un slogan rentable sur bien des plans, mais ne semble pas toujours refléter les besoins réels des individus.

Dans l'identification des problèmes de santé, deux grandes approches peuvent être utilisées par les planificateurs [2]. Une première approche se base sur la notion épidémiologique du besoin mesuré par l'état de santé de la population. Diverses mesures de mortalité et de morbidité peuvent alors être utilisées comme indicateurs sanitaires pour estimer les besoins de santé d'une population. Évidemment, cette approche correspond plutôt à une perspective professionnelle des besoins.

Une seconde approche possible dans l'étude du besoin correspond à la notion psychosociologique des besoins. Essentiellement, cette démarche consiste à identifier les problèmes d'une population et à déterminer leur importance à partir de la perception qu'ont les individus de leurs problèmes. Toutefois, ce type d'enquête qui suppose une base de recherche est coûteux et laborieux. À cause de ces contraintes, il est souvent impossible d'entreprendre de tels sondages. C'est l'argument que les planificateurs emploient bien souvent pour justifier une approche normative et professionnelle dans l'identification des problèmes et la détermination des priorités. Ils oublient que le recours à des sondages scientifiques n'est pas toujours nécessaire pour identifier les problèmes d'une population. La consultation avec les organismes représentatifs et la participation des individus à l'identification de leurs problèmes sont aussi des moyens efficaces pour

s'assurer que les besoins identifiés correspondent à des préoccupations réelles de la population.

Après avoir décidé d'une approche de la détermination des besoins, il faut définir la population pour laquelle les programmes seront établis. La population-cible doit être définie selon les variables de sexe, d'âge, de géographie et de niveau socio-économique.

En second lieu, les problèmes sont identifiés selon l'approche ou les approches choisies. Il faut tenter de déterminer la relation entre les divers problèmes identifiés et les facteurs éventuellement en cause. Cette démarche est très importante puisqu'elle détermine jusqu'à un certain point la validité des programmes de santé qui seront établis. Il n'est pas toujours facile d'expliquer pourquoi l'incidence d'une maladie, par exemple, est plus élevée dans une région que dans les autres. Pour porter des jugements éclairés, on pourra recourir à des experts. De même, les résultats de certaines recherches pertinentes peuvent se révéler utiles à cette fin.

Au terme de cette démarche, il se peut que pour des problèmes de santé jugés importants, les solutions les plus susceptibles d'améliorer la situation se trouvent en dehors du système de santé. Dans ce cas, même si ces solutions ne tombent pas sous la responsabilité du planificateur de santé, il faut que celui-ci intervienne dans les autres secteurs pour tenter de sensibiliser d'autres organismes aux problèmes et demander leur collaboration et leur participation aux programmes qui seront développés. Il en est bien souvent ainsi des programmes de santé pour la population d'âge scolaire.

2 — ÉTABLISSEMENT DES PRIORITÉS

Après avoir identifié les problèmes, il reste à déterminer les priorités en matière d'intervention. Suivant la perspective épidémiologique, on peut poser en principe que les problèmes de santé qui causent les plus grandes pertes de vie productive et qui sont susceptibles d'être améliorés ou éliminés par des programmes de santé commandent une action prioritaire. En pratique, deux critères peuvent donc être utilisés pour déterminer les priorités d'intervention: (1) l'importance de la perte résultant de la présence d'une maladie et (2) la sensibilité de cette maladie à un programme de santé [8].

À la base de la méthode proposée de pondération des priorités se trouve l'idée que la protection sanitaire a pour objet: (1) de réduire ou, si possible, d'éliminer le risque existant, et (2) de rétablir autant que possible la santé compromise avec les moyens disponibles en redonnant au malade sa place dans la société. L'importance de la perte causée par la maladie peut s'exprimer soit par le raccourcissement de la vie résultant de la maladie ou par le degré d'invalidité causé par un processus morbide chronique. Il est à noter que l'emploi de l'une ou de l'autre de ces mesures peut changer la pondération des priorités. Le critère, perte résultant du décès, attribue évidemment une importance plus grande aux maladies mortelles de l'enfance qu'à celles de l'âge avancé. De plus, ce même critère attribue peu d'importance aux maladies chroniques qui causent un grand degré d'invalidité sans toutefois être mortelles. Chacun de ces gains sanitaires, résultant de la baisse de la mortalité et de la morbidité étant qualitativement différent de l'autre, il est difficile de les additionner directement, même s'ils sont exprimés dans les mêmes unités de mesure. En d'autres termes, cela reviendrait à dire qu'une année de gain résultant de l'élimination d'une maladie mortelle équivaut à une année de gain résultant de la réduction d'invalidité causée par les maladies chroniques. Le coefficient dont il faut affecter chacune de ces mesures doit donc être déterminé par des jugements de valeur correspondant à certaines valeurs fondamentales d'une société.

Le deuxième critère employé dans la pondération des priorités est la sensibilité de la maladie à un programme de soins. Ce critère peut être déterminé de plusieurs façons: (1) normativement, en se basant sur des jugements d'experts et de professionnels de la santé; (2) empiriquement, en se basant sur les expériences réalisées dans d'autres pays ou régions; et enfin, (3) de façon opérationnelle, en estimant la quantité et surtout le coût des ressources qui devront être mobilisées pour obtenir les gains sanitaires escomptés. Cette dernière méthode correspond essentiellement à l'analyse rendement-coût. Ainsi, plus le coût d'un programme pour atteindre un objectif déterminé est bas, plus le problème est sujet à une intervention efficace.

En principe, on peut donc identifier quatre types de décisions pouvant résulter de l'interaction des deux critères, importance du gain sanitaire et sensibilité. Ces quatre possibilités sont représentées au tableau suivant.

Comme l'indique ce tableau, une priorité faible pour l'établissement d'un programme de soin peut constituer une priorité élevée pour un programme de recherche dans le cas où les gains sanitaires à obtenir sont considérables. À titre d'exemple, le tableau suivant présente certaines maladies classées en fonction du coût et de l'efficacité d'un programme de santé relatif à ces maladies.

TABLEAU I

PONDÉRATION DES PRIORITÉS DANS LES PROGRAMMES DE SANTÉ SELON L'IMPORTANCE DE LA PERTE ET LA SENSIBILITÉ

Importance de la perte résultant d'une maladie	*Sensibilité*	*Décision*
Considérable	Elevée	Priorité 1 pour l'établissement du programme
Considérable	Faible	Priorité 1 pour la recherche
Peu considérable	Elevée	Priorité 2 pour l'établissement du programme
Peu considérable	Faible	Priorité 2 pour la recherche Priorité la plus faible pour l'établissement du programme

Source: Donabedian, A. *Aspects of Medical Care Administration: Specifying Requirements for Health Care.* Harvard University Press, 1973, p. 170.

TABLEAU II

EXEMPLES DE MALADIES CLASSÉES SELON LE COÛT ET L'EFFICACITÉ DU TRAITEMENT

Coût	*Gain*	
	Elevé	*Faible*
Faible	Diabète sucré Appendicite Poliomyélite Chorio-épithéliome	Arthrose déformante Accident cérébro-vasculaire Emphysème pulmonaire
Elevé	Infarctus coronarien dans le cas de soins intensifs	Lésion transverse Sclérose en plaque
	Greffe du rein	Chirurgie à cœur ouvert Imbécillité Leucémie

Source: O.M.S. *Méthodes modernes de gestion et organisation des services de santé;* Cahiers de Santé publique No 55. Genève, 1974, p. 88.

L'ordre de priorité entre les quatre catégories de maladie va évidemment de « gain élevé-coût faible » vers « gain faible-coût élevé ».

La méthode que nous avons proposée pour pondérer les priorités des programmes de santé est évidemment une simplification de la réalité. Les situations concrètes sont souvent fort complexes et les critères utilisés pour la prise de décision difficilement quantifiables. Il est donc important de mettre en garde contre ce qu'on pourrait appeler le « réductionisme » épidémiologique. En effet, l'importance des problèmes de santé ne doit pas être déterminée par la seule possibilité de quantifier ces phénomènes et de mesurer leur impact sur l'état de santé. Il est évident que l'impact de certains programmes de santé peut être mesuré par une modification de certains indicateurs sanitaires. Cependant, tel n'est pas le cas pour tous les programmes de santé. Par exemple, il est difficile d'établir avec précision l'effet qu'un programme visant à faciliter l'accès aux services de première ligne peut avoir sur l'état de santé d'une population. Néanmoins, le problème d'accessibilité aux soins peut être une priorité pour certaines populations. Bref, l'approche proposée ne constitue qu'un outil d'analyse et un cadre de référence qu'il faut utiliser avec discernement et relativité.

3 — DÉTERMINATION DES OBJECTIFS

Une fois le choix de certaines priorités d'intervention arrêté, il faut définir les objectifs qu'on veut atteindre. Cette partie comprend d'abord une définition du problème que le programme vise à résoudre comme, par exemple, la prévention de la tuberculose, la diminution des accidents de la route ou la réduction de la mortalité périnatale.

En second lieu, la définition des objectifs comprend une définition précise de la population visée par le programme. Dans certains cas, la population-cible du programme correspondra à toute la population dont l'organisme de santé, D.S.C. ou C.L.S.C., est responsable. Pour la plupart des programmes, cependant, la population-cible ne comprendra qu'un groupe d'âge spécifique de cette population.

La détermination des objectifs est compliquée par la grande confusion qui règne au sujet du concept même d'objectif. La terminologie dans ce domaine est variée. Par exemple, les termes objectifs, mission et buts sont employés comme synonymes. De plus, certains auteurs établissent une distinction très nette entre fins et moyens, conformément au sens courant de ces termes, alors que Simon, un théoricien de l'organisation, souligne qu'ils constituent une succession dans laquelle chaque étape est considérée comme l'objectif de la précédente et le moyen de la suivante [4]. Selon cet auteur, la distinction entre fins et moyens est artificielle. Cependant, encore faut-il que les moyens dont il est question correspondent à des étapes intermédiaires dans la poursuite d'un état final et non à des activités qu'il faut accomplir pour atteindre chacune de ces étapes. Autrement, le fait de ne pas distinguer objectifs et activités ne fait qu'augmenter la confusion.

Il n'y a donc pas de règle fixe et immuable pour établir une distinction entre ces termes. L'important c'est de définir de façon précise les termes utilisés. À ce sujet, Deniston et coll. proposent une ligne de conduite qui paraît utile à suivre [5]. Ces auteurs suggèrent qu'à tous les niveaux dans l'élaboration d'un programme de santé, les objectifs représentent un état ou un comportement des individus qui com-

posent la population-cible. Les autres états qui correspondent aux comportements des professionnels de la santé participant au programme devraient être désignés comme des activités. Bref, selon cette formulation, les moyens correspondent aux activités et les objectifs aux états à atteindre chez la population-cible.

Les objectifs ou états à atteindre peuvent ensuite être organisés de façon hiérarchique. On a l'habitude en administration de répartir les objectifs en trois classes: (1) l'objectif général ou ultime, (2) les objectifs intermédiaires ou sous-objectifs et (3) les objectifs spécifiques ou opérationnels. L'important ici n'est pas la terminologie mais plutôt l'analyse systématique qu'exige l'organisation hiérarchique des objectifs.

Pour illustrer le principe de l'organisation hiérarchique des objectifs, citons l'exemple que présentent Greenberg et Mattison. Il s'agit d'un programme visant à réduire la morbidité et la mortalité périnatales [6].

Voici comment les objectifs pour les différentes étapes du programme sont énoncés;

1. réduction du taux de morbidité et de mortalité de X%;

2. x proportion de femmes qui devraient suivre tout le programme pendant leur grossesse et atteindre certaines normes quant à leur régime, leur poids, les connaissances acquises sur la physiologie de la grossesse et de l'accouchement et sur les soins à donner aux nouveau-nés;

3. x proportion de femmes qui sans atteindre les normes établies, devraient modifier leur comportement;

4. x proportion de femmes qui devraient modifier leurs attitudes;

5. x proportion de femmes qui devraient améliorer leurs connaissances;

6. x proportion de femmes qui devraient suivre le programme;

7. x proportion de femmes qui devraient connaître l'existence du programme;

8. x proportion de femmes qui devraient recevoir l'information relative au programme.

Cet exemple est très détaillé. Évidemment, comme le souligne Rochon, la prévision des détails d'un programme au-delà d'un certain point peut devenir un exercice purement intellectuel [7]. D'autre part, l'exemple cité permet de comprendre l'importance de l'analyse systémique d'un programme dans lequel la réalisation d'un objectif dépend de celle de l'objectif qui se trouve à un niveau inférieur et contribue également à la réalisation d'un objectif supérieur. L'importance d'une telle analyse apparaîtra plus clairement lors de l'évaluation du programme.

4 — DÉTERMINATION DES ACTIVITÉS ET DES RESSOURCES

Cette étape comprend: (1) l'identification des activités qui devront être entreprises pour réaliser les objectifs, et (2) la détermination des ressources qui seront affectées au programme.

Les principales activités à remplir dans un programme de santé communautaire peuvent être classées de la façon suivante:

a) activités d'éducation sanitaire: séances d'information, dissémination de l'information par divers media, etc.

b) activités de prévention: immunisations, conditionnement physique, etc.

c) activités de dépistage: examens pour dépister certaines maladies chez des individus présentant un risque élevé, etc.

d) activités de diagnostic et de traitement.

Pour chaque type d'activités, il convient ensuite de spécifier la quantité et la qualité des ressources qui y seront affectées. Ces ressources peuvent être:

a) humaines: effectif requis et proportion du temps consacré aux diverses activités;

b) financières: estimation du budget et prévision des sources de paiement possibles;

c) matérielles: équipement, locaux, etc.

Enfin, les modalités administratives du programme doivent être prévues, comme par exemple, les mécanismes de coordination entre les différents services et les différentes catégories de personnel.

5 — ÉVALUATION

L'étape de l'évaluation est très importante puisque tous les éléments du programme y sont remis en question et jugés en fonction de certains critères. Ces critères ou normes peuvent être définis de façon empirique en se basant, par exemple, sur les expériences rapportées dans d'autres milieux. Ils peuvent aussi être définis de façon normative, à partir des jugements de professionnels et d'experts. Dans les deux cas il faut s'assurer de la fiabilité et de la validité des critères utilisés. [8].

Comme le montrent la Fig. 1 et le Tableau III, l'évaluation d'un programme peut comprendre au moins trois aspects qui correspondent aux objectifs,

TABLEAU III

MESURES D'ÉVALUATION ET MÉTHODES DE COLLECTE DE L'INFORMATION SELON LES DIFFÉRENTS TYPES D'ÉVALUATION

Eléments du programme	*Objectifs*			*Activités*	*Ressources*
Type d'évaluation	Résultats (outcomes)			Déroulement (Process)	Structuration (structure)
	Rendement (efficacy)	Efficacité (effectiveness)	Efficience (efficiency)		
Indices et mesures d'évaluation	Nombre d'individus rejoints par le programme	Impact du programme sur la population-cible, e.g. proportion des individus rejoints	Rapport entre les résultats et les ressources déployées	Méthodes et techniques utilisées	Quantité et qualité des ressources humaines, matérielles et financières
Méthodes de collecte de l'information		Epidémiologiques, psychosociologiques, sociologiques		Administratives	Comptables

aux activités et aux ressources du programme. De plus, comme le montre le Tableau III à chaque aspect de l'évaluation correspondent des indices spécifiques de mesure et des méthodes particulières de collecte de l'information. Examinons chacun de ces types d'évaluation de plus près [9].

1 — ÉVALUATION DE LA STRUCTURATION DES RESSOURCES

Ce premier type d'évaluation mesure la quantité et la qualité des ressources qui ont été mobilisées. Une question type consiste à déterminer s'il y a du personnel qualifié en nombre suffisant. Cette méthode d'évaluation ne renseigne pas sur le degré de réussite du programme. Elle fournit cependant des indications sur certaines déficiences qui peuvent être la cause de l'échec partiel ou même total du programme.

2 — ÉVALUATION DU DÉROULEMENT DES ACTIVITÉS

L'évaluation du déroulement vise les activités qui ont été accomplies pour atteindre les objectifs du programme. Elle s'attache plus précisément aux méthodes utilisées pour réaliser ces objectifs. Par exemple, a-t-on tenu un nombre suffisant de séances d'information? A-t-on fait un nombre suffisant de visites? Est-ce que le contenu de ces séances et de ces visites était adéquat? Tel est le type de questions que comporte l'évaluation du déroulement. Cette méthode peut donc mettre en évidence les raisons directes de succès ou d'échec d'un programme.

3 — ÉVALUATION DES RÉSULTATS (OUTCOMES)

Dans l'évaluation des résultats, l'analyse porte sur les objectifs du programme et vise à déterminer le degré de sa réussite. L'utilité de maintenir une distinction entre les objectifs d'un programme exprimés en terme d'états ou de comportements chez les individus de la population-cible et les actions accomplies par les professionnels qui participent à la réalisation du programme apparaît plus clairement à ce stade. En effet, l'évaluation des résultats mesure une modification dans les états ou les comportements des individus composant la population-cible. Elle porte évidemment avant tout sur l'objectif ultime du programme exprimé en termes d'atteinte d'un niveau de santé défini, mais également sur les objectifs intermédiaires correspondant à divers états ou comportements observés chez les individus de la population-cible. Ainsi, dans l'exemple de Greenberg et Mattison cité plus haut, l'évaluation des résultats vise d'abord à déterminer le degré de réduction de la mortalité et la morbidité périnatales. De plus, elle mesure le degré d'atteinte de certains états intermédiaires, comme par exemple la proportion de femmes qui ont suivi tout le programme d'éducation sanitaire pendant leur grossesse.

Jusqu'ici, l'évaluation des résultats s'est préoccupée plus particulièrement de mesurer la réalisation des objectifs, c'est-à-dire de déterminer l'efficacité du programme selon les catégories de résultats présentées au Tableau III. Une autre façon d'évaluer les résultats d'un programme consiste à en mesurer le rendement. Cette méthode vise essentiellement à identifier le nombre d'individus touchés par le programme. Ce peut être, par exemple, le nombre de femmes qui ont suivi tout le programme d'éducation sanitaire pendant leur grossesse. Cette mesure fournit donc une indication de l'ampleur des résultats obtenus.

Enfin, la troisième mesure de résultat est l'efficience d'un programme qui s'exprime par le rap-

port entre l'efficacité ou le rendement et les ressources déployées. Ce type d'évaluation permet de déterminer le rendement d'un programme par rapport à d'autres. La question alors est de savoir si les ressources mobilisées pour obtenir un résultat défini ont été convenablement utilisées.

En résumé, on peut voir que les diverses méthodes d'évaluation présentées répondent à des questions différentes. Pour le planificateur, il ne s'agit pas d'en préférer une et de négliger les autres. Chacune d'elles porte sur un élément particulier du programme. Le planificateur qui veut le succès de son programme et qui veut connaître les raisons de la réussite plus ou moins grande qu'il a obtenue aura intérêt à inclure toutes ces méthodes dans son protocole d'évaluation.

Le but de cet exposé est de sensibiliser ceux qui doivent participer au développement de programmes de santé aux principaux concepts utilisés dans l'élaboration de tels programmes. Nous réalisons qu'en voulant synthétiser la documentation très abondante qu'on retrouve sur le sujet nous avons pu simplifier à outrance des aspects par ailleurs complexes qui auraient mérité une analyse plus approfondie. Nous pensons notamment à certains éléments de la méthodologie de l'évaluation que nous n'avons fait qu'effleurer. Néanmoins, nous espérons que ces quelques réflexions pourront être utiles à ceux qui ont la responsabilité de planifier des programmes de santé.

Résumé

La gestion par programme peut être utilisée avec avantage dans le domaine de la santé communautaire. Le planificateur de la santé soucieux d'atteindre des résultats significatifs se doit d'identifier clairement les divers éléments d'un programme de santé et d'en spécifier les différentes étapes d'élaboration.

Summary

Program management is an administrative tool that can be advantageously used in the field of community health. To achieve measurable results, the health planner must clearly identify the various components of a health program and specify the successive stages of its development.

BIBLIOGRAPHIE

1. Taylor, C.E.: "Stages of the Planning Process", dans *Health Planning: Qualitative Aspects and Quantitative Techniques*, W.A. Reinke (éd.), The Johns Hopkins University, School of Hygiene and Public Health, 1972, pp. 20-34.
2. Donabedian, A.: *Aspects of Medical Care Administration: Specifying Requirements for Health Care*, Cambridge, Mass.: Harvard University Press, 1973, pp. 58-192.
3. Donabedian, A.: *op. cit.*, pp. 164-192.
   O.M.S.: *Méthodes modernes de gestion et organisation des services de santé*, Cahiers de Santé publique No 55, Genève, 1974, pp. 84-89.
4. Simon, H.A.: *Administrative Behavior*, New York: MacMillan, 1961.
5. Deniston, O.L. et coll.: "Evaluation of Program Effectiveness", *Public Health Reports*, 83 (avril 1968), pp. 323-335.
6. Greenberg, B.G. et Mattison, B.F.: "The Whys and Wherefores of Program Evaluation", *Canad. J. of Pub. Health*, 46: 1955, pp. 293-299.
7. Rochon, J.: "Stratégie d'évaluation", dans *Méthodes d'évaluation des soins de santé*, D.L. Sacket et M.S. Baskin (éd.), McMaster University, 1971, pp. 10-11.
8. Donabedian, A.: *A Guide to Medical Care Administration, Volume II: Medical Care Appraisal — Quality and Utilization*. The American Public Health Association Inc., New York, N.Y., 1969.
9. Pour une analyse plus détaillée de ces divers types d'évaluation on pourra consulter les sources suivantes:

   James G.: "Evaluation in Public Health Practice", *Am. J. of Public Health*, 52: juillet 1962, pp. 1145-1154.

   Rochon, J.: "Measurements of Quality of Care: Concepts and Strategies for Evaluation", *Annals of the Royal College of Physicians and Surgeons of Canada*, 7: octobre 1974, pp. 301-311.

   Suchman, E.A.: Evaluative Research: Principles and Practice in Public Service and Social Action Programs, Russell Sage Foundation, New York, N.Y., 1967.

Fiche DD

Valeurs sociales influençant les choix en planification de la santé

	Valeurs	Positions libérales (importance de la réussite personnelle et de la liberté individuelle face aux pouvoirs de coercition politiques)	Position égalitaire (importance de l'égalité dans les opportunités)
1.	Responsabilité personnelle	toute récompense doit être méritée	certaines commodités sont des droits plus que des privilèges
2.	Conscience sociale	la charité, la philantropie, sont des moyens d'aider ceux qui ne réussissent pas	mécanismes sociaux doivent être prévus pour éviter le recours à la charité car dégradant et incontrôlable
3.	Liberté	intervention de l'Etat entrave libertés individuelles	Etat doit s'assurer que les libertés individuelles peuvent s'exprimer i.e. qu'il soit possible pour les individus de faire des choix
4.	Egalité	égalité devant la Loi	opportunités égales pour que les individus se réalisent

DE:

Donabedian A. Aspects of medical Care Administration, 1973

LA PLANIFICATION DES SERVICES DE SANTÉ : UNE PERSPECTIVE ÉPIDÉMIOLOGIQUE

Cet article ouvre un échange permanent entre «Gestions Hospitalières» et «Administration Hospitalière et Sociale» animée par nos amis québecois. Selon les accords prévus, des échanges de communications seront périodiquement pratiqués entre les deux Revues.

«Gestions Hospitalières» se réjouit d'apporter à ses lecteurs une ouverture sur le Monde de la Santé.

R. PINEAULT
Chef de Département de santé communautaire
Cité de la Santé de Laval
Professeur adjoint
Département de médecine sociale et préventive
Université de Montréal

L'épidémiologie comme discipline a connu un essor considérable et une popularité sans précédent ces dernières années grâce à ses applications dans le domaine des soins et des programmes de santé («medical care organization», pour les anglo-saxons). Des publications récentes témoignent de cette situation (1-5). La plupart des pays occidentaux connaissent une augmentation de leurs dépenses en matière de santé, plus rapide que le développement économique et cette situation inquiète beaucoup les responsables dans ce domaine. L'épidémiologie peut précisément fournir une méthode et des outils utiles pour une planification et une évaluation permettant des décisions plus éclairées et plus rationnelles. Dans cet article, nous nous proposons de discuter des principales applications de l'épidémiologie au domaine des soins et des programmes de santé.

LA PLANIFICATION DE SANTE : UNE PERSPECTIVE EPIDEMIOLOGIQUE

Tout d'abord, décrivons brièvement ce qu'est la planification des soins de santé dans une perspective épidémiologique (6).

Les étapes que le planificateur suit dans l'élaboration d'un programme sont représentées à la figure 1.

Les deux premières étapes : identification des problèmes de santé et détermination des priorités d'intervention, sont préliminaires à la planification par programme. Elles constituent le cadre général à l'intérieur duquel les priorités entre les différents problèmes identifiés et, par conséquent, les actions qu'ils justifient, peuvent être déterminées. En d'autres termes, la planification des programmes de santé doit se situer dans le cadre plus vaste d'une planification d'ensemble. L'établissement de l'objectif concrétise la direction vers laquelle le programme veut s'orienter. L'objectif doit être exprimé en terme d'amélioration de l'état de santé de la population-cible. Par la suite, les activités pour l'atteindre doivent être spécifiées de même que les ressources nécessaires à leur production. Enfin, l'évaluation du programme vise à déterminer si les résultats escomptés sont atteints et, en cas d'insuccès, jusqu'à quel point le déroulement des activités et la production des services sont responsables d'un tel résultat, ou encore, jusqu'à quel point les ressources elles-mêmes sont insuffisantes ou inadéquates.

Cette brève description du processus de planification met en évidence les principales applications de l'épidémiologie en planification des soins de santé. En effet, si l'on réfère à la figure 1, on peut identifier ces applications aux étapes suivantes : 1) identification des problèmes de santé ; 2) établissement des priorités ; 3) déter-

mination des activités et des ressources requises ; 4) évaluation des résultats. Nous allons maintenant discuter chacune de ces applications.

IDENTIFICATION DES PROBLEMES ET DES BESOINS DE SANTE

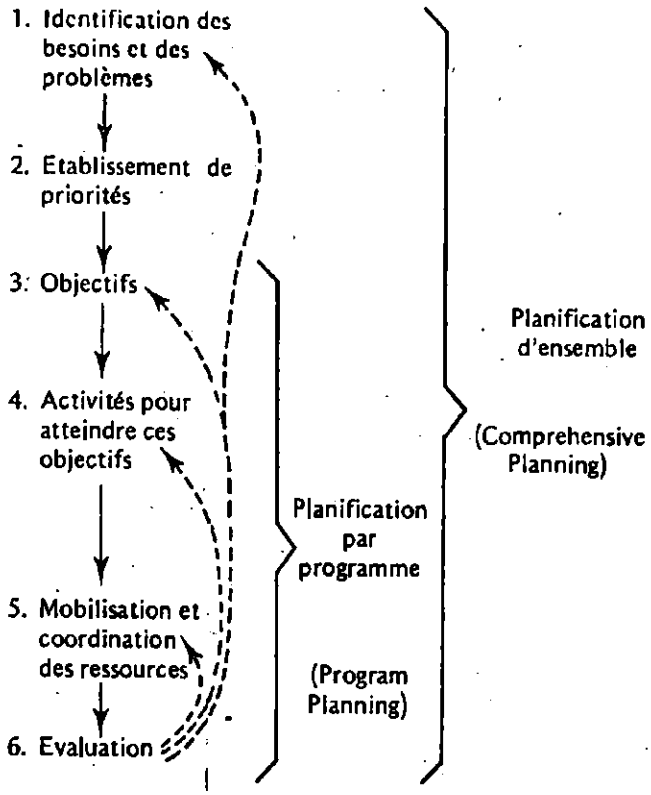
La mesure de l'état de santé d'une population permet d'identifier les différents problèmes de santé. Plusieurs personnes ont exprimé l'espoir qu'on puisse mesurer la santé de façon positive et non seulement par son absence, telle que la maladie ou la mort. Cependant, sur le plan opérationnel, on se voit contraint d'employer des indicateurs «négatifs» de la santé, notamment, la mortalité, la morbidité et les facteurs de risque (7).

Les indicateurs de mortalité les plus utiles pour l'identification des problèmes de santé sont bien sûr les taux spécifiques de mortalité selon l'âge, le sexe et la cause parce qu'ils permettent de découvrir des conditions particulières pouvant faire l'objet d'une intervention ou encore d'identifier une population-cible particulièrement défavorisée sur le plan de la mortalité. De plus, comme il faut comparer cette population avec d'autres populations, la standardisation en fonction de la structure d'âge et de sexe des différentes populations est de première importance. La mortalité ne renseigne que sur les maladies mortelles. Elle ne fournit pas d'information sur le nombre d'individus malades et, en définitive, sur l'importance de la maladie dans une population. Les indicateurs de morbidité sont donc un complément essentiel si on veut obtenir un tableau plus complet de l'état de santé d'une population. Comme les maladies à évolution chronique constituent une partie de plus en plus importante de nos problèmes de santé, leurs mesures de prévalence sont les plus utiles pour fins de planification. Par contre, les mesures d'incidence nous renseignent mieux sur les maladies à évolution plus courte.

La difficulté majeure en ce qui concerne la morbidité demeure la partialité des informations recueillies. Sur la figure 2, on voit qu'il est possible d'identifier quatre types de morbidité selon le procédé d'enquête (8). La morbidité ressentie correspond à la perception qu'à l'individu des symptômes et des malaises qu'il associe à un mauvais état de santé. Une partie seulement de cette morbidité ressentie va être légitimée lors du contact avec l'appareil de soins et devenir ainsi une morbidité diagnostiquée. La partie diagnosticable de la morbidité dépasse nettement la morbidité diagnostiquée en laissant toutefois une fraction de la morbidité ressentie. Cette dernière ne pourra pas être confirmée par l'appareil de soins, du moins avec les moyens technologiques dont dispose la médecine actuellement,

Figure 1

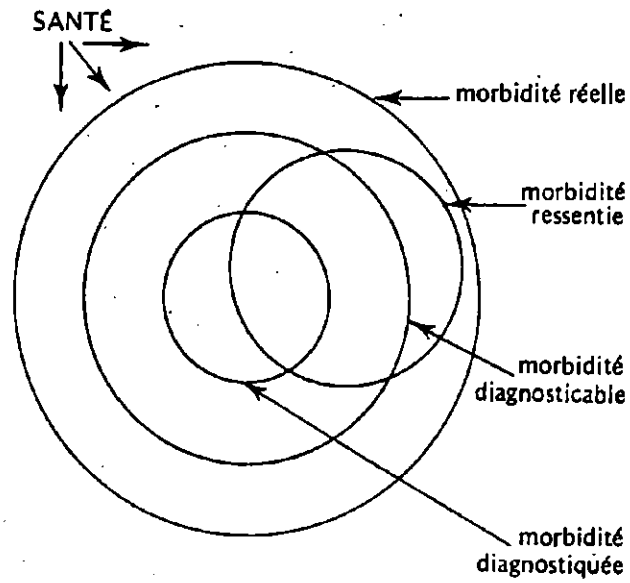
ETAPES DE LA PLANIFICATION DES PROGRAMMES DE SANTE



Source : Pineault, R. «Eléments et étapes d'élaboration d'un programme de santé communautaire» Un. Méd. du Can., 1976, 105, p. 1209.

Figure 2

CODAGE DE L'ETAT DE SANTE ET TYPES DE MORBIDITE



Source : Adapté de Levy, E. et al. Economie du système de santé, Paris : Dunod, 1975.

l'état de santé d'une population et, plus précisément, à l'écart qui peut exister à un moment donné entre l'état de santé observé et l'état de santé désirable ou désiré (1). En d'autres termes, le besoin de santé se traduit par un objectif collectif et constitue le fondement des politiques sociales en matière de santé.

Le degré de succès dans l'atteinte de l'objectif de santé s'établit généralement en prenant comme norme la situation observée dans d'autres collectivités qui présentent une situation sanitaire meilleure et des caractéristiques socio-économiques semblables.

En résumé, les données concernant la mortalité, la morbidité et les facteurs de risque fournissent un tableau assez complet de l'état de santé d'une population. Elles permettent d'identifier les problèmes de santé et d'orienter l'action vers certains problèmes. C'est le point de départ pour le planificateur.

ETABLISSEMENT DES CHOIX ET DES PRIORITES

Après avoir identifié les problèmes, il faut, dans un contexte de rareté des ressources, faire des choix, et établir des priorités d'intervention. A l'étude du processus de planification, l'épidémiologie peut apporter une contribution importante. Suivant la perspective épidémiologique, on peut poser comme principe que les problèmes de santé qui causent les plus grandes pertes de vie productive et qui sont susceptibles d'être améliorés ou éliminés par des programmes de santé exigent une action prioritaire. En pratique, deux critères peuvent être utilisés pour déterminer les priorités d'intervention : 1) l'importance de la perte résultant de la présence d'une maladie ; et 2) la sensibilité de cette maladie à un programme de santé (1,6).

A la base de la méthode proposée, se trouve l'idée que la protection de la santé a pour objet de réduire et, si possible, d'éliminer le risque existant, de rétablir autant que possible la santé compromise en redonnant au malade sa place dans la société. L'importance de la perte causée par la maladie peut s'exprimer soit par le raccourcissement de la vie résultant d'un décès, soit par le degré d'invalidité causé par un processus morbide. On peut également combiner les pertes résultant de la mortalité et de la morbidité en les exprimant par des mesures de temps perdu. Dans ce cas, l'unité de temps considéré peut changer la pondération des priorités. Ainsi, les pertes calculées par rapport à l'espérance de vie ont tendance à accorder une importance plus grande aux maladies mortelles de l'enfance qu'à celles de l'âge avancé. De plus, ce même critère attribue peu d'importance aux maladies chroniques responsables d'invalidités sans toutefois être mortelles. C'est la base que le Gouvernement canadien a choisi pour déterminer ses priorités dans le domaine de la santé alors que la mesure choisie est le nombre d'années potentielles de vie perdues entre 1 et 70 ans (10-11).

Une attitude plus égalitaire est d'utiliser une période de temps plus limitée par exemple, un an, et de calculer les pertes résultant de la mortalité et de la morbidité pendant cette période de temps. Cependant, même dans ce cas, le problème de la combinaison de mortalité et de la morbidité demeure entier. Chacune des pertes sanitaires résultant de la mortalité ou de la morbidité étant qualitativement différente, il est difficile de les additionner directement, même si elles sont exprimées dans les mêmes unités de mesure. En d'autres termes, cela revient à dire qu'une année perdue par un décès équivaut à une année perdue en raison d'invalidité causée par une maladie chronique. Bref, le cœfficient dont il faut affecter chacune de ces mesures de mortalité et de morbidité doit être établi à partir de certaines valeurs fondamentales de la société.

En ce qui concerne l'importance relative des facteurs de risque, il faut d'abord pouvoir identifier et quantifier la contribution du facteur aux diverses morbidités. Pour chiffrer cette contribution, l'outil classique en épidémiologie est la fraction ou le risque attribuable. Cette méthode a été employée par le Gouvernement canadien pour déterminer les priorités et les stratégies d'une action préventive en matière de santé(12). Il s'agit de calculer les pertes résultant de la mortalité et de la morbidité imputables à chaque facteur de risque. On peut ainsi chiffrer la contribution de chacun des facteurs aux diverses maladies et déterminer leur importance relative par rapport aux autres facteurs. Évidemment, la validité de cette méthode dépend entre autres de la validité des résultats fournis par les études épidémiologiques pour établir le risque attribuable.

Le deuxième critère employé dans la pondération des priorités est la sensibilité de la maladie à un programmes de soins. Le critère peut être déterminé de plusieurs façons : 1) normativement, en se basant sur des jugements d'experts et de professionnels de la santé ; 2) empiriquement, en se basant sur les expériences réalisées dans d'autres pays ou régions ; 3) enfin, de façon opérationnelle, en estimant la quantité et surtout le coût des ressources qui devront être mobilisées pour obtenir les gains sanitaires escomptés. Cette dernière méthode correspond essentiellement à l'analyse coût-bénéfice. Ainsi, plus le coût d'un programme pour atteindre un objectif déterminé, est bas, plus le problème est sujet à une intervention efficace.

compte tenu de l'évolution naturelle de certaines maladies qui demeurent encore fort mal connues. Enfin, on peut concevoir une morbidité réelle, dont les limites se situeraient bien au-delà des morbidités ressenties, diagnostiquées et diagnosticables. Cette morbidité réelle engloberait la morbidité ressentie mais non diagnostiquée et non diagnosticable, de même que les états non ressentis et non diagnosticables actuellement, mais néanmoins réels, dont l'évolution ne tardera pas à se manifester. Il est difficile d'apprécier la morbidité réelle puisque sa mesure est liée à la connaissance de l'évolution naturelle de plusieurs états morbides mal connus, à la capacité diagnostique des méthodes d'investigation et enfin aux contraintes des coûts. Au-delà de la morbidité réelle se situe théoriquement la santé.

Cette représentation quelque peu théorique n'est pas sans intérêt pour la planification car elle illustre bien le caractère très partiel de l'information dont nous disposons sur la morbidité. Dans la plupart des pays, seule la morbidité diagnostiquée est enregistrée, et bien souvent ce n'est qu'une fraction correspondant à la fréquentation hospitalière et non pas la consommation globale des soins. Les facteurs reliés à la consommation des soins introduisent un biais important dans la mesure de la morbidité lorsque la consommation des soins est utilisée comme mesure de morbidité. Ainsi, il y a certainement une surestimation de la morbidité pour des catégories socio-économiques les plus favorisées et une sous-estimation pour les autres. Cette situation est particulièrement embarrassante lorsqu'il s'agit de mesurer les besoins de l'ensemble de la population.

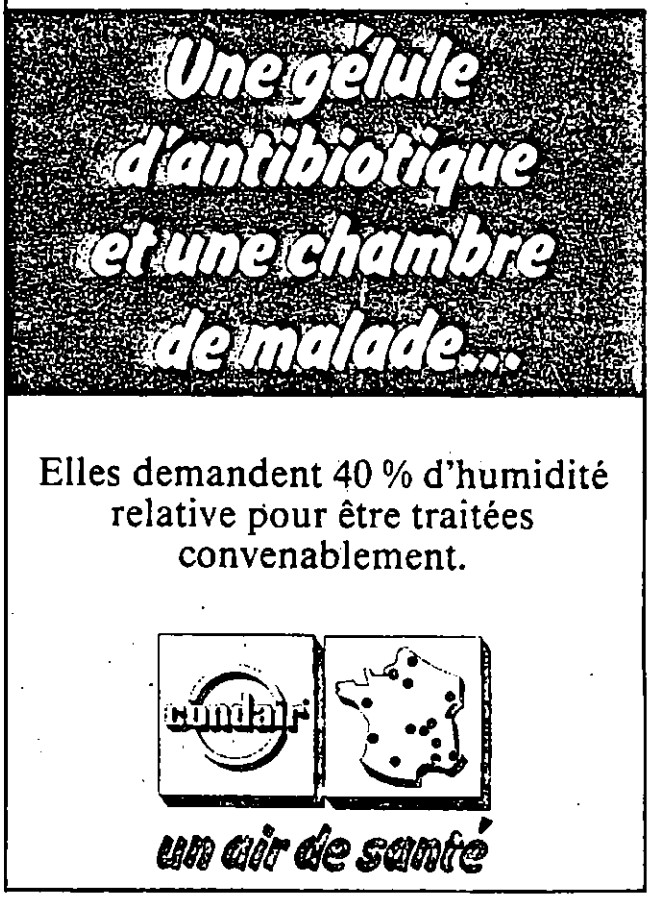
Pour ce qui est de la morbidité ressentie et diagnosticable, seulement quelques pays sont en mesure de produire une telle information grâce à une enquête permanente sur l'état de santé. De plus, il faut noter que les moyens d'investigation mis en œuvre dans de telles enquêtes ne peuvent déceler qu'une très faible partie de la morbidité diagnosticable et il ne peut en être autrement étant donné les coûts astronomiques que nécessiterait une recherche clinique et technique plus poussée. En dépit de ces restrictions, les informations les plus utiles qui proviennent de ces enquêtes sont sans doute celles ayant trait aux maladies chroniques et débilitantes. En effet, elles permettent 1) de calculer le nombre de journées où l'activité a été limitée pendant une année et le degré de cette incapacité, 2) de déterminer l'impact de la morbidité sur une population dans des termes qui peuvent également exprimer la mortalité et, 3) en combinant les informations précédentes, de représenter l'importance globale de chacune des maladies.

Un troisième aspect de l'état de santé, relié à la morbidité, porte sur les facteurs de risque. La distinction entre morbidité et facteurs de risque est artificielle : l'hypertension, par exemple, peut être considérée ou bien comme une morbidité en soi, ou bien comme un facteur de risque puisque cet état est précurseur de conditions cardio-vasculaires. En pratique, il est utile d'identifier les facteurs de risque puisqu'un même facteur peut être relié à plusieurs conditions morbides (e.g. le tabac) et qu'une action dirigée contre les facteurs de risque peut réduire la morbidité et la mortalité.

Pour obtenir une information valide et complète sur les facteurs de risque présents, il faut faire des enquêtes sur les habitudes et le mode de vie dans la population ainsi que des examens cliniques et diagnostiques en rapport avec certains paramètres biologiques. L'enquête sur les habitudes alimentaires et l'enquête en cours sur l'état de santé des canadiens correspondent à ce type d'étude.

A un état de santé correspond des besoins de santé. La notion de besoin a été utilisée de façon restrictive par les économistes pour exprimer dans le contexte du modèle micro-économique de la consommation des soins, la morbidité ressentie par l'individu, celle-ci pouvant s'exprimer éventuellement en une demande de soins (9).

Dans la perspective épidémiologique, la notion de besoin se rapporte d'abord et avant tout à



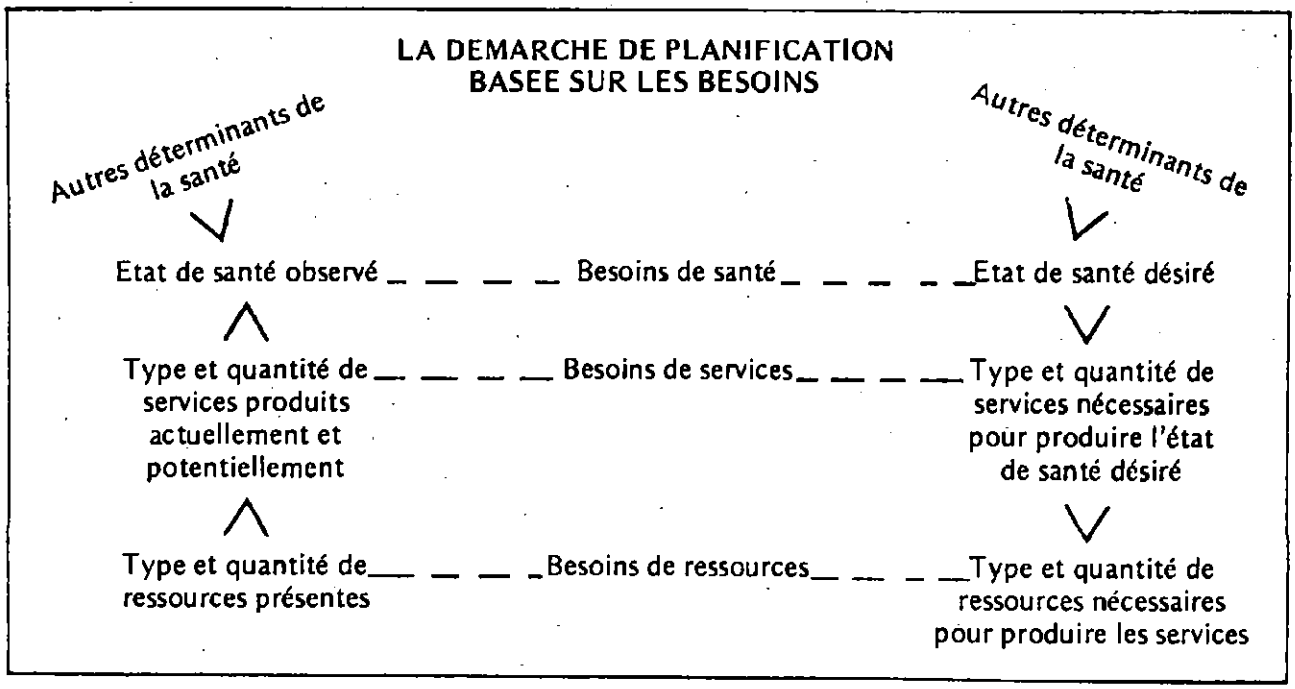
DETERMINATION DES ACTIVITES ET DES RESSOURCES REQUISES

L'épidémiologie peut servir aussi de point de départ au processus de planification dans la prévision des programmes de soins dont a ou aura besoin une population. Nous avons déjà précisé le concept de besoin pour désigner un état de santé, c'est-à-dire un écart entre l'état de santé désiré et l'état de santé actuel d'une population. Parallèlement au besoin de santé, il existe également des besoins de services et des besoins de ressources, comme le montre la figure 3, représentant le processus de planification (1). Par besoins de services, nous entendons l'écart qui existe entre les services existants et les services requis pour satisfaire les besoins de santé. Les besoins de ressources représentent l'écart entre les ressources actuellement disponibles et les ressources nécessaires pour produire les services requis. Le schéma de la figure 3 montre bien que la démarche du planificateur est double : la partie gauche du schéma correspond à l'analyse de la situation actuelle, et la partie droite représente l'analyse prévisionnelle. L'analyse de la situation actuelle consiste essentiellement à établir l'adéquation entre les ressources actuelles et le niveau de santé observé. Il s'agit ici de déterminer si les ressources ont été employées de façon adéquate pour atteindre le niveau de santé observé. La démarche prévisionnelle procède à partir d'un objectif de santé désiré, plutôt qu'à partir des ressources comme l'indique le sens des flèches sur la figure 3. Il s'agit maintenant d'estimer la quantité et la nature des ressources et des services nécessaires pour produire l'état de santé désiré. L'écart entre les données de gauche et de droite détermine l'ajustement qu'il faut effectuer si l'on veut atteindre le niveau de santé escompté.

Notons que ce modèle de planification est fondé essentiellement sur l'identification des besoins de santé. Il n'existe pas de consensus concernant le point de départ pour une telle démarche (15). Les économistes tendent à établir leurs prévisions à partir de consommation actuelle de soins sous prétexte qu'elle est plus réaliste à court terme. Les planificateurs qui adoptent une perspective épidémiologique prétendent qu'une telle approche tend à reproduire les disparités qu'on observe dans la consommation des services de santé, c'est-à-dire que les plus nantis consomment plus de soins, souvent même au-delà de leurs besoins. Par ailleurs, la démarche basée sur les besoins semble reposer sur une conception plus égalitaire de la société (1).

La planification basée sur les besoins a donc été utilisée par les planificateurs dont l'orientation est épidémiologique. On retrouve des exemples classiques dans la littérature (1, 16). De notre côté, nous utilisons actuellement un modèle de ce type, pour déterminer les besoins en programmes de soins des personnes âgées au Québec (17). L'application de cette approche comporte évidemment certains problèmes majeurs qu'il faut surmonter. Il est facile, par exemple, d'en arriver à une surestimation des besoins en ressources. En effet, la traduction des besoins de santé en besoins de services, puis des besoins de services en besoins de ressources, doit forcément se faire de façon normative en utilisant des jugements d'experts dans les domaines de soins concernés. Chaque expert tend bien entendu à estimer les services requis en fonction de sa propre disciplI-

FIGURE 3



En principe, on peut donc identifier quatre types de décision pouvant résulter de l'interaction des deux critères, importance de la perte résultant de la maladie et sensibilité du problème à un programme de santé (tableau 1). Comme l'indique le tableau 1, une priorité faible pour l'établissement d'un programme de soins peut constituer une priorité élevée pour un programme de recherche dans le cas où les pertes résultant de la maladie sont importantes. Par ailleurs, lorsque le problème est peu important et que sa sensibilité à un programme de soins est faible, il devient une priorité secondaire et pour les soins et pour la recherche.

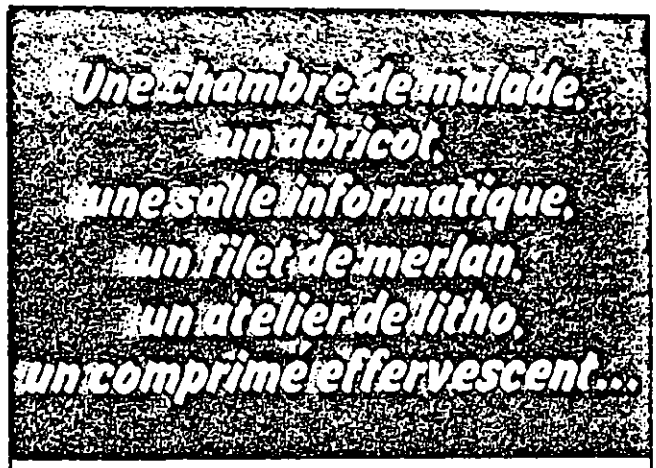
Cette méthode pour pondérer les priorités des programmes de santé a servi de base à la construction de plusieurs indices combinant des mesures de morbidité et de mortalité, dont l'objet est précisément de déterminer les priorités parmi les problèmes de santé d'une population donnée(13). A un autre niveau mais en nous basant sur ce schéma, nous avons proposé une méthode pour décider des priorités entre diverses techniques diagnostiques (14). Les diverses techniques peuvent être classées dans l'un des quatre quadrants du schéma en fonction de leur spécificité et de leur sensibilité, de l'utilité (efficacy) du traitement pour les conditions auxquelles s'applique le diagnostic et, enfin, de l'incidence ou de la prévalence de ces conditions.

TABLEAU 1

PONDERATION DES PRIORITES DANS LES PROGRAMMES DE SANTE SELON L'IMPORTANCE DE LA PERTE ET LA SENSIBILITE DU PROGRAMME

Importance de la perte résultant d'une maladie	Sensibilité du programme	Décision
Considérable	Elevée	Priorité 1 pour l'établissement du programme de santé
Considérable	Faible	Priorité 1 pour la recherche
Peu considérable	Elevée	Priorité 2 pour l'établissement du programme de santé
Peu considérable	Faible	Priorité 2 pour la recherche Priorité la plus faible pour l'établissement du programme de santé

Source : Donabedian, A. Aspects of Medical Care Administration : Specifying Requirements for Health Care. Cambridge : Harvard University Press, 1973, p. 170.



Ils ont tous besoin d'eau pour bien se porter.
L'humidificateur à vapeur Condair fournit à chacun le degré hygrométrique voulu.
Il peut intervenir dans toutes les applications tertiaires et industrielles pour maintenir les meilleures conditions de travail, de production et de stockage. En toute simplicité :
• le distributeur S.I.B. humidifie l'air en utilisant une source de vapeur existante,
• le générateur à vapeur E.S. « auto-adaptatif » se contente de l'eau ordinaire et évite un dispositif de prétraitement. En toute sécurité :
Condair est distribué par un réseau national solide et étendu.

CLYMA :
Siège Social :
32, rue de la Favorite - B.P. 6
69246 LYON CEDEX 1
Tél. (7) 825.32.61 - Télex : 330462
Paris :
85, avenue Jean-Jaurès
92120 MONTROUGE
Tél. (1) 656.02.32 - Télex : 204286
Clermont :
4, avenue de Fontmaure
63400 CHAMALIERES
Tél. (73) 37.90.51
Toulouse :
50, avenue Jules-Julien - B.P. 4001
31028 TOULOUSE CEDEX
Tél. (61) 52.21.14 - Télex : 531934
Bordeaux :
2, rue Jules-Verne
33270 FLOIRAC
Tél. (56) 86.85.94

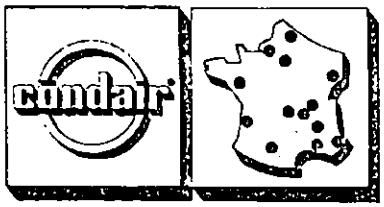
STERC MAISSIAT :
2, rue de Sévigné - B.P. 312
44010 NANTES CEDEX
Tél. (40) 71.45.75 - Télex : 700627

DEGRYSE ET LESAGE :
Lille :
CRT, rue du Petit-Quinquin
59810 LESQUIN
Tél. (20) 95.01.35 - Télex : 130553
Rouen :
Rue François-Arago - Z.I.
76300 SOTTEVILLE LES ROUEN
Tél. (35) 73.13.75 - Télex : 180776

STEN :
4, boulevard Magallon
13015 MARSEILLE
Tél. (91) 64.81.48 - Télex : 440469
Nice :
Euro 3000, avenue de Verdun
06700 SAINT LAURENT DU VAR
Tél. (93) 07.24.80

COMAFRANC :
Service climatisation - B.P. 49
90300 BELFORT VALDOIE
Tél. (84) 21.22.31 - Télex : 360616

EMAFU :
Saint-Etienne :
36, rue de la Croix de Mission - B.P. 137
42012 SAINT ETIENNE CEDEX
Tél. (77) 32.12.07 - Télex : 300013
Valence :
52, avenue Jean-Jaurès
26500 BOURG LES VALENCE
Tél. (75) 43.43.46 - Télex : 345634



un air de santé

Yves Saurav Conseil

ne sans tenir compte de la substitution possible entre les divers services et ressources. Il faut donc procéder à des séances de confrontation ou, du moins, essayer de donner plus d'importance aux jugements des professionnels qui, dans la pratique quotidienne auprès des personnes âgées, ont à décider des programmes de soins requis par elles.

EVALUATION DES PROGRAMMES DE SOINS DE SANTE

Une quatrième application de l'épidémiologie, et non la moins importante, est l'évaluation des programmes de soins de santé. Si l'on se réfère au schéma de la planification à la figure 1, on peut voir que l'évaluation n'est qu'une rétroaction sur les principales étapes de la planification. On distingue donc trois types d'évaluation : l'évaluation des ressources, celle des activités et celle des résultats. Comme le montre le tableau 2, la contribution de l'épidémiologie porte essentiellement sur l'évaluation des résultats (18). L'application de l'épidémiologie à l'évaluation des programmes de santé se situe à deux niveaux. D'abord, sur le plan méthodologique, le «design» par excellence de l'évaluation demeure l'essai clinique contrôlé avec répartition des sujets au hasard (19-20). En dépit de nombreuses difficultés d'application de ce protocole dans la recherche portant sur les humains, il est surprenant de voir qu'il a inspiré plusieurs études évaluatives dans le domaine des soins médicaux comme dans les programmes de dépistage et de prévention (21-26).

L'épidémiologie a également contribué à l'évaluation par les mesures de résultats qu'elle peut fournir et également par la précision qu'elle apporte en ce qui concerne le dénominateur utilisé. L'évaluation des résultats se rapporte aux objectifs du programme, exprimés en termes d'un état de santé défini pour une population-cible. Plus précisément, l'évaluation des résultats sanitaires peut viser trois aspects selon qu'on cherche à évaluer l'efficacité (effectiveness), l'utilité (efficacy) ou le rendement (efficiency), comme le montre le tableau 2. Lorsqu'il s'agit de déterminer l'impact d'un programme sur la population-cible, on parle d'efficacité du programme. Par ailleurs, le terme utilité s'applique à la population effectivement rejointe par un programme. Le rendement d'un programme s'exprime par le rapport entre les résultats sanitaires enregistrés et les activités et les ressources déployées. Les résultats obtenus se rapportent à un aspect mesurable de l'état de santé qui peut être influencé par un élément particulier des soins médicaux. Les résultats obtenus peuvent donc s'exprimer par des mesures comme la mortalité et la longévité, la mortalité périnatale, la létalité, les courbes de survie liées à des maladies chroniques (y compris le cancer, les affections aiguës, l'incapacité physique, etc.). Les données nécessaires à ce type d'évaluation sont essentiellement tirées de statistiques épidémiologiques, démographiques et sociologiques. D'une part, la recher-

TABLEAU 2

INDICATEURS SERVANT A L'EVALUATION ET METHODES DE COLLECTE DES DONNES SELON LES DIFFERENTS ASPECTS EVALUES DANS LE PROGRAMME

COMPOSANTES DU PROGRAMME SANITAIRE	Ressources	Activité et services	Objectifs de santé		
EVALUATION GENERALE	Structure	Processus	Résultat		
EVALUATION SPECIFIQUE	Suffisance (adequacy)	Justesse (appropriateness)	Efficacité (effectiveness)	Utilité (efficacy)	Rendement (efficiency)
INDICATEURS SERVANT A L'EVALUATION	Quantité et qualité des ressources humaines, matérielles et financières	Méthodes et techniques appliquées	Effects sanitaires au niveau de la population-cible	Effets sanitaires au niveau des individus rejoints par le programme	Rapport entre les résultats sanitaires et les ressources déployées
METHODE DE COLLECTE DES DONNEES	Comptabilité Administration	Evaluation confraternelle	Epidémiologie, Démographie, Sociologie	Epidémiologie, Recherche clinique	Epidémiologie, Recherche opérationnelle, Economique

Source : Pineault, R. «Rationalisation de l'évaluation et du contrôle dans les organismes de santé» Adm. hosp. et soc., 1977, 23 (mars-avril), p. 24.

che clinique peut contribuer à déterminer l'utilité de diverses méthodes et, d'autre part, la recherche opérationnelle et l'économique sont d'un précieux concours dans l'évaluation du rendement (18).

Il est intéressant de noter que l'évaluation des résultats, traditionnellement réservée aux programmes de santé publique, trouve de plus en plus d'applications dans le domaine des soins médicaux et dans l'évaluation des établissements de santé (28, 29). La difficulté majeure de l'évaluation des résultats dans les établissements de santé est la précision du dénominateur. En effet, comment définir la population servie par une ressource sanitaire ?

Alors que les programmes de santé publique visent tous les membres d'une population-cible, les organismes sanitaires desservent une population dont seulement un petit nombre des membres bénéficiera de leurs services pendant une certaine période à la suite d'une maladie. Les paramètres de la population-cible des programmes de santé publique peuvent et doivent être définis avant que ces derniers soient mis en œuvre. Au contraire les organismes sanitaires sont des entités vivantes et dynamiques et la population qu'ils desservent n'est pas statique de sorte qu'il faut la redéfinir constamment pour pouvoir expliquer les changements survenus au niveau de l'utilisation.

En dépit de ces difficultés, les planificateurs sanitaires ont mis au point des méthodes qui permettent de définir la population desservie par un organisme de santé (30). Fondamentalement, ces méthodes consistent à déterminer les pourcentages de gens dans des régions géographiques déterminées qui s'adressent à divers organismes sanitaires au cours d'une certaine période. Ces pourcentages sont appelés indices d'attraction (relevance indices). Ils peuvent être calculés pour toutes les zones géographiques dont l'organisme sanitaire tire sa clientèle. Connaissant les caractères démographiques des malades admis dans l'établissement et les caractéristiques de la population régionale, on peut présumer que la population desservie par l'organisme dans chaque zone géographique correspond au pourcentage de gens de la région qui ont effectivement utilisé ses services. En d'autres termes, connaissant la proportion des gens de chaque région qui a fait usage des différents services de santé et connaissant leurs caractéristiques, on peut déduire les caractéristiques et la structure de l'ensemble de la population desservie, c'est-à-dire non seulement des individus qui ont fait usage des services mais de ceux qui n'en ont pas fait usage.

On peut procéder à une identification plus poussée de la population desservie par un organisme de santé en définissant des populations différentes pour chacun des programmes mis en œuvre par cet organisme. Cela peut avoir une certaine importance dans le cas des organismes qui fournissent une vaste gamme de services, allant de services de soins généraux à des services ultra-spécialisés. De toute évidence, la population desservie par des programmes spécialisés est beaucoup plus dispersée que la population servie par des programmes de soins de première ligne. Si l'on veut évaluer l'efficacité des divers programmes, il faut en établir correctement la population-cible.

CONCLUSION

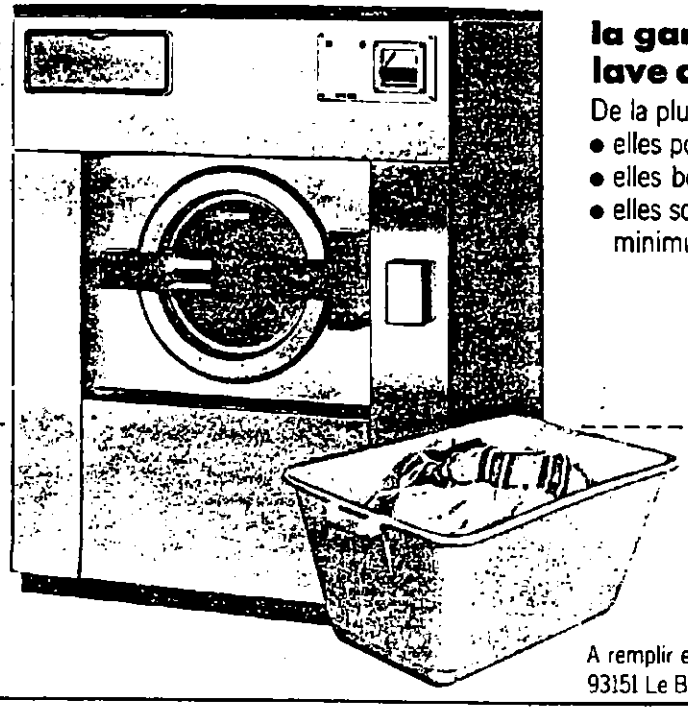
Dans cet article nous avons passé en revue les diverses applications de l'épidémiologie aux programmes et soins de santé. Ces applications sont nombreuses et importantes. Il est intéressant de noter que ce domaine s'est développé grâce aux contributions des sciences sociales et de l'épidémiologie. Il a particulièrement bénéficé de la contribution épidémiologique dans le secteur de la planification alors que les méthodes et les outils épidémiologiques fournissent une base empirique solide pour un processus décisionnel éclairé concernant l'allocation des ressources en matière de santé.

REFERENCES BIBLIOGRAPHIQUES

1) **Donabedian, A.** *Aspects of medical Care Administration : Specifying Requirement for Health Care,* Cambridge, Harvard University Press, 1973.
2) **Holland, W.W. and Gilderdale, S. (eds).** *Epidemiology and Health.* London : Henry Kimpton Pub., 1977.
3) **Ipsen, J.,** *«Epidemiology and Planning»,* Int. *J. of Health Serv.,* 1971, 1, 149-153.
4) **Morris, J.N.** *Uses of Epidemiology.* London : E.S. Livingstone, 1970.
5) **White, K.L. and Henderson, M.M. (eds).** *Epidemiology as a Fundamental Science - Its use in Health Services Planning, Administration and Evaluation.* New-York : Oxford University Press, 1976.
6) **Pineault, R.** «Eléments et étapes d'élaboration d'un programme de santé communautaire». *L'Un. méd. du Can.,* 1976, 105, 1208-1214.
7) **Haynes, M.A.** *«Assessment of Personal Health Problems : An Epidemiologic Approach» dans Health Planning-Qualitative Aspects and Quantitative Techniques,* W.A. Reinke (ed.), The John Hopkins University, 1972, 158-178.
8) **Levy, E.** *et al. Economie du système de santé,* Paris : Dunod, 1975.
9) **Boulding, K.** *«The Concept of Need for Health Services» Milbank Mem. Fund Quarterly,* 1966, 10, 202-223.
10) **McWhinnie, J.R.** *et al. Indicateurs du domaine de la santé.* Ottawa : Ministère de la Santé nationale et du Bien-être social, 1976.
11) **Romeder, J.M. et McWhinnie, J.R.** *«Le développement des années potentielles de vies perdues comme indicateur de mortalité prématurée».* Notes de recherche, planification à long terme, Ministère de la Santé

nationale et du Bien-être social, Ottawa, Février 1977.
12) Romeder, J.M. *«Priorités et stratégies d'actions de prévention : développement et application d'un cadre conceptuel»*. Notes de recherche, Planification à long terme, Ministère de la Santé nationale du Bien-être social, Ottawa, mars 1977.
13) Sullivan, D.F. *«A Single Index of Mortality and Morbidity», HSMHA Health Reports*, 1971, 86, 347-354.
14) Pineault, R. *«Cadre conceptuel et opérationnel pour l'évaluation et le choix de nouvelles méthodes diagnostiques» Adm. hosp. et soc*, 1977, 23 (mai-juin), 9-16.
15) McEwin, R. et Hall, J. *«Planning Health Services - Demand or Need ?»* Med. Journal of Australia, 1976, 2, 350-352.
16) Schonfeld, H.K. et al. *«Number of Physicians Required for Primary Medical Care»*. N. Engl. J. of Med. 1972, 286, 571-576.
17) Tilquin, C., Pineault, R., Sicotte, C. et Audette, L.M. *«Administration d'un réseau de services socio-sanitaires pour les personnes âgées»* Adm. hosp. et soc., 1977, 23 (mai-juin), 26-32.
18) Pineault, R. *«Rationalisation de l'évaluation et du contrôle dans les organismes de santé»* Adm. hosp. et soc., 1977, 23, (mars-avril), 14-26.
19) Byar, D.P. et al. Randomized Clinical Trials : Perspectives on Some Recent Ideas» *N. Engl. J. of Med.*, 1976, 295, 74-80.
20) Spitzer, W. *«What is a health care trial»* JAMA, 1975, 233, 161-163.
21) Sackett, D.L. et al. *«The Burlington Randomized, Trial of the Nurse Practitioner : Health Outcomes of Patients» Annals of Int. Med.*, 1974, 80, 137-142.
22) Mather, H.G. et al. *«Myocardial Infarction : A Comparison between Home and Hospital Care for Patients». Brit. Med.* Journal, 1976, 4, 925-929.
23) Nobrega, F.T. *«Quality Assessment in Hypertension Analysis of Process and Outcome Methods» N. Engl. J. of Med.*, 1977, 296, 145-148.
24) Hutter A.M. et al. *«Early Hospital Discharge after Myorcardial Infarction» N. Engl. J. of Med.*, 1973, 288, 1141-1144.
25) Shapiro, S. *«Evidence of Screening for Breast Cancer from a Randomized Trial»* Cancer, 1977, 39, 2772-2782.
26) Olsen, D.M. et al. *«A Controlled Trial of Multiphasic Screening» N. Engl. J. of Med.*, 1976, 294, 925-930.
27) Roemer, M.I. et al. *«A Proposed Hospital Quality Index : Hospital Death Rates Adjusted for Case Severity» Health Serv. Res.*, 1968, 3, 96-118.
28) Williamson, J.W. et al. *«Health Accounting : An Outcome - Based System of Quality Assurance : Illustrative Application to Hypertension» Bull. N.Y. Acad. of Med.*, 1975, 51, 727-738.
29) Pineault, R. *«Dynamique de l'évaluation de l'acte médical et détermination des besoins éducatifs : évaluation des résultats» La Vie méd. au Can. franc.*, 1977, 6, 813-823.
30) Martin, J.R. *Comprehensive Health Planning : Analytic Concepts, Blue Cross Association*, 1974.

INFAILLIBLEMENT



la gamme des machines Miele lave de 4,5 à 22 kg de linge par cycle.

De la plus petite à la plus grande :
- elles possèdent la qualité et la solidité légendaires Miele;
- elles bénéficient d'une assistance technique permanente;
- elles sont assurées en pièces de rechange pendant 15 ans minimum, après la fin de la série.

Miele
La passion de l'exigence

Précisez la capacité des machines qui vous intéresse :
4,5 à 7 kg ☐ 10 à 13 kg ☐ 15 à 22 kg ☐
nous vous enverrons gratuitement la documentation correspondante :
Nom _____
Adresse _____
Ville _____ Code postal ☐☐☐☐☐

A remplir et à renvoyer à MIELE - 9, av. Albert-Einstein - Z.I. du Coudray 93151 Le Blanc-Mesnil.

GH BW-25 11

LE PROCESSUS D'UTILISATION DES SERVICES DE SANTE

par: André-Pierre CONTANDRIOPOULOS

Le processus d'utilisation des services de santé

Le schéma du Tableau II illustre le fait que fondamentalement, à un moment donné, dans un endroit donné, l'utilisation des services médicaux résulte d'une interaction entre le comportement du consommateur qui décide de faire appel au système de santé et le comportement du médecin qui après avoir établi un diagnostic, propose un plan de traitement puis entreprend le traitement.

L'ensemble du processus prend naissance par la reconnaissance qu'il existe des problèmes de santé (1) et que le système de soins peut efficacement contribuer à les résoudre.

Le consommateur lorsqu'il ressent un malaise, une douleur (morbidité ressentie), peut, après avoir recherché des informations supplémentaires auprès de parents, d'amis, de professionnels de la santé autre que des médecins, décider soit de se traiter lui-même ou de faire appel au réseau parallèle de soins (chiropraticiens, naturopathes, acupuncteurs, etc.) soit de recourir au système de soins. Cette dernière décision se concrétise par la recherche d'un contact avec un médecin, c'est-à-dire rendez-vous dans un cabinet privé, dans un C.L.S.C., visite en salles d'urgence, etc. Lorsque cette démarche aboutit, le consommateur rencontre un médecin et lui exprime sa demande pour des soins.

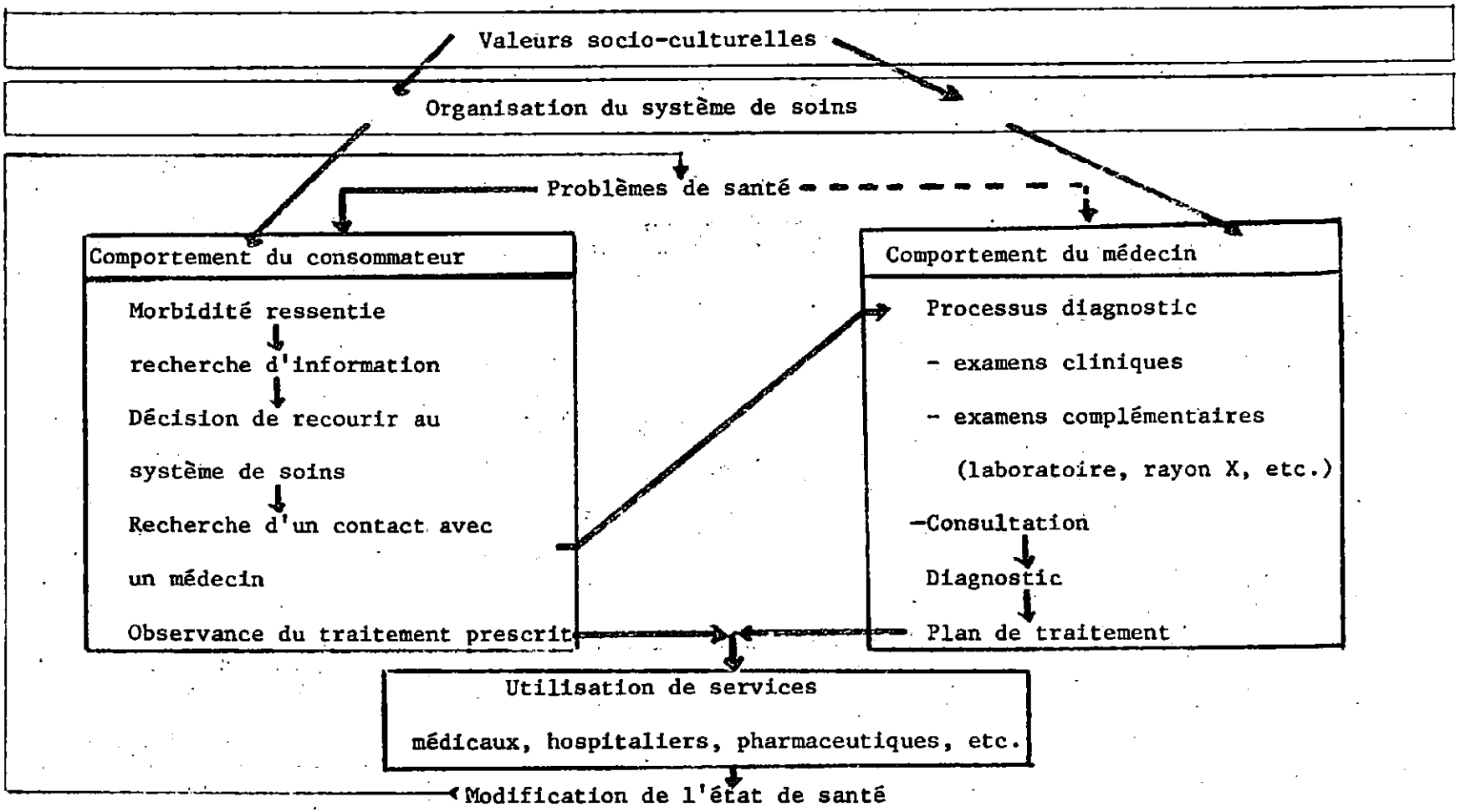
L'ensemble de ces décisions correspond au comportement des consommateurs.

---

(1) La discussion sur la définition de la santé ainsi que sur sa mesure est menée brièvement plus loin. Elle est reprise de façon beaucoup plus détaillée dans les éléments de sociologie de la santé.

TABLEAU II

Processus d'utilisation des services de santé



SOURCE: Aspects of Medical Care Administration, A. Donabedian, Harvard University Press, 1973.

Le comportement des médecins peut se résumer ainsi: le médecin confronte ses connaissances aux informations que lui fournit le patient, les examens cliniques, les résultats des examens complémentaires, les opinions que les médecins consultants peuvent lui donner et établit un diagnostic. Puis, à la lumière de ce diagnostic, propose au patient un plan de traitement qui, si ce dernier l'accepte, se traduit en utilisation de services.

La contre partie monétaire de l'ensemble de ces services que l'on peut appeler curatif, ainsi que de ceux qui ont permis d'établir le diagnostic correspond au coût des services de santé (1).

Ces coûts dépendent donc des comportements des consommateurs et des comportements des médecins. Comportements qui sont eux-mêmes influencés par les modalités organisationnelles du système de santé et plus généralement encore par les valeurs socio-culturelles de la société.

Mais alors qu'autrefois les décisions des médecins portaient principalement sur les services médicaux, aujourd'hui ces décisions sont à l'origine de la plus grande partie des coûts du système de soins.

Il a été estimé que plus de 90% du coût des services médicaux au Québec étaient liés aux décisions que les médecins prennent lors du processus diagnostic et lors du traitement (2).

Il est bien entendu qu'il n'est pas question de dire que ces décisions étaient abusives ou inutiles, il est certain que la plus grande partie d'entre

---

(1) L'analyse de ces coûts ainsi que de leur évolution fait l'objet de la Section II.

(2) Les consommateurs et les coûts de la santé, Boutin & Bisson, Régie de l'Assurance Maladie du Québec, Québec, 1977.

elles sont prises pour soulager ou guérir la population. Il n'en demeure pas moins que ces décisions ont des conséquences monétaires très importantes que les médecins d'aujourd'hui ne peuvent plus négliger. Il faut qu'ils comprennent que leurs décisions sont prises dans un monde où les ressources sont limitées, ce qui implique qu'il y a des choix à faire. De la même façon l'économiste et le planificateur doivent comprendre que le médecin dans son activité professionnelle journalière, ne peut faire autrement que d'offrir à chacun de ses patients tous les services que leur état de santé requiert.

SECTION III - Les facteurs déterminants des coûts du système de santé

Il convient en premier lieu de bien comprendre que ce que l'on appelle les coûts du système de santé sont en fait le chiffre d'affaire d'un secteur de l'activité économique. Et on parle de coût plutôt que d'accroissement du volume de production principalement parce qu'il s'agit d'un secteur financé en très grande partie par l'Etat et qu'il fournit des services et non des biens.

Partant de cette constatation et du fait que la presque totalité des ressources du système de soins sont utilisées pour payer des salaires et des honoraires, il est possible de bâtir la double identité suivante:

$$\text{(nombre de personnes)} \times \text{(salaires)} \cong \boxed{\text{coûts du système de soins}} \equiv \text{(volume des services)} \times \text{(Prix)}$$

Cette double relation permet de comprendre qu'une réduction des coûts du système de soins n'est possible que si d'une part il y a une diminution du volume des services utilisés ou du prix des services et que d'autre part il y a une réduction dans le nombre des personnes qui travaillent dans le système de soins ou dans leur salaire. Toute politique qui a comme objectif une réduction des coûts du système de santé ou un ralentissement du taux de croissance de ces coûts devra être analysée en fonction de son impact probable à la fois sur les services utilisés et sur les ressources employées.

Cette analyse sera facilitée par une bonne compréhension des facteurs qui sont à l'origine de l'utilisation des services de santé et par conséquent des coûts du système de soins. Ils pourront être regroupés en 3 catégories: les facteurs reliés aux comportements des consommateurs, ceux reliés aux comportements des médecins et ceux que l'on peut associer au cadre organisationnel. Cette classification découle directement de l'analyse du système de soins qui a été faite dans le module précédent et qui est schématisée sur le Tableau II.

Les facteurs reliés aux comportements des consommateurs

L'état de santé de la population est bien entendu le facteur essentiel. Mais comme nous l'avons déjà mentionné, la santé est une notion complexe, difficile à cerner de façon précise, et que l'on ne sait pas vraiment mesurer. C'est pourquoi quand on veut essayer de connaître l'influence du comportement des consommateurs sur l'utilisation des services, on ne considère pas la santé ou la maladie directement mais plutôt des variables indirectes, reliées à l'état de santé, et faciles à mesurer.

Elles sont de deux types: les caractéristiques démographiques et les caractéristiques socio-économiques.

- Caractéristiques démographiques. Il s'agit essentiellement de l'âge et du sexe.

L'influence de ces deux variables sur l'utilisation des services médicaux est illustrée sur le Tableau X.

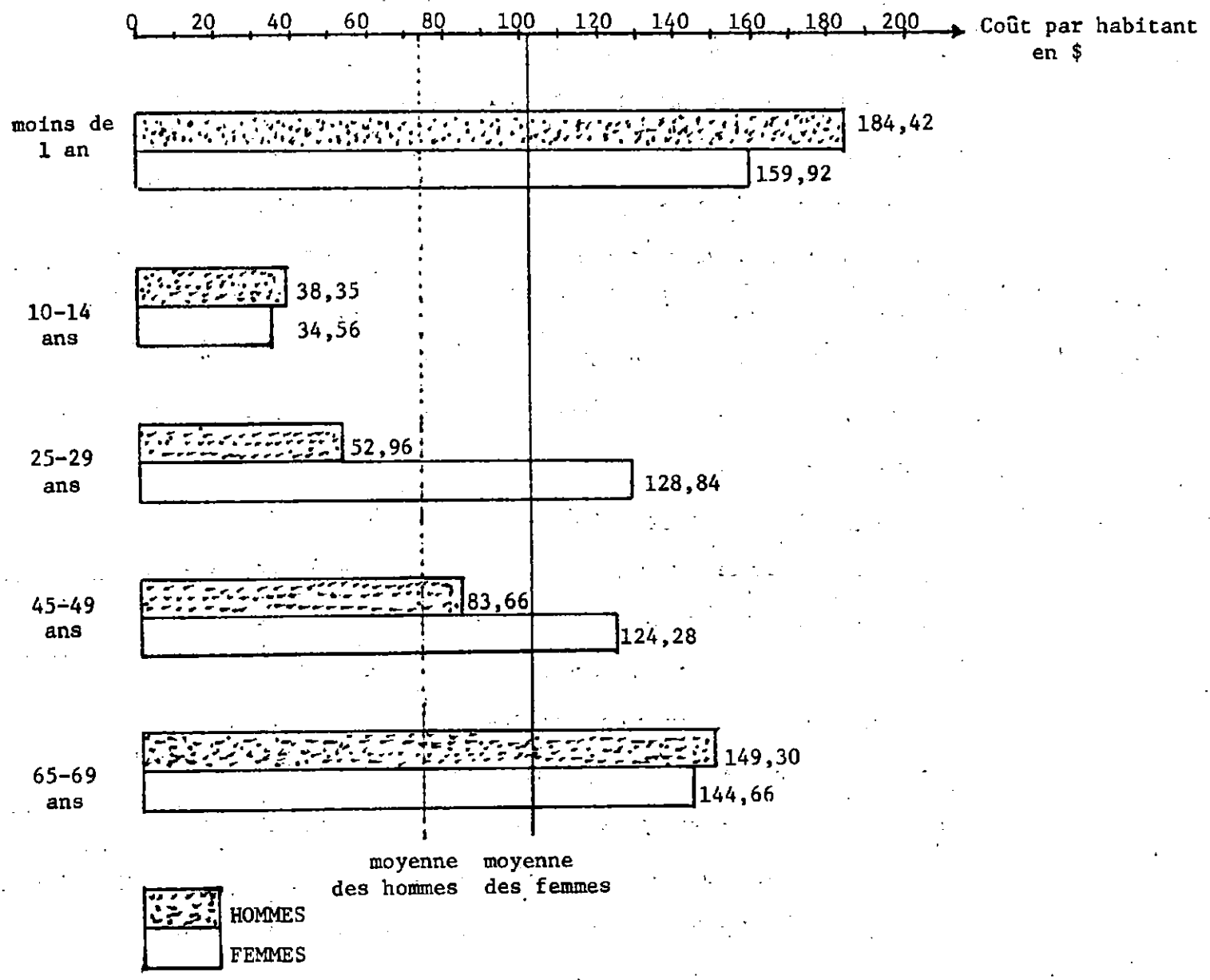
On constate qu'en moyenne les femmes utilisent plus de services que les hommes, et qu'à partir de 55-60 ans aussi bien pour les femmes que pour les hommes, l'utilisation des services médicaux augmente rapidement. Le même type de relation existe pour les services hospitaliers et les médicaments, ce qui implique que le vieillissement de la population auquel on assiste au Québec, (Tableau XI et XII) aura comme effet même si la population augmente faiblement, d'accroître les coûts du système de soins.

- Caractéristiques socio-économiques. Il s'agit de savoir quel est l'effet du niveau de revenu, de l'éducation, c'est-à-dire le niveau socio-économique, sur l'utilisation des services de santé quand en général les modalités organisationnelles restent constantes et plus particulièrement quand les prix des services ne varient pas.

Le niveau socio-économique influence l'utilisation de deux façons, d'une part et en général plus le niveau socio-économique est faible plus le niveau de santé est mauvais et d'autre part les gens de niveau socio-économique élevé sont en général plus sensibles aux problèmes de santé et sont ainsi portés à y consacrer plus de temps et d'argent. Les deux types d'influence sont illustrés sur les Tableaux XIII et XIV.

TABLEAU X

Coût des services médicaux par habitant selon le sexe et certains groupes d'âge, Québec 1978.



SOURCE: Régie de l'assurance-maladie du Québec, Statistiques annuelles 1978, p. 37, Québec 1979.

Tableau XI- Effectifs et augmentation de la population totale selon le sexe, Québec, 1951-2001

	HOMMES		FEMMES		ENSEMBLE	
	Population	Augmentation	Population	Augmentation	Population	Augmentation
1951	2022127		2033554		4055681	
		> 30.2%		> 29.2%		> 29.7%
1961	2631856		2627355		5259211	
		> 13.8%		> 15.4%		> 14.6%
1971	2994545		3003215		6027764	
		> 5.6%		> 7.9%		> 6.2%
1981	3162400		3240800		6403200	
		> 5.3%		> 6.0%		> 5.7%
1991	3331400		3435200		6766600	
		> 1.5%		> 2.0%		> 1.8%
2001	3380800		3404400		6885200	

Sources: Statistique Canada, 1971, 1.2 (Bulletin 1.2-3) tableau 7
Statistique Canada, Projection démographique pour le Canada et les provinces 1976-2001, catalogue (91-250) hors-série, Projection 3.

Tableau XII- Proportion de la population âgée de 65 ans et plus par rapport à l'ensemble de la population selon le sexe, Québec, 1951-2001

	HOMMES	FEMMES	ENSEMBLE
1951	5.6%	5.8%	5.7%
1961	5.5%	6.2%	5.8%
1971	6.0%	7.7%	6.9%
1981	7.4%	10.1%	8.8%
1991	8.9%	12.7%	10.8%
2001	9.9%	14.6%	12.3%

Sources: voir tableau précédent.

TABLEAU XIII

Nombre de contacts médecin-patient par personne, par an, à Montréal, selon le niveau du revenu familial (1971-72).

Revenu familial annuel	Nombre de contacts
- 3 000	7.8
3 000 - 4 999	6.0
5 000 - 8 999	4.7
9 000 -14 999	4.6
15 000 et plus	4.8

SOURCE: McDonald et al. "Études sur l'assurance-maladie du Québec" Sociologie et sociétés, vol 9, No. 1, avril 1979.

TABLEAU XIV

Utilisation des services de santé selon le niveau socio-économique

Montréal 1974.

	Niveau socio-économique		
	faible	moyen	élevé
% de la population adulte ayant vu un médecin durant les 4 dernières semaines	34.1	38.7	32.6
% des enfants non malades ayant vu un médecin durant les 4 dernières semaines	11.6	18.4	20.5
% des adultes non malades ayant vu un médecin durant les 4 dernières semaines	19.4	22.8	20.3
% des adultes ayant été hospitalisés durant la dernière année	12.6	17.4	9.9
% des adultes ayant eu une ordonnance de médicaments lors de leur plus récente visite chez le médecin	60.4	49.2	46.5

SOURCE: Siemiatycki, B., Richardson, L. "Social correlates of health care utilisation under "free" medical care in Montreal", Institut Armand Frappier, 1979.

2 Les facteurs reliés aux comportements des producteurs de services

D'une façon générale plus le nombre de médecins, de dentistes, de lits d'hôpitaux, etc. est grand pour une population donnée, plus l'utilisation des services de santé est importante (Tableau XV).

L'utilisation des services est aussi influencée par les caractéristiques des producteurs. Ainsi le volume des services offerts augmente avec l'âge pour atteindre un maximum vers 40 ans et diminuer ensuite. Les femmes médecins, toutes choses étant égales par ailleurs, consacrent en moyenne par semaine, moins d'heures à leurs activités professionnelles et réclament pour chaque heure un montant inférieur à celui de leurs confrères. Les médecins qui emploient du personnel auxiliaire offrent une plus grande quantité de services par heure que les autres, etc. (*)

3 Les facteurs reliés à l'organisation du système de soins.

Le shéma du fonctionnement du système de soins (Tableau II) a permis de montrer que les comportements des consommateurs et des producteurs de services sont influencés par la façon dont est organisé le système de soins.

On a déjà plusieurs fois indiqué que l'introduction des régimes d'assurance hospitalisation puis d'assurance maladie étaient associés à la croissance des coûts. Ce qu'il importe de comprendre maintenant c'est que l'influence des modalités organisationnelles et de leurs changements, sur le coût du système de soins, est explicable par des

---

(*) CONTANDRIOPOULOS, A.P. et LANCE, J.M.,: Un modèle de prévision de la main d'oeuvre médicale. Université de Montréal, Montréal 1976.

TABLEAU XV

Variations dans l'utilisation et le coût des services médicaux de 1971 à 1978.

	Variations en % de 1971 à 1978
Population	3,9
Nombre de médecins	42,5
Nombre de contacts-patients par habitant	32,6
Nombre de contacts-patients par médecin	- 3,4
Nombre d'examens en cabinet	45,9
Nombre d'examens à domicile	-30,5
Nombre d'examens en cliniques externes et urgence	100,5
Coût par contact-patient	47,9
Coût total des services médicaux	103,6

SOURCE: Statistiques annuelles 1971, 1972, 1973 et 1978, Régie de l'assurance-maladie du Québec.

modifications des comportements des consommateurs et des producteurs. Sans faire une analyse détaillée de l'influence de chaque modalité organisationnelle, nous en isolerons quelques unes tout en sachant qu'elles sont toutes interdépendantes.

Introduction d'un régime public d'assurance maladie.

La mise en place en 1970 d'un régime public universel et complet d'assurance maladie au Québec, associée à une réforme du système de distribution des services de santé (loi sur les services de santé et les services sociaux) a eu comme effet d'améliorer l'accessibilité socio-économique aux services de santé (Tableau XVI). L'accessibilité géographique mesurée par le nombre de médecins pour 10 000 habitants s'est améliorée entre 1972 et 1978. Mais en 1978 il subsiste des écarts importants entre les régions (Tableau XVII). La liberté du choix du lieu de pratique ainsi que la certitude pour les médecins d'être payés a permis à de nombreux médecins de pratiquer dans les régions urbaines du Québec, ce qu'ils n'auraient probablement pas pu faire si le régime d'assurance maladie n'avait pas été instauré.

La qualité des services mesurée par leur globalité, leur continuité et leur conformité aux normes de la bonne pratique ne s'est pas améliorée autant que les auteurs de la réforme l'avaient espéré.

En effet, le maintien du paiement à l'acte dans un système public avec tiers payant ainsi que le manque de participation des médecins à l'organisation et au fonctionnement du système de soins, ont favorisé l'apparition d'une médecine ponctuelle, orientée vers la quantité des services plus que vers le contact entre le médecin et son patient.

TABLEAU XVI

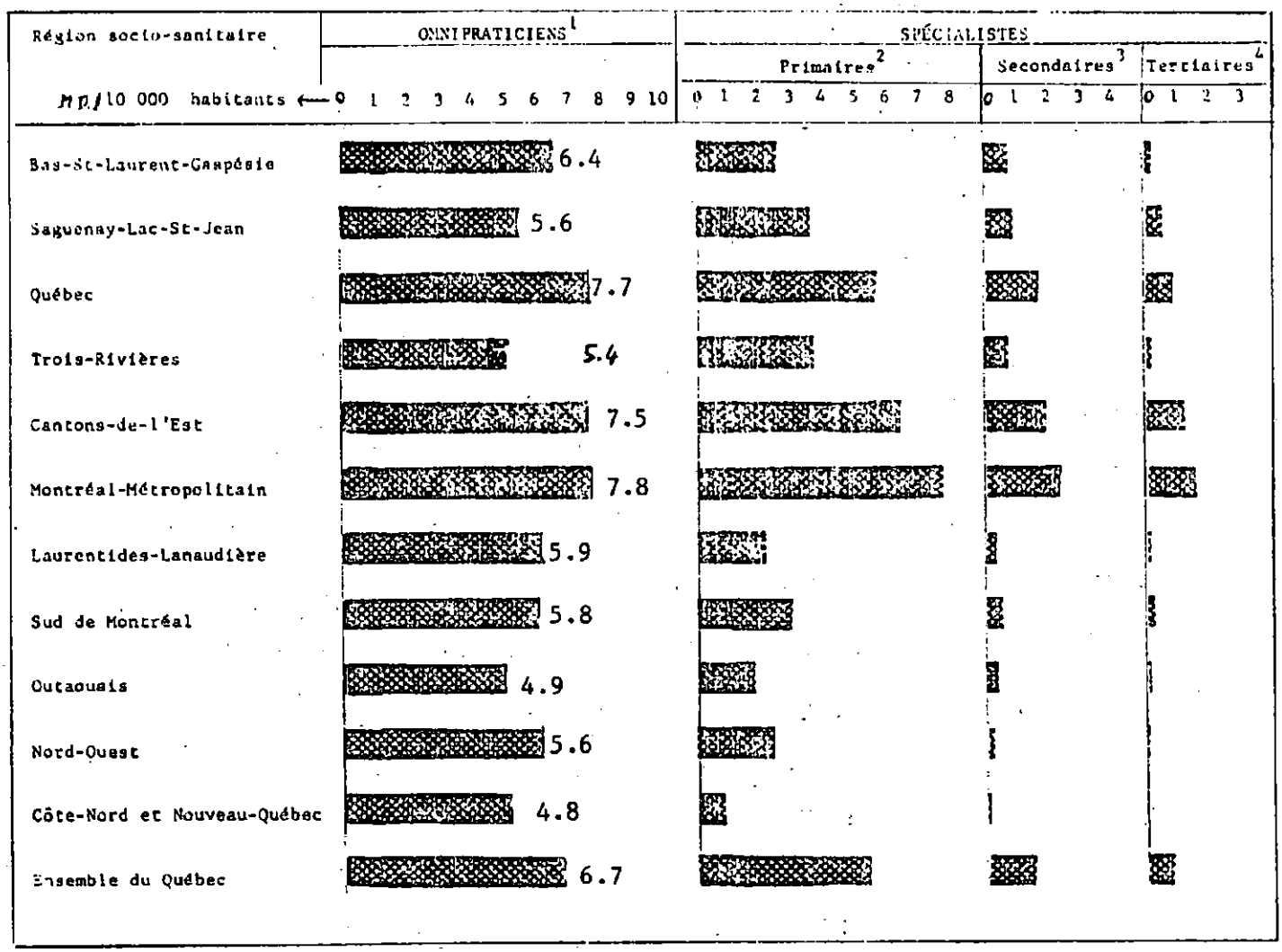
Nombre de consultations par individu avant et après l'introduction de l'assurance-maladie (Montréal 1969-70 et 1971-72).

Revenu familial annuel en $	Nombre de consultations en 1969-70	Nombre de consultations en 1971-72	Variations en %
moins de 3 000	6,6	7,8	18,2
3 000 - 4 999	5,5	6,0	9,1
5 000 - 8 999	4,7	4,7	0
9 000 - 14 999	5,1	4,6	- 3,9
15 000 et plus	5,3	4,8	- 9,4

SOURCE: McDonald, A.D. et al. "Études sur l'assurance-maladie du Québec". Sociologie et Sociétés, vol. 9, No. 1, p. 58, Avril 1977.

TABLEAU XVII

Nombre de médecins équivalents plein temps par 10 000 habitants selon certaines catégories de médecins et la région socio-sanitaire, Québec, 1978.



1 Cette catégorie regroupe les médecins omnipraticiens et les spécialistes en hygiène publique.

2 Cette catégorie regroupe les médecins spécialistes en anesthésie-réanimation, cardiologie, médecine interne, pédiatrie, psychiatrie, chirurgie générale, chirurgie orthopédique, gynécologie-obstétrique, ophtalmologie et radiologie diagnostique.

3 Cette catégorie regroupe les médecins spécialistes en dermato-syphiligraphie, gastro-entérologie, neurologie, pneumologie, oro-rhino-laryngologie, urologie, anatomo-pathologie et hématologie.

4 Cette catégorie regroupe les médecins spécialistes en allergie, endocrinologie, néphrologie, physiatrie, rhumatologie, chirurgie plastique, chirurgie cardio-vasculaire, neurochirurgie, biochimie médicale, médecine nucléaire, microbiologie et radiologie thérapeutique.

Source : Contandriopoulos, André-Pierre et Al. Les Effectifs médicaux au Québec : Situation de 1976 à 1978 et projection pour 1982, Montréal, Corporation professionnelle des médecins du Québec, 1979, tableaux XVII et XXVIII.

Une médecine où le recourt à la technologie a remplacé le dialogue et l'examen clinique. Une médecine enfin, ou la prévention, la santé continue, la santé au travail, la formation médicale continue, la recherche, la participation à des activités administratives est découragée.

Cette discussion est reprise dans la section suivante sur les modes de rémunération.

Structure tarifaire

La structure tarifaire qui a été utilisée de 1970 à 1976, ainsi que celle qui est actuellement en vigueur, contiennent des stimulants économiques auxquels les médecins réagissent. Ils ajustent le volume et l'éventail des actes qu'ils réclament en fonction, entre autre, du caractère plus ou moins lucratif des actes.

Cette manifestation du pouvoir des médecins sur la quantité et l'éventail des services offerts est illustrée par le Tableau XVIII.

Ce tableau a été construit à partir de la cohorte des médecins qui pratiquaient à plein temps en 1971 et 1973. Durant toute cette période les tarifs de remboursement de la Régie de l'assurance-maladie sont restés fixes. Or nous constatons que les revenus des médecins ont augmenté et que la quantité des actes posés a variée, cette variation s'est accompagnée d'une évolution dans le profil de pratique des médecins. Les services additionnels fournis par les médecins améliorent-ils la santé de la population ? Les quelques chiffres suivants permettent d'en douter.

Pour les généralistes de 1971 à 1973, l'augmentation de l'importance relative par rapport à la valeur totale des services fournis, de l'exci-

TABLEAU XVIII

Évolution des rémunérations et des services fournis pour la cohorte des médecins en exercice de 1971 à 1973 au Québec (a)

	1971	1973	% de variation 1971-1973
Médecins généralistes			
Rémunérations versées $	87,707,288	92,385,053	5.33%
nombre d'actes posés	17,156,322	16,273,997	-5.14%
nombre de médecins	1850	1850	-
Spécialités médicales			
Rémunérations versées $	57,741,122	62,679,456	8.55%
nombre d'actes posés	8,047,442	9,071,513	12.72%
nombre de médecins	1449	1449	-
Spécialités chirurgicales			
Rémunérations versées $	90,691,793	92,806,338	2.33%
nombre d'actes posés	5,884,216	5,761,434	-2.09%
nombre de médecins	1321	1321	-

SOURCE: Statistiques fournies par le R.A.M.Q. dans le cadre du projet "Evolution du profil de pratique des médecins du Québec depuis l'introduction de l'assurance-maladie" (A.P. Contandriopoulos, 1976).

(a) Cette cohorte est composée des médecins qui ont perçu plus de $10,000.00 en 1971 et en 1973. Durant toute la période les tarifs de remboursement sont restés fixes.

sion de verrues plantaires a été de 56%, l'augmentation des excérèses de cérumen de 11%, l'augmentation des injections de substances sclérosantes dans les veines des membres inférieurs (traitement des varices) de 202%, l'augmentation des examens complets de l'oeil en supplément d'un examen physique complet de 71%, l'augmentation de la dilatation du prépuce de 42%.

Pour les spécialités chirurgicales notons que l'augmentation relative des appendicectomies simples a été de 59%, et de plus, alors qu'en 1971, dans 73% des cas l'appendicectomie était considérée comme l'intervention principale, en 1973 il n'y en a plus que 28% qui sont considérées comme intervention principale.

Pour les spécialités médicales notons que l'importance relative des injections de substances sclérosantes dans les veines des membres inférieurs (traitement des varices) a été multipliée par dix; pour les extractions de comédons, l'augmentation de l'importance relative a été de 12%, pour l'excérèse de cérumen de 22%.

Il est embarrassant de soutenir que l'accroissement de l'importance relative de ces quelques actes puisse être justifiée par l'amélioration du niveau de santé de la population. Cet accroissement nous semble indiquer que les médecins utilisent effectivement le pouvoir qu'ils ont pour influencer l'utilisation des services afin de tenter de maintenir leur revenu en jouant sur le type et la quantité des actes qu'ils posent.

Cette impression est largement confirmée par l'analyse qu'ont eu les modifications négociées en 1976. L'augmentation des tarifs des visites à domicile s'est traduite par un renversement radical de la tendance observée de 1971 à 1976.

Le plafonnement des revenus des médecins omnipraticiens a modifié le comportement des médecins ayant un haut revenu, diminution du volume d'activités, augmentation du nombre de jours de vacance, etc.(*)

Les modes de rémunération

De ce que nous venons de voir il apparaît qu'un moyen important qui permettrait de modifier le système de soins réside dans des changements de mode de rémunération des professionnels.

Quels sont les divers modes de rémunération du corps médical utilisés à travers le monde ? On peut les regrouper en cinq types: le paiement à l'acte, le salariat, la capitation, la vacation et le paiement au cas. Le Tableau XIX fournit une classification de ces modes de rémunération en fonction d'une double interrogation: ce qui est rémunéré et le type de paiement.

Le paiement à l'acte est une rationalisation et une systématisation de l'honoraire professionnel. Il est relié à un service spécifique donné à un patient en particulier. On fixe un tarif aux différents actes que peut poser un médecin et le coût du traitement correspond à la somme de ces tarifs. Ceux-ci peuvent être établis à partir d'une liste très détaillée des actes ou à partir de rubriques qui peuvent contenir plusieurs actes. Généralement on établit des tarifs uniformes pour les services tels la consultation, la visite au cabinet, à domicile, le jour, le soir aux-

---

(*) BOUTIN, J.G. Effets de certaines mesures de la deuxième entente des médecins omnipraticiens sur leurs activités dans le cadre du régime d'assurance maladie. Régie de l'assurance maladie, Québec 1979.

TABLEAU XIX

Les modes de rémunération des professionnels

Ce qui est rémunéré	Type de paiement			
	honoraires	forfait	salaire	vacation
des actes définis dans une nomenclature	paiement à l'acte			
les soins requis pour un problème de santé donné		paiement au cas		
la fonction exercée dans un poste d'une organisation administrative			paiement à salaire (année)	
la responsabilité d'une personne assurée		paiement par capitation (mois)*		
le temps				paiement à la vacation (3 heures)

* Entre parenthèses, se trouve la durée pendant laquelle on fixe habituellement la valeur unitaire de base de la rémunération.

quels on ajoute une liste de services spécifiques pour lesquels un tarif spécial est déterminé: accouchements, amygdalectomies, etc.

Pour la majorité des associations médicales nord-américaines, seul le paiement à l'acte permet une pratique médicale individuelle et concurrentielle qui motive favorablement le médecin à l'égard du patient, puisque son revenu dépend de la satisfaction de ce même patient. Selon ce mode de rémunération, le revenu du médecin est directement relié aux types, au nombre et aux tarifs des actes que pose le médecin.

Parmi les nombreux inconvénients que comporte le paiement à l'acte, signalons: (1) qu'il tend à axer toute la pratique médicale sur la phase active de la maladie, au détriment de la prévention et de la réadaptation; (2) il peut entraîner une répartition inégale des effectifs médicaux par spécialités, au détriment de spécialités comme la santé publique, par exemple, qui se prêtent mal à une identification précise de chacun des actes posés; (3) la relation médecin-patient ne peut être considérée comme une relation d'affaires, le malade étant dans une situation d'infériorité, d'ignorance et d'inquiétude et il ne désire aucunement marchander sa santé; un des avantages du régime actuel d'assurance-maladie est que cette relation d'affaires n'existe plus; (4) on peut mettre en doute qu'actuellement l'indépendance du praticien soit chose désirable, en raison des développements rapides de la science médicale et de la technologie, en raison de l'apparition de spécialisations médicales de plus en plus étroites et en raison aussi de l'absence généralisée de politiques de recyclage pour les médecins.

Le salariat est essentiellement un paiement pour les fonctions exercées dans un poste à l'intérieur d'une organisation. Le revenu du médecin varie selon ses qualifications professionnelles: expérience, spécialisations, renommée, degré de responsabilités, etc. Ce mode de rémunération doit être exercé dans le cadre d'une organisation. Le salariat: (1) est d'une administration beaucoup plus simple que le paiement à l'acte, (2) il favorise la prévention, (3) il représente la meilleure façon d'assurer la présence permanente de médecins dans les centres de santé, (4) il favorise la pratique en groupe et la consultation entre praticiens, (5) il facilite la continuité des soins, (6) il permet une meilleure répartition des tâches au sein de l'équipe de santé, (7) il comporte des avantages marginaux que le médecin doit se payer lui-même sous un régime à l'acte, etc.

Parmi les désavantages que l 'on reconnaît au salariat: (1) il peut affecter de façon indésirable la relation patient-médecin, (2) il pourrait étouffer chez certains médecins l'initiative personnelle, (3) il suppose d'importants déboursés par l'Etat, à court terme, pour racheter l'équipement appartenant aux médecins en pratique privée, etc. (4) il est associé à une perte de libertés professionnelles et place le médecin dans une situation de subordination administrative.

Sous la capitation, le médecin est rémunéré pour la responsabilité qu'il a de fournir les services requis aux personnes inscrites sur sa liste. Autrement dit le médecin touche un revenu en fonction d'une liste qui contient un certain nombre de personnes auxquelles il pourrait éventuellement donner des soins. En Angleterre, par exemple, tous les omnipraticiens touchent essentiellement une capitation. Parmi les avantages de la capitation notons: (1) elle permet d'assigner à chaque individu un

médecin qui s'en occupera, (2) elle favorise une relation durable entre le patient et son médecin de famille, (3) elle assure au médecin un revenu stable, (4) elle ne relie pas le niveau de rémunération au nombre d'actes posés, (5) elle favorise la prévention et l'éducation sanitaire, etc.

Le principal inconvénient de la capitation est que ce mode de rémunération postule implicitement que la somme de travail exigée du médecin est directement proportionnelle au nombre de patients dont il est responsable; donc ce mode ne tient pas compte de l'incidence variable des conditions morbides parmi la population. Il constitue aussi une entorse à la liberté de la population de choisir le professionnel de son choix et à la liberté du médecin de refuser un patient.

Tout comme le salariat, la vacation est un paiement pour une période de temps. On établit par exemple la durée en heures d'une vacation à laquelle on accorde ensuite un taux de rémunération quelconque. Ce mode de rémunération est généralement utilisé dans le cadre d'une institution pour les médecins qui y remplissent, à temps partiel, des fonctions médicales, administratives, d'enseignement ou de recherche; ce mode laisse aux médecins la possibilité dêtre rémunérés à l'acte dans la pratique privée. Dans ces circonstances, la vacation comporte à la fois les avantages et les inconvénients tant du salariat que du paiement à l'acte. Le principal avantage additionnel est que la vacation assure aux institutions que seront remplies les tâches que seuls les médecins sont en mesure d'accomplir.

La rémunération au cas; étant donné le diagnostic posé, on fixe une somme globale pour le traitement fourni par le médecin impliqué. Le revenu du médecin est alors relié au nombre de cas qu'il traite et au prix

Lisi DD

A.P. Contandriopoulos
R. Pineault

Lundi le 6 juin 1983

INTRODUCTION

- Cadre conceptuel pour l'analyse du système de santé et la planification
- Perspective de santé communautaire en planification de la santé
- La planification et la programmation en santé communautaire
  - . approche technique VS pragmatique
  - . types de planning
  - . la notion de programmes
  - . étapes de la planification des programmes de santé

fixé pour chacun de ces cas. Le tant par cas (1) favorise une relation permanente et durable entre le médecin et son patient, (2) il conserve la liberté du patient de choisir son médecin et vice-versa, (3) il stimule le médecin à écarter les traitements non nécessaires ou peu utiles, (4) le malade n'est pas fractionné en autant de parties qu'il y a d'actes médicaux à poser, car chaque cas correspond à un certain processus de traitement, suite à un diagnostic.

Le plus grand inconvénient du tant par cas résulte de la difficulté de définir un cas et de rendre justice aux variations dans le nombre, le type et la durée des services médicaux requis par les divers cas.

En terminant et pour permettre une discussion plus juste sur les mérites relatifs des différents modes de rémunération il faut préciser ce qu'est l'activité professionnelle pour laquelle les médecins sont payés.

L'activité professionnelle est un ensemble complexe d'inter-actions entre le professionnel, les patients et l'environnement dans lequel il exerce. Cette activité complexe ne peut pas être décomposée de façon satisfaisante en unités aux définitions précises. L'acte tel que nous le connaissons n'est qu'une approximation grossière de l'activité professionnelle qui se cache derrière chaque réclamation. Ce décalage entre l'activité professionnelle réelle et la définition des actes a eu deux conséquences; premièrement la tendance naturelle a été de décomposer l'activité professionnelle en unités de plus en plus fines et de plus en plus spécifiques pour tenter de reconstruire sa diversité en combinant des unités aussi précises que possible. Cet émiettement de l'activité professionnelle a d'une part favorisé les actes techniques

aux contours bien précis, défavorisé le dialogue entre le professionnel et son patient et éliminé les activités difficilement décomposables en actes (la prévention auprès de populations, les tâches administratives) et d'autre part largement contribué à l'apparition d'une activité professionnelle normalisée, impersonnelle où les différences de compétence entre les professionnels sont ignorées tout comme les différences qui se cachent derrière le même acte posé à des patients différents dans des conditions différentes.

La deuxième conséquence du décalage entre l'activité professionnelle réelle et la définition des actes est liée à l'impossibilité de décomposer en actes spécifiques plusieurs dimensions de l'activité professionnelle: (prévention collective, formation médicale continue, enseignement, recherche, responsabilités administratives, etc.).

La rémunération actuelle avec référence au temps de présence (vacation, psychothérapie de 30', 50', etc.) témoigne de cette incapacité de cerner sous la nomenclature des actes certaines dimensions de l'activité professionnelle. D'autres activités sont rémunérées totalement ou partiellement par une forme de salariat (enseignement, responsabilités administratives). Enfin, certaines composantes de l'activité professionnelle (éducation médicale continue, tâches administratives, consultations téléphoniques) sont actuellement exclues du champ des activités rémunérées.

Cette coexistence du paiement à l'acte avec un paiement au temps, une forme de salariat ou une absence de paiement, n'est pas sans poser de problèmes; notons à titre d'illustration les exemples suivants:

- le salariat apparaît comme un mode de rémunération nécessitant des contrôles administratifs bien structurés pour ne pas entraîner d'abus. Ces contrôles sont perçus par les professionnels comme des atteintes

inacceptables à leurs libertés, libertés qu'ils considèrent comme nécessaires à la dispensation de services de qualité. Dans ces conditions, les médecins à salaire peuvent facilement être dévalorisés par rapport à leurs confrères à l'acte;

- manque d'étanchéité entre la rémunération au temps et la rémunération à l'acte qui peut entraîner une double rémunération durant une même période de temps;
- dévalorisation de toutes les activités non-rémunérées;
- dans le cas des vacations, difficulté de connaître et donc d'évaluer le volume de la partie de l'activité payée de cette façon;
- difficulté pour un professionnel de maintenir un équilibre entre des activités rémunérées à l'acte et celles qui sont rémunérées au temps.

Sans faire le procès du paiement à l'acte il faut cependant reconnaître qu'il est difficile dans le régime actuel d'assurance maladie de freiner ou de prévoir la croissance des coûts avec un mode comme le paiement à l'acte. Par ailleurs sous d'autres modes comme le salariat ou la capitation, le revenu du médecin n'étant pas relié aux types et au volume des actes qu'il pose mais étant plutôt le fruit de négociations entre les médecins et l'employeur, il est alors davantage possible de freiner la croissance des coûts des services médicaux. Mais comme ces autres modes comportent aussi leurs inconvénients et que, de plus, la génération actuelle des médecins n'est pas prête à abandonner les libertés professionnelles associées au paiement à l'acte, la décision d'adopter un autre mode n'est pas facile à prendre.

On doit cependant noter que plusieurs facteurs aideront, dans les années qui viennent, un changement dans les modes de rémunération

si ce changement apparaît indispensable.

(1) la croissance rapide dans les effectifs médicaux (notons à ce sujet que l'accroissement annuel du nombre de médecins est de 8 à 10 fois plus rapide que l'accroissement de la population),

(2) l'importance accrue de la pratique en milieu institutionnel par opposition à la pratique indépendante, suite aux développements de la médecine comme science, aux développements technologiques et à l'acceptation du concept d'équipes de santé multidisciplinaires pour favoriser des soins complets et continus.

(3) la croissance rapide du nombre des femmes médecins qui sont en général plus prêtes que leur confrère à pratiquer sous d'autre mode de rémunération que l'acte.

LE SYSTÈME DE SANTÉ ET LA PRÉVISION DE L'OFFRE DE SERVICES MÉDICAUX

A.P. CONTANDRIOPOULOS (*) Ph.D.

in : "Modélisation et Maîtrise des systèmes" AFCET 1977
Hommes et Techniques
Paris 1977

I.- UN SCHÉMA DU SYSTÈME DE SANTÉ

Le schéma présenté sur le Tableau I est axé principalement sur les services médicaux (2) il peut se décomposer en 3 sous-modèles.

Le sous-modèle 1 illustre les relations existantes entre l'état de santé de la population, la demande de soins et l'utilisation de services médicaux.

L'état de santé d'une population à un moment donné dépend d'une multitude de facteurs que nous regroupons en 4 catégories: facteurs prédisposants (âge, sexe, héritage génétique, etc.); les facteurs détériorants de la santé (accidents, maladies, habitudes négatives de vie, etc.); les facteurs de promotion de la santé (hygiène, prévention primaire, etc.) et les facteurs de restauration de la santé (services médicaux, services hospitaliers, médicaments, etc.). Le système de santé est consacré principalement à la production et à la distribution des facteurs de restauration de la santé, il contribuera donc au maintien de l'état de santé mais sera probablement impuissant à réellement améliorer cet état (3). La mise en route du processus d'utilisation des services médicaux peut être conceptualisée de la façon suivante:

C'est la morbidité ressentie par une personne qui l'incitera, après avoir éventuellement demandé des conseils à des parents, des amis, un pharmacien, etc. (réseau non professionnel d'orientation et réseau professionnel non médical d'orientation), à décider de demander des soins à un médecin. Cette décision se concrétise par le processus de recherche de soins qui aboutit à un contact avec un médecin. C'est lors de cette visite que le médecin établit seul, ou en faisant appel à des collègues ou à d'autres professionnels, un diagnostic et qu'il propose un plan de traitement. C'est alors que le patient compte tenu des recommandations du médecin et de ses caractéristiques sociales, psychologiques, économiques, etc., exprimera une demande de services spécifiques qui en fonction des services disponibles se traduira par une utilisation de services (4).

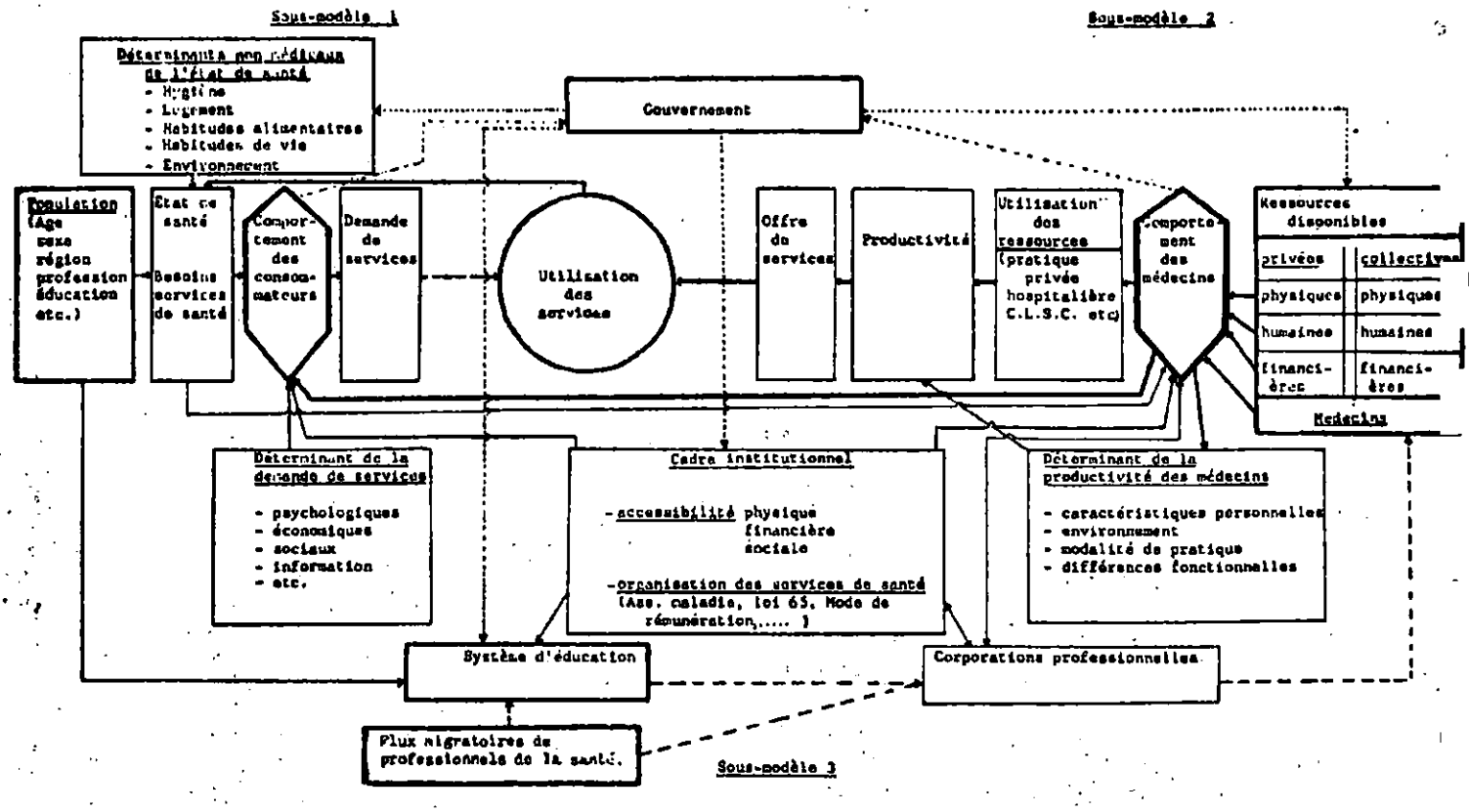
Nous constatons que la demande de services dépend fortement des conseils du médecin puisque le patient ne dispose pas des informations nécessaires pour établir un diagnostic et pour décider tout seul du traitement à suivre. Si le médecin, qui à la fois conseille et fournit le service conseillé, ne peut faire entièrement abstraction de ses intérêts personnels quand il propose un plan de traitements à ses patients, alors il faut reconnaître que la demande de services spécifiques est dépendante du comportement du médecin et donc de l'offre de services, c'est-à-dire du sous-modèle 2 (5).

Le sous-modèle 2 est celui qui schématise le processus de production des services médicaux (6). La mise en oeuvre des ressources physiques humaines et financières du système de santé pour produire des services médicaux dépend des décisions des médecins. La quantité des services médicaux offerts dépend de la productivité avec laquelle les ressources sont utilisées. Cette productivité est fonction des caractéristiques des médecins, de la façon dont ils pratiquent et du lieu où ils exercent.

Le sous-modèle 3 porte sur l'offre des médecins. Les caractéristiques de ce sous-modèle sont: l'existence de mécanismes limitant le nombre des médecins

TABLEAU 1

Un modèle schématique du système de santé



formés, le fait que l'autorisation de pratiquer la médecine est contrôlée par un organisme professionnel et que les médecins immigrants ne sont autorisés à pratiquer qu'après des contrôles sévères de leurs connaissances.

Le système de santé que nous venons de décrire est tout entier influencé par les valeurs socio-culturelles de la société et par le cadre institutionnel dans lequel il est placé.

C'est au gouvernement à qui revient, en tant qu'exécuteur des choix collectifs, d'une part, la tâche de décider quelle proportion des ressources de la société doit être affectée au secteur de la santé et aux autres secteurs de l'économie et d'autre part le rôle de fixer le cadre institutionnel (accessibilité aux soins, régime d'assurance, mode de rémunération des médecins, etc.) du système de santé.

Cette tâche, le gouvernement ne peut pas l'exercer de façon rationnelle si les relations indiquées par des flèches sur le Tableau I ne sont pas spécifiées et quantifiées et s'il n'est pas en mesure de prévoir comment évoluera le système dans le futur.

Etant donné d'une part que le sous-modèle 2 joue dans le système de santé un rôle dominant, ce sont en effet les décisions des médecins qui sont à l'origine de la plus grande partie des coûts du système de santé (7), et d'autre part que les problèmes conceptuels liés à la définition et à la quantification de l'état de santé ainsi qu'à l'influence relative des différents facteurs déterminants de la santé sur le niveau de santé sont loin d'être résolus, nous avons limité notre recherche aux sous-modèles 2 et 3, c'est-à-dire à l'offre de services.

LECTURE POUR MARDI LE 7 JUIN 1983

Rev. Epidém. et Santé Publ., 1979, 27, 51-68.

Revue générale.

Indicateurs de santé et « sanométrie » : les aspects conceptuels des recherches récentes sur la mesure de l'état de santé d'une population.

PREMIÈRE PARTIE

M. GOLDBERG [1], W. DAB [1], J. CHAPERON [2], R. FUHRER [3], F. GRÉMY [1].

Paris (France), Boston (U.S.A.).

Mots-clefs : Mesure de la santé, Indice d'état de santé, Mortalité, Morbidité, Incapacité, Adaptation, Santé positive, Etat de santé global, Population.

Key-words : Health measurement, Health status index, Mortality, Morbidity, Disability, Adaptation, Positive health, Global state of health, Population.

Health indexes and « sanimetry » : conceptual aspects of recent research on the measurement of a population's state of health.

SUMMARY. *The major health indexes based on the classical concepts of mortality and morbidity are reviewed while emphasizing their limitations : underestimation of health problems socio-economic impact, inadequacy for the evaluation of certain main objectives of health care delivery systems, purely negative measurement of health. This paper investigates the recent attempts to improve these aspects: extension of the concept of morbidity by the introduction of disability measurements and the development of a " positive health " concept. Finally, a discussion is presented of the major health indexes whose development is based on the use of these concepts and the main conceptual, methodological problems they give rise to.*

RÉSUMÉ. Après avoir rappelé les principaux indicateurs de santé reposant sur les concepts classiques de mortalité et de morbidité, les auteurs soulignent leurs insuffi-

1. Groupe I.N.S.E.R.M. U. 88, G.E.R.S.S., 91, bd de l'Hôpital, F-75634 Paris Cedex 13.
2. Ministère de la santé, G.E.R.S.S.
3. Harvard University School of Public Health, Boston.
— *Tirés-à-part :* M. GOLDBERG.
— Texte reçu le 30 octobre 1978.

sances majeures : sous-estimation de l'impact socio-économique des problèmes de santé, impossibilité d'évaluer certaines activités importantes des systèmes de soins, mesure purement négative de l'état de santé.

Cet article expose les tentatives récentes pour répondre à ces insuffisances, extension du concept de morbidité par introduction de mesures de l'incapacité et développement du concept de « santé positive ». Sont ensuite passés en revue et discutés les principaux indices d'état de santé construits à partir de ces notions nouvelles, ainsi que certains des problèmes conceptuels, méthodologiques et pratiques qu'ils posent.

Historiquement, la mesure de l'état de santé s'est élargie du phénomène le plus solidement établi, la mort, pour inclure dans son champ les atteintes morbides les plus graves, puis les atteintes moins sévères et finalement la santé. La route ainsi parcourue est devenue progressivement plus difficile et incertaine.

M. TERRIS [135].

1. POURQUOI MESURER L'ÉTAT DE SANTÉ D'UNE POPULATION ?

Le besoin de mesure de la santé d'une population peut être comparé à celui de l'appréciation de la santé d'un individu. Un médecin examinant un patient doit tout d'abord apprécier globalement son état de santé, puis il cherchera une ou plusieurs causes susceptibles d'expliquer une éventuelle détérioration de l'état de santé ; il proposera le cas échéant un traitement et en appréciera les effets. En transposant ce schéma au niveau collectif, on remplacera le patient par une population et le médecin par le système de santé servant cette population. On peut alors dire que le système de santé ou plus exactement ceux qui en ont la charge (qu'on les appelle les décideurs, les planificateurs, les responsables de la Santé publique) doivent être capables d'apprécier le niveau de santé de la population, de repérer et d'évaluer l'ampleur de problèmes sanitaires particuliers, notamment en mettant en évidence des différences entre sous-populations, de choisir et d'optimiser l'allocation des moyens à mettre en œuvre pour améliorer l'état de santé général ou pour résoudre un problème particulier et d'évaluer les résultats des actions entreprises. En bref, mesurer l'état de santé d'une population, c'est se donner un instrument pour définir une politique de santé au niveau de la collectivité ainsi que pour en surveiller les résultats.

Toutes ces activités représentent le domaine d'application de la mesure de l'état de santé d'une population. Toutes nécessitent, pour être entreprises et menées à bien, la connaissance de faits et de chiffres représentant d'une façon ou d'une autre l'état de santé de la population, ainsi que l'ont indiqué deux rapports de l'Organisation mondiale de la santé (O.M.S.) [101, 105].

De façon plus synthétique, Bickner [15] décrit trois utilisations majeures de la mesure de l'état de santé d'une population : l'information du public, la gestion du système de santé, la science médicale. K. L. White [140], quant à lui, divise le domaine de la mesure de la santé d'une population en six secteurs : la recherche étiologique, l'étude de l'utilité des actions sanitaires, l'étude des résultats de ces actions (« effectiveness »), de leur efficience (« efficiency »), l'évaluation, terme générique appliqué aux trois précédents secteurs et l'éducation des personnels de santé comme du public.

Le but de cet article est de tenter de faire le point des recherches de ces dernières années dans le domaine de la mesure de l'état de santé d'une population. En effet, il s'agit selon nous d'un aspect de la recherche en Santé publique dont l'importance théorique, méthodologique et pratique est fondamentale à beaucoup d'égards, mais qui est très peu connu et étudié dans notre pays alors que de nombreuses revues scientifiques étrangères, notamment anglo-saxonnes, y consacrent un intérêt considérable et grandissant. Cependant, depuis peu, des voix se font entendre qui plaident en faveur du développement de la recherche dans ce domaine : Lévy *et al.* [81, 82], Sournia [128] notamment.

Nous espérons, par cet article, contribuer à faciliter l'accès à l'abondante littérature internationale, surtout anglo-saxonne, concernant les problèmes de la mesure de la santé dans une population.

Il est important de souligner que nous ne cherchons, dans cet article, qu'à faire le point de la recherche dans le domaine de la mesure de la santé d'une population. Dans la panoplie de ce que les économistes appellent les « indicateurs de santé » nous n'envisagerons que les « indicateurs de résultat », domaine déjà fort vaste, et nous ne parlerons pas des indicateurs « d'influence exogène », ni des indicateurs de coût (pour utiliser la classification d'Andréani [4]), ni des « indicateurs de système » (Lévy *et al.* [81, 82]). Notre approche du domaine de la santé est celle d'épidémiologistes, non celle d'économistes de la santé, même s'il nous arrive de faire allusion à certains problèmes de nature économique lorsque cela est nécessaire.

2. Les concepts classiques de la mesure de l'état de santé d'une population.

Jusqu'à une époque récente, l'état de santé d'une population était exprimé exclusivement en termes de *mortalité*. Cela se justifiait pour deux types de raisons : pratiques d'abord, parce que la mort est un phénomène qui ne se prêtait pas à ambiguïté dans sa constatation, fonctionnelles aussi parce que le problème sanitaire qui se posait à la communauté était avant tout celui de la mort prématurée. Un travail comme celui de Hollingsworth [66] est à cet égard significatif. Il a étudié l'augmentation de l'espérance de vie à la naissance des enfants de la haute noblesse britannique nés de 1330 à 1954. Les résultats, d'autant plus probants qu'ils concernent une fraction particulièrement privilégiée de la population du point de vue des conditions de vie comme des soins médicaux, montrent bien que la mortalité précoce a été le problème essentiel pendant des siècles.

Antonovsky [5], compilant des données de mortalité éparses remontant au milieu du XVII[e] siècle jusqu'à nos jours, dans divers pays du monde occidental, a pu montrer que l'espérance de vie à la naissance de l'ensemble de la population a assez régulièrement augmenté durant les deux derniers siècles à partir d'un niveau très bas.

Une étude récente de l'Institut national d'études démographiques [69] portant notamment sur l'évolution de l'espérance de vie générale en France de 1750 à nos jours montre des résultats comparables.

Si les premières tables de mortalités dues à l'astronome Haley, concernant la ville de Breslau, furent publiées en 1693 [132] ce n'est que beaucoup plus tard, en 1842, que le concept de *morbidité* fut utilisé par Chadwick pour décrire l'état de santé d'une population dans une étude visant à comparer l'état de santé de travailleurs couverts par un système d'assurance à celui de prisonniers [28]. L'indicateur utilisé était le nombre annuel de jours de maladie. C'est encore plus tard, à la suite des travaux de Pasteur que s'est développée l'épidémiologie « classique » et qu'ont été imaginés les principaux indicateurs de morbidité, incidence et prévalence, outil de contrôle des maladies transmissibles. L'utilisation de ces indicateurs a été ensuite progressivement étendue aux maladies non infectieuses en même temps que le développement de la connaissance scientifique des maladies permettait l'établissement de classifications cohérentes dont l'actuelle Classification internationale des maladies de l'Organisation mondiale de la santé représente l'aspect le plus connu [103].

A l'heure actuelle, la mesure de la santé d'une population est exprimée dans la plupart des pays, dont le nôtre, exclusivement au moyen d'indicateurs variés, plus ou moins spécifiques, mais reposant entièrement sur les concepts *de mortalité et de morbidité, celle-ci étant définie par l'existence de maladies codifiées par la science médicale.*

3. Les principaux indicateurs et indices reposent sur ces concepts.

Afifi *et al.* [1] définissent les termes d'indicateur et d'indice comme des descripteurs de l'activité ou de l'état d'un système. Les *indices* sont des mesures quantitatives agrégées dont l'objectif est de décrire l'activité totale du système, les *indicateurs* peuvent être ou non quantitatifs et ne reflètent habituellement que des aspects limités du système. Dans ce qui suit, nous utiliserons ces termes dans cette acceptation, en la complétant par les recommandations de Berg [11, p. 253] qui réserve l'emploi du terme *indicateur* pour l'expression de mesures spécifiques et de celui d'*indice* (ou *d'index*) pour des mesures composites combinant plusieurs indicateurs.

En 1955, l'Organisation mondiale de la santé chargea un groupe d'experts de définir des indicateurs permettant d'évaluer le niveau de santé de la population d'un pays [101]. Ils proposèrent trois indicateurs, tous basés sur la mortalité : espérance de vie à la naissance et à un an, taux brut de mortalité et taux de mortalité

proportionnelle à 50 ans et plus. L'espérance de vie à un an a été individualisée afin d'éviter l'influence de la mortalité infantile ; le taux de mortalité proportionnelle a été proposé par Swaroop et Uemura [133] afin de s'affranchir de la nécessité de connaître la structure par âge de la population comme c'est le cas pour l'espérance de vie. Plus récemment, un rapport de l'Organisation de coopération et de développement économique (O.C.D.E.) a proposé, pour ne pas masquer des modifications éventuelles des taux de mortalité aux différents âges, de calculer également l'espérance de vie à 20, 40 et 60 ans [108, pp. 64-65]. La mortalité infantile, définie par les décès survenus pendant la première année de vie et rapportée au nombre de naissances (vivantes ou non suivant le cas) est également un indicateur très utilisé en raison de l'importance accordée à cette période de la vie et parce que pour beaucoup d'auteurs, elle est un bon reflet de l'état de santé global de la population.

A partir de ces indicateurs de base, de nombreux indicateurs et indices ont été construits à des fins diverses, soit en désagrégeant les indicateurs de base et en proposant des mesures spécifiques par âge et par sexe et par cause notamment [72, pp. 33-60], soit au contraire, en fabriquant des indices composites. On peut citer à titre d'exemple les travaux de Damiani en France [42] qui a combiné différents taux de mortalité pour obtenir plusieurs indices qu'il appelle Indicateur sanitaire de mortalité, Indice de classement des départements, Indicateur lié aux facteurs sanitaires (ce dernier comprenant d'ailleurs d'autres types de données).

Un autre exemple est celui des « Années potentielles de vie perdue » [93, 119] utilisée par le ministère de la santé du Canada : le but de cet indicateur est de montrer l'impact socio-économique des principales causes de décès, en soustrayant l'âge de survenue des décès de 70 ans, âge choisi arbitrairement. On obtient de cette façon un tableau de la mortalité par causes, classées par ordre d'importance, qui est bien différent de celui basé sur la fréquence simple des décès. Ceci a bien entendu des implications directes sur la planification des priorités sanitaires.

Il en est de même de « l'index Q » [91] dont l'objectif est de permettre d'établir des priorités sanitaires pour des populations défavorisées (en l'occurrence les Indiens des Etats-Unis) ou « l'index G » proposé par M. K. Chen [30] pour remédier à certaines de ses insuffisances. Ces indices combinent des données de mortalité, de morbidité et d'utilisation de services sanitaires, ceci pour la population-cible et une population de référence simultanément. Les résultats obtenus permettent de classer par ordre de priorité les pathologies sur lesquelles on se propose d'agir.

Pour ce qui concerne spécifiquement la morbidité, les indicateurs classiques reposent sur l'*incidence* définie comme la fréquence dans une population donnée des nouveaux cas de maladies pendant une période définie, et sur la *prévalence* instantanée ou de période, nombre de cas présents à un moment donné ou durant la période [122]. Là aussi, comme pour la mortalité, on a construit par agrégation ou désagrégation suivant de nombreux critères, des indicateurs à usage spécifique. Un exemple français est l'Indicateur de morbidité régionale ou départementale de Damiani [43] qui est composé des taux de prévalence théorique de certaines maladies calculés par sexe et par unité géographique, destiné à des études comparatives

de morbidité. On trouvera d'autres exemples dans des ouvrages d'épidémiologie tels que celui de Jenicek [72, pp. 61-71].

Bien que ne reposant pas sur la morbidité proprement dite, mais sur le *concept voisin d'anomalies biologiques ou physiologiques*, on peut citer aussi l' « Indice biologique de santé » de Suchet [134] destiné à permettre une comparaison du niveau de santé de groupes de population reposant sur la réponse à des tests biologiques. Les sujets de chaque groupe sont soumis à une batterie de tests biologiques ; l'indice de santé est calculé pour chaque groupe comme une fonction du pourcentage de réponses normales et anormales dans le groupe. Une version un peu plus raffinée introduit une pondération traduisant l'importance des anomalies constatées (un poids différent est attribué selon que le test donne une réaction « suspecte », « pathologique » ou « très pathologique»).

Le Centre d'hygiène appliquée Doria [3] calcule un « Index omaphthoralogique » utilisé pour mesurer le degré d'usure physiologique des individus d'une collectivité, à des fins de comparaisons entre collectivités humaines. Cet indice est établi à partir des résultats à des tests pathologiques (existence de troubles cliniques) des tests de sénescence et des tests sociaux. Il est lui aussi calculé à partir des pourcentages de réponses anormales dans chaque groupe de population. L'indice est utilisé notamment pour comparer l'état de santé (ou plutôt l'état d'usure, selon les auteurs) de différentes catégories socio-professionnelles.

4. Insuffisances des indicateurs classiques.

Tels qu'ils sont, les indicateurs que nous avons brièvement décrits ont été et sont encore extrêmement utiles, d'un point de vue aussi bien cognitif qu'opérationnel. On constate d'ailleurs, avec juste raison, un regain d'intérêt des épidémiologistes pour l'étude de la mortalité, qui peut être perçu par la multiplication des publications scientifiques sur ce sujet, depuis quelques années. Mais, surtout dans le contexte des pays industrialisés, ils souffrent d'insuffisances qui leur interdisent de rendre compte d'un certain nombre de phénomènes pour lesquels l'intérêt des professionnels comme du public va grandissant. Au cours des dernières années, un certain nombre de faits sont venus mettre en doute, de façon parfois radicale comme chez Illich [68], l'efficacité du système de santé des nations dites développées. Ainsi l'accroissement très important des dépenses de santé, constaté surtout dans la dernière décennie, s'accompagne d'un rendement marginal en terme de santé très faible, selon le mot de Fuchs [55]. Pire encore, l'espérance de vie générale, après avoir longtemps crû régulièrement, s'est stabilisée ces dernières années dans les pays industrialisés et a même commencé de décroître [85, 100] (encore que les tous récents résultats des Etats-Unis dans ce domaine semblent montrer une remontée légère). Un certain nombre d'études récentes, dont plusieurs françaises [50, 56, 77, 79] ont montré, malgré la diversité des méthodes utilisées, qu'on ne pouvait établir aucune corrélation significative entre l'utilisation du système de santé et le niveau de santé de la population mesuré par les indicateurs classiques. Enfin, le poids des atteintes

chroniques tant en termes de mortalité et de morbidité qu'en termes de dépenses de santé est maintenant largement prépondérant dans le tableau sanitaire des nations industrialisées.

L'accumulation de ces faits soulève toutes sortes de problèmes. Nous nous contenterons ici de montrer que l'étude de ces phénomènes ne peut passer que par l'amélioration de nos outils de mesure de la santé d'une population.

On peut regrouper en trois grands chapitres les critiques à formuler aux indicateurs reposant sur les concepts classiques de mortalité et de morbidité : 1) ils ne permettent d'apprécier que très partiellement l'impact socio-économique des problèmes de santé ; 2) ils ne permettent pas d'évaluer correctement l'efficacité du système de santé ; 3) enfin, et peut-être surtout, ils ne donnent de la santé d'une population qu'une mesure négative.

4.1. *L'impact socio-économique des problèmes de santé.*

Les indicateurs de mortalité ne peuvent rendre compte du poids des problèmes de santé, au niveau individuel et/ou collectif, que pour des phénomènes pathologiques se traduisant par un décès prématuré. Les indicateurs basés sur la notion d'Années potentielles de vie perdues que nous avons déjà citée sous-estiment donc l'importance relative croissante des maladies chroniques. Quant aux mesures de morbidité reposant sur l'incidence et la prévalence, elles ne font que compter le nombre de cas de maladies ou d'anomalies biologiques ou physiologiques, mais elles ne permettent pas de différencier, pour une même affection, les cas qui s'accompagnent de troubles fonctionnels interdisant l'activité habituelle du malade de ceux qui sont compatibles avec une poursuite de cette activité. Ce n'est pas seulement le volume des répercussions socio-économiques qui est ainsi mal mesuré, c'est aussi leur nature : le retentissement psychologique et social de la maladie pour l'individu n'est pas abordé du tout.

4.2. *L'évaluation de l'efficacité du système de santé.*

Les administrateurs et les économistes disposent d'un ensemble fourni d'indicateurs décrivant aussi bien les ressources des systèmes de santé (équipements, personnels, etc.) que leur fonctionnement, notamment les dépenses qu'ils occasionnent. Pendant longtemps, on a utilisé ces « indicateurs de moyens » selon la terminologie de Lévy *et al.* [82] comme mesure indirecte de l'état de santé d'une population. Cette approche correspond à ce que les chercheurs du CEREBE [27] qualifient de « confusion entre les inputs et les ouputs du système » ; elle repose sur l'hypothèse qu'il existe une relation directe entre l'activité du système de santé et le niveau de santé de la population bénéficiant des ressources de celui-ci. La construction d'indices tels que « l'Indicateur lié aux facteurs sanitaires » de Damiani [42] qui combine données de mortalité et de ressources, ou « l'Indice de sous-médicalisation » développé à l'Université du Wisconsin [64] reflètent bien cette école de pensée.

La critique majeure que l'on peut faire de cette approche est que l'hypothèse de départ, à savoir la relation directe entre niveau de ressources sanitaires et

niveau de santé, ne semble pas être vérifiée. De nombreux auteurs la mettent actuellement en doute avec des arguments convaincants, que Winkelstein [142] a résumés de façon très claire.

Mais il est frappant de constater que cette mise en doute de l'efficacité des systèmes de santé, mesurée par l'élévation du niveau de santé des populations repose sur des études, telles que celles que nous avons citées, qui utilisent comme critères des indicateurs de mortalité et morbidité classiques. Or, dans beaucoup de cas, et notamment pour les maladies chroniques, ni l'incidence, ni la prévalence, ni la mortalité ne peuvent être un reflet de l'efficacité des soins, ainsi que l'a noté Sullivan [129] ; en effet, si les techniques médicales peuvent atténuer les troubles entraînés par ces maladies, cet effet n'est pas exprimé par les indicateurs utilisés. Pire même, ils peuvent au contraire montrer une aggravation de la situation sanitaire ; ainsi Sanders [124] a montré qu'une amélioration de la qualité des soins pouvait augmenter la prévalence des maladies chroniques dans une population en allongeant la durée de vie des patients. De même, la diminution de la mortalité infantile obtenue en France ces dernières années peut s'accompagner d'une augmentation du nombre de handicapés [29].

Il n'est donc pas interdit de penser que l'impossibilité actuelle de mettre en évidence une action positive des systèmes de santé en termes de niveau de santé globale des populations provient de ce que l'instrument de mesure utilisé n'est pas adapté et ne permet pas de rendre compte du fait qu'une partie de plus en plus importante des ressources sanitaires des pays industrialisés est orientée vers l'atténuation des conséquences de la maladie et l'amélioration de la « qualité de vie » des patients, notamment grâce aux activités de réhabilitation.

4.3. *Une mesure négative de l'état de santé d'une population.*

La critique la plus fondamentale que l'on puisse apporter à une mesure de la santé d'une population entièrement fondée sur les concepts de mortalité et de morbidité est sans doute que ce type de mesure n'est pas valide, en ce sens, qu'il ne rend pas compte du phénomène qu'il prétend étudier. En effet, « la santé » pour reprendre la définition de l'O.M.S., la plus couramment utilisée [102] « n'est pas seulement l'absence de maladies ou d'infirmité mais un état de total bien-être, mental et social ». Cette définition est critiquée par de nombreux auteurs (et nous y reviendrons) mais l'unanimité se fait, pour en admettre la première partie ; en mesurant la mortalité et la morbidité, on ne mesure par la santé d'une population, mais sa « mauvaise santé » (« ill-health » des anglo-saxons), ses maladies. En bref, on ne pratique qu'une mesure négative de la santé. Cette constatation ne s'inscrit pas uniquement dans le cadre d'une discussion théorique, conceptuelle et philologique, mais a de profondes répercussions pratiques dans au moins deux domaines : 1) l'évaluation de l'impact socio-économique des problèmes de santé comme des résultats d'actions sanitaires par une mesure négative ne peut être que près partielle, ainsi que nous l'avons déjà montré ; 2) et surtout, l'utilisation exclusive de la mortalité et de la morbidité, catégorisées en « causes médicales » pour reprendre la

terminologie habituelle, indique implicitement un point de vue médical : c'est le médecin, et lui seul, qui peut décider si on est en bonne santé ou non, c'est lui, et lui seul, qui peut définir les problèmes sanitaires et choisir les moyens de remédier à une situation jugée mauvaise, non seulement au plan individuel, mais aussi, et c'est ce qui nous préoccupe particulièrement ici, au plan collectif, même s'il a besoin de l'aide d'autres professionnels (administrateurs, économistes, sociologues, épidémiologistes, statisticiens, etc.) pour définir les modalités précises de leur mise en œuvre et les appliquer. Mesurer la santé d'une population de façon purement négative, c'est aussi contribuer à perpétuer le monopole du corps médical sur la santé individuelle comme sur la santé publique, c'est faire obstacle à une prise en charge « globale » des problèmes sanitaires d'une collectivité par des moyens qui ne seraient pas exclusivement médicaux.

5. Vers une mesure opérationnelle de l'état de santé d'une population.

Pour surmonter ces insuffisances, des recherches se sont développées depuis plusieurs années dans deux directions complémentaires : 1) l'extension du concept de morbidité, 2) l'introduction du concept de « santé positive ».

5.1. *L'extension du concept de morbidité.*

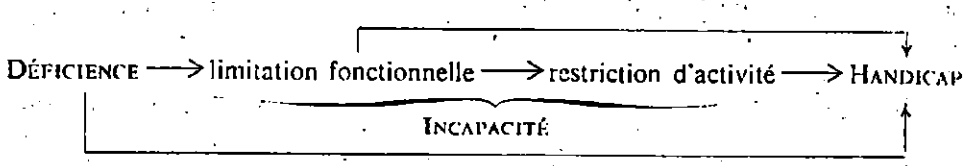
Le type de mesure de morbidité actuellement utilisé possède deux caractéristiques fondamentales qui sont étroitement liées : l'auteur de la mesure est le médecin et le phénomène qui est mesuré (en nombre de décès, en incidence ou en prévalence) est la maladie, pathologie codifiée par la science médicale, classée au moyen d'un système de rangement dont le type est la Classification internationale des maladies de l'O.M.S. De plus, les instruments et les unités de mesure utilisés ne permettent d'appréhender la maladie qu'en terme de « tout ou rien », sans pondération reflétant son importance. Les travaux concernant la mesure de la morbidité qui sont proposés par les chercheurs actuels constituent des approches différentes pour l'un et/ou l'autre de ces deux aspects.

a) Pour ce qui concerne les *phénomènes mesurés*, un grand nombre de recherches ont été entreprises pour obtenir une mesure de la morbidité qui ne repose pas sur la notion de maladie, présente ou absente, mais plutôt sur les *conséquences de la maladie*. Dans divers domaines, l'intérêt de cette approche est de répondre aux deux premières objections que nous avons formulées plus haut concernant la mesure traditionnelle de la santé (évaluation de l'impact socio-économique de la maladie et de l'efficacité du système de soins dans la lutte contre les conséquences de la maladie).

Ce type de mesure repose sur l'appréciation des *limitations fonctionnelles* entraînées par un mauvais état de santé. Depuis de nombreuses années, on utilise à des fins individuelles des classifications de différents types de limitation : ainsi, la réparation des accidents du travail et des maladies professionnelles en France est

basée sur une estimation du préjudice subi à partir d'une classification des limitations occasionnées [89] ; les gériatres ont aussi depuis longtemps essayé de classifier les actes de la vie quotidienne pour l'accomplissement desquels une aide est nécessaire à leurs patients (Skinner et Yett [127] ont fait une revue générale des principales échelles utilisées à cet effet). Le concept d'incapacité engendrée par un mauvais état de santé a fait également l'objet de mesures dans de larges populations : ainsi, les enquêtes de santé nationales qui existent dans différents pays prennent en compte, de façon variable mais le plus souvent très grossière (journées passées au lit ou encore en institution, par exemple), les conséquences les plus visibles de la maladie (pour plus de détails concernant les enquêtes de santé nationales existantes, on peut lire les articles de Linder [84], Armitage [7], sur les enquêtes dans la Communauté européenne, et A. Mizrahi [94] pour l'enquête en cours d'élaboration au Canada).

Plus récemment, Wood, dans le cadre de la 9e révision de la Classification internationale des maladies de l'O.M.S., a proposé une classification générale dont une version française provisoire est disponible [143]. Il s'agit d'une tentative de synthèse reposant sur trois concepts complémentaires : 1) la déficience, définie comme « un terme générique qui englobe toute ingérence ou trouble de structure ou de fonction, contrariant le fonctionnement normal de l'organisme, y compris celui des fonctions mentales » ; 2) l'incapacité, qui est « la perte ou la réduction d'une aptitude fonctionnelle et d'une activité conséquence d'une déficience » ; 3) le handicap, « inconvénient qui résulte de la déficience ou de l'incapacité ». Les trois concepts sont liés logiquement de la façon suivante (*ibid.*, p. 19) :



Afin de permettre une mesure différenciée des besoins sanitaires (et une évaluation de l'utilisation des services) d'une part, et des conséquences socio-économiques des phénomènes pathologiques d'autre part, la classification de Wood est double : elle comporte un code I (pour les déficiences et incapacités) et un code H (pour les handicapés) ; ce dernier comporte six dimensions distinctes et complémentaires : non-dépendance physique, mobilité, activités habituelles, intégration sociale, indépendance financière et évolution future.

Cette classification des déficiences, incapacités et handicaps a été conçue pour servir de base à une mesure des mauvais états de santé dans une population qui ne serait donc pas liée exclusivement à un décompte des cas de maladies répertoriées.

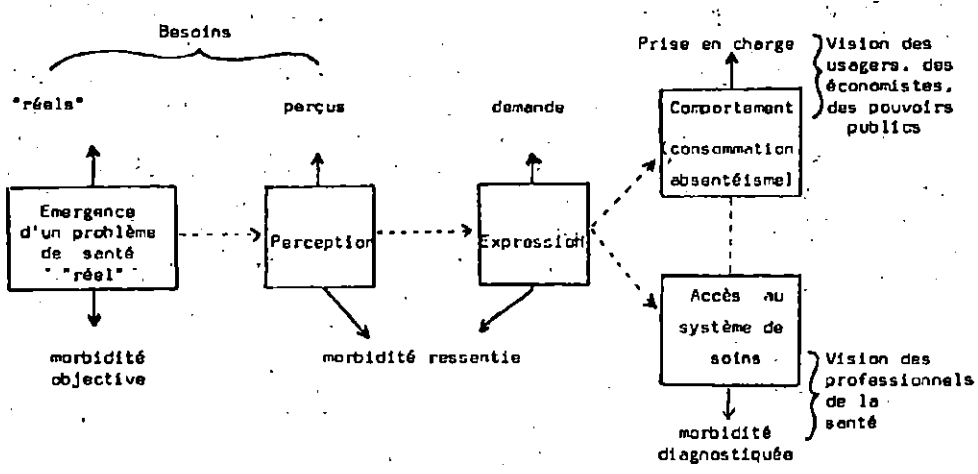
Il faut signaler que le travail de Wood n'a pas la prétention d'être définitif et de nombreux chercheurs continuent de travailler dans différents domaines de la mesure de l'incapacité dans une population : Willliams *et al.* [141], Bynder et

Kong-Ming New [24] sont cités à titre d'exemple, parce que leurs articles contiennent une intéressante bibliographie sur les problèmes de la mesure de l'incapacité pour le premier, sur ses dimensions psycho-sociales pour le second.

b) Pour ce qui est *des méthodes de mesure*, on doit considérer que l'appréciation par les médecins de la morbidité n'est qu'une mesure partielle. Ainsi, Chen et Bryant [32] proposent de distinguer la morbidité, reportée par les patients eux-mêmes, et celle qui est observée par un professionnel, cliniquement ou à l'aide d'examens complémentaires.

Lévy *et al.* [82, pp. 19-99] distinguent morbidité ressentie (par les patients), morbidité diagnostiquée (par les professionnels de la santé) et morbidité objective (telle qu'elle peut être détectée au moyen d'examens systématiques de type bilan de santé). Sullivan [129] fait une distinction entre la morbidité selon qu'elle est appréciée cliniquement (ce qui regroupe la morbidité diagnostiquée et objective), subjectivement par les patients ou par le biais d'un comportement : absentéisme, consommation médicale, etc., c'est-à-dire par un tiers qui n'est ni le patient, ni un professionnel médical. Le schéma suivant est un essai de synthèse destiné à représenter les différents types de morbidité mesurée à la fois selon le stade d'évolution d'un problème de santé « réel », selon l'auteur de la mesure et selon la source des données.

(*Remarque importante :* les guillemets qui entourent l'adjectif « réel » signifient ici que la réalité de problème de santé est contingente à certaines normes ainsi que nous le verrons plus loin. Cette réalité ne peut être appréhendée avec une objectivité au sens d'Auguste Comte, « l'exacte représentation du monde réel », chimère qui masque un sens commun pétri de préjugés. Elle ne peut l'être qu'avec une « objectivité technique » résultant de l'application correcte de l'instrument de mesure dans le cadre des conventions retenues).



A l'évidence, ces différents types de mesure de la morbidité ne sont pas équivalents. Zola [147] a mis en évidence les différences obtenues entre médecins d'une part, et des patients appartenant à plusieurs groupes ethniques d'autre part ; les résultats sont très différents dans chaque cas. L'enquête Boulogne-Billancourt de l'I.N.S.E.R.M. [70] a retrouvé des phénomènes semblables en comparant morbidité déclarée et morbidité mesurée par un examen systématique. Dingle [47], compilant de nombreuses données de mortalité et de morbidité d'origine diverse, fait le point de façon très claire dans un article remarquablement documenté sur les différences rencontrées selon la source des informations. Il constate combien « elles dépendent, à un degré surprenant, du fait que les maladies sont perçues par les patients, les médecins ou les spécialistes des statistiques sanitaires ».

De nombreux biais structurels expliquent souvent de façon évidente les différences observées en fonction des auteurs des mesures. Mais, ce qu'il faut surtout souligner, c'est que dans chaque cas on n'a qu'une vision partielle des phénomènes dans une collectivité, qu'on ne mesure pas la même chose suivant que l'observateur est le patient, un médecin, un système complexe de bilan de santé ou un organisme administratif chargé de régler les prestations sociales occasionnées par la maladie. De même, les méthodes de mesure (statistiques administratives ou de routine médicale, enquêtes par sondage, examens systématiques de groupes de volontaires) ainsi que la population de référence (personnes fréquentant une institution, volontaires d'un centre de bilan de santé, population générale dont est tiré un échantillon) sont extrêmement diversifiés. Purola [115], notamment, a montré combien il est important de pouvoir disposer simultanément de mesures effectuées par des sources différentes : la morbidité ressentie permet essentiellement d'approcher la notion de besoin, donc de demande vis-à-vis du système de santé ; la morbidité diagnostiquée est avant tout le reflet du fonctionnement de l'appareil de soins ; la morbidité objective, quant à elle, veut être une mesure de la prévalence réelle des phénomènes morbides dans une population, en fonction de normes proposées par l'état des connaissances médicales ; enfin, la morbidité comportementale (absentéisme, consommations, etc.) reflète les implications socio-économiques des problèmes de santé ainsi que les comportements face à ces problèmes. Haro [62, 63] et Reinke *et al.* [118] partent des mêmes conclusions et proposent, en se basant sur des expériences réalisées dans des contextes parfois très différents (pays industrialisés et pays en voie de développement), la mise en place de systèmes de recueil de données sanitaires dans une population, mettant en jeu les instruments de mesure les plus variés dans le but d'obtenir une information « totale ». C'est bien sûr la morbidité diagnostiquée qui a fait l'objet du plus grand nombre de travaux, car elle reflète le fonctionnement du système de soins et, à ce titre, sa connaissance est un instrument précieux de gestion de ce système. Ces dernières années ont vu la publication de travaux reflétant un souci de différenciation de certains aspects de la morbidité diagnostiquée, mesure de l'activité de sous-ensembles spécifiques du système de soins. Ils ont en commun une remise en cause de la Classification internationale des maladies de l'O.M.S. sous sa forme actuelle, considérée comme

peu adaptée aux besoins particuliers de tel ou tel secteur de soins. Ceci explique l'apparition de classifications plus spécifiques : le système Snomed [40], correspondant aux besoins hospitaliers, permet de coder un diagnostic suivant cinq dimensions simultanément (topographie, étiologie, morphologie, troubles fonctionnels et procédures entreprises) ; la Classification internationale des problèmes de santé dans les soins primaires [99] établie pour mesurer l'activité des services de soins de premier contact : la Classification des symptômes pour les malades ambulatoires [98] qui a pour but de permettre un suivi des malades non hospitalisés.

Bien que deux au moins de ces classifications existent en version française, et qu'elles fassent l'objet d'une utilisation parfois déjà ancienne et très étendue dans les pays industrialisés (notamment SNOMED), elles sont pratiquement inconnues en France et n'y sont utilisées par personne à notre connaissance.

Par contre, des classifications à visée non plus générale sur le plan du champ pathologique couvert, mais spécifiques à une discipline médicale, ont eu parfois un meilleur accueil, bien que limité habituellement à quelques services hospitaliers universitaires.

Cependant, il faut souligner que l'ensemble des types de mesure de la morbidité que nous avons décrit ont une caractéristique commune : ils reposent sur des observations individuelles, ces observations pouvant être ensuite agrégées de façon et à un niveau variables pour fournir une mesure de la morbidité d'un groupe ou d'une population. Or, de nombreux auteurs, notamment sociologues, ont mis en évidence le fait que les problèmes de santé apparaissant dans un groupe ou une population ne sont pas simplement la somme des problèmes individuels qui s'y rencontrent. Ainsi, Lerner [78] insiste sur la spécificité apportée aux problèmes de santé d'une communauté par les liens sociaux institutionnels, géographiques, etc. Kelman a montré [75] la nature sociale de la définition des problèmes de santé. Très peu d'applications opérationnelles d'une mesure collective de la morbidité d'une population existent actuellement ; il faut cependant signaler la très intéressante expérience italienne menée par des syndicalistes dans certaines entreprises. Elle repose sur l'appréciation collective des problèmes de santé de groupes d'ouvriers par ces groupes eux-mêmes. Cette expérience est décrite dans quelques publications rédigées en français [45, 136].

5.2. *L'introduction du concept de « santé positive ».*

Les efforts pour étendre le concept de morbidité (en y introduisant les notions de déficience, d'incapacité, et de handicap, en multipliant les classifications spécialisées, soit dans un domaine pathologique, soit dans un secteur institutionnel du système de soins, en acceptant des mesures effectuées à différents niveaux sociaux et institutionnels de l'expression des phénomènes morbides) ont pour but de pallier certaines insuffisances des indicateurs classiques de la santé d'une population. Mais il n'en reste pas moins qu'ils n'en permettent toujours qu'une mesure négative, ainsi que nous l'avons signalé.

Si la santé « n'est pas la simple absence de maladie ou d'infirmité » comme l'affirme l'O.M.S. dans la définition déjà citée, il faut alors définir un niveau supérieur qui est le domaine de la « santé positive ». Pour Sullivan [129], ce domaine commence à partir de l'absence de toute morbidité apparente ; de plus, pour être l'objet d'une mesure, il est nécessaire de pouvoir identifier des niveaux distincts à l'intérieur même du domaine de la santé positive. Il résulte de cette double constatation que pour parvenir à définir une mesure opérationnelle du niveau de santé d'une population, qui englobe également le champ de la santé positive, deux étapes doivent être parcourues : 1) conceptualisation de la notion de santé avec introduction de niveaux positifs ; 2) classification des états de santé positive et négative.

1. *Conceptualisation de la notion de santé.* La définition de la santé de l'O.M.S. ouvre la voie à une mesure positive en indiquant les trois dimensions, physique, mentale et sociale, dans lesquelles celle-ci peut s'exprimer. Mais elle reste restrictive en ceci qu'elle n'envisage « qu'un état de complet bien-être ». Or, une étude de l'abondante littérature consacrée au problème de la définition de la santé montre qu'en fait on peut distinguer d'une façon schématique, chez les différents auteurs, trois modes d'abord, souvent complémentaires mais distincts au moins dans leurs conséquences pratiques : un abord *perceptuel*, c'est-à-dire fondé essentiellement sur la notion de bien-être, un abord *fonctionnel*, et un abord centré sur le *concept d'adaptation*.

Parmi les définitions de la santé qui privilégient le côté *perceptuel*, on peut classer celle de Sigerist [126] : « La santé n'est pas simplement l'absence de maladie, c'est quelque chose de positif, une attitude joyeuse envers la vie et une acceptation réjouie des responsabilités que la vie impose à l'individu », ou celle de Hoyman [67] : La santé est une aptitude personnelle optimale pour une vie pleine, fructueuse et créatrice. » On peut également ranger la définition de l'O.M.S. dans cette catégorie.

L'abord *fonctionnel* se caractérise par une approche statique de la définition de la santé comme capacité de l'individu à fonctionner : ainsi, pour Lifson [83], « la santé est le degré avec lequel les fonctions d'un être humain (sens, traitements des informations, déplacement...) sont accomplies alors que la douleur est absente » ; pour le comité d'experts de l'O.C.D.E., « la notion de santé positive signifie un peu plus que l'absence de restrictions imposées aux activités quotidiennes et couvre alors les autres activités qui n'appartiennent pas nécessairement à la vie quotidienne » [108, p. 63] ; Parsons [109] considère la santé comme « l'état de capacité optimale d'un individu pour l'accomplissement efficace des rôles et des tâches pour lesquels il a été socialisé » ; Bonnevie [18] estime que « la santé est déterminée par une capacité comportementale incluant des composantes biologiques aussi bien que sociales à accomplir des fonctions fondamentales ».

Un autre groupe d'auteurs, quant à lui, propose une approche de la santé qu'on pourrait qualifier de « mixte », qui *rassemble des éléments perceptuels* et *fonctionnels* dans une même définition. Terris [135] propose une version modifiée

1.28

de la définition de l'O.M.S. : « La santé est un état de bien-être physique, mental et social et de capacité à fonctionner et non simplement l'absence de maladie ou d'infirmité. » Pour Blum [17], « la santé consiste : 1) en la capacité de l'organisme à maintenir un équilibre approprié à son âge et à ses besoins sociaux, dans lequel il est raisonnablement indemne de profonds inconforts, insatisfaction, maladie ou incapacité et 2) à se comporter d'une façon qui permette la survie de l'espèce aussi bien que l'accomplissement personnel ou le contentement de l'individu ».

Enfin, l'approche reposant sur le *concept d'adaptation*, déjà présent dans la définition de Blum, est prônée par plusieurs auteurs. Rossdale [120] considère la santé comme « un produit d'une relation harmonieuse entre l'homme et son écologie » ; Gordon [59] est plus restrictif puisque, pour lui, la santé est exprimée par « des niveaux de résistance à la maladie » ; Canguilhem [26] a une conception voisine : « La santé, c'est la marge de tolérance des infidélités du milieu (...), c'est le volant régulateur des possibilités de réactions (...) ; être en bonne santé, c'est pouvoir tomber malade et s'en relever, c'est un luxe biologique. » Pour Wylie [144], « la santé est l'ajustement parfait, permanent d'un organisme à son environnement. Inversement, la maladie serait un ajustement imparfait ». Enfin, Bonnevie [18] complète la définition déjà citée en affirmant que les « fonctions fondamentales » qu'il considère ne peuvent être accomplies que grâce à « l'adaptation, définie comme une interaction réciproque entre l'être humain expliqué de façon biopsychosociologique et ses conditions de vie déterminées socio-écologiquement ».

En dehors de l'imprécision et des aspects discutables de la plupart des définitions citées, il est un problème de fond commun à toutes : quelle que soit la définition adoptée, on n'échappe pas à la nécessité d'introduire des *normes* dans la définition, donc la mesure, de la santé. C'est ce qui est indirectement suggéré par l'emploi d'adjectifs tels que « optimal », « approprié », « raisonnable », etc. Les différents problèmes liés à une attitude normative ont été discutés par de nombreux auteurs qui mettent en relief les difficultés liées à la variabilité de toute norme sociale. Ce qui est perçu comme problème de santé dans tel contexte culturel, idéologique ou professionnel ne l'est pas nécessairement ailleurs. Nous avions déjà rencontré ce problème à propos des différents types de morbidité mesurés. Il prend d'autant plus d'importance ici qu'il apparaît aussi bien au niveau de la définition même de la santé qu'à celui de l'opérationnalisation de la mesure de celle-ci : comment décider que tel état de santé est meilleur que celui-là ? Après tout, selon l'adage populaire, « le plus grand des nains à la même taille que le plus petit des géants ».

Pour le lecteur intéressé par cet aspect, nous citerons quelques travaux caractéristiques de la réflexion sur les attitudes normatives dans le domaine de la santé. Majone [88] discute du problème général et des techniques utilisables pour l'établissement de normes de santé ; pour ce qui concerne l'établissement de normes dans le domaine biologique et physique (au sens de la définition de l'O.M.S.), Cochrane [39] et Chen [31] donnent des points de vue intéressants, et Jénicek [72, pp. 29-33] en fait une revue assez large ; dans le domaine des normes subjec-

tives (normes du bien-être correspondant à l'abord perceptuel), Knoppreth *et al.* [76] proposent une revue générale très claire des techniques utilisables lorsqu'on veut établir de telles normes. Arkhipoff [6] et Tornstam [137] abordent des problèmes plus théoriques concernant la possibilité d'établir des normes acceptables, sinon idéales ; enfin, pour l'aspect social, Dreitzel [48] et surtout Kelman [75] l'analysent de façon détaillée, aussi bien pour ce qui concerne la nature des activités sociales prises en compte dans une définition de la santé, que pour la variation dans la conception même de la santé liée à différents types de sociétés. On peut, en guise de conclusion, citer le mot de Dumont et Latouche [49] qui résume le point de vue des auteurs qui ont exploré ce domaine : « Vécue individuellement, la santé se définit socialement. Une définition de la santé ne peut être intemporelle. » Cependant, nous n'avons pas l'intention, dans le cadre de cet article, de discuter plus au fond les différentes approches décrites, et les définitions de la santé qui s'y rattachent. Ce qui nous concerne ici, ce sont les implications que des conceptions aussi différentes imposent au niveau du type de mesure différent aussi bien pour ce qui est des phénomènes mesurés que des méthodes de mesure.

2. *Classification des états de santé.* Malgré leurs différences essentielles, les trois types d'approche d'une définition de la santé que nous avons individualisés ont en commun de ne pas reposer sur la notion de maladie ; or, celle-ci offrait l'avantage pratique indéniable de fournir un cadre référentiel commode pour une classification des phénomènes de mauvaise santé. Vouloir dépasser ce stade impose donc la création de nouveaux cadres référentiels de classification, dont le résultat sera la fabrication de catégories de ce qu'il est devenu courant d'appeler des « *états de santé* » (les adjectifs « bon » ou « mauvais » pouvant y être accolés suivant l'occasion afin de les distinguer nettement des entités morbides qui peuvent les conditionner).

Il faut souligner qu'il s'agit là d'une étape de niveau plus synthétique que celle qui consiste à établir une classification de troubles isolés, comme la Classification des déficiences, incapacités et handicaps de Wood par exemple, que nous avons décrite plus haut : en effet, ce qui est en question ici, c'est la définition d'*états de santé globaux*, ces états de santé globaux pouvant être définis, bien entendu, en fonction de troubles reconnus qui, isolément ou en association, définissent des états de santé globaux individualisés. Sous-jacente à cette approche apparaît l'idée explicite ou implicite introduite par Reed [116] qu'il existe un *continuum des états de santé* allant du meilleur au plus mauvais et permettant d'établir des listes ordonnées d'états de santé.

Une telle façon de faire semble être la voie obligée pour obtenir une mesure opérationnelle de niveaux de santé positive, rien n'empêchant alors de distinguer parmi des états de santé globaux un ou plusieurs états distincts se situant au-dessus d'une limite où les individus sont libres de toute morbidité apparente. De plus, en dehors de la possibilité théorique d'appréhender la notion de santé positive, un avantage essentiel du concept d'état de santé réside dans son caractère global, synthétisant l'ensemble des phénomènes caractérisant la santé d'un individu.

Bien entendu, l'établissement d'une classification des états de santé globaux dépend au premier chef de la définition de la santé adoptée par l'auteur de la classification. On devrait donc s'attendre à retrouver trois types distincts de classification des états de santé globaux, basés respectivement sur une approche fonctionnelle, une approche perceptuelle et une approche « adaptative ». Or, dans la pratique, les choses ne sont pas si bien tranchées et on trouve essentiellement des classifications assez peu homogènes de ce point de vue, où les aspects perceptuels et fonctionnels sont mêlés, avec cependant l'accent porté plus fortement sur l'un ou l'autre de ces aspects selon les auteurs. Quant à l'approche « adaptative », il est curieux de constater que malgré la place importante qu'elle tient dans bon nombre de définitions de la santé, elle n'a pas fait l'objet (sauf omission de notre part) de travaux visant à baser une mesure de la santé sur les phénomènes d'adaptation. Il existe certes une abondante littérature relatant des recherches dans certains domaines de l'adaptation, notamment en ergonomie et en psychologie. Mais à notre connaissance aucun indicateur de santé qui reposerait sur une mesure des capacités d'adaptation n'a été proposé, bien que pour Terris [135] par exemple, une « épidémiologie de la santé » ne puisse se développer qu'à travers des études portant notamment sur les notions voisines de performance, de capacité de performance et d'obstacles aux performances.

Les auteurs ayant établi des classifications d'états de santé globaux utilisent le procédé suivant. On établit une liste d'*activités* et/ou de *situations* et/ou de *jugements*, qui peuvent être positifs ou négatifs, puis chaque individu est confronté avec cette liste. Dans certains cas, la liste est ordonnée allant de l'état le meilleur au plus mauvais, et chaque rubrique représente un état de santé défini par les activités (ou situations ou jugements) correspondantes ; dans d'autres cas, la liste ne contient pas d'ordre, mais on affecte un score pour chaque rubrique, le total obtenu par un individu permettant de la classer dans un groupe correspondant à un état de santé.

Cette description est bien sûr très schématique, et on trouve dans la littérature un très grand nombre de variantes, aussi bien en ce qui concerne la façon d'établir la liste de référence, que son contenu, la façon d'affecter chaque patient à une catégorie et de la situer par rapport à chaque élément de la liste, etc. Nous verrons plus loin quelques exemples typiques des différentes approches essayées. D'autre part, il est évident qu'une telle façon de faire repose sur l'arbitraire le plus total, chaque auteur imposant implicitement ou explicitement sa conception de la santé et le choix des phénomènes observables qui la reflètent. Cela condamne-t-il pour autant la méthode ? Sans doute pas si on en admet clairement les limites, qui sont essentiellement liées, ainsi que nous l'avons déjà dit, à l'aspect normatif de toute mesure de la santé. A l'heure actuelle, il n'existe pas d'accord général sur ce que doit être une classification satisfaisante des états de santé, et au fond il n'y a pas de raison qu'on y parvienne jamais, puisque pour cela il faudrait d'abord qu'il y ait accord sur une définition de la santé. Or, nous avons vu que cela n'était pas réaliste, ni peut-être souhaitable en raison de la contingence inhérente à une telle

définition. Mais ceci étant admis, à l'intérieur du cadre offert par une définition explicitement établie, la méthode consistant à établir une classification des états de santé peut s'avérer précieuse sur le plan opérationnel, notamment lorsqu'on cherche à exprimer quantitativement le niveau de santé d'une population.

Sur le plan pratique, une mesure de la santé indépendante de la nature précise des phénomènes morbides permet en effet d'évaluer des situations complexes : ainsi, par exemple, si on prend le cas d'un malade atteint de tuberculose que l'on parvient à guérir, comment évaluera-t-on, s'il est secondairement victime d'un ulcère duodénal, l'éventuel « gain sanitaire », si on se cantonne à l'utilisation de catégories diagnostiques pour décrire son état de santé ? Une mesure basée sur l'appréciation de l'état de santé global n'offre pas les mêmes difficultés ; cependant, on sent bien que d'autres problèmes se posent dans ce dernier cas pour ce qui concerne la quantification (peut-on dire que le malade souffrant d'un ulcère duodénal est en meilleure ou plus mauvaise santé que lorsqu'il avait une tuberculose pulmonaire ?) ; nous discuterons de ces difficultés plus loin *.

* La bibliographie complète paraîtra à la fin de la seconde partie de cet article.

© Masson, Paris, 1979

Rev. Epidém. et Santé Publ., 1979, 27, 133-152.

Revue générale

Indicateurs de santé et « sanométrie » : les aspects conceptuels des recherches récentes sur la mesure de l'état de santé d'une population.

SECONDE PARTIE *

M. GOLDBERG, W. DAB, J. CHAPERON, R. FUHRER, F. GRÉMY.

Paris (France), Boston (U.S.A.).

6. L'EXPRESSION QUANTITATIVE D'UNE MESURE DU NIVEAU DE SANTÉ D'UNE POPULATION NE REPOSANT PAS SUR LES CONCEPTS « CLASSIQUES ».

Nous n'avons pas l'intention de faire ici une véritable revue générale exhaustive des très nombreux indices proposés récemment par les chercheurs ; nous nous contenterons d'en décrire quelques-uns que nous estimons particulièrement intéressants et représentatifs des principales écoles de pensées.

Cependant, pour le lecteur intéressé par des aspects plus techniques et plus précis, ou à la recherche d'un éventuel indice correspondant à ses besoins, il existe une assez abondante littérature critique et bibliographique dans ce domaine, se présentant sous différentes formes : livres, rapports, publications périodiques, revues générales. Nous citerons quelques ouvrages particulièrement intéressants et facilement accessibles. Parmi ceux de nature bibliographique, Reeder, Ramacher, Gorelnik [117] présentent une analyse des principaux indices du domaine de la santé, avec un chapitre spécialement consacré aux indices d'état de santé ; le Bureau de la planification sanitaire du Ministère fédéral de la Santé et du Bien-être des Etats-Unis publie périodiquement des références bibliographiques contenant des résumés d'articles concernant les méthodes de mesure du niveau de santé dans une communauté [20] ; il existe également aux Etats-Unis, depuis 1973, un Bureau de documentation sur les indices de Santé, installé au Centre national de statistiques sanitaires dont l'adresse est donnée aussi en référence [37] : ce bureau publie trimestriellement une bibliographie annotée qui est adressée sur simple demande. Plusieurs conférences scientifiques sur les indices d'état de santé se sont

* La première partie de cette revue générale est parue dans Rev. Epidem. et Santé Publique n° 1, 1979 (p. 51-68).

déjà tenues : le compte rendu d'au moins une d'entre elles a été publié [11] et nous en avons déjà cité dans cet article plusieurs contributions individuelles. Un numéro spécial d'une importante revue a été récemment consacré à ce sujet [52]. L'O.M.S., de son côté, a également organisé des conférences et créé des groupes d'experts sur le sujet. Nous les avons déjà cités plus haut [101, 104, 105].

Il existe également un assez grand nombre d'études synthétiques critiques qu'on peut classer en deux catégories selon qu'elles traitent des problèmes généraux des indices de santé et notamment de leur validité, ou qu'elles se focalisent sur certains points particuliers. Dans la première catégorie, les travaux les plus intéressants sont certainement ceux de Sullivan [129], Bickner [15], Goldsmith [57, 58], Rosser [121], Balinski et Berger [8], Sackett *et al.* [123], Siegmann et Elinson [125], Lévy *et al.* [81], Jazairi [71]. L'ensemble des problèmes conceptuels, méthodologiques et techniques sont abordés par ces auteurs.

Dans le second groupe, on trouve des travaux tels que ceux de Miller [92] ou de Chen et Bryant [32] qui proposent une classification opérationnelle des indices de santé dans le but de faciliter un choix en fonction des besoins de l'utilisation. D'Agostino [2] expose les difficultés de choix, de recueil, de traitement et d'analyse statistique des données nécessaires pour calculer des indices généraux ; Torrance [138] traite les aspects mathématiques de la construction d'indices d'état de santé et propose un modèle général permettant de décrire formellement les indices proposés par différents auteurs ; Chen, Bush et Zaremba [34] passent en revue les problèmes posés par l'utilisation d'index de santé pour des études de mesure d'efficacité et analysent plusieurs indices proposés ; enfin, certains travaux traitent des index de santé utilisables dans un environnement particulier. A titre d'exemple, Baylet et Benyoussef [9] s'intéressent aux indices nécessaires dans les pays en voie de développement ; un groupe d'étude de l'O.M.S. [107] a rédigé un rapport sur les indices de santé de la famille qui contient une abondante bibliographie.

Dans un but descriptif, nous séparerons grossièrement en trois groupes les indices d'état de santé d'une population non fondés uniquement sur les concepts de mortalité et de morbidité « classique » qui sont proposés dans la littérature : 1) ceux qui utilisent les catégories diagnostiques habituelles mais y ajoutent des éléments de pondération permettant d'obtenir un indice unique qui fournit une mesure synthétique ; 2) ceux qui proposent une mesure unique reposant à la fois sur la mortalité et la morbidité « étendue », c'est-à-dire exprimée par les notions d'incapacité et de handicap, indépendamment de toute catégorie diagnostique ; 3) ceux qu'on peut nommer du terme générique d'« Index d'état de santé globale » qui sont fondés sur une tentative d'appréhension globale du niveau de santé des individus.

Ce découpage en trois groupes ne représente pas une véritable tentative de classification des indices d'état de santé. Il n'est en effet basé que sur un seul critère, la nature des phénomènes mesurés, qui nous semble le plus important au niveau de la généralité où nous voulons rester dans le cadre de cet article.

6.1. *Indices établis à partir de la morbidité diagnostiquée pondérée.*

Dans ce groupe, on trouve des indices proposés pour remédier partiellement aux insuffisances d'une mesure basée sur un simple comptage des maladies diagnostiquées. En faisant intervenir certains éléments de pondération, chaque auteur espère parvenir à une expression quantitative de la morbidité qui reflète au moins un aspect des problèmes posés par la maladie.

Magdelaine *et al.* [87] calculent pour chaque maladie, ou association de maladies, d'un patient un degré de morbidité qui repose simultanément sur l'appréciation de trois composantes de la gravité : invalidité, risque vital, durée des diverses affections dont souffre l'individu. Un score final compris entre 0 et 6 est attribué à chaque patient par combinaison de ces divers éléments. Culyer *et al.* [41] ont proposé un index très comparable.

Wyler *et al.* [145, 146] se proposent de mesurer un aspect particulier de l'impact de la maladie : le degré de gravité tel qu'il est globalement perçu par les individus. Pour cela, ils ont dressé une liste de 126 maladies soigneusement choisies et ont demandé à un groupe de médecins et de non-médecins d'attribuer à chacune une note indiquant quel est d'après eux son degré relatif de gravité. En combinant les notes attribuées par l'ensemble des participants, ils obtiennent une note générale pour chaque maladie qui est l'indice de gravité recherché.

6.2. *Indices établis à partir de la morbidité « étendue » ou d'une combinaison de la mortalité et de la morbidité « étendue ».*

Il s'agit, dans ce groupe, d'indices reflétant la mortalité et/ou la morbidité, non plus en fonction des catégories médicales habituelles mais en fonction de leur impact exprimé par la notion d'incapacité (ce que nous avons appelé plus haut morbidité « étendue »), la mort n'étant considérée dans cette optique que comme l'expression d'une incapacité totale et définitive. Il faut noter tout d'abord que la distinction entre les concepts d'incapacité et de handicap, clairement établie par Wood dans le travail que nous avons cité plus haut, ne se retrouve pas dans les différents indices proposés dans cette catégorie, sauf erreur de notre part. Ces deux concepts ne font pas encore l'objet de mesures distinctes et ne sont pas différenciés dans la construction d'indices de santé. Ceci tient sans doute au caractère récent des recherches dans ce domaine et aux multiples difficultés d'ordre conceptuel et pratique qu'elles rencontrent. Dans ce qui suit, nous utiliserons donc le terme d' « incapacité » en sachant qu'il peut recouvrir également la notion de handicap.

Collishaw [38] propose que deux approches distinctes de l'incapacité soient prises en compte pour la construction d'indices : une approche basée sur l'aspect temporel et une approche basée sur l'aspect fonctionnel, ces deux approches n'étant pas exclusives, mais au contraire complémentaires. Pour l'auteur, en effet, qui considère que l'élément commun à toutes les situations d'incapacité est l'interruption d'une activité sociale normale, le premier aspect reflète mieux les problèmes

posés par les incapacités temporaires de courte durée et le second, ceux liés aux cas d'incapacité permanente ou de très longue durée.

Cependant, si cette distinction est fondamentale sur le plan conceptuel, elle ne peut être faite de façon totale dans la pratique. En effet, pour rendre sensible une mesure de l'incapacité, l'aspect temporel doit être pondéré par l'importance de l'aspect fonctionnel et vice versa.

On sent bien qu'un autre problème de fond, pour construire des indices d'incapacité, est la définition des différents degrés d'incapacité exprimant l'atteinte fonctionnelle. Ce problème a été attaqué par plusieurs auteurs que nous avons déjà cités plus haut [24, 127, 141, 143] ainsi que par Sullivan [131] qui aborde le sujet spécifiquement du point de vue de la construction d'un indice de santé. Il n'existe pas cependant à l'heure actuelle de consensus dans ce domaine. Le comité d'experts de l'O.C.D.E. pour les indicateurs de santé le reconnaît et insiste sur l'importance et l'urgence des recherches concernant ce sujet.

Malgré l'état provisoire des résultats concernant les recherches dans le domaine de la mesure de l'incapacité, différents indices reposant sur elle ont été proposés. Bien que leur validité soit sujette à caution, pour la raison que nous venons d'indiquer, ils montrent bien quel type d'expression quantitative pourrait être utilisé si on parvient à une mesure opérationnelle de l'incapacité au niveau individuel. C'est pourquoi nous citerons les principaux d'entre eux, sans discuter la nature et la validité de la mesure de l'incapacité sur laquelle ils reposent et qui est très variable d'un auteur à l'autre en l'état actuel des choses.

Un exemple d'indicateur construit pour rendre compte de l'incapacité temporaire est donné par le Centre national des statistiques sanitaires des U.S.A. [97] qui, à partir de son enquête de santé nationale par interview, calcule trois indicateurs différents d'incapacité, qui ne sont pas basés directement sur l'appréciation quantitative directe des limitations fonctionnelles, mais qui reflètent globalement une hiérarchie dans l'importance de cette limitation : 1) le nombre de jours de restriction d'activité, définis comme les jours où une personne réduit partiellement ses activités habituelles ; 2) le nombre de jours de perte d'activités majeures, où une personne, qui habituellement va à l'école, travaille, ou effectue des tâches ménagères, ne peut accomplir ces activités majeures ; 3) le nombre de jours passés au lit, à domicile ou à l'hôpital.

Pour ce qui concerne l'incapacité permanente dans une population, le comité d'experts de l'O.C.D.E. déjà cité propose l'élaboration de deux indicateurs statistiques [108, pp. 67-68] : 1) la proportion prévue de leur vie future que passeront, en état d'incapacité, les individus ne souffrant pas de déficiences permanentes à l'âge de 1 an, 20 ans, 40 ans et 60 ans ; 2) proportion, dans certaines tranches d'âge, de personnes souffrant d'incapacité du fait d'une déficience permanente.

L'O.M.S. [106] préconise, dans le cadre des études coût-efficacité nécessaires pour le choix et l'évaluation des résultats des programmes de santé, l'emploi de trois indicateurs distincts destinés à mesurer le « gain sanitaire » attendu : 1) les Années d'espérance de vie (A.E.V.) gagnées grâce au programme s'il s'agit de

lutter contre une maladie dont l'issue peut être fatale ; 2) les Années-personne de bonne santé (A.B.S.) pour les maladies dont on peut attendre une guérison totale ; 3) les Années de validité (A.V.) dans les cas de maladies laissant des séquelles permanentes quelle que soit l'efficacité du traitement. Ces trois indicateurs sont calculés par cumul des années gagnées pour l'ensemble de la population-cible du programme par rapport à ce qu'on aurait observé sans le programme. Le dernier de ces indicateurs (A.V.) repose bien entendu sur une mesure de l'incapacité dans la population avec et sans programme.

Diderot [46] vers la fin de sa vie était conscient, en faisant des projets d'avenir, de la nécessité de retrancher à son espérance de vie les années perdues par incapacité : « J'ai peut-être encore une dizaine d'années au fond de mon sac. De ces dix années, les fluxions, les rhumatismes et le reste de cette famille incommode en prendront deux ou trois. » Il définissait ainsi le principe d'un indice unique reflétant à la fois la mortalité et l'incapacité, préconisé par bon nombre d'experts actuels.

Ainsi, Sullivan [130] propose le calcul de deux indices synthétiques complémentaires : l'Espérance de vie sans incapacité et l'Espérance d'incapacité. La méthode de calcul du premier est la même que pour l'Espérance de vie mais en retranchant le temps perdu pour incapacité qui forme ainsi le second indice. L'auteur a calculé ces deux indices à titre expérimental à partir des données de l'enquête nationale de Santé des Etats-Unis pour les informations concernant l'incapacité. En fait, il a essayé deux versions distinctes des indices, avec une définition plus ou moins extensive de l'incapacité dans chaque cas : les résultats sont bien entendu différents et mettent en évidence l'importance de l'établissement d'une définition non ambiguë de l'incapacité.

C'est également la conclusion à laquelle arrivent les experts de l'O.C.D.E. [108, p. 72 et 206] à propos de l'Espérance de vie en bonne santé calculée sur des données japonaises et américaines : les données d'incapacité dans ces deux cas sont dissemblables et interdisent toute comparaison. Mais pour eux comme pour Sullivan, le calcul d'un tel indice est possible et pourvu d'un grand intérêt, car il complète très utilement l'espérance de vie simple.

Une autre façon d'utiliser simultanément des données de mortalité et de morbidité « étendue » est l'indice proposé par Morrow *et al.* [96]. Il s'agit d'une estimation du nombre de jours de bonne santé perdus chaque année pour les principales maladies du fait des décès prématurés et des incapacités temporaires ou permanentes. La publication citée donne un exemple de calcul de cet indice pour le Ghana, à des fins d'établissement de priorités sanitaires. Il est dans ce sens très proche des Années potentielles de vie perdues [119] déjà cité mais celui-ci ne reposait que sur des données de mortalité.

Un autre indice de ce groupe, calculé simultanément à partir de l'incapacité et de la mortalité, a été développé dans un but plus ponctuel. Il s'agit de l'Indice de gravité des accidents du travail [51] qui permet une mesure synthétique de la gravité des accidents du travail dans une collectivité à des fins de comparaisons

diverses. Cet indice est le rapport au nombre d'heures total de travail pour l'année du nombre de jours « conventionnels » d'arrêt de travail pendant une année dus aux accidents. Ce nombre de jours « conventionnels » est lui-même obtenu en tenant compte de l'incapacité de courte durée (nombre de jours d'arrêt réel), de l'incapacité permanente (taux d'invalidité permanente) et de la mortalité (accidents mortels) qui font l'objet d'une pondération différente dans le calcul des jours « conventionnels ».

6.3 *Indices d'état de santé globale.*

Dans ce groupe, on trouve des indices qui ont en commun deux caractéristiques complémentaires : 1) ils reposent sur la définition d'états de santé globale, au sens où nous l'avons défini plus haut, et sont donc indépendants des catégories diagnostiques médicales habituelles (contrairement aux indices du premier groupe) ; 2) ils reposent sur l'hypothèse de l'existence d'un continuum des états de santé, allant du meilleur au pire : tantôt cette hypothèse est exprimée par la constitution *a priori* d'une liste ordonnée d'états de santé et il s'agit d'affecter chaque patient à un état de la liste (nous parlerons dans ce qui suit « *d'états de santé d'emblée globale* ») et tantôt elle se traduit par l'affectation de scores aux résultats des mesures partielles effectuées, suivie de la mise en œuvre d'une méthode d'agrégation des scores partiels pour obtenir un indice global dont la valeur se situe dans les limites de l'échelle de notation représentant le continuum des états de santé.

Mises à part ces deux caractéristiques communes, les index proposés sont extrêmement variés selon plusieurs critères, ainsi que nous l'avons signalé plus haut. Nous allons décrire quelques-uns de ces index, choisis parmi les plus représentatifs des principales écoles de pensée, puis nous examinerons les principaux problèmes de fond qu'ils suscitent (nous renvoyons le lecteur au début du paragraphe 6 pour une bibliographie générale concernant les très nombreux indices d'état de santé globale décrits dans la littérature).

Nous ferons dans la description qui suit une distinction entre les modèles d'indices d'état de santé individuels et les modèles d'indices d'état de santé de population, construits soit par agrégation d'indices individuels d'états de santé globale, soit à partir de mesures individuelles non globales, la méthode d'agrégation permettant dans ce cas de définir des états de santé globale au niveau de la population elle-même.

a) *Modèles d'indices d'état de santé globale individuels.*

On peut schématiquement classer en trois catégories les indices individuels selon la conception qu'ont les auteurs de la santé globale d'un patient, des phénomènes qui la représentent et des méthodes de mesure de ces phénomènes : 1) les indices reposant sur des mesures physiologiques « objectives » ; 2) les indices reposant sur les mesures d'incapacité fonctionnelle ; 3) les indices reposant essentiellement sur un abord perceptuel.

1. Les indices reposant sur des mesures physiologiques.

L'indice « -H » de M. K. Chen [31] est un modèle théorique destiné à permettre une mesure quantitative de l'état de santé physiologique d'un individu appartenant à un groupe de population homogène bien défini. Il se justifie par la relativité du résultat des mesures physiologiques en fonction du groupe auquel appartient un individu. Le modèle proposé prend donc en compte les résultats de l'ensemble du groupe aux diverses mesures, et permet une expression agrégée des résultats individuels pondérés par ceux-ci et par la taille du groupe de référence. Tel qu'il est présenté, l'indice -H constitue un type de réponse au problème de la normativité que nous avons évoqué plus haut ; mais cette réponse n'est que très partielle car elle ne concerne que les normes à appliquer aux résultats de mesure et non les normes conceptuelles d'une définition de la santé.

2. Les indices reposant sur des mesures d'incapacité fonctionnelle.

Un des indices les plus classiques de ce groupe est celui de Grogono et Woodgate [61] ; il a, de plus, le mérite d'être simple à la fois sur le plan conceptuel et sur celui de sa mise en œuvre, ce qui explique qu'il ait servi de modèle de base à de nombreux auteurs. Dix activités ou aspects de la vie quotidienne ont été choisis de telle sorte qu'ils soient exhaustifs tout en évitant des redondances évidentes : travail, loisirs, souffrance physique, souffrance mentale, communication, sommeil, dépendance d'autrui, nutrition, excrétion et activité sexuelle. Pour chacune de ces rubriques, un observateur entraîné affecte une note qui peut être 1 (normal), 0,5 (gêne) ou 0 (incapacité) ; le total des notes partielles, divisé par dix, est l'indice global d'état de santé. Cet indice a été utilisé, à titre expérimental, pour classer par ordre de gravité un certain nombre de patients hospitalisés atteints de diverses affections.

Skinner et Yett [127] ont construit leur *Index de débilité* pour répondre à un besoin précis : celui de la gestion d'institutions de soins pour malades chroniques. Il s'agit là d'une des applications les plus anciennes des indices d'état de santé globale, pour laquelle de nombreux indices basés sur les activités de la vie quotidienne ont été développés ; ces indices dont le principe a été établi par Katz [73, 74] sont appelés dans la littérature anglo-saxonne du terme générique d' « ADL indexes » (Activities of Daily Life Indexes). L'Index de débilité de Skinner et Yett est particulièrement intéressant dans ce domaine, à la fois parce qu'il en représente l'état le plus achevé et parce que les problèmes méthodologiques de construction et de validation de l'index sont sérieusement traités. Il comporte cinq types d'activités. Pour chacune d'elles, différentes possibilités sont listées de façon à obtenir une échelle de Gutman (échelle cumulative où une réponse positive à une possibilité implique une réponse également positive à toutes celles qui suivent). Les différentes combinaisons des réponses sont regroupées pour définir un certain nombre de « classes de débilité » qui sont autant d'états de santé globale. Il s'agit donc d'une mesure d'emblée globale. Mais, en toute rigueur, on ne peut parler

d'index car il n'y a pas d'expression quantitative globale pour chaque individu qui est simplement affecté à une classe : il faudrait attribuer à chaque classe un score qui serait l'indice d'état de santé recherché.

C'est ce que propose Bush. Dans une première publication avec Fanshel [54], les auteurs indiquent les principes de construction d'un index d'état de santé d'emblée globale applicable au niveau individuel, mais destiné à être l'élément de base d'un indice de population. Nous ne décrirons ici que la composante individuelle de l'indice, son extension à un indice de population étant décrite plus loin.

L'élément central de l'indice individuel est ici, comme pour l'Index de débilité de Skinner et Yett, une échelle ordonnée d'états fonctionnels construite par combinaison d'activités élémentaires. Les activités élémentaires ont été choisies de façon à s'appliquer théoriquement à n'importe quel type de population et notamment à une population de « bien-portants ». Dans la première version décrite dans la publication citée, les auteurs définissent ainsi onze états fonctionnels distincts ; ce nombre a été porté à trente par la suite [111].

Chaque patient peut être affecté à un de ces états, repérant ainsi son « statut » fonctionnel. Reste à établir une méthode de quantification des états fonctionnels permettant le calcul d'un indice d'état de santé globale. Elle repose sur le concept de « préférence sociale » : il s'agit d'attribuer des poids (des « utilités » dans le langage de la théorie de la décision) à chacun des états fonctionnels. Pour rendre réaliste cette opération de pondération, une matrice où les états fonctionnels sont distribués selon cinq classes d'âge et quarante-deux « complexes symptômes/problèmes » choisis pour être l'origine des troubles fonctionnels a été établie. Les différentes techniques, relevant des méthodes de la théorie de la décision, utilisées pour mesurer les « niveaux de bien-être », c'est-à-dire attribuer des utilités exprimant le niveau de santé globale correspondant à chaque catégorie, sont décrites en détail dans plusieurs publications [16, 111].

3. Les indices reposant essentiellement sur un abord perceptuel.

Nous avons regroupé dans cette catégorie des indices d'état de santé globale construits à partir des réponses des individus à des questions concernant leur état de santé. Certaines des questions peuvent concerner des phénomènes factuels se prêtant relativement peu à une interprétation « subjective » de la part du sujet ; mais il existe toujours un ensemble d'items, parfois prépondérant, lui permettant d'exprimer son « ressenti », son appréciation personnelle de son niveau de bien-être.

Les indices de ce groupe sont certainement les plus variés, l'abord perceptuel se prêtant facilement à des conceptualisations de la santé très diverses. Nous décrirons trois indices de ce type, représentant chacun une approche particulière de la mesure de la santé globale, choisis à la fois parce qu'ils résument les principales écoles de pensée et parce qu'ils ont fait chacun l'objet d'une expérimentation sur le terrain à une assez vaste échelle.

Breslow [19], dans le cadre de l'expérience du Laboratoire de population humaine (où la population de tout un comté de Californie est considérée comme

sujet d'étude), a développé un indice destiné à permettre une expression quantitative de la définition de la santé de l'O.M.S. Il repose sur un questionnaire comportant trois catégories d'items correspondant respectivement à la santé physique, la santé mentale et la santé sociale. Pour la santé physique, les questions portent sur les déficiences, les incapacités, les états chroniques, les symptômes ressentis et le « niveau d'énergie » ; cet aspect est décrit en détail dans un article de Belloc *et al.* [10]. La santé mentale fait l'objet d'un ensemble de questions permettant d'exprimer des sentiments négatifs et positifs ; Berkman [14] présente une description de l'Index de bien-être psychologique développé dans ce cadre. Enfin, l'aspect social de la santé est exploré par des items portant sur la vie professionnelle, la satisfaction conjugale, la sociabilité, l'activité au sein de la communauté.

Un indice est calculé ensuite par la méthode de l'analyse des Ridits (« Relative to an Identified Distribution ») qui permet, comme son nom l'indique, d'obtenir pour chaque sujet une valeur établie en fonction de la distribution des résultats de l'ensemble des sujets du groupe auquel il appartient. Il s'agit donc à la fois d'un indice individuel, donnant pour chaque sujet une note globale de niveau de santé, et d'un indice de population : en effet, il est possible de calculer de tels scores pour des groupes définis selon des critères divers. Ainsi, dans les publications citées, la méthode est utilisée pour établir et comparer les indices de santé globale des groupes définis par le sexe, l'âge, le revenu, l'ethnie, etc.

Le « Sickness Impact Profile » (S.I.P.) que l'on peut traduire approximativement par « profil de l'impact des troubles de santé » est un indice développé par Bergner *et al.*, [12], destiné à mesurer les dimensions perceptuelles et comportementales de l'état de santé. Le questionnaire est divisé en 14 catégories (activité sociale, locomotion, sommeil, nutrition, activité ménagère, mobilité, mouvements du corps, fonctionnement intellectuel, vie familiale, hygiène personnelle, loisirs, émotions, communications, travail) représentant au total 235 réponses possibles. Chacune d'elles est affectée d'un poids attribué par un groupe de 25 experts. La somme des valeurs correspondantes aux réponses faites par un sujet aux questions posées constitue son indice d'état de santé globale. Cet indice a été expérimenté sur une vaste échelle aux Etats-Unis et des questionnaires auto-administrés ont été mis au point et largement utilisés. Sa validité a été étudiée par Bergner *et al.* [13] et sa fiabilité par Pollard *et al.* [114]. Une version abrégée du S.I.P. réduite à 85 items a été développée à l'Université de Nottingham par Martini *et al.* [90] et expérimentée sur 14 500 patients âgés de 18 à 60 ans en Angleterre. L'objectif à atteindre, aussi bien pour l'équipe américaine que pour l'équipe britannique, est la mise au point d'un indice apte à traduire quantitativement l'effet des soins médicaux sur l'état de santé globale des patients.

L'objectif de Ware *et al.* [139] est différent ; il s'agit de développer une « mesure perceptuelle de l'état de santé », c'est-à-dire une mesure de la perception par les individus de leur niveau de santé. Ceci se traduit par l'établissement d'une « échelle de mesure de perceptions concernant l'état de santé » comprenant 6 dimensions traduites chacune par un indice spécifique : indices d'état de santé

actuel, d'état de santé antérieur, d'état de santé futur, d'anxiété à propos de l'état de santé, de susceptibilité et de résistance à la maladie, d'acceptation et de rejet de la maladie. Pour chacune des dimensions individualisées, une série de jugements est proposée aux sujets auxquels il est demandé de donner leur avis, sous forme d'une note variant de 1 à 5 selon leur degré d'accord avec chaque affirmation de la liste. Une méthode de calcul assez simple est ensuite utilisée pour totaliser les notes attribuées pour chaque dimension et les agréger en un indice unique qui exprime donc, selon les auteurs, le niveau de santé globale perçue par les patients.

b) *Modèles d'indices d'état de santé globale de population.*

Bien entendu, tous les indices individuels d'états de santé peuvent constituer l'élément de base de la construction d'indices de population représentant la mesure de l'état de santé globale de la population à laquelle appartiennent les individus. Dans certains cas, la construction d'indices de population correspondant au modèle d'indice individuel est directe, soit par simple sommation des valeurs individuelles suivie d'une division par l'effectif, soit parce que l'indice individuel est établi à partir des valeurs de la population comme pour l'indice de Breslow décrit plus haut [19]. Ailleurs, la méthode d'agrégation est plus sophistiquée afin de tenir compte de phénomènes n'apparaissant qu'au niveau de la population, ou bien de la non-homogénéité d'une importante population.

Ainsi, M. K. Chen, dans le travail déjà cité [31], propose l'index de population « -T » calculé à partir de l'index individuel « -H ». L'hétérogénéité d'une vaste population ne permettant pas une agrégation directe des indices -H établis à partir de sous-populations homogènes, l'auteur définit -T comme une moyenne pondérée des mesures effectuées dans toutes les sous-populations, les poids étant l'effectif de ces sous-populations.

Bush et ses collaborateurs [33] ont construit à partir de leur indice individuel un indice de population en procédant de la façon suivante. Dans un premier temps, les indices individuels sont agrégés de façon directe pour former l'indice W appelé « Indice de statut fonctionnel » défini ainsi :

$$W = \frac{\sum_j N_j W_j}{N} \quad \text{avec } 0 \leqslant W \leqslant 1 \text{ ; où :}$$

N est l'effectif de la population ;
N_j est l'effectif dans l'état fonctionnel j ;
W_j est l'utilité attribuée à l'état fonctionnel j ;
j est l'indice des états fonctionnels 0, 1, 2, . . ., m.

Le concept de *pronostic*, c'est-à-dire ici de probabilité de transition d'un état fonctionnel dans un autre au cours du temps, est ensuite introduit (le taux de mortalité étant considéré comme la probabilité de passer de l'état de mort venant d'un état différent). Un modèle markovien, décrit en détail, par ailleurs [22] est utilisé pour calculer « l'espérance d'état fonctionnel », Y, qui n'est autre que la

distribution de l'espérance de vie parmi les différents états fonctionnels (comme le notent les auteurs, l'indice Y est comparable par sa construction à l'Espérance de vie sans incapacité et à l'Espérance d'incapacité définie par Sullivan que nous avons cité plus haut).

Finalement l'indice Q, appelé « Espérance de vie ajustée par valeur », est obtenu à partir des espérances d'états fonctionnels (Y_j) et des utilités attribuées à ces états (W_j) :

$$Q = \sum_j W_j Y_j$$

Q est donc l'indice d'état de santé globale de la population recherché. Il a été utilisé effectivement comme critère de résultat dans deux importantes études de coût-efficacité, l'une concernant un programme de dépistage de la phényl-cétonurie [21] et l'autre un programme de dépistage de la tuberculose [23].

Chiang, dans un travail publié en 1965 [35] et repris avec quelques modifications en 1973 [36] a développé un modèle théorique d'indice de santé globale de population très comparable à celui de Busch et coll. qui s'en sont largement inspiré.

c) *Les principaux problèmes soulevés par les indices d'état de santé globale.*

Un groupe d'experts américains s'est réuni en 1972 pour faire le point sur les indices d'état de santé globale. Cette conférence s'est terminée par la rédaction de conclusions en forme de recommandations pour des recherches futures [11, p. 253-258]. Bien que six années se soient écoulées, ces conclusions semblent toujours valables ; nous allons résumer celles qui concernent les problèmes conceptuels toujours en attente d'une solution satisfaisante, en écartant les aspects purement techniques.

Bien évidemment, le premier, et sans aucun doute le plus important, des problèmes à résoudre est celui du *choix des composantes d'un indice d'état de santé globale* qui soient valides et possibles à collecter dans les conditions réalistes, notamment sur le plan économique. Nous n'insisterons pas sur ce point l'ayant déjà largement traité plus haut, sinon pour redire qu'en ce domaine la définition d'un indice universellement accepté est une utopie : il n'y a pas d'inconvénient à notre sens, à définir l'indice qu'on désire, étant entendu qu'il doit refléter une conception précise et opérationnelle de la santé dans les conditions satisfaisantes de validité et de faisabilité.

Si cette étape du choix des composantes d'un indice d'état de santé globale est réalisée, un autre problème de fond se pose aussitôt : faut-il introduire une *pondération des diverses composantes* et le cas échéant, comment ? Certains, dont Grogono [60] résume les arguments, sont opposés au principe même de la pondération. Pour d'autres, au contraire, elle est nécessaire pour valablement comparer des situations, individuelles et collectives, où les états de santé et la nature des problèmes sont différents. Il faut dans ce cas faire des choix concernant différents problèmes : quel(s) critère(s) de pondération choisir (critère économique, bien-être, par exem-

ple) ; à qui doit-on demander de déterminer les utilités (médecins, malades, usagers potentiels des services de santé, etc.) et comment ? ; peut-on valablement (et comment ?) agréger les utilités déterminées par différents individus ou groupes ; doit-on appliquer une dégressivité (« discount » des économistes) des utilités dans le temps et comment déterminer les taux à appliquer. Cet ensemble de questions pose de nombreux problèmes, tant méthodologiques que pratiques. La multiplication des expériences dans ce domaine est d'autant plus souhaitable que celui-ci concerne très directement le problème de la normativité dont nous avons vu plus haut combien il est fondamental dans un tel contexte. Nous renvoyons le lecteur à la discussion, abordée plusieurs fois dans les pages précédentes, liée à l'existence de systèmes de valeurs propres à chaque contexte socio-culturel : on peut concevoir autant de systèmes de pondération que de groupes sociaux.

Une fois choisies les composantes d'un indice, définies (ou refusées) les utilités correspondantes, il faut construire un *modèle d'indice global*. Le problème ici est d'obtenir un indice utilisable, c'est-à-dire facilement compréhensible et interprétable par un non-expert et qui inspire confiance. Il n'est pas évident, là non plus, que c'est le cas pour tous les indices que nous avons décrits.

Enfin se pose le problème des *données à collecter :* définir un indice d'état de santé globale idéal serait de peu d'utilité si les données correspondantes ne peuvent être réunies correctement, pour des raisons de faisabilité et/ou de coût.

7. Conclusions.

Ce tour d'horizon des tendances des recherches actuelles dans le domaine de la mesure de l'état de santé d'une population amène à formuler quelques remarques générales.

Tout d'abord, les limites des indices classiques ont amené, après Moriyama [95], la quasi-totalité des spécialistes des statistiques sanitaires à réclamer la mise en place des méthodes de mesure de la santé plus sensibles aux modifications des facteurs influençant celle-ci : facteurs de l'environnement économique et social surtout, et aussi action du système de soins. Il est évident que nous quittons l'époque du triomphalisme médical où la relation « plus de soins médicaux entraîne une amélioration proportionnelle du niveau de santé » était admise par tous. Nous entrons dans l'*ère critique de l'évaluation* et pour beaucoup il s'agit de relever le défi de démontrer qu'il existe effectivement une relation positive entre niveau de santé de la population et utilisation des services de santé, que ce soit au niveau global ou qu'il s'agisse de services particuliers, comme l'hôpital par exemple, et d'évaluer cette relation.

Quelles sont les voies ouvertes dans cette direction par les recherches que nous avons examinées ? Incontestablement leur apport essentiel est d'élargir le champ de la santé, d'y introduire de nouvelles dimensions et de proposer des moyens de mesurer des phénomènes qui étaient passés sous silence par les méthodes antérieures. On peut espérer ainsi pouvoir apprécier quantitativement d'éventuelles améliorations dans différents aspects de la « qualité de la vie » qu'amènerait au niveau

d'une population l'utilisation du système de soins. Toute une recherche est à mettre en place pour vérifier cette hypothèse de travail. Dans le contexte particulier de la France, il faut cependant remarquer qu'une amélioration des outils de mesure de la santé (indicateurs de « résultats ») n'est pas suffisante pour permettre des études évaluatives satisfaisantes : en effet, l'évaluation consiste à mettre en rapport des coûts et des résultats. Or, dans le domaine des coûts, nous ne sommes guère plus avancés que dans celui des résultats. Chacun connaît l'exemple du système de tarification hospitalière français qui interdit de connaître les coûts véritables des soins.

Il est intéressant de noter qu'une des retombées des travaux cités est l'abolition du monopole médical sur l'appréciation des problèmes de santé, en même temps qu'on a vu apparaître l'individu isolé ou en groupe, comme source de référence crédible de jugement. La nécessité d'obtenir son opinion personnelle sur un grand nombre de phénomènes pour lesquels il est seul capable de fournir des informations pertinentes a également eu une conséquence méthodologique et pratique très importante : c'est la place fondamentale que prennent les enquêtes par entrevue individuelle directe ou par questionnaires auto-administrés dans la panoplie des « épidémiologistes de la santé ». On assiste ainsi à l'apparition, à côté de l'approche épidémiologique classique, d'une « approche psycho-sociologique » de la santé [Pineault, 112].

Une autre constatation que ne peut manquer de faire l'observateur attentif est que nous assistons au développement simultané de la construction d'indices d'états de santé dans deux directions opposées, mais complémentaires : d'une part, des indices très synthétiques, applicables au niveau de vastes populations dont l'exemple type est l'Espérance de vie en bonne santé, et d'autre part, des « micro-indices » destinés à mesurer des phénomènes extrêmement variés dans des groupes souvent très restreints. Cette double tendance semble parfaitement justifiée quand on songe à la fois aux limites de l'utilisation d'indices synthétiques et à la multiplicité des problèmes de santé en fonction de la population-cible, qui peut se définir par des critères variés (âge, lieu géographique, profession, niveau de revenu, pathologie particulière, lieu d'observation-hôpital, par exemple, etc.) : dans ces conditions, on comprend que la diversité des phénomènes à appréhender et des instruments de mesures adaptés conduira à la mise en place à terme, de « batteries », d'indicateurs et d'indices comme le recommandent notamment Lévy *et al.* [81]. Cet état de chose repose sur la reconnaissance évidente, explicitée ou non, que toute mesure ne représente qu'un aspect partiel et biaisé du phénomène étudié, la « santé ». Celle-ci, nous pensons l'avoir montré, est un concept multiforme, variable, susceptible d'approches très différenciées. On peut dire, de ce point de vue, qu'objectivement, elle est une entité n'ayant qu'une existence virtuelle : un concept unificateur commode.

Parmi la panoplie des instruments que nous avons passés en revue, les indicateurs et indices « classiques » conservent leur place : ils sont en effet irremplaçables pour de nombreux aspects de l'identification des problèmes de santé et de la gestion du système de soins. On peut d'ailleurs affirmer avec Elinson [53] que ces

indicateurs et indices sont nettement sous-utilisés et que beaucoup peut encore être fait dans ce domaine. Mais les indicateurs et indices de mortalité et de morbidité classiques ne doivent jouer, à l'instar du diagnostic en médecine clinique, qu'un rôle pivot, correspondant à un temps opérationnel, mais inaptes à servir de critère exclusif d'évaluation, qu'il s'agisse d'estimer des services ou de mesurer des résultats. Nous espérons avoir montré qu'ils doivent absolument être complétés par un autre type de mesure, seul susceptible de permettre de peser l'importance des problèmes de santé, de les comparer entre eux pour définir des priorités, et finalement, de fournir des critères corrects pour l'évaluation des résultats des actions entreprises. Seuls les indicateurs et indices, expression quantifiée d'un niveau de santé ne reposant pas sur les catégories morbides habituelles, peuvent fournir le support d'informations nécessaires pour sortir de l'ornière actuelle de la « spécialisation médicale » de la Santé publique, où on constate que les ressources du système sont allouées non pas en fonction du poids des problèmes de santé à résoudre, mais bien plus selon le poids respectif des différentes structures médicales existantes.

Enfin, il est intéressant de noter que nous n'avons jamais rencontré un problème qui nous semble pourtant fondamental aussi bien sur le plan conceptuel et méthodologique que sur le plan pratique : *quelle est la nature de la relation entre santé des individus et santé d'une population ?* Tous les indices que nous avons examinés font l'hypothèse implicite que le niveau de santé d'une population n'est que la somme du niveau de santé des individus qui la composent. Or, des considérations théoriques (la théorie des systèmes nous a appris que le comportement de ceux-ci ne pouvait se déduire simplement du comportement de chacun de leurs éléments) et pragmatiques (l'efficacité individuelle de certaines thérapeutiques ne semble pas se vérifier à l'échelle d'une collectivité, par exemple) font douter qu'on se trouve devant un phénomène aussi simple. Il existe là de multiples voies de recherches passionnantes qui ne semblent pas, à l'heure actuelle, avoir été véritablement abordées, quoique indirectement évoquées par des auteurs tel Lerner [78] déjà cité, qui note la spécificité des problèmes de santé d'une collectivité introduite par les caractéristiques propres de la communauté.

Nous sommes encore loin de disposer des indicateurs et indices d'état de santé idéaux, qui pour le Comité d'experts des statistiques sanitaires de l'O.M.S. [104] doivent posséder les qualités suivantes : faciles à obtenir, représentatifs de la population visée, stables, universels dans la représentation des phénomènes qui affectent le niveau de santé, simples à calculer, acceptables (c'est-à-dire faciles à comprendre et non contestables dans leur principe), reproductibles, spécifiques, sensibles, fidèles.

Espérer pouvoir un jour obtenir une batterie d'indicateurs et indices d'état de santé possédant simultanément toutes ces qualités est certes une utopie. Mais chaque pas dans cette direction est un progrès nécessaire.

Le fait que la France est loin du peloton de tête des nations qui consentent un important effort de recherche dans le domaine de la mesure de la santé d'une population, constitue une raison supplémentaire d'inquiétude pour l'avenir de la

Santé publique dans notre pays. Force est de reconnaître avec Desrosières [44] que les lacunes de l'observation statistique révèlent « en creux » les sujets que notre société a préféré ne pas approfondir, et que, comme l'indique Letourmy [80], on ne s'est pas donné les moyens de juger efficacement la cohérence d'une politique de santé.

BIBLIOGRAPHIE

[1] Afifi A. A., Bigelow J. H., Glass N.: Some terminology revisited. *In : « Systems aspects of Health Planning »*, N.T.J. Bailey and M. Thompson, eds. North Holland Publ., 1974, 342.

[2] D'agostino R. B.: Social indicators: a statistician view. *Social Indicators Research*. 1975, *1*, 459-484.

[3] Aillaud Y.: *Compte rendu d'activité pour l'année 1975*. Centre d'Hygiène Appliquée « Doria ». Tome 3 : Rapport technique. Fédération des Organismes de Sécurité sociale de la Région du Sud-Est, 108-125.

[4] Andréani E.: Des instruments pour la politique de la santé, les indicateurs sociaux. *Economie et statistique*, *71* et *72*. 1975.

[5] Antonovsky A.: Social class, life expectancy and overall mortality. *Milbank Memorial Fund Quarterly*, 1967, *45*, 31-73.

[6] Arkhipoff O.: *Peut-on mesurer le bien-être national ?* 190, Série C, 41, I.N.S.E.E., Paris, 1976.

[7] Armitage P.: National health survey systems in the European Community. *Int. J. Epidemiol.*, 1976, *5*, 4, 321-326.

[8] Balinski W., Berger R.: Review of the research on general health status indexes. *Med. Care*, 1975, *13*, 4, 282-293.

[9] Baylet R., Ben Youssef A.: Santé et développement : priorités, planification et indicateurs sanitaires. *Soc. Sci. Méd.*, 1975, *9*, 69-73.

[10] Belloc N. B., Breslow L., Hochstim J. R.: The measurement of physical health in a general population survey. *Am. J. Epidemiol.*, 1971, *93*, 328-336.

[11] Berg R. L.: *Health status indexes*. Hospital Research and Educational Trust, Chicago, 1973.

[12] Bergner M., Bobbitt R. A., Kressel S., Pollard W. E., Gilson B. S., Morris J. R.: The Sickness Impact Profile : conceptual formulation and methodology for the development of a health status measure. *Int. J. Health Serv.*, 1976, *6*, 417.

[13] Bergner M., Bobbitt R. A., Pollard W. E., Martin D. P., Gilson B. S.: The sickness impact profile : validation of a health status measure. *Med. Care*, 1976, *14*, 1, 56-57.

[14] Berkman P. B.: Measurement of mental health in a general population survey. *Am. J. Epidemiol.*, 1971, *94*, 105-111.

[15] Bickner R. E.: Measurements and indices of health. *In : Methodology of identifying, measuring and evaluating outcomes of health services programs, systems and sub-systems*. C.E. Hopkins Publ. DHEW, PHS, Rockville (Maryland), 1969.

[16] Blischke W. R., Bush J. W., Kaplan R. M.: Successive intervals analysis for preference measures in a health status index. *Health Services Research*, 1975, *10*, 2, 181-198.

[17] Blum H. L.: A working definition of health for planners : merging concepts. 22-23. University of California School of Public Health, Berkeley, 1971.

[18] Bonnevie P.: The concept of health. A sociomedical approach. *Scand. J. Soc. Med.*, 1973, *1*, 2, 41-43.

[19] Breslow L.: A quantitative approach to the World Health Organisation definition of health : physical, mental and social well-being. *Int. J. Epidemiol.*, 1972, *1*, 4, 347-355.

[20] Bureau of Health planning and resources development. *Selected Bibliographic references on methodologies for community health status assessment*. Health Planning bibliography Series, US DHEW-PHS-HRA, Rockville (Maryland), 1977.

[21] Bush J. W., Chen M. M., Patrick D. L. : Cost effectiveness using a health status index. Analysis of the New York State PKU screening program. *In : Health Status Indexes*. R. L. Berg. éd., 172-208. Hospital Research and Educational Trust, Chicago, 1973.

[22] Bush J. W., Chen M. M., Zaremba J. : Estimating health program outcomes using a markow équilibrium analysis of disease development. *Am. J. Publ. Health*, 1971, *61*, 12, 2362-2375.

[23] Bush J. W., Fanshel S., Chen M. M. : Analysis of a tuberculin testing program using a health status index. *Socio-Econ. Plan. Sci.*, 1972, *6*, 49-68.

[24] Bynder H., Kong Ming New P. : Time for a change : from micro-to macro-sociological concepts in disability research. *J. Health and Social Behavior*, 1976, *17*, 45-52.

[25] Campbell A. : Subjective measures of well being. *Am. Psychologist, 1976*, 117-124.

[26] Canguilhem G. : *Le normal et le pathologique*, 150-132, P.U.F., Paris, 1966.

[27] Cerebe : *Pratiques médicales et système hospitalier*, Cerebe, Paris, 1974.

[28] Chadwick E. : *Report on the sanitary condition of the labouring population of Great Britain*, 280, Edinburgh University Press, Edinburgh, 1842, Réimpression 1965.

[29] Chapalain M. T. : La construction d'indicateurs nouveaux de santé. *Cahiers de Sociologie et de Démographie Médicales*. XV[e] année, 1975, *2*, 69-73.

[30] Chen M. K. : The G index for program priority. *In : Health Status Indexes*, R. L. Berg éd., 28-34. Hospital research and Educational Trust, Chicago, 1973.

[31] Chen M. K. : Aggregated physiological measures of individual and group health status. *Int. J. Epidemiol.*, 1975, *4*, 2, 87-92.

[32] Chen M. K., Bryant B. E. : The measurement of health. A critical and selective overview. *Int. J. Epidemiol.*, 1975, *4*, 4. 257-264.

[33] Chen M. M., Bush J. W., Patrick D. L. : Social indicators for health planning and policy analysis. *Policy Sciences*, 1975, *6*, 1, 71-89.

[34] Chen M. M., Bush J. W., Zaremba J. : Effectiveness measures. *In : Operations Research in Health Care*. A critical analysis, L. S. Shuman, R. Speas, J. Young eds., 276-301, John Hopkins University Press, Baltimore, 1975.

[35] Chiang C. L. : *An index of health : mathematical models*. Vital and Health statistics, Series 2, 5, Washington National Center for Health Statistics, 1965, 1-19.

[36] Chiang C. L., Cohen R. D. : How to measure health : a stochastic model for an index of health. *Int. J. Epidemiol.*, 1973, *2*, 1, 7-13.

[37] Clearing house on health indexes. *Division of Analysis*. National Center for Health Statistics. HRA, Dhew, Rockville (Maryland).

[38] Collishaw N. : Canada Health Survey : health problems. A content proposal. Ministère de la Santé Nationale et du Bien-être Social, Ottawa, 1976.

[39] Cochrane A. L. : The history of the measurement of ill-health. *Int. J. Epidemiol.*, 1972, *1*, 89-92.

[40] Cote R. : The development of a Systematized Nomenclature of Medicine (SNOMED). *In : Journées d'Informatique Médicale de Toulouse*, 1973, 3, 151-156, IRIA, Paris, 1973.

[41] Culyer A. J., Lavers R. S., Williams A. : Social indicators : health. *Social Trends*, 1971, 2.

[42] Damiani P. : Contribution à la mesure du niveau de santé : définition d'un indicateur lié aux facteurs sanitaires. *Int. Stat. Rev.*, 1974, *42*, 1, 29-37.

[43] Damiani P. : Etude de la liaison entre la morbidité et la mortalité et définition d'un indice de morbidité. *Bull. Trim. de l'Institut des Actuaires Français*, 1975, 292, 191-205.

[44] Desrosières A. : Eléments pour l'histoire des nomenclatures socio-professionnelles. *In : Pour une histoire de la statistique*, 1, INSEE, Paris, 1977.

[45] Devers L., Silberstein P. : Eléments d'intervention pour une critique pratique des nuisances du travail. *Fracture*, 1977, *1*, 37-47.

[46] Diderot D. : *Correspondance*. Tome XIV, dernière lettre à Sophie Volland. Ed. de Minuit, Paris, 1968.

[47] Dingle J. H. : The illnesses of man. *In : Life and Death and Medicine*. 49-56. A Scientific American Book, W. H. Freeman and C[o], San Francisco, 1973.

[48] Dreitzel H. P. : *In : The Social Organization of Health*, H. F. Dreitzel ed., 5-17, Mac Millan, New York, 1971.

[49] Dumont J., Latouche J. : *L'hospitalisation malade du profit*. 199, Editions Sociales, Paris, 1977.

[50] Dupuy J.-P. : *Relations entre dépenses de santé, mortalité et morbidité*. Cerebe, Paris, 1973.

[51] Electricité de France. Service Prévention et Sécurité. *Statistiques accidents agents E.D.F.*, E.D.F., 30, Paris, 1975.

[52] Elinson J. : Socio-medical health indicators. *Int. J. Health Serv.*, 1976, *6*, 3.

[53] Elinson J. : Insensitive health statistics and the dilemna of HSAs. *Am. J. Publ. Health*, *67*, 5, 417-418.

[54] Fanshel S., Bush J. W. : A health status index and its application to health services outcome. *Operations Research*, 1970, *18*, 6, 1021-1066.

[55] Fuchs V. R. : *Who shall live ? Health, Economics and Social choice*, 31, New York, Basic Books Inc., 1974.

[56] Gervois M. et S. : *La santé sociale*. Son expression quantitative. Lille, Université du Droit et de la Santé, 1976.

[57] Goldsmith S. B. : The status of health status indicators. *Health Services Reports*, 1972, *87*, 3, 212-220.

[58] Goldsmith S. B. : A reevaluation of health status indicators. *Health Services Reports*, 1973, *77*, 10, 937-941.

[59] Gordon J. E. : The world, the flesh, and the devil as environment, host and agent of disease. *In : The epidemiology of health*, Gladstone, ed., N. Y. Health Education Council, 1953.

[60] Grogono A. W. : Measurement of ill health : a comment. *Int. J. Epidemiol.* 1973, *2*, 1, 5-6.

[61] Grogono A. W. : Woodgate D. J. : Index for measuring health. *Lancet II*, 1971, 1024-1026.

[62] Haro A. S. : Planning an information system for health services. *Meth. Inform. Med.*, 1972, *11*, 1, 1-8.

[63] Haro A. S. : Measurement of need. *In : Proceedings of the Sixth International Scientific Meeting of the Epidemiological Association*, 49-59, Primosten (Yougoslavie), 1972.

[64] Health services research group : Development of the Index of Medical Underservice. *Health Services Research*, 1975, 168-180.

[65] Himatsingani C. : Approaches to health and personnal social services planning in the National Health Service and the place of health indices. *Int. J. Epidemiol.*, 1973, *2*, 1, 15-21.

[66] Hollingworth G. W. : A demographic study of the british Ducal families. *In : Population history London*, D.V. Glass and D. E. C. Eversley ed., 358, Edward Arnold, 1965.

[67] Hoyman H. S. : *Our modern concept of health*. Communication à l'American Public Health Association, 16, 1, Detroit (Mich.), 1961.

[68] Illich I. : Nemesis médicale. L'expropriation de la santé. Ed. du Seuil, Paris, 1975.

[69] INED : Sixième rapport sur la situation démographique de la France. *Population*, 1977, *2*, 253-338.

[70] INSERM : *Santé Publique*. Problèmes de santé dans une agglomération urbaine en mutation : Boulogne-Billancourt. Paris, INSERM, 1975.

[71] Jazairi N. T. : Différentes approches pour l'élaboration d'indicateurs de santé. *Programme d'Elaboration des Indicateurs Sociaux de l'OCDE. Etudes Spéciales*, 2, OCDE, Paris, 1976.

[72] Jenicek M. : *Introduction à l'épidémiologie*. Sainte Hyacinthe, Edisem, Québec, Maloine, Paris, 1976.

[73] Katz S., Ford A. B., Moskowitz R. W., Jackson B. A., Jaffe M. W. : The index of ADL : a standardized measure of biological and psychological function. *J. Am. med. Assoc.*, 1963, *184*, 914, 21.

[74] Katz S., Downs T. D., Cash H., Grotz R. C. : Progress in development of the index of ADL. *Gerontologist*, 1970, *10*, 20.

[75] Kelman S. : The social nature of the definition problem in health. *Int. J. Health Services*, 1975, *5*, 4, 625-642.

[76] Kneppreth N. P., Gustafson D. H., Rose J. H., Leifer R. P. : Techniques for the assessment of worth. *In : Health Status Indexes*, R. L. Berg ed., 228-240, Hospital Research and Educational Trust, Chicago, 1973.

[77] Lebart L. : *Recherche sur le coût de la protection de la vie humaine dans le domaine médical*. CREDOC, Paris, 1970.

[78] Lerner M. : Conceptualization of health and social well being. *In : Health Status Indexes*, R. L. Berg ed., 1-6, Hospital Research and Educational Trust, Chicago, 1973.

[79] Letourmy A. : Some aspects of the relationships between mortality, environmental conditions and medical care. *In : Systems Aspects of Health Planning*, N. T. J. Bailey and M. Thompson eds, 267-278, North Holland Pub., 1975.

[80] Letourmy A. : Intérêt et limites des indicateurs de santé en matière de planification. *Cahiers de Sociologie et de Démographie Médicales*, 1977, XVII[e] année, 2, 66-72.

[81] Lévy E., Bungener M., Duménil G., Fagnani F. : *Indicateurs de santé et analyse du système français*. Commissariat Général du Plan, Paris, 1973.

[82] Lévy E., Bungener M., Duménil G., Fagnani F. : *Economie du système de santé*. Dunod, Paris, 1975.

[83] Lifson M. W. : *Definitions and terminology*. California Center for Health Services Research, Los Angeles, 1969.

[84] Linder F. E. : Les enquêtes nationales sur la santé par interrogatoire direct. *In :* Etude de la morbidité et de la mortalité : tendances actuelles. *Cahiers de Santé Publique*, 1967, 97, 85-120, OMS, Genève.

[85] Longone P. : Mortalité et morbidité. *Population et Sociétés*, 1972, 43.

[86] Mc Dowell I., Martini C. J. M. : Problems and new directions in the evaluation of primary care. *Int. J. Epidemiol.*, 1976, 5, 3, 247-250.

[87] Magdeleine M., Mizrahi A. A., Rösch G. : Un indicateur de morbidité appliqué aux données d'une enquête sur la consommation médicale. *Consommation*, 1967, 2.

[88] Majone G. : On the logic of standards setting in health and related fields. *In : Systems Aspects of Health Planning*, N. T. J. Bailey and M. Thompson eds, 279-291, North Holland Pub., 1975.

[89] Mayet, Rey, Mathieu : *Barème accidents du travail et maladies professionnelles*. Lamarre-Poinat, Paris, 1975.

[90] Martini C. J. M., Allan G. J. B., Davison J., Backett E. M. : Health Indexes Sensitive to Medical Care Variation. *Int. J. Health Services*, 1977, 7, 2, 293-309.

[91] Miller J. E. : An indicator to aid management in assigning program priorities. *Health Serv. Rep.*, 1970, 85, 725.

[92] Miller J. E. : Guidelines for selecting a health status index : suggested criteria. *In : Health Status Indexes*, R. L. Berg ed., 243-247, Hospital Research and Educational Trust, Chicago, 1973.

[93] Ministère de la Santé nationale et du bien-être social. *Indicateurs du domaine de la santé*, 21-24, Ottawa, 1976.

[94] Mizrahi A. et A. : *L'enquête sur la santé des Canadiens*. Rapport de mission, CREDOC, Paris. Ministère de la Santé Nationale et du Bien-être Social, Ottawa, 1975.

[95] Moriyama I. M. : Problems in the measurement of health status. *In : Indicators of social Change*, E. B. Sheldon and W. E. Moore eds, Russel Sage Foundation, New York, 1968.

[96] Morrow R. H., Smith P. G., Nimo K. P. : Health care priorities for less developed countries : results of a new analytic approach in Ghana. *Int. J. Epidemiol.* (A paraître).

[97] National center for health statistics : *Disability days*. United States, July 1965, June 1966, Vital and Health Statistics PHS Publication, 1000, Series 10, 47, 1968.

[98] National center for health statistics : *The national ambulatory medical care survey : symptom classification*. Series 2, 63, DHEW, Publication (HPA) 74-1337, Washington D.C., Government Printing Office, 1974.

[99] Oncan : *Classification internationale des problèmes de santé dans les soins primaires*. Version française : Statistique-Canada, Ottawa, 1974.

[100] Organisation mondiale de la Santé : *Rapport de statistiques sanitaires mondiales*, 29, 1, 37-39.

[101] Organisation mondiale de la Santé : *Groupe d'étude sur la mesure des niveaux de santé*, Document OMS, PHA/LEV, HLTH/1-17, OMS, Genève, 1955.

[102] Organisation mondiale de la Santé : *Constitution de l'OMS*. Annexe 1. OMS, Genève, 1958.

[103] Organisation mondiale de la Santé : *Classification internationale des maladies (adaptée)*, 8e révision, OMS, Genève, 1968.

[104] Organisation mondiale de la Santé : *Comité d'experts sur les statistiques sanitaires*. Indicateurs statistiques pour la planification et l'évaluation des programmes de santé publique. 14e rapport. Série de Rapports Techniques, 472, OMS, Genève, 1971.

[105] Organisation mondiale de la Santé : *Nouvelles conceptions en matière de statistiques sanitaires*. Série de Rapports Techniques, 559, OMS, Genève, 1974.

[106] Organisation mondiale de la Santé : Méthodes modernes de gestion et d'organisation des services de santé. *Cahiers de Santé Publique*, 55, OMS, Genève, 1974.

[107] Organisation mondiale de la Santé : *Indices statistiques de la santé de la famille*. Série de Rapports Techniques, 587, OMS, Genève, 1976.

[108] Organisation de coopération et de développement économiques : Mesure du bien-être social. Progrès accomplis dans l'élaboration des indicateurs sociaux. OCDE, Paris, 1976.

[109] Parsons T. : *In : Patients, Physicians and Illness*. E. Garthy Jaco, ed., 107-127, Free Presse, New York, 1972.

[110] Patrick D. L., Bush J. W., Chen M. M. : Toward an operational definition of health. *J. of Health and Social Behaviour*, 1973, *14*, 1, 6-23.

[111] Patrick D. L., Bush J. W., Chen M. M. : Methods for measuring levels of well being for a health status index. *Health Services Research*, 1973, *8*, 3, 228-245.

[112] Pineault R. : Eléments et étapes d'un programme de santé communautaire. Union Médicale du Canada, 1976, 105, 1208-1224.

[113] Pole J. D. : The use of outcomes measures in health service planning. *Int. J. Epidemiol.*, 1973, 2, 1, 23-30.

[114] Pollard W. E., Bobbitt R. A., Bergner M., Martin D. P., Gilson B. S. : The Sickness Impact Profile : reliability of a health status measure. *Med. Care*, 1976, *14*, 2, 146-155.

[115] Purola T. : A systems approach to health and health policy. *Med. Care*, 1972, *10*, 373-379.

[116] Reed L. J. : Principles applying to the collection of information on health as related to socio-environmental factors. *In : Backgrounds of social medicine*, 24-32, Milbank Memorial Fund, New York, 1949.

[117] Reeder L. G., Ramacher L., Gorelnik S. : *Handbook of scales and indices of health behaviour*. Pacific Palisades : Goodyear Publishing, 1976.

[118] Reinke W. A., Taylor C. E., Parker R. L. : Functional analysis of health needs and services. *In : Proceedings of the Sixth International Scientific Meeting of Epidemiological Association*, 60-85, Primosten (Yougoslavie), 1972.

[119] Romeder J. M., Mc Whinnie J. R. : Potential years of life lost between ages 1 and 70 : an indicator of premature mortality for health planning. *Int. J. Epidemiol.*, 1977, *6*, 2, 143-151.

[120] Rossdale M. : Health in a sick society. *New Left Review*, 1965, 34, 82-90.

[121] Rosser R. M. : Recent studies using a global approach to measuring illness. *Med. Care*, 1976, *14*, 5, Supplément, 138-147.

[122] Rouquette C., Schwartz D. : *Méthodes en Epidémiologie*, 98, Flammarion, Paris, 1970.

[123] Sackett D. L., Chambers L. W., Mc Pherson A. S., Psych D., Goldsmith C. H., Mc Cauley R. E. : The development and application of indices of health : general methods and a summary of results. *Am. J. Publ. Health*, 1977, *67*, 5, 423-428.

[124] Sanders B. S. : Measuring community health levels. *Am. J. Publ. Health*, 1964, *54*, 7, 1063-1070.

[125] Siegman A. E., Elinson J. : Newer socio-medical health indicators : implications for evaluation of health services. *Med. Care*, 1977, *15*, 5, supplément, 84-92.

[126] Sigerist H. E. : *Medicine and human welfare*, 100; Yale University Press, New Haven, 1941.

[127] Skinner D. E., Yett D. E. : Debility index for longterm care patients. *In : Health status indexes*, R. L. Berg ed., 69-82, Hospital Research and Educational Trust, Chicago, 1973.

[128] Sournia J.-C. : Ces malades qu'on fabrique. La médecine gaspillée, 11-12 et 193-194, Ed. du Seuil, Paris, 1977.

[129] Sullivan D. F.: Conceptual problems in developing an index of health. PHS Pub. 1000, Vital and Health Statistics, Series 2, 17. National Center for Health Statistics, 1966, 1-18.

[130] Sullivan D. F.: A single index of mortality and morbidity. HSMHA, *Health reports*, 1971, *86*, 4, 348-354.

[131] Sullivan D. F.: *Disability components for an index of health*. DHEW, PHS, HRA, National Center for Health Statistics, Vital and Health Statistics, Series 1971, 2, 42.

[132] Susser M.: *Causal thinking in the Health Sciences* (Citation des travaux de Halley, 17). Oxford University Press, New York, London, Toronto, 1973.

[133] Swaroop S., Uemura K.: Proportional mortality of 50 years and above. *Bull. Wld. Hlth. Org.*, *17*, 459-481, 1957.

[134] Szabason F.: L'évaluation de la santé dans un groupe humain par « l'indice biologique de santé » du Docteur Suchet. *Thèse Médecine*, Paris XI, 1975.

[135] Terris M.: Approaches to an epidemiology of health. *Am. J. Publ. Health*, *65*, 10, 1037-1045, 1975.

[136] Theureau J., Tort B.: *Contribution des recherches ergonomiques à la constitution d'une méthodologie d'enquête statistique sur les conditions de travail*. Ministère du Travail, Paris, 1977.

[137] Tornstam L.: Health and selfperception. *Gerontologist*, 1975, *15*, 3, 264-270.

[138] Torrance G. W.: Health status index models: a unifed mathematical view. *Management Science*, 1976, *22*, 9, 990-1001.

[139] Ware J. E., Wright W. P., Snyder M. K.: *Measures of perceptions regarding health status: preliminary findings as to scale reliability, validity and administration procedures*. Technical report MHC, 74-13, Carbondale, Ill., Southern Illinois University school of medicine, 1-86.

[140] White K. L.: Contemporary Epidemiology. *Int. J. Epidemiol.*, 1974, *3*, 4, 295-303.

[141] Williams R. G. A., Johnston M., Willis L. A., Bennet A. E.: Disability: a model and measurement techniques. *Br. J. Prev. Soc. Med.*, 1976, *30*, 71-78.

[142] Winkelstein Jr. W.: Epidemiological considerations underlying the allocation of health and disease care resources. *Int. J. Epidemiol.*, 1972, *1*, 1, 69-74.

[143] Wood P. H. N.: *Edition française du projet de Classification des déficiences, incapacités et Handicaps*. WHO/ICD. 9/Rev. Con 7/75.15. Centre OMS pour la classification des maladies. INSERM, Paris, 1975.

[144] Wylie C. M.: The definition and measurement of health and disease. *Public Health Rep.*, *85*, 101-104, 1970.

[145] Wyler A. R., Masuda M., Holmes T. H.: Seriousness of illness rating scale. *J. Psychosomatic Research*, 1968, *11*, 363-374.

[146] Wyler A. R., Masuda M., Holmes T. H.: The seriousness of illness rating sascale: reproductibility. *J. Psychosomatic research*, 1970, *14*, 59-64.

[147] Zola I.: Culture and symptoms: an analysis of patients presenting complaints. *Am. Sociol. Rev.*, 1966, *31*, 615-630.

© Masson, Paris, 1979

152

Program planning in a small community health care setting

Carol Clemenhagen
and
Francois Champagne

By imaginative and rational use of a combination of techniques available within the needs assessment framework, it is possible for community health agencies to document needs without placing inordinate pressure on resources.

0361-6274/82/0071-0047$2.00
© 1982 Aspen Systems Corporation

The systematic assessment of health care needs and the development and implementation of health care programs in small community settings have inherent problems—notably, resource constraints in terms of personnel, finances and planning time. However, the needs assessment and planning process can be a focus of creativity. By imaginative and rational use of a combination of techniques available within the needs assessment framework, it is possible for community health care agencies to document needs without placing inordinate pressure on their current resources. A case study of a small community health care clinic's effort to plan a program for the elderly illustrates this.

A CASE STUDY

The first step in considering how to plan a health care program for the elderly, to be set up and delivered in a small community health care setting, was to establish the framework in which planning was to be carried out. As Hilleboe has pointed out, "Much confusion about planning can be eliminated by labeling every health planning effort appropriately as to scope, geographic area and focus."[1]

The clinic had several program goals that set the focus for the planning process. These goals were as follows:

1. to determine the health and social needs of the elderly (aged 65 and older) population in the clinic's catchment area;
2. to establish a program to respond to assessed needs;
3. to increase the number of catchment area elderly residents using the clinic's health services; and
4. to ensure community input and participation in a broad community planning process.

The program was to be evaluated after an unspecified period of operation.

At the outset of the planning process, the community health care clinic concerned was serving ap-

Carol Clemenhagen, *M.H.A., is Management Consultant at Ottawa General Hospital, Ottawa Health Sciences Centre, in Ottawa.*

Francois Champagne *is a Ph.D. candidate in the Department of Health Administration at the University of Montreal.*

Work for this article was carried out in the Department of Epidemiology & Community Medicine, University of Ottawa, including the assistance of the department's R.A. Spasoff and J. M. Last. The opinions expressed herein are entirely those of the authors.

proximately 12 percent of the total senior citizen population in the catchment area. The catchment area, located in the heart of a city, had a population of around 15,000 persons, approximately 2,500 of whom were over 65 years of age. The clinic was the only one of its type operating in the catchment area. Senior citizens represented approximately 10 percent of the regular users of its health services. Clinic personnel involved in the planning process included a nurse-coordinator, physician, community development worker and board member. Acting as consultants to this group were another physician (specialist in community medicine) and a research officer with training in health administration.

Three constraints had to be considered before any planning strategy could be suggested. First, the financial resources that could be dedicated to the planning process were limited. Second, it was hoped that the program would be developed and initiated within a short period of time. Last, clinic personnel could not devote much time to the planning process.

PROFESSIONAL AND CONSUMER DEFINITIONS OF NEED

To understand the concept of health care need, it is necessary to take account of its various dimensions. The unitary explanation of needs as "some disturbance in health and well-being" is insufficient.[2] Consideration should also be given to the different perspectives from which need is defined. Definitions of need are conditioned by value judgments.[3] Differing value judgments affecting perception of need can lead to conflict between consumer and professional definitions.

Practitioners involved in the direct provision of health care services know that programs based on the identification of needs that do not correspond to the community's perceptions are generally doomed to failure.

> Health professionals primarily consider health services in technical terms. Consumers on the other hand, often use very different criteria when considering health services, placing greater emphasis on the non-technical components of service delivery from which they expect to derive values or benefits.[4]

In a paper presented at the 1979 Canadian Public Health Association Conference, Lynch and Clemenhagen illustrated the frequent gap between professional and consumer defined health care needs.[5] On the basis of data from community-requested health surveys in two different settings, existing facilities and services appeared to be adequate for the level of need detected by the survey instruments. Nevertheless, the community interest groups that had requested the surveys definitely perceived a need for additional or alternative services. Lynch and Clemenhagen suggested that community-based surveys such as the *World Health Organization International Collaborative Study of Medical Care Utilization* or the *Canadian Health Survey* may not be adequate for determining community health care needs when it is necessary to incorporate consumer as well as professional definitions. These types of surveys define needs within a professional framework and are not generally designed to measure consumer values, preferences and motivation to attain alternative modes of health care.

The differing values and perspectives of health care professionals and consumers are likely to focus their needs perception at different levels of the hierarchy of community health care needs.

Maslow defines a need as being a lack of something, a deficit state (while a motive is a conscious desire, a felt impulse).[6] Just as Maslow described a hierarchy of individual needs, it may be relevant to acknowledge a hierarchy of community needs. The differing values and perspectives of professionals and consumers are likely to focus their needs perception at different levels of this hierarchy. Professionals are probably more concerned about and centered around first level needs (medical care requirements), while community pressure groups would tend to center around higher order needs such as satisfaction with the manner in which care is given or preference for alternative care settings.

HEALTH PROGRAM PLANNING

The identification of health care needs takes place within a planning process.[7] Health program planning has been operationally defined as the orderly process of defining health problems and unmet needs, of assessing available and potential resources, of establishing priority goals, of formulating necessary administrative action to achieve program goals; and of evaluating the program.[8] Thus a program, which is a prearranged set of activities specifying the means to achieve a defined human service goal,[9] is based on three fundamental elements:

1. the desired results or objectives;
2. the activities and resources necessary to accomplish these objectives; and
3. the definition of the population for which the program is provided.[10]

Figure 1 outlines the steps involved in program planning. Needs identification and the assessment of priorities among needs are the necessary preliminary steps. It is within these two steps that the basis for program responsiveness to both consumer and professional definitions of need is established. The field of interaction between client and professional perspectives is where varying viewpoints coincide, conflict or compromise. Community needs assessment methods must make creative use of this field of interaction if they are to integrate consumer and professional definitions of need. Rossi points out that:

> The outcome of such surveys and varying perspectives on needs may demonstrate the need for reconceptualizing the problem or the prospective intervention or perhaps may indicate the advisability of abandoning the program especially if the varying perspectives turn out to be intensely held and highly contradictory.[11]

In any case, needs assessment should attempt to blend consumer and professional participation in the planning process and should be undertaken at the community or regional level. Naturally, it should be a "rationalizing force in the health planning process."[12]

NEEDS ASSESSMENT APPROACHES

Needs assessment concerns itself with setting priorities among identified competing needs, taking into account the resources and service technology available.[13] The steps in planning a needs assessment study have been documented as follows:

1. securing a commitment to use assessment data for planning purposes;
2. establishing a steering committee;
3. locating information sources and fiscal resources; and
4. selecting needs assessment approaches that are most appropriate for the agency or community conducting the assessment project.[14]

As can be expected, approaches vary in complexity, cost, time necessary for completion and relative effectiveness.[15] Three functions are used to characterize needs assessment approaches. These are compilation (collecting data from already existing sources), development (producing new information) and integration (synthesizing information originating from inside and outside the system boundaries).[16]

Table 1 provides a summary of the needs assessment methods documented in the literature. Three approaches are defined: the indicator approach, the survey approach and the consensus-reaching approach.

Indicator approaches

Health and social indicators

Since health care needs are based on the health status of the population, indicators used to measure health status can also be used to measure needs. Such indicators are developed through analysis of morbidity, mortality and, more recently, disability data. Sources of this information include hospital admission and release reports; notifiable disease statistics; maternal, neonatal and infant mortality statistics; life expectancy tables and disability indexes for specific populations.[17]

Social indicators are relevant to health care needs because they correlate with health care utilization.[18] As with health indicators, need is inferred from measurements of the social condition of the population. Social indicators are generally used only as rough pointers since the relationship between social factors and health may be tenuous.[19] Sources of information for construction of social indicators include statistics on age, sex, education, ethnic background, housing, employment and food consumption.

Extrapolation/assumption

The extrapolation/assumption method applies epidemiologic data on the prevalence and incidence of certain health conditions in a reference population to the demographic profile of a larger study population; the purpose is to estimate expected health needs associated with these same conditions in the larger population. The quality of the estimates depends on the validity of the rates calculated for the reference population, and on their applicability to the study population.

Survey approaches

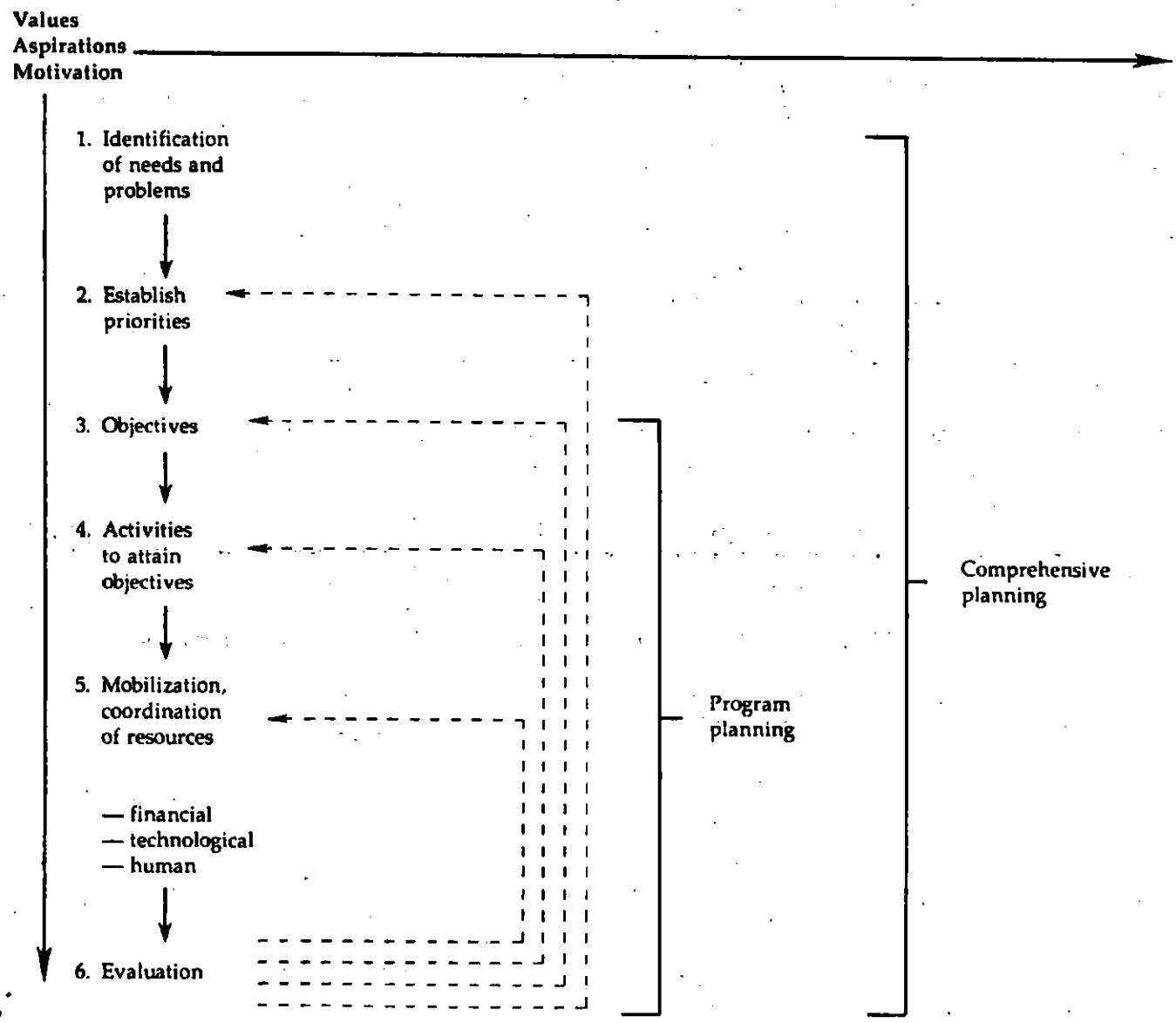
Analysis of utilization

The analysis-of-utilization method of needs assessment examines need in terms of the "demand" for services. Demand is measured by the types and amounts

155

FIGURE 1

STEPS IN HEALTH PROGRAM PLANNING



Sources: Pineault, R. "Elements et etapes d'elaboration d'un programme de sante communautaire." *Union Medicale du Canada* 105:8 (1976) p. 1208-1214.
Blum, H.L. *Planning for Health* (New York: Human Sciences Press 1974).

156

TABLE 1

NEEDS ASSESSMENT APPROACHES AND METHODS

Approach	Method	Information processing function	Measurement expertise needed	Time and resources needed
Indicator approach				
Health indicators	analysis of statistics on life expectancy, morbidity, mortality and disability	compilation	moderate to high	moderate
Social indicators	analysis of social statistics related to health and health care utilization	compilation	moderate to high	moderate
Extrapolation/ assumption	need extrapolations based on epidemiological data from reference population	compilation & integration	moderate	minimal
Survey approach	analysis of service utilization or rates under treatment	compilation	moderate	moderate
	survey sample of labor and service facilities	compilation & development	moderate	moderate
	survey sample of general population	development	high	extensive
	survey sample of service or provider population	development	high	moderate
Consensus-reaching approach	community forum	integration	low	moderate
	nominal group	development	moderate	minimal
	key informants	development	moderate	minimal
	Delphi technique	development & integration	moderate	moderate
	community impressions	development, compilation & integration	moderate	minimal

Source: Siegel, L.M., Attkisson, C.C. and Carson, L.G. "Need Identification and Program Planning in the Community Context." In Attkisson C.C. et al., eds. *Evaluation of Human Service Programs* (New York: Academic Press, 1978).
Chambers, I.W., Woodward, C.A. and Dok, C. *Guide to Health Needs Assessment: A Critique of Available Sources of Health Care Information* (Ottawa: Canadian Public Health Association, 1980).
MacStravic, R.E. *Determining Health Needs* (Ann Arbor: Health Administration Press 1978).

of services actually utilized. In its narrowest interpretations, this method assumes that no need exists that does not result in utilization of services and that services utilized totally meet the need expressed.[20]

More broadly, analysis of utilization compares utilization among income, ethnic or other groups. One reason for this comparison is to gain insight into possible barriers to service which may affect particular segments of the population.

Rates under treatment

The rates-under-treatment method looks specifically at instances of service utilization. For example, a survey of service encounters, e.g., clinic visits, during a specified time period may be undertaken. Data such as client characteristics, services received, health status, transportation problems and waiting times may be collected in specially designed encounter forms or

abstracted from agency records. Information is obtained through the organization providing services.[21] Needs assessment is based on the characteristics documented in the encounter survey. This measure of need is biased toward heavy service users.

Labor and service facilities

The labor and service facilities method of needs assessment is based on the assumption that those individuals receiving care in fact need it. The number of service providers and facilities in the setting under study is compared with the volume of known service utilization. The extent to which providers and facilities cannot cope with existing utilization represents the degree of need.[22] Need is extrapolated from the *supply* of health manpower and service facilities.

Sample surveys

Sample surveys of the general population assess needs by collecting data on health problems, disability and needs perception directly from respondents, often in their own homes. An alternative approach consists of interviewing service users at the point of service utilization. This approach collects information from a group that has had at least some contact with health services. It differs from the rates-under-treatment method in that the service agency is not directly involved. Like the rates-under-treatment method, this method misses nonusers. A survey of persons involved in the provision of services would collect data on their perceptions of client needs. These opinions would of course reflect the professional perspective.[23]

Consensus-reaching approaches

Community forum

The consensus-reaching approach focuses on means by which lay and professional views of health care needs can be assessed in participative group discussions. The community forum is an open meeting to which *all* community members are invited to present views on the needs of their community. This method is justified as a supplement to more thorough methods and is used to verify findings and build "supportive consensus."[24]

Nominal group

The nominal group process consists of a very structured, multiphased meeting of individuals who are closely associated with the problem area being assessed. For example, a small target group (7 to 10 individuals) of health care consumers, administrators and organization staff may meet to define the nature of health care needs in their community.

A preset, orderly procedure is followed so that ideas are defined individually and independently at first, then later enumerated and clarified in the group setting through a round-robin process. The group then rates these ideas by secret ballot. The result is a rank ordering of needs defined by the group.[25]

It appears that no single approach can adequately assess health needs. A combination of approaches, taking account of the purposes of a given needs assessment, is more productive.

Key informants

In the key informant method, interviews are conducted with community residents or local workers having extensive firsthand experience in the community under study. Interview questions relate to existing services and demographic characteristics of the community. The data collected from these interviews are collapsed to obtain an overall picture of the community—from the key informant's perspectives.[26]

Delphi technique

The Delphi technique has been defined as a "method for systematic solicitation and collection of judgements on a particular topic through a set of carefully designed sequential questionnaires interspersed with feedback of opinions derived from earlier responses."[27] This technique aims to collect and refine judgments from experts through a reiterative process. It ultimately produces a collective agreement as to the nature and range of present community health care needs or a prediction of future needs. Opinions are usually submitted anonymously. The nature of the topic being investigated affects the organization of the Delphi process. Variations in question structure, rules governing aggregation of judgments and the interaction of respondents, among other things, determine its specific form.[28]

158

Community impressions

The community impressions method integrates information collected in interviews with small groups of key informants with an as-wide-as-possible range of existing indicator, utilization or survey data. The scenario of needs developed in this manner is then validated through the community forum process.[29]

APPLICATION OF NEEDS ASSESSMENT APPROACHES AND TECHNIQUES: THE CASE STUDY

It appears that no single approach can adequately assess health care needs. A combination of approaches, taking account of the purposes of a given needs assessment, is more productive.[30] With this in mind, the consultants to the community health care clinic in the case study proposed a five-step strategy to plan its health care program for the elderly.

The five steps

Analysis of service utilization data from the clinic's records would be carried out to document the demographic characteristics of senior citizen users and the types of health and social problems resulting in consultation. This corresponds to the analysis-of-utilization and rates-under-treatment methods discussed earlier.

The health needs and requirements of the aged have been extensively documented in the literature.[31] In terms of demographic data, there was no reason to expect that the elderly residents of the clinic's catchment area had special characteristics. It was therefore agreed that a literature review of the elderly's health care and social needs would provide an acceptable catalogue of possible needs of the senior citizens in the catchment area. This step corresponds to the indicator approach to needs assessment and uses a somewhat modified extrapolation/assumption method.

In order to rank needs according to the situation existing in the catchment area, methods by which local experts and senior citizens themselves could participate in the planning process were explored. The first proposal suggested using a structured questionnaire to interview local experts (physicians, public health nurses, pharmacists, clergy, organization personnel and senior citizen leaders involved in the community) about their perceptions of health and social needs of the catchment area's elderly population. This would correspond to the key informant method of needs assessment.

This step was subsequently modified somewhat to lessen the amount of time it would require and to integrate more of the concrete data that had been collected. It was agreed to replace the interviews with a workshop at which participants would use the literature review and utilization data in conjunction with their own perceptions to enumerate and rank the needs they believed existed in the catchment population. Workshop participants would include a physician and nurse from the clinic, a public health nurse working in the area, two members of the local clergy, a local pharmacist, representatives from the Senior Citizen's Council and local senior citizen leaders. The workshop would be led by a member of the consulting team to focus discussions and to ensure a satisfactory pace. The consulting team would prepare a working document based on the results of the workshop discussions. This method might correspond to the nominal group method, although it is less structured and includes use of specific data (previously collected) to provide a framework for discussion.

In addition to the workshop, a public meeting of senior citizens from the catchment area was planned. Every effort would be made to publicize this meeting and to encourage as much senior citizen participation as possible. The meeting would present the results of the workshop discussions, thus allowing for agreement or disagreement with the ideas collected. This corresponds to the community forum method of needs assessment.

Finally, using the information collected and taking account of financial, personnel and technological resources, the clinic's planning committee would establish the objectives and content of a program to meet the needs of the elderly residents of the catchment area. Clinic staff would be responsible for initiating program elements, for ensuring community awareness and for program evaluation.

The overall strategy developed in the case study corresponds to the community impressions method, which uses key informants, integration of existing data and the community forum to arrive at a reasonable depiction of community needs.[32]

Results

The clinic's board of directors readily accepted the proposed action strategy. However, administrative problems and personnel conflicts arose, greatly affecting the clinic's stability and, though unrelated to the planning project, jeopardizing the planning effort.

159

Several longstanding board members and the administrator of the clinic resigned. The community development worker who had been extensively involved in the planning project had left the clinic to return to school prior to the management crisis.

The consulting team sought to remain detached from the clinic's management problems while continuing to support the proposed planning strategy. It is unknown at this time whether the clinic will reestablish the planning effort. This outcome demonstrates the hazards of the consultant role in planning.

DISADVANTAGES AND ADVANTAGES OF NEEDS ASSESSMENT TECHNIQUES

All of the needs assessment methods discussed are subject to problems of reliability and validity. There is a danger that exclusive reliance on any one method will yield a biased or impressionistic view of needs.

> Evaluation and interpretation of human service need is influenced by the vested interests and values of those formulating program goals; the diffuse and interrelated nature of social and health needs; the rapidly changing character of human service needs; and the capabilities and interests of staff . . .[33]

One advantage of using not one but several needs assessment methods is that they are complementary and can be used to produce what has been termed a "convergent analysis." Convergent analysis is a methodology in which information from a variety of sources, collected by various strategies, is identified, defined, evaluated and assigned priority. It is then integrated to form a cumulative description of the needs in a particular social area.[34] Another clear advantage of a convergent analysis is the potential for savings in terms of time and money. Existing information that might have limited usefulness on its own can be an effective component within a larger framework, thus saving data collection time and expenses.

The importance of the method of program planning proposed in the case study is that it illustrates the potential—particularly in small community health settings—for integrating consumer and professional definitions of need, and for providing a framework in which different concepts of need can be examined as part of a creative and rational planning process. Needs assessment methods must, of course, be adapted to each planning situation. For program planners in small community health care settings, the approach described here can be an economical and useful alternative.

REFERENCES

1. Hilleboe, H.E. "Health Planning on a Community Basis." *Medical Care* 6:3 (1968) p. 203-214.
2. Donabedian, A. *Aspects of Medical Care Administration: Specifying Requirements for Health Care* (Cambridge: Harvard University Press 1973).
3. Blum, H.L. *Planning for Health* (New York: Human Sciences Press 1974).
4. Flexner, W.A. and Berkowitz, E.N. "Marketing Research in Health Services Planning: A Model." *Public Health Reports* 94:6 (1979) p. 503-513.
5. Lynch, G. and Clemenhagen, C. "A Re-examination of Community-Based Measurements of Health." Paper presented at The 1979 Canadian Public Health Association Conference, Winnipeg, Manitoba, June 1979.
6. Maslow, A.H. "A Theory of Human Motivation." *Psychological Review* 50:4 (1943) p. 370-396.
7. MacStravic, R.E. *Determining Health Needs* (Ann Arbor: Health Administration Press 1978).
8. Hilleboe. "Health Planning on a Community Basis."
9. Delbecq, A.L., Van de Ven, A.H. and Gustafson, D.H. *Group Techniques for Program Planning—A Guide to Nominal Group and Delphi Processes* (Glenview, Ill.: Foresman, 1975).
10. Pineault, R. "Elements et etapes d'elaboration d'un programme de sante communautaire." *Union Medicale du Canada* 105:8 (1976) p. 1208-1214.
11. Rossi, P.H., Freeman, H.E. and Wright, S.R. *Evaluation: A Systematic Approach* (Beverly Hills: Sage Publications 1979).
12. Siegel, L.M., Attkisson, C.C. and Carson, L.G. "Need Identification and Program Planning in the Community Context" in Attkisson, C. C. et al., eds. *Evaluation of Human Service Programs* (New York: Academic Press 1978).
13. Ibid.
14. Bell, R.A. et al. "Service Utilization, Social Indicator, and Citizen Survey Approaches to Human Service Need Assessment." In Attkisson et al., eds. *Evaluation of Human Service Programs.*
15. Ibid.
16. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
17. Committee on Health Sciences. Science for Health Services. Science Council of Canada Report no. 22 (Ottawa: Science Council of Canada 1974).
18. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
19. Rossi, Freeman and Wright. *Evaluation: A Systematic Approach.*
20. MacStravic. *Determining Health Needs.*
21. Chambers, L.W., Woodward, C.A. and Dok, C. *Guide to Health Needs Assessment: A Critique of Available Sources of*

160

Health Care Information (Ottawa: Canadian Public Health Association, 1980).
22. Ibid.
23. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
24. Rossi, Freeman and Wright. *Evaluation: A Systematic Approach.*
25. Delbecq, Van de Ven and Gustafson. *Group Techniques for Program Planning.*
26. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
27. Delbecq, Van de Ven and Gustafson. *Group Techniques for Program Planning.*
28. Ibid.
29. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
30. Chambers, Woodward and Dok. *Guide to Health Needs Assessment.*
31. Ontario Council of Health. *Health Care for the Aged* (Toronto: Council of Health 1978).
32. Siegel, Attkisson and Carson. "Need Identification and Program Planning in the Community Context."
33. Ibid.
34. Ibid.

161

GUIDE TO HEALTH NEEDS ASSESSMENT: A CRITIQUE OF AVAILABLE SOURCES OF HEALTH AND HEALTH CARE INFORMATION

by

Larry W. Chambers B.A. (Hons.), M.Sc., Ph.D.
Assistant Professor

Christel A. Woodward B.S., M.A., Ph.D.
Associate Professor

Caroline Dok
Research Assistant

Department of Clinical Epidemiology and Biostatistics
Faculty of Health Sciences
McMaster University
Hamilton, Ontario

Published by
The Canadian Public Health Association

First Edition
1980
Second Printing

PRETEST FOR POTENTIAL USERS OF THIS GUIDE

1. The meaning of the phrase "Health needs assessment" is widely agreed upon in the health and health care evaluation and planning literature.

   1. TRUE
   2. FALSE

2. Sources of health and health care information which can be accessed at any time with minimal cost include:

   1. MORBIDITY/MORTALITY STATISTICS
   2. UTILIZATION/RATES UNDER TREATMENT STATISTICS
   3. STATISTICS ON MANPOWER AND HEALTH CARE FACILITIES
   4. SOCIAL INDICATOR STATISTICS
   5. NONE OF THE ABOVE

3. Health and health care information can be generated by developing a method of collecting data from:

   1. KEY INFORMANTS
   2. SERVICE AGENCY STAFF
   3. SERVICE AGENCY RATES UNDER TREATMENT
   4. SURVEY OF THE PERSONS SERVED BY THE AGENCY(S) UNDER REVIEW
   5. A PUBLIC FORUM
   6. SURVEY OF THE POPULATION
   7. SURVEY OF REFERRAL SOURCES
   8. NONE OF THE ABOVE

(Answers on page 3)

Suggestions on How to Use this Guide

I Do the pretest to initially assess your knowledge of the contents of this Guide.

II This Guide addresses only the issue of availability (Is it reaching people who need it?) of health manoeuvres, procedures or services.

Information on availability is of little value unless preceded by, and based on, research demonstrating efficacy ("Can it work?" Does the health manoeuvre, procedure or service do more good than harm to people who fully comply with associated recommendations or treatment) and effectiveness ("Does it work" Does the manoeuvre, procedure or service do more good than harm to those people to whom it is offered?)
Methodologic standards for assessing efficacy and effectiveness are described in Sackett DL. On the evaluation of health services. In: Last J, ed. Preventive Medicine and Public Health, 11th Edition. New York, Appleton Century, Crofts(in press).

III The Guide is divided into seven sections:

	Page
1. Different Conceptual Approaches to Health Needs Assessment.	4
- definitions of need, want, demand, use, supply	
2. Existing Sources of Health and Health Care Information.	5
- for each source the Guide	
(1) suggests methods of gathering information,	
(2) reviews the advantages and disadvantages of using the source,	
(3) gives examples of use of the information source in addressing need, demand, want, use and/or supply problems	
- the existing sources are:	
- mortality/morbidity statistics	5
- utilization/rates under treatment statistics	7
- manpower and health care facilities	9
- social indicators	11
3. New Sources of Health and Health Care Information.	13
- alternative sources	
1. New Information from key Informants	13
2. New Information from Service Agencies	16
(a) Survey of Agency Staff	16
(b) Rates under Treatment	17
3. New Information by Survey of the Service Population	19
4. New Information from Referral Sources	20
5. New Information from the General Population	21
(a) Open Forum	21
(b) Population Survey	22
- for each source	
(1) methods,	
(2) advantages and disadvantages,	
(3) examples	
4. Alternative Tactics For Asking Questions when Obtaining New Information.	23
5. Suggested Further Readings:	24
- on "needs assessment"	24
- selected readings on prioritizing needs	24
- further methodologic papers regarding use of existing data sources	24
- further readings on survey research	26
IV Post-test on the relationship between sources of information and type of health and health care information provided.	26

IV STOP READING!

Before reading further we suggest you write out the purposes for which you need health and health care information.

After you have documented your objectives, review all the approaches suggested in the Guide and decide on their suitability for your objectives.

Note:

No single approach will provide all the health and health care information required for an adequate assessment of "needs"

→ Use of a combination of approaches (with differing strengths and weaknesses) is suggested as the best tactic.

V The methods for deciding on priorities for new or existing health programs to meet identified "needs" are beyond the scope of the Guide.

→ see selected readings regarding priority setting on page 26.

Answers to Pretest

1 = FALSE
2 = 1,2,3,4
3 = 1,2,3,4,5,6,7.

1. DIFFERENT CONCEPTUAL APPROACHES

Possible Definitions of the Term "Need" (Types of Health and Health Care Information)

1. Want or "expected" services
   - -the perceptions by the public of their own wants, expectations, rights
   - -public's perception of health status may differ from health professional's view of optimal health status
   - -to an economist, wants apply to consumers' desires for health services prior to any knowledge of the costs or prices involved

2. Need or "ought to be" services
   - -services that should be provided to the public on the basis of the perceptions of experts
   - -this includes interpretation of need for health services on the basis of health status information about target groups

4. Use or health services actually utilized

3. Demand or the types and amounts of health services requested or desired by the public once they know the costs and prices involved
   - -wants change with additional knowledge of what it will cost (or what alternatives would be foregone) to fulfill them
   - -frequently demand is expressed in public action by informed consumers

5. Supply or "get-at-able" health service
   - -either quantitatively considering the numbers and distributions of facilities and health personnel relative to the populations they serve, or qualitatively considering the quality of care provided by these services.

When the Health Care System is in Equilibrium:

1) Want = Need = Demand = Use = Supply and Quantity = Quality, and information abour one can be used as a substitute for information about any other.

2) However, the health care system is usually not in equilibrium:

eg. Want > Need
- -when consumers disagree over:
  - eg. the necessity for house calls
  - eg. whether antibiotics are required for mild upper respiratory infections

eg. Need > Demand
- -when the poor fail to request additional health services despite the lowering of economic barriers to access

eg. Demand > Need
- -when communities after building health care facilities still find it difficult to attract and keep primary care clinicians

eg. Use < Supply
- -some authorities have suggested that elective surgical procedures such as cholecystectomies are more common in locales with large numbers of surgeons

eg. Quantity > Quality
- -when a profusion of unsolicited laboratory screening test results are shown not to benefit patients admitted to hospital

*

Before proceeding further in this guide, consider whether the information you require for your program or study relates to: want, need, demand, use or supply (quantity and/or quality) or a combination of these.

The following sections of this guide refer to two basic methods to obtaining information about whether the health manoeuvre/procedure/service is reaching those who need it. The first accesses existing health and health care information. The second develops new health and health care information. The following table summarizes these information sources and their adequacy in measuring want, need, demand, use or supply (quantity and/or quality).

* For further examples see: Sackett DL. On the evaluation of health services. In: Last J, ed. Preventive Medicine and Public Health, 11th Edition. New York, Appleton Century, Crofts (in press).

2. EXISTING SOURCES OF HEALTH CARE INFORMATION

In this Guide "existing data" are defined as routinely reported data which is made available to all interested users usually in annual printed reports or computer data tapes by federal and provincial governments and profit and non-profit private health agencies.

The general method employed in utilizing sources of existing data involves:

1. Determining the type of data required in order to answer your questions.
2. Obtaining copies of the report or data tape with this information (confidentiality may pose problems)
3. Analyzing the data
4. Writing a report of your findings

Sources of existing data will be divided into four types according to broad subject areas: i.e. mortality/morbidity data; utilization/rates under treatment; manpower and health care facilities; and, social indicators.

1. Mortality/Morbidity Statistics

a) Sources of Information

1) National Data

Statistics Canada reports which provide this type of information include: (1) vital statistics(e.g. deaths, life tables, disease specific mortality); (2) public and psychiatric hospital morbidity statistics (e.g. hospital separations by diagnosis, hospital admissions by diagnosis); (3) notifiable disease reports (e.g. weekly and annual); (4) new primary sites of malignant neoplasms (e.g. based on provincial cancer registry reports); (5) tuberculosis (e.g. occasional reports), infectious hepatitis (e.g. occasional reports); and (6) illness and health care (e.g. 1951 Canada Sickness Survey and the 1979 Canada Health Survey)

> See the Statistics Canada Catalogue, available free on request from Statistics Canada, Ottawa, K1A 0T6, which lists all current publications available under a wide range of topics. Refer especially to the section listing those publications under Education,Health and Welfare. Section 82-Public Health, Section 83-Hospitals and Health Manpower, and Section 84-Vital Statistics will list those most applicable.

The Department of Health and Welfare Canada has published a number of special reports on mortality and morbidity due to smoking, alcohol use and motor vehicle traffic accidents.

In addition to routine Canadian reports, a number of national surveys in the United States provide further information on acute conditions, chronic conditions, chronic limitation of activity and mobility, impairments, psychologic measurements, physiological measurements, use of special aids and usual activity status. These data are derived from: the U.S. National Interview Survey, the U.S. National Health Examination Survey, the U.S. National Health and Nutrition Examination Survey, and the U.S. National Ambulatory Care Survey.

> National Center for Health Statistics: Current listing and topical index to the vital and health statistics series 1962-1977. Hyattsville, Md. : National Center for Health Statistics. (DHEW publication no: (HSM) 72-1036).

2) Provincial Data

Provincial governments in Canada are responsible for maintaining vital statistics records and the provision of health services by hospitals, physicians and some community agencies (e.g. public health units). Mortality and morbidity data reports are usually available from each province covering most of the areas reported at the federal level with the main distinction that smaller geographic areas and target groups are reported on (e.g. provincial health planning areas, district health councils, health unit regions, cities and counties etc.).

> See for example the annual reports on Hospital Statistics and Vital Statistics published by the Ontario Ministry of Health.

3) Profit and Non-Profit Private Health Agencies

These include nursing home and special disease associations (e.g. heart, cancer, multiple sclerosis, mental retardation, physically handicapped) which collect primarily morbidity statistics for facilities they operate or special disease registries.

b) Advantages and Disadvantages

Advantages

1. Easily communicated to health professionals
2. Direction of improvement in measures often straight forward
3. Usually amenable to mathematical index (rate) construction
4. Readily available and inexpensive to access

Disadvantages

1. Suffer from community/population under enumeration because in most instances (exception is mortality) based on health agency records
2. Residence classification unreliable (e.g. location of death may not be place where lived most of life)
3. Problems of repeatability in assigning codes (e.g. ICDA diagnostic codes)
4. Lack of continuity in trends and changes in coding rules resulting from revisions of the ICDA
5. Variations in medical terminology
6. Differences in throughness in completing medical cause of death certification and morbid events in hospitals and communities
7. Difficulties in establishing the correct cause of death and diagnosis

Examples of Use of Exisiting Health and Health Care Information on Mortality and Morbidity

The Lalonde Report (Need)

This report, entitled "A New Perspective On The Health Of Canadians" relied heavily on mortality statistics to emphasize that diseases of the heart, malignant neoplasms, cerebrovascular diseases and accidents of all types are the four leading causes of death in Canada. In addition, these analyses were further refined by ranking causes of death according to potential years of life lost, method which gives more weight to deaths which are "premature". Death before the age of 70 was arbitrarily considered "premature". Using this method, motor vehicle accidents, ischaemic heart disease, other accidents and respiratory diseases and lung cancer rank as the four most important causes. The report went on to argue that health care services would not have an impact on these diseases which were associated with "unhealthy life styles" eg. cigarette smoking, overly stressful environments and living, drinking and driving, underuse of automobile seatbelts, etc.

1. Lalonde M. A new perspective on the health of Canadians, Ottawa: Health and Welfare Canada, Government of Canada, 1974.

You may wish to track down these examples and the examples that follow in order to reach your own opinion about the quality of the data.

Prediction of Hospital Bed Requirements of the Elderly (Supply)

An examination of the current patient use of in-patient care for the elderly was conducted in a Scottish health district where pre-admission assessment visits by the Consultant in Geriatric Medicine ensures that only those who require to be in hospital are admitted. Bed utilization and census projections were used to derive an age, sex, related formula for the calculation of bed requirements.

Morris JB, Cohen C. A formula for estimating the bed requirements for a department of geriatric medicine. Health Bulletin (May) 1978; 36: 3: 111-114

2. Utilization/Rates under Treatment Statistics

Utilization and rates under treatment statistics can include a number of indicators regarding the demand and use of health services. Common statistics include: number of health services (volume of services), number of persons receiving services, and services per patient rates. In health economic analyses especially the time taken by health professionals to provide services, waiting time of patients, broken appointment rates and time in the day or week when the services are provided are also useful in order to arrive at services per unit of time statistics. These are sometimes summarized by type of service, age, sex, diagnosis, socio-economic status and other social and health characteristics of persons using the health services.

a) Sources of Information

1) National Data

Statistics Canada routine (annual or more frequently) utilization and rates under treatment reports include: (1) public and psychiatric hospital statistics (eg. occupancy rates, number and type of surgical procedures and treatment, movement of patients, laboratory, radiology and other special and supportive services, and hours of work); (2) 1951 Canada Sickness Survey (eg. reports of hospital, medical, dental and social services); (3) 1979 Canada Health Survey (eg. reports of hospital, medical, dental and social services); and (4) therapeutic abortions

The Department of Health and Welfare Canada also publishes reports on medical services (eg. cost and frequency) paid for through federal/provincial cost sharing and northern medical and nursing services.

2) Provincial Data

Provincial governments in Canada are responsible for providing health services by hospitals chronic care facilities (hospitals and nursing homes), physicians and some community agencies (eg. public health units, workmen compensation) and routinely report utilization and rates under treatment statistics on a provincial and local level.

3) Profit and Non-Profit Private Health Agencies

Non-government health agencies providing statistics on use of and demand for services include many nursing homes, homes for the aged, sheltered workshops, Red Cross, meals on wheels, home nursing serivces, drop-in centers for alcoholics and rehabilitation centers for the physically and mentally handicapped.

b) Advantages and Disadvantages

Advantages

1. Availability and low cost.
2. Usually reliable if financial incentive to complete forms and/or to report performance
3. Can be used to compare utilization accross facilities, regions and time

Disadvantages

1. May be inaccessible or poorly reported (under-reporting is common)
2. Coding may vary from center to center making comparison difficult

Example of Use of Existing Utilization/Rates Under Treatment Data

Management of Government Financing in Ontario Using the Relative Stay Index (Use, Demand, Supply)

In 1969 the Ontario Hospital Services Commission introduced a program of financial incentives for hospitals based on a Relative Stay Index (RSI) Report. The RSI Report is derived from patient data routine reported by hospitals in Ontario to the provincial government. "The RSI Report examines all discharges from groups of active treatment hospitals in Ontario, based on bed size or teaching affiliation, during a year or other specified period of time. Each discharge is classified according to age, sex and primary diagnosis of the patient where primary diagnosis is classified as one of 41 broad diagnostic categories. This information is used to develop comparative reports for each hospital within the group. Through these comparative statistics, a hospital may learn that its patients in a particular diagnostic category remained in hospital for a significantly greater length of time than the experience of other similar hospitals would suggest is required. The concept used to express this idea is the Relative Stay Index (RSI)

"The RSI shows the percentage by which a hospital's average length of stay, for a group of patients of specified age, sex, and diagnostic mix, is greater or less than would be expected on the basis of the experience which other hospitals of similar size had with the same mix of patients."

The following table shows the guidelines of the financial incentive program for hospitals:

Commission Policy on Changes in Patient - Days

Patient-days	No. of Patients	Length of Stay (RSI)	Commission Policy
Up or no change	Up	No change or down	Would agree to additional funds
Down	No change or up	Down	Would not reduce budget
Up	Down	Up	Would not automatically agree to additional funds

Migué J-L, Bélanger G. The Price of Health. Toronto Macmillan 1974, p. 144

3. Manpower and Health Care Facilities

The supply of health manpower and facilities to house equipment and personnel is often the easiest health data to obtain and the most inadequate in terms of providing information on whether those who require health services are receiving them.

a) Sources of Information

1) National Data

Statistics Canada routine reports include data on: (1) public and psychiatric hospital statistics (eg. number, bed capacity, by province, type of hospital service and size of hospital; personnel; training facilities); (2) manpower statistics (registered nurses, physicians and surgeons, radiological technicians, physiotherapists); and, (3) list of special care facilities (eg. nursing homes, homes for the aged, physically handicapped, the blind, deaf, the mentally handicapped, emotionally distrubed children, alcoholics and/or drug addicts, delinquents, unmarried mothers).

The Department of Health and Welfare Canada also publishes manpower and health care facilities reports.

2) Provincial Data

Provincial governments also publish manpower and facilities reports

3) Profit and Non-Profit Private Agencies

Professional association at national, provincial and local levels, such as the Canadian Medical Association, the Canadian Nurses Association, the Royal College of Physicians and Surgeons, the Canadian Physiotherapist Association and many others (usually one for each type of health care speciality) all keep data on the numbers of their members who are in training or in the work force.

The Canadian Hospital Association, and its provincial counterparts can provide data on acute and chronic care facilities. Nursing home associations also exist in most provinces.

b) Advantages and Disadvantages

Advantages	Disadvantages
1. Availability and low cost	1. May be inaccessible or poorly reported (manpower reports become dated quickly)
2. Can be used to compare manpower and facilities across provinces, regions and time	2. Name of facility often does not accurately reflect the type of services it provides - Similarly, health professionals trained to do one thing may be actually involved in activities not directly related to their training or professional affiliation

Medical Manpower Studies (Supply)

Spaulding and Spitzer (1972) conducted a study of primary (general practitioners and family physicians) physicians in the Province of Ontario between 1961 and 1971 using Canadian Medical Directories publications and lists of the College of physicians and Surgeons of Ontario and census of population reports of Statistics Canada. Population per physician ratios and population per primary physician ratios were calculated using the routinely collected data obtained from these three sources. The following table is an example of their findings:

	Population per Primary Physician		
	1961	1966	1971
Southern Ontario	1689	1708	1694
Northern Ontario	2158	2072	2038
Ontario	1731	1741	1723

This table shows the supply of primary physicians improved slightly for the province, with a gain having been made in the north but the northern ratio continued to differ substantially for the south. With this and other data the authors argue that increases in medical manpower from immigration and Ontario medical schools would be unlikely to overcome the north/south discrepancy and, indeed the south would become even more "over doctored" without changes in government policy

2. Spaulding WB, Spitzer Wo. Implications of medical manpower trends in Ontario 1961-1971. Ontario Medical Review (September) 1972:39:9: 527-533

4. Social and Indicators

Health professionals use social indicators to evaluate health services and plan new health services.

a) Sources of Information

1) National Data

Statistics Canada routinely publishes the following social indicators which are often used to assess the need for health services: (1) vital statistics (eg. births, marriages, divorces); (2) criminal and correctional statistics; (3) census of the population (eg. population density, age composition, income, education, occupation, housing characteristics, ethnicity, natural origin, length of residence, marital status, family status, employment), transportation statistics (eg. motor vehicle accidents by location, time of day, driving conditions, age and sex of victims); and employment and compensable injury statistics.

The Department of Health and Welfare Canada also publishes reports on many social indicators (eg. demographic characteristics of smokers and alcoholics.

2) Provincial Data

Provincial governments also produce numerous reports on similar social indicators

3) Profit and Non-Profit Private Agencies

Most communities in Canada have social planning and research councils which publish many reports at a local level. In addition, they usually begin by compiling a list of social and health agencies in the community from which information regarding various social indicators can be obtained.

b) Advantages and Disadvantages

Advantages

1. Readily available and inexpensive to assess
2. Can be used to monitor effects of intervention

Disadvantages

1. May be grossly outdated as most data is collected at 10 year intervals (the census) (1951,61,71,81)
2. "Hard" measures like age, sex, date of birth are often accurate "Soft" or frequently inaccurate and complex measures include income, education, employment
3. Often only incidence (new characteristics/problems of persons at a given point in time) data available and not prevalence (total number of persons with characteristics/ problems at a given point in time)
4. Relationships between health and social indicators often tenuous
5. Comparisons across jurisdictions often impossible because of overlapping jurisdictions or lack of comparable data about each jurisdiction

Example of Use of Social Indicators

Monitoring the Effect of Traffic Safety Legislation (need)

In 1969 the Criminal Code of Canada was amended to give police the power to arrest a driver on the basis of a positive breathalyser test, or, for refusal to submit to such a test. Accident rate data for Canada were dervied from Statistic Canada's Quarterly Motor Vehicle Traffic Accident Reports and Population Census Reports. Incidence, case-fatality and mortality rates were used to extrapolate trends of pre-1969 Breathalyser Amendment rates for the post breathalyser years 1970, 1971 and 1972 in Canada. The difference between actual and predicted rates after the introduction of the 1969 Breathalyser Amendment was explained by initial reductions in the incidence of injury (the greatest reductions in incidence occurred at "heavy drinking" periods - 6 p.m. to 12 midnight and 12 midnight to 6 a.m.) and initial reductions in the severity of the resulting injury, producing a magnified downward movement in mortality rate. However, this early "shock" effect lasted less than one and a half years with 1971 and 1972 mortality rates exceeding pre-breathalyser levels. This report and others have contributed to the introduction of further measures such as the requirement that drivers be tested at the roadside and confiscation of the drunk driver's licence for up to two days after a positive roadside test.

Chambers LW, Roberts RS, Voelker CA. The epidemiology of traffic accidents and the effect of the 1969 Breathalyser Amendment in Canada. Accident Analysis and Prevention 1976; 8: 201-206

3 NEW SOURCES OF HEALTH and HEALTH CARE INFORMATION

In this Guide "new sources of health and health care information" are defined as information which is not available from existing sources but which must be collected prospectively. New sources of information include data obtained from key informants, service agencies, surveys of a population served by an agency or information from the general population.

1. New Information from Key Informants

Key informants are people you, or persons familiar with the community, identify as possessing considerable knowledge of and experience in the community, its people and their health and health care wants, needs, demand, use of services and supply (quantity and quality) of services.

i) Methods

1. Write out objectives

2. Identify key informants - those most familiar with the problem
   e.g. physicians, clergy, public health nurses, directors of health agencies (e.g. Multiple Sclerosis Society, Heart Fund)
   e.g. community leaders (formal elected or appointed positions as reported in directories OR informal leaders based on survey of opinions of citizens)

3. Decide on method of contacting key informants
   e.g. one-to-one contacts (telephone or face-to-face interview, mailed questionnaire)
   e.g. group contacts (meeting of more than one key informant at a time) can take the form of:
      1. questions to each informant in the group in presence of others
      2. requesting the group to discuss program and produce a report (preferably written out - long or short depending on feasibility) which summarizes the group's "answers".

4. Contact survey or convene group meetings of key informants
   - Be sure to have questions ready before the first meeting (or meetings if all the issues cannot be covered in one meeting).
   - Sometimes it is better to record replies of key informants after the meeting.
   - Tape recording can be used as long as anonymity and confidentiality are ensured.

5. Write report based on face-to-face or group responses.

ii) Advantages and Disadvantages

Advantages of Key Informant Approach

1. Easy to do
2. Inexpensive
3. Broad range of information can be elicited
4. Secondary benefit of "preparing" key informants for later involvement in planning and carrying out program to provide services to these who need it
5. If meetings used, bring key health and social (caring) services together to discuss common concerns (may open new lines of communication)

Disadvantages of Key Informant Approach

1. Key informants views on want, need, demand, use and supply may be limited to only the "visible"
2. Personal bias of key informants may distort information (eg. they may describe "the world" only in terms of their organization)
3. Poorly chosen questions may limit the study
4. Dependent on cooperation of key informants chosen.
5. Unless skillful interviewers are chosen or trained much data may be lost or distorted.
6. If meetings used, the loudest voice or most persuasive person may prevail.

Examples of Use of Key Informants

Health Care in the East End of Hamilton

A study was designed to describe the way health services were being made available in the east end of Wentworth County and to identify the problems experienced by its residents in receiving care and by providers in delivering care. Family doctors and community nurses were identified as potential key informants familiar with the area and were surveyed to help assess the health needs of this community. It should be noted that in this study, at least three other approaches to needs assessment were utilized i.e. a consumer survey, a survey of various agency staff and application of historical data.

Report of the Steering Committee for the study of health care in the East End of Hamilton-Wentworth, June 1978.

1974 Health Planning Task Force (need, supply)

A committee was asked by the Ontario Ministry of Health to develop proposals for a comprehensive plan to meet the health needs of the people of Ontario. The committee consisted of a dean of a health sciences center, a director of public health nursing in a health unit, a health economics professor, a chartered accountant, rector of a university, an assistant deputy minister of health, a physician, the president elect of the Canadian Medical Association and two senior staff members in the Ministry of Health. The committee was asked to develop a plan for organizing, within available fiscal resources, a comprehensive health plan for Ontario. The 77 page report made the following recommendations:

Regionalization

Within 6 regions, District Health Councils to be established; Area Boards within Councils to coordinate institutional management; primary care groups to provide continuous complete care; secondary care groups to provide specialist services.

Decentralization

District Councils to have policy and planning role; to coordinate health and educational programs; to determine needs and objectives; to evaluate district programs; to make recommendations for Ministry approval; health districts may be larger than regional government areas

Local Health Boards

Elements of personal health care to move to primary care sector (promotion, innumization, child guidance). Boards to be responsible for environmental concerns (industrial health, food safety, information on communicable diseases). M.D.H. to become district advisor in community health

Consumerism in Decision-Making

District councils to consist of 10 impartial members appointed by Lt. Governor - in - Council from lists of nominees submitted by organizations and individuals in district and up to 5 representatives of municipal government in district Professional and administrative personnel to be advisory to council.

Primary Care Organization

24 hr. 7 day-a-week continuous comprehensive care must be available task sharing and delegation of authority for optimal efficiency; primary care level is point of contact for almost all major and minor ailments. Primarily care to be delivered by "groups" (teams). A group could serve 10-15,000 using 5-6 physicians, 5-6 nurses.

(Cont'd on next page)

Health Promotion

Programs of health promotion might be used by District Councils to improve health status; programs based in primary groups. Lifestyle modification programs should be emphasized

Control of Cost Increases

Selective bed closings; alteration of hospitals or hospital wings; payment of professionals to be advised through remunerations advisory committee; need for alteration of fee-for-service system; benefit-for-service schedule advocated to permit unit billing by primary groups regardless of professional delivering service; ultimate savings through a health care system stressing primary care.

Allocation of Funds to Sectors

More funds to be channeled into primary care sector. Program budgeting systems to be employed. In initial stages, health councils would not have power to authorize expenditure; budget recommendations for funding must pass through district council.

Development and Enforcement of Standards

Minister of health ultimately responsible; should coordinate and finance an evaluation system. Standardized record system should be established. A Council of Health to advise Minister and a Health Disciplines Board to oversee professions.

Evaluation of Quality

(a) On-going outcomes to objectives; monitoring resources used (b) Audit extension of medical audit of patient records to primary care groups and secondary care programs (c) Review: External peer review by teams including layman; use of audit and OHIP data; internal peer review mechanisms to be established for groups and programs.

Education

All health professions should have increased exposure to primary care settings; clinical training should be based on district services rather than teaching hospitals, training in team approach essential; continuing education programs must be fostered; advancement and upgrading programs, flexible in nature, must be provided.

Rules

Flexibility rather than rigidity in role definitions; individuals to be treated by person most capable of handling problems; alternative career patterns (Ladders) to be established. Needs in new system: skilled administrators, health researchers and planners.

Utilization

Types and numbers of health professionals for a district to be set by district board. Other professionals, especially physicians, could practice in district but would be excluded from OHIP.

Some of the recommendations have been implemented whereas others will perhaps never be implemented in Ontario.

Ontario Ministry of Health. Report of the health planning task force. Toronto: Ontario Ministry of Health, 1974.

2. New Information from Service Agencies

Agencies providing services in the health area under investigation may be asked to contribute new health and health care information by a survey of agency staff or by obtaining information regarding "rates under treatment." The agency(s) itself may have been chosen (1) because of its service objectives, (2) the type of staff, (3) its catchment area, (4) the type of clientele, or (5) the type of services it provides.

(a) Survey of Agency Staff

Agency staff could include both administrators and providers of services in the agency (or agencies).

i) Methods

1. Define objectives of survey
2. Decide whether to contact staff independently (face-to-face, a telephone interview or self-administered, mail questionnaire) or in group interview.
3. Select appropriate sampling strategy
4. Design questionnaire
5. Implement survey
6. Analyze data
7. Write report

ii) Advantages and Disadvantages

Advantages

1. Obtain health and health care information from the providers perspective
2. Number of providers is usually small thus, the cost is relatively low
3. Location of providers usually known (eg. medical directory)

Disadvantages

1. Agency staff may reflect the biases of the organization
2. Fairly difficult to implement if require agency permission and professional organization approval
3. Magnitude of problems may be underestimated as agency staff usually have contact primarily with those who seek services and find them accessible

Example of Survey of Agency Staff

Voluntary Health Agency
Survey of District Health Council (Need, Supply)

The Thames Valley District Health Council developed a question to administer to 63 voluntary agencies in the Health District to obtain data regarding each agency's (1) organization, (2) funding, (3) expenses, (4) availability of services, (5) length of time in operation, (6) identity of groups served, (7) age groups (child, adolescent, adult) served (8) source of referrals, (9) staffing and (10) types of services provided. The survey results showed the groups served by the 63 voluntary agencies were: blind, deaf, emotionally disturbed, mentally retarded, unwed mothers, dependent children, alcoholics, drug abusers, physically disabled, senior citizens, other neurologically handicapped, unspecified illness and other. The agency representatives were also asked to identify gaps in services in the Health District. For example, lack of psychiatric services for the mentally retarded and the deaf were mentioned by three agencies. Two agencies noted a shortage of volunteers. Government regulations excluded the chronically ill and the elderly from services altogether or on an insufficient basis.

Boucher MP, Pablo RY, Roberts RDT. Voluntary health agencies - who, what and why? A survey of the Thames Valley District Health Council Area. Can J Public Health (May/June) 1979; 70: 183-198

2. New Information from Service Agencies

(b) Rates Under Treatment

This term usually refers to the following kinds of information about the population served by an agency:

- who is receiving care (by age, sex, geographic location, education status, etc.)
- their health status
- their transportation needs
- the kinds of services provided
- the number and cost of services per person served
- waiting lists and waiting times.

Although some of this health and health care information will be routinely maintained by agencies, in most instances the routinely collected information is incomplete.

i) Method

1. Outline survey objectives
2. Prepare a standard client encounter form or agency record abstracting form
3. Decide on a particular time period (future or past) for data collection
4. Determine whether all clientele will be surveyed or only a probability sample of clientele
5. Implement survey
6. Write report

ii) Advantages and Disadvantages

Advantages

1. Provides complete and uniformly collected utilization information
2. Relatively inexpensive
3. May reveal over-utilization or redundant services

Disadvantages

1. May require cooperation of numerous agencies
2. May be difficult to convince all staff to complete standard forms and/or record additional information
3. Provides no information on persons who have not accessed the agency

Examples of "New" Rates under Treatment

Emergency department use at two Hamilton hospitals

In 1971, 2608 consecutive emergency department visits during a 3 week period at St. Joseph's Hospital and 1360 similar visits made at Hamilton General Hospital during a 2 week period were reviewed to determine the age and sex distribution of users, their socioeconomic status; characteristics of use-such as time of arrival, ambulance use, use of radiology and laboratory facilities, duration of visits, urgency of visits, chief complaint and duration and severity of symptoms.

A total of 40 visits per night were registered at both hospitals. By extrapolation about 60 emergency department visits are made between midnight and 8 a.m. on any given day with approximately three being true emergencies. At both hospitals only 1 visit in 20 was designated as an emergency. Therefore, there are approximately 10 serious accidents or trauma cases in each 24 hour period in the city. These findings had implications for regionalization of emergency services and increased organization of primary care services.

Vayda E, Gent M, Hendershot A. Emergency department use at two Hamilton hospitals. Can Med Assoc J 1975; 112: 961-965

Coronary Care Unit Facilities in a Rural Area (Need, Supply)

In an attempt to determine the most appropriate type of facility for the treatment of myocardial infarction (heart attacks) in a rural community, the five-year experience of all (approximately 10,000) persons living in the catchment area of a rural hospital/health centre were reviewed. Geographic isolation resulted in persons receiving all their first contact health care in one facility where inpatient and ambulatory records were stored together as one chart and maintained for every citizen. The records were reviewed of all those cases where there was mention of some form of heart disease or where there was no other clear diagnosis given. Where the clinical chart did not contain enough information the relatives were interviewed. This population was found to have a low incidence of myocardial infarction, a low death rate from the condition both in hospitalized patients and in the general population, and a low incidence of complications in the survivors. None of the deaths either in or out of hospital were judged by a medical reviewer to be likely to have been prevented through the use of specialized facilities. It was concluded that no such facilities were necessary in the catchment area.

Black DP. Management of myocardial infarction in a rural area. Can Med Assoc J 1973; 103: 863-867.

3. New Information by Survey of the Service Population

Unlike the rates under treatment where the information is obtained through the agency, the surveys of the service population are usually done separately from the agency.

i) Method

1. Define objectives of survey
2. Define service population through:
   -referral agencies (eg. family physicians)
   -service agencies
   -consumer advocate groups
3. Decide on questionnaire format (face-to-face or telephone interview or mail questionnaire)
4. Develop questionnaire
5. Contact subjects
6. Analyse data
7. Write report

ii) Advantages and Disadvantages

Advantages	Disadvantages
1. Comprehensive method of obtaining health and health care information from group who have used services at least once	1. Fear of loosing needed service can "positively" bias people surveyed
2. Can offer insights on how to improve service: to make them more accessible and appealing	2. May not be representative of group of people in community in need of same services
	3. Volunteers for the survey may be atypical of the user group.
	4. Direct contact may be more expensive than settling for rates under treatment approach.

Example of Survey of Service Users (Use, Demand, Want)

Emergency department use at Hamilton hospitals

A random sample (stratifying for day of week and time of day) of patients who used emergency services at either St. Joseph's Hospital or Hamilton General Hospital were interviewed to identify the medical socioeconomic and demographic determinants of emergency department use in Hamilton and to develop recommendations regarding future organization of services.

Although one hospital served a lower socioeconomic and industrial population and the other a predominantly suburban population, the patterns of use were found to be similar. Over one third of the visits at both hospitals were classified as non-urgent. A perceived lack of availability and accessibility of primary care services in the evenings and on weekends and convenience appeared to be the major reason for the heavy "off-hours" emergency department use with approximately one-sixth of all visits being made between midnight and 8 a.m. and with Saturday being among the busiest days at both hospitals and Sunday the busiest day at St. Josephs.

This study indicated that long-term solutions for Hamilton could require regionalization of emergency services and increased organization of primary care services.

*Vayda E, Gent M, Hendershot A. Emergency department use at two Hamilton hospitals. Can Med Assoc J 1975; 112: 961-965

4. New Information by Survey of Referral Sources

Agencies directly servicing people are often asked whether there is a need for change in the level of service provided and to give estimates of manpower needs. As they may have an interest in maintaining or increasing service levels, the answers given are sometimes viewed with suspicion. A clearer picture of need for services may emerge by surveying the sources of referral (schools, family doctors, public health nurses) to the agency(s).

i) Method

1. Define objectives of survey
2. Define population of referral sources and specify sampling technique
3. Decide on format of survey (face-to-face or telephone interview or mail questionnaire)
4. Develop questionnaire
5. Contact referral sources
6. Analyse data
7. Write report

ii) Advantages and Disadvantages

Advantages

1. Useful way of obtaining information on ease of referral, waiting times, accessibility of service agencies, ability, to meet demand of referral sources for service.

Disadvantages

1. Referral sources may be unaware of some service agencies which could be used to meet client needs.
2. Only the service needs of clients who have already tried to access the health care system for help will be assessed.

Example of Survey of Service Referral Agencies

Assessment of Manpower Needs in Psychiatry (need, supply)

A manpower survey of Ontario psychiatrists requesting information regarding professional background, type of practice, workload, waiting times, referrals not accepted, etc. is proposed . To compliment this survey, a parallel survey of a random sample of Ontario family doctors is proposed. This survey of referral sources would obtain information on the availability and accessibility of psychiatric services for the physicians' patients. Information regarding length of time needed to obtain emergency and routine psychiatric consultation, etc. will be requested.

This study is currently in its planning stages. The information will be used to help make decisions regarding the number of postgraduate training slots in Psychiatry needed to meet manpower needs for the 1980's.

5. New Information from the General Population

Contacting citizens, whether or not they have been served by different agencies, often is the only way to determine an agency's relevance to the community and to obtain first hand health and health care information from citizens. Information from citizens can be obtained by open forums or by surveys.

(a) Open Forum

Open forums refer to a series of public meetings not unlike the "hearings" which are often held by government Royal Commissions or task forces.

i) Method

1. Define objectives
2. Arrange locations(s) for forum
3. Organize publicity to attract a cross-section of persons in the community
4. Train moderator(s) i.e. to cover the questions
5. Hold meetings
6. Tabulate responses
7. Write report

ii) Advantages and Disadvantages

Advantages

1. Relatively easy to arrange
2. Inexpensive
3. If attendence high, usually obtain wide spectrum of views
4. May identify previously unknown community people or agencies interested in planning and developing services

Disadvantages

1. Difficult to insure attending group is representative of the community
2. Vulnerable to individuals or groups with special points of view or grievances taking over meeting (creates bias and is upsetting to others attending)
3. Unrealistically high expectations may be communicated leading to rejection of later, more modest programs
4. Data collected is highly impressionistic. Usually not amenable to statistical treatment

(b) Population Survey

As this approach to obtaining new health and health care information usually involves contacting a large number of individuals, it is often useful to have determined the range of issues to be covered in the survey by other approaches outlined previously in order to not waste time or leave out important issues.

i) Method

1. Define objectives
   - reports of illness episdoes, symptoms or disease
     (e.g. Kohn R, White KL. Health care: an international study. London: Oxford University Press, 1976)
   - assessments of physical function, social function, emotional function
     (e.g. Stewart A, Ware JE, Brook RH. The meaning of health: understanding functional limitations. Med Care 1977; 11: 939-952)
   - want/demand for service, use of services, and/or satisfaction
     (e.g. Batchelor GM, Spitzer WO, Comley AE, Anderson GD. Nurse practitioners in primary care IV. Impact of an interdisciplinary team on attitudes of a rural population. Can Med Assoc J 1975; 113: 1415-1420)
   - ratings of preference for health status
     (e.g. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chron Dis 1978; 31: 967-704)

2. Define the target population
   i.e. the group of persons to which you wish to generalize the results of your survey

3. Define the sampling units
   e.g. patients
   e.g. a household
   e.g. a family

4. Define the sampling frame
   e.g. voting lists
   e.g. census lists
   e.g. city directory
   e.g. telephone directory

5. Specify sampling techniques
   - simple random
   - stratified
   - cluster
   - systematic

6. Develop or select questionnaire

7. Plan field procedures
   - pretesting
   - interviewer training (if interviews)
   - follow-up techniques for non-responders in mail surveys

8. Data collection

9. Data analysis

10. Write report

ii) Advantages and Disadvantages

Advantages

1. As long as response rate 80% or better, best method of estimating true community health and health care problems
2. People themselves (not professionals, politicians, or agencies) provide the health and health care data
3. Flexible approach - can be used to assess wide range of community needs and/or supplement approaches to gathering new health care information outlined previously

Disadvantages

1. Requires expertise and support staff to execute properly
2. Relatively expensive
3. Often time to collect and report data too long leading to outdated report
4. Unless questions are relevant and carefully worded and pretested (if interviewers used, they should be trained in interviewing techniques), quality of information will be poor

Example of General Population Survey

Survey of the Jewish aged population in Hamilton (assessments of physical and social function and wants and use)

Under management of the Regional Service Program at McMaster University, a 50 per cent simple random sample of the Jewish population in Hamilton over the age of 60 was surveyed to assess their social, physical, functional and financial status. The study was designed to also obtain data to help the Hamilton Council of Jewish Organizations (CJO) determine whether there was a need for a residential home for older Jewish residents of the area. The survey results indicated there were not enough Jewish persons in the community who would benefit from the construction of a home. A home was not built, despite political pressure. The survey results also revealed a need for other services for the Jewish Aged population. Since the survey was completed, a meals on wheels service and sheltered workshop have been put into operation.

Vayda E, Baskin M, Goldsmith C, Gordon C. The Jewish aged of Hamilton: reports of a survey. Can J Pub Health 1972; 53: 414-423

Survey of Handicapped Persons in London

A questionnaire designed for use by volunteer interviewers was used in a door to door survey of persons living in a London borough. If, at the door, the interviewer determined their was a handicapped person in the household, the interviewer would go on to complete the questionnaire. The questionnaire consists of sections on (1) description of the handicap, (2) health status measured in terms of a check list of "difficulties experienced" with "getting out and about", "moving around at home", "communication", "home care", "family care", "self care", "special self care need", "self organization", "employment", "education" and "other activities" (each of these categories has a number of specific subcategories) and (3) social and health service requirements as a result of the handicap in terms of a checklist of "has" or "wants" services related to "personal help", "advice on", "regular personal help for", "frequency of help needed", "special centre needed", "equipment needed or its maintenance for mobility, housing and home facilities, communication or personal equipment." (each category has a number of specific subcategories). The survey results have be used planning and identifying health services for the handicapped in London. The Ministry of Health is planning to use a similar checklist questionnaire in a survey of the handicapped in Ontario.

Agerholm M. Outset 1977. London (Mare Street): St. Joseph's Hospice, Heenan House (Residential Rehabilitation and Resource Centre for the Physically Handicapped), 1977.

4. Alternative Tactics For Asking Questions When Obtaining New Information

Alternative Tactics	Comments
-structured (predetermined)questions versus unstructured questions	Structured questions make it possible to statistically analyze the data collected while unstructured (open-ended and/or loose agenda) questions allow other, sometimes important, dimensions of the problem to emerge.
-request information about ideal or necessary minimum in providing services	Depending on the tactic, very different assessments and/or information may emerge
-stressing services needed or health status (levels of physical function, social function and emotional function)	The health status tactic requires the use of experts to interpret the service needs.
-questioning in presence of others versus independent reports	Information provided in presence of others may result in reports of the "others" rather than those of the respondents originally chosen
-anonymous versus labelled (name attached) reply	Given the assurance of anonymity, people may voice opinions they would otherwise not express
-self-selected versus random selection	Volunteers are usually healthier and "health motivated" and therefore differ from patient groups the general population, key information etc.

See also: Woodward C, Chambers LW. Guide to Questionnaire Construction and Question Writing. Hamilton, Department of Clinical Epidemiology and Biostatisitcs, McMaster University, August,1979.

6 FURTHER READINGS

1. Further Readings on "Needs Assessment"

Rossi PH, Freeman HE, Wright SR. Evaluation: a systematic approach. London: Sage Publications, 1979.
- new data sources : 92-105
- existing data sources : 105-115

Spiegel AD, Hyman HH. Basic health planning methods. Germantown: Aspen Systems Corporation, 1978
- new data sources : 23-31; 34-37
- existing data sources : 31-34; 52-79

Bell RA, Warheit GJ, Schwab JJ. Needs assessment: a strategy for structuring change (Chapter 4). In: Coursey RD, ed. Program evaluation for mental health: methods, strategies and participants. New York: Grune and Stratton, 1977
- new data sources : 67-72
- existing data sources : 73-76

2. Selected Readings on Prioritizing Needs

Spiegel AD, Hyman HH. Basic health planning methods. Germantown: Aspen Systems Corporation, 1978
- see chapter 5 "Priority Determination": 179-237

Delbecq AL, Van de Ven AH. The nominal group as a research instrument for exploratory health studies. Am J Public Health 1972: 52: 337-342.

Drake J, McCann P, Adams S, Issacs J. Methods for priority setting in area wide health health planning. Washington: Arthur Young and Company, 1970: 1-40

3. Further Methodological Papers on Existing Data Sources

a) Mortality/Morbidity Data

Woolsey TD. Cause-of-death data. Hyattsville, Md. : National Center for Health Statistics, 1978 (Statistical notes for health planners, no. 6) (DHEW publication no. (PHS) 78-1237).

Klebba JA, Dolman AB. Comparability of mortality statistics for the Seventh and Eighth Revisions International Classification of Diseases. Rockville, Md. : National Center for Health Statistics, 1975 (Vital and health statistics. Series 2 : Data Evaluation and Methods Research, no. 66) (DHEW publication no. (HRA) 76-1340.

b) Utilization/Rates Under Treatment

Roos LL, Roos NP, Nicol P, Johnson C. Using administrative data banks for research and evaluation (Chapter 14). In: Cook T, ed, Evaluation studies review annual: volume 3. London: Sage Publications, 1978.

c) Social Indicators

Sheldon E, Freeman HE. Notes on social indicators: promises and potential. Amsterdam: Elsevier, 1970.

Wilcox LD, Brook RM, Beal GM, Klonglan, GE. Social indicators and societal monitoring. Amsterdam: Elsevier, 1972.

Salley D. Federal clearinghouse inventory. Ottawa: Statistics Canada, 1978 (New Surveys - Notes on statistical survey activity within the federal government. Volume 4, no. 2) (Statistics Canada catologue no. 11-006, quarterly).

4. Further Readings on Survey Research

Statistics Canada. Social concepts directory for statistical surveys. Ottawa: Statistics Canada, 1977 (Special report of the Census Division).

Brown P, Marsden F. Effectiveness and cost of interviewer call backs. Ottawa: Statistics Canada, 1978 (New Surveys-Notes on statistical survey activity within the federal government. Volume 4, no. 1) (Statistics Canada catalogue no. 11-006, quarterly).

Warwick DP, Lininger CA. The sample survey: theory and practice. New York: McGraw-Hill, 1975.

Dillman DA. Mail and telephone surveys: the total design method. New York: John Wiley and Sons, 1978.

187

IV POST-TEST

The Relationship Between Source of Information and Type of Health Care Information Provided.
Indicate (with an 'X') the type of health and health care information usually obtained from each source.

Sources of Health and Health Care Information	WANT ("expected" services)	NEED ("ought to be" services)	DEMAND (want plus info on cost)	USE (service utilization)	SUPPLY quantity (a) quality (b) (includes satisfaction)	
					(a)	(b)
Existing Sources						
1. Mortality/ Morbidity Data (eg. provincial registrars, hospitals, disease registries, national summaries)						
2. Utilization/Rates Under Treatment/ (eg. hospital, clinic, community agencies)						
3. Manpower and Health Care Facilities (eg. beds, public health nurse desks, CAT scanners)						
4. Social Indicators (eg. census, accident, crime, marriages)						
New Sources						
1. Key Informants (eg. community leaders)						
2. Service Agency (eg. hospital, health unit)						
(a) Survey of Agency Staff						
(b) Rates under Treatment						
3. Service Population (eg. patients in hospital, patients visited by public health nurses)						
4. Referral Sources						
5. General Population (people living in a defined geographic area or people with "target" characteristics eg. children, the poor, women, alcoholics)						
(a) Open Forum (sample of convience at a public meeting)					X	
(b) Survey of Citizens (sample on scientific basis)					X	

LECTURE POUR MERCREDI LE 8 JUIN 1983

ADV. NEUROL., 19: 17-42, 1978.

Descriptive Epidemiology

Bruce S. Schoenberg

Section on Epidemiology, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20014

Descriptive epidemiology characterizes observations of disease in human populations. To best illustrate the techniques of descriptive epidemiology and the application of these methods to the investigation of a neurological disorder, let us examine a hypothetical example.

Suppose that a clinician who has been practicing in a small town (Hometown) for 30 years becomes concerned after having seen several cases of neurological disease in the preceding year. His usual experience has taught him to expect only 1 or 2 patients with neurological problems every 2 or 3 years. In systematically analyzing this problem, we shall follow the procedures outlined in Table 1.

TABLE 1. *Procedures for descriptive neuroepidemiology*

I. Define population at risk
II. Define disease under investigation
III. Calculate epidemiologic indices
  A. Characterize disease by
    1. Person
      a. Age
      b. Sex
      c. Race
      d. Occupation
      e. Personal characteristics
    2. Place
      a. Home
      b. Locality
      c. Urban/rural
      d. Geographic differences
    3. Time
      a. Seasonal
      b. Cyclic
      c. Trend
IV. Compare with known indices of disease

DEFINING THE POPULATION AT RISK

The first step in an epidemiologic study is to define the population to be examined, thereby precisely outlining the boundaries of the investigation. These boundaries may be geographic, such as a particular city or state. The population may also be defined by occupation (e.g., workers in a particular industry), by membership in an organization (e.g., members of a professional society), etc. For purposes of our hypothetical example, we shall restrict our investigation to the inhabitants of Hometown. Thus our boundaries are outlined on the basis of geography and place of residence. The characteristics of Hometown are briefly described in Table 2. In Hometown there are 1,000 individuals. When dealing with such a small group, many of the problems inherent in defining the population at risk are greatly reduced. With much larger communities, the task is not as easily accomplished. Counting all the

TABLE 2. *Characteristics of Hometown*

Population = 1,000	
No migration (in or out)	
Industries:	Chemical manufacturing (insecticides)
	Nuclear power plant

inhabitants of a large city is a difficult and expensive task; one must decide how to deal with transient residents or migrants. In a situation in which a large medical center exists in the community of interest, one must consider those individuals who moved to that region because they were ill and sought treatment at a particular hospital or clinic. Just as one may define the population at risk in terms of geographic boundaries, one may also define the group by time interval. For purposes of our hypothetical study, we shall restrict the investigation to the preceding year.

DEFINING THE DISEASE UNDER INVESTIGATION

During the preceding year the physician in Hometown saw 5 individuals with neurological problems. Their case histories are outlined in Table 3. He also carefully checked with other doctors, hospitals, and nursing homes serving the residents of Hometown to be certain he had not missed any other patients with neurological dysfunction. For purposes of discussion, we shall limit our attention to primary intracranial neoplasms. But how does one establish minimal diagnostic criteria for such tumors? How restricted this definition should be depends on the purpose of the investigation and the nature of the disorder. Once established, case definition must be adhered to.

Clinical features, supplemented by the results of various laboratory procedures, are employed to determine the diagnosis. Although it may seem that

TABLE 3. *Summary of patients with neurological signs/symptoms: Hometown*

Patient 1:	A 40-year-old white male auto mechanic who presented with headaches, nausea, and vomiting the preceding December. He was not seen by any physician; he died rather suddenly shortly after the end of the year. An autopsy was refused, and the cause of death remained unknown.
Patient 2:	An 85-year-old retired mailman who had a low-grade intracranial astrocytoma partially removed at craniotomy 3 years earlier. His neurologic symptoms (focal motor seizures) recurred in March. He died in May, and the brain tumor recurrence was demonstrated at autopsy.
Patient 3:	A 65-year-old white female school teacher who in April developed right hemiplegia. She has had numbness of both feet for several years, together with a mildly elevated blood glucose level. An angiogram suggested the presence of a meningioma. The tumor was removed at craniotomy, and she is still living.
Patient 4:	A 75-year-old white male retired carpenter who in preceding March developed progressive difficulty with his speech and an incomplete right hemiparesis. An arteriogram showed a brain tumor. The primary neoplasm of the brain could not be completely removed, and he died shortly after New Year's Day. The histologic type of the tumor was glioblastoma multiforme.
Patient 5:	A 55-year-old white housewife with chronic atrial fibrillation; she developed sudden onset of difficulty with speech and right hemiplegia the preceding December. She did not get worse, and the temporal profile of the neurological dysfunction suggested a vascular etiology.

one should define acceptable cases with criteria so precise that the chance for a wrong positive diagnosis is close to zero, this strategy would exclude patients who were not subjected to intensive clinical investigation. This introduces the adequacy and completeness of health care as one of the determinants of whether or not a patient satisfies certain diagnostic criteria. Usually one must accept some compromises in defining cases, with the realization that a small percentage of individuals without the disease will satisfy the diagnostic criteria.

Before exploring these problems in more detail, we should consider variations that exist when any measurement is made on a population. There are two broad categories of variation: biologic variation and experimental error. Biologic variation refers to true differences in test results that are determined by physiologic or pathologic host factors. Inter-individual differences are the variations between people. Intra-individual differences are variations found when the measurement is repeated on the same individual; these are due to the lability of the individual's responses or physiologic mechanisms.

Experimental error, on the other hand, includes all the sources of variation inherent in the test itself or its interpretation. The accuracy or validity of a test is a measure of how well or closely a test measures the "true" or "actual" value of what it is intended to measure. For example, how well does a laboratory procedure designed to determine the blood glucose level provide a result that is close to the "true" concentration of glucose in the blood? In applying a diagnostic procedure to a given disease, the test should provide an indication of which individuals have the disorder and which do not.

When assessing the accuracy of a diagnostic test, one must consider two

qualities: sensitivity and specificity. Sensitivity is the ability of a test to designate as positive those with the disease. Specificity is the ability of a test to designate as negative those without the disease. Both measurements fall between the limits of 0.0 and 1.0. Table 4 further explains these parameters. The headings "Positive" and "Negative" refer to the results of a given test that may be employed in the diagnosis of a particular disease. The table compares the test results with whether or not the patient actually has the condition or disorder of interest.

The letter A denotes those individuals found positive on the test who have the condition or disorder being studied (i.e., the true positives). The group labeled B includes those who have a positive test result but who do not have the disease (the false positives). Group C includes those with negative test results who in fact have the disease (false negatives). Finally, those with negative test results who do not have the disease are included in group D (true negatives).

With reference to Table 4, sensitivity and specificity can be represented mathematically. The ability of a test to designate as positive those with the disease may be defined as the ratio of the true positives (A) to the number of tested individuals who have the disease (A + C). Therefore, sensitivity = A/(A + C). The ability of a test to designate as negative those without the disease may be represented by the ratio of the true negatives (D) to the total number of those tested who do not have the condition (B + D). Specificity = D/(B + D).

Let us now explore how these concepts apply to the diagnosis of primary intracranial neoplasms. To accomplish this we shall examine two diagnostic procedures currently employed for patients suspected of having a primary brain tumor. (The terms *primary intracranial neoplasm* and *primary brain tumor* will be used interchangeably in this discussion.)

TABLE 4. *Definitions of sensitivity and specificity*

Test result	True diagnoses: Disease present	True diagnoses: Disease absent
Positive	A	B
Negative	C	D
Total	A + C	B + D

True positives = A
False positives = B
False negatives = C
True negatives = D

$$\text{Sensitivity} = \frac{A}{A + C}$$

$$\text{Specificity} = \frac{D}{B + D}$$

The first of these tests is electroencephalography (EEG). Let us assume that we evaluate 360,040 individuals with this procedure and that 40 of these people are later shown, on the basis of more definitive tests, to have primary brain tumors. Using the hypothetical figures in Table 5A, it is possible to calculate the sensitivity and specificity of EEG as a test to diagnose primary intracranial neoplasms. Of the 40 individuals with primary brain tumors, 36 had EEG findings suggestive of such a lesion. In the remaining 4, the EEG was negative. The calculation for sensitivity is illustrated in Table 5A and yields a value of 0.900. On the other half of the table, there were 360,000 additional people who had EEG but who did not have primary brain tumors. Of this group, 54,000 had findings suggesting the presence of a primary brain tumor, while 306,000 did not have any such abnormalities. The calculation for the specificity of this test yields a value of 0.850 (see Table 5A).

The second diagnostic procedure examined in this exercise is computerized tomography (CT scan). Combining X-ray and computer technology, it is a noninvasive way to reveal normal and abnormal brain structures. Let us

TABLE 5. *Calculation of sensitivity and specificity for tests commonly used in the diagnosis of primary intracranial neoplasms (hypothetical results)*

A. Electroencephalography (EEG)

EEG results	Brain tumor	
	Present	Absent
Positive	36	54,000
Negative	4	306,000
Total	40	360,000

$$\text{Sensitivity} = \frac{36}{40} = 0.900$$

$$\text{Specificity} = \frac{306{,}000}{360{,}000} = 0.850$$

B. Computerized tomography (CT)

CT results	Brain tumor	
	Present	Absent
Positive	39	18,000
Negative	1	342,000
Total	40	360,000

$$\text{Sensitivity} = \frac{39}{40} = 0.975$$

$$\text{Specificity} = \frac{342{,}000}{360{,}000} = 0.950$$

assume that the 360,040 individuals who were examined by EEG also had CT scans. On the basis of the hypothetical figures in Table 5B, of the 40 who actually had primary brain tumors, 39 had positive CT scans suggesting the presence of such a lesion. In the much larger group of 360,000 without primary brain tumors, 18,000 had positive CT scans, while 342,000 had negative scans. From the calculations shown in Table 5B, the sensitivity is 0.975 and the specificity is 0.950. On the basis of the hypothetical figures in Table 5, the CT scan is both more sensitive and more specific than EEG for diagnosing primary brain tumors. However, since CT scanning still has limited availability, if we were to define acceptable cases on the basis of a positive scan, we would eliminate a large number of patients for whom this procedure was not available.

In dealing with diagnostic tests that yield a quantitative result (cerebrospinal fluid protein, blood glucose, etc.), it is possible to vary the sensitivity and specificity simply by altering the definition of an abnormal level. An example is the 2-hr postprandial blood glucose level as a measure of diabetes mellitus. As noted earlier in Table 3, patient 3 from Hometown had a reported history of elevated blood glucose levels. Did this patient have diabetes mellitus, which might account for some of her neurologic symptoms?

Figure 1A provides the hypothetical distribution of 2-hr postprandial blood glucose levels in the normal (healthy) and the diseased (diabetic) populations. Note that the two curves overlap. Some patients with diabetes mellitus have lower values than healthy individuals without the disease. Suppose that we choose level A as the upper limit of normal. Then all values above A will be considered abnormal. This provides a very sensitive, but very nonspecific, test. All diabetics will fall above A. Thus the test will be positive for all individuals who have the disease. However, the shaded area of Fig. 1B indicates the large number of individuals in the healthy population who will yield a positive test result, with blood glucose levels above A. The test therefore has low specificity, since a large number of individuals without the disease will be positive.

If we change the definition of positive and extend the upper limit of normal to point B, the situation changes entirely. The test will lose sensitivity but gain specificity. Now the test has maximal specificity. All individuals without the disease have a negative test result. However, as noted in the shaded area of Fig. 1C, a considerable percentage of the diabetic population will also have a negative test result. In this particular example, as one quality is improved, the other is weakened.

Which of these two qualities (i.e., sensitivity and specificity) is more important in epidemiologic studies? No categorical answer can be given. However, it should be noted that if the disorder being measured is rare (as diagrammed in Fig. 1D), the use of a relatively insensitive but highly specific test may be preferable where no confirmatory test is to be done. Otherwise, if a very sensitive nonspecific test were chosen (using point A as the upper

limit of normal), the relatively small number of true positives (area under the diseased population curve) would be seriously diluted by the large number of false positives (area under the normal population curve above point A). When the test is being used for case finding (i.e., to bring cases of the disease under medical care), a greater degree of sensitivity, even at the expense of specificity, is desirable. In this situation, subsequent clinical examination of the suspects can be relied on to rule out the disease in the false positives. There are limits, however, to the tolerable number of false positives. It is always necessary to evaluate the penalties of both kinds of misassignment and decide on the best balance. A useful index in making this

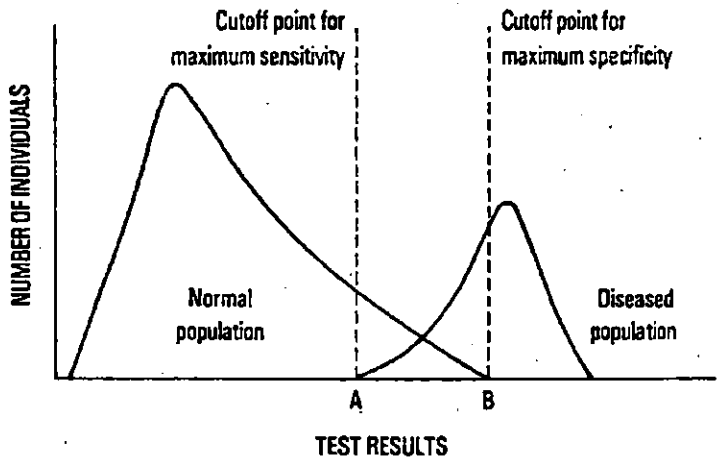


FIG. 1A. Hypothetical distribution of 2-hr postprandial blood glucose levels in normal (healthy) and diseased (diabetic) populations.

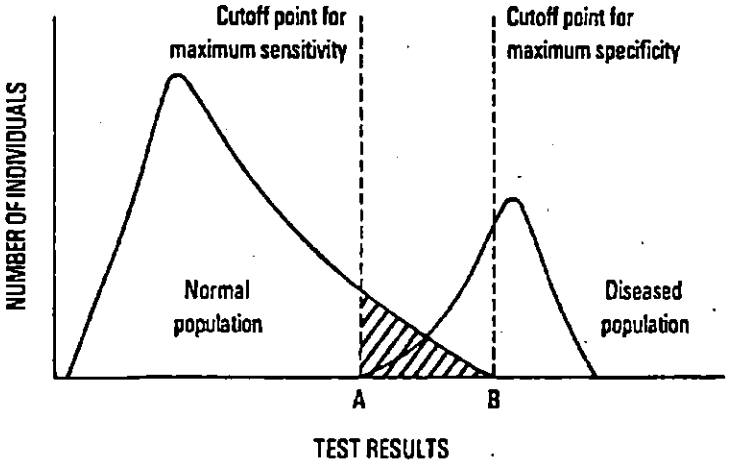


FIG. 1B. Point A is chosen as upper limit of normal. Shaded area of graph indicates false positives (i.e., those with positive test results who do not have the disease).

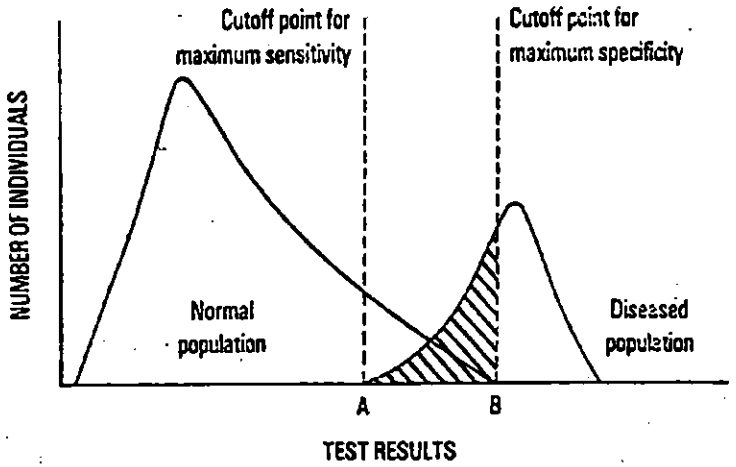


FIG. 1C. Point B is chosen as upper limit of normal. Shaded area of graph indicates false negatives (i.e., those with negative test results who have the disease).

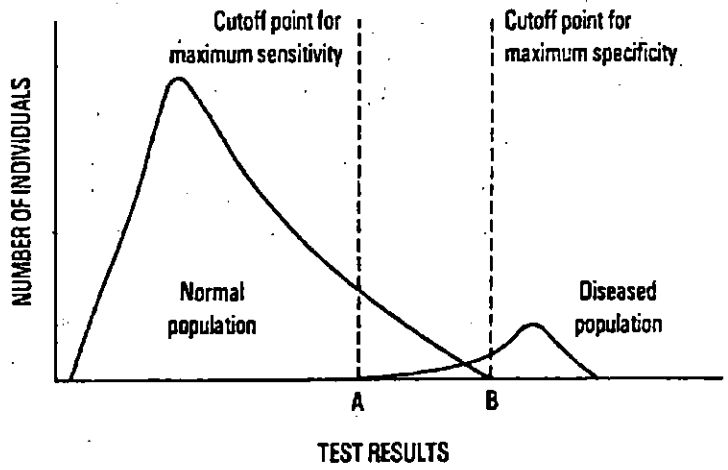


FIG. 1D. Hypothetical distribution of laboratory values in normal (healthy) and diseased populations. In this instance, the disease is relatively rare in the community.

decision is the predictive value of a positive test (1). This index measures the percentage of positive results that are true positives; it is a function of the sensitivity and specificity of the test, as well as the frequency of the disease. The predictive value of a positive test can be calculated from the following formula:

$$\frac{\text{prevalence} \times \text{sensitivity}}{\text{prevalence} \times \text{sensitivity} + (1 - \text{prevalence}) \times (1 - \text{specificity})}$$

Prevalence is an indicator of disease frequency and is precisely defined later in this chapter.

In choosing from among a variety of tests, gain in sensitivity may not be at the expense of specificity. We observed this in comparing CT scanning with EEG in the diagnosis of primary brain tumors.

Every part of a clinical neurologic examination has a certain sensitivity and specificity in diagnosing a particular disease. The experienced clinician goes through a series of decisions based on the significance of hyperreflexia, loss of sensation, etc. The patient's personal characteristics are also important. Thus, diminished vibratory sensation in the lower extremities of a 20-year-old man is very different from the same finding in an 85-year-old man, where it is often seen in the absence of other recognizable neurological symptoms.

Finally, in evaluating a particular diagnostic test one should also consider the problem of precision. This is a measure of the variability or reproducibility of the test itself or its interpretation. For example, this index measures the reproducibility of a result when a laboratory test is repeated on the same tissue specimen or the same patient. It also measures the variability when the same test result (e.g., an X-ray) is interpreted repeatedly either by the same observer or by different observers.

With these considerations in mind, we must now establish minimal diagnostic criteria for primary intracranial neoplasms. For purposes of this discussion, we shall define primary intracranial neoplasms on the basis of confirmatory microscopic examination of a biopsy or autopsy specimen. In the absence of such documentation, angiographic, CT scan, surgical, or autopsy evidence of such a lesion in an individual who is otherwise free of malignant neoplasms, would also constitute an acceptable case. We may now apply these criteria to the five cases in Hometown outlined in Table 3.

With regard to the first patient, although he may have had a brain tumor, the details are too vague and the diagnostic rules are not fulfilled. Patients 2, 3, and 4, however, had definite confirmation of the presence of primary intracranial neoplasms. Patient 5 did not meet our diagnostic criteria and was thought to have had a stroke.

CALCULATING EPIDEMIOLOGIC INDICES

The magnitude of the disease burden in the population is usually expressed in terms of certain epidemiologic indices such as mortality, prevalence, and incidence. To say how many people had or died of a particular disease has little meaning unless one also states how many people were at risk of having or dying of that specific disease. One hundred patients dying of stroke in a town of 1,000 people is quite different than 100 patients dying of stroke in a city of 1 million inhabitants. To adjust for this, the epidemiologist usually

expresses disease magnitude as a rate or ratio, e.g., number of cases per 100,000 population. With such rates or ratios, the frequency of disease (numerator) is related to the population at risk of having disease (denominator). Determining the population at risk is so important that epidemiologists have been referred to as investigators in search of a denominator. In observational or descriptive epidemiologic studies, the disease is further characterized by better defining the persons who develop the disease, the particular places or locales in which the disease occurs, and the changes in the disease over a period of time; e.g., a disease may be most common among elderly white women (person) living in southern cities in the United States (place) during the winter months (time).

In defining acceptable cases and the population at risk, we have the necessary data to calculate the standard epidemiologic indices. Figure 2 provides a temporal profile of the three acceptable cases from Hometown; it will provide a useful guide in the subsequent calculations. Patient 2 had symptoms prior to the year of interest in our study, and he died during the course of that year. Patient 3 began to have symptoms in April, and she is still alive. Patient 4 began to have symptoms in March, and he died shortly after the end of the year. There were no births and only one death in Hometown during the year of interest; there was no migration either into or out of the community. Now we shall apply standard epidemiologic indices to these observations.

Mortality is defined in Table 6, together with the formula for calculation of the mortality rate. There was one death (due to primary brain tumor) during the course of the year. At the beginning of the year, the population of Hometown was 1,000. At the end of the year the population was 999 (after the one death). Since the date of death is known precisely, it is possible to calculate the exact population at risk of dying of primary brain tumors during the course of the year. In most instances it is sufficient to estimate the popu-

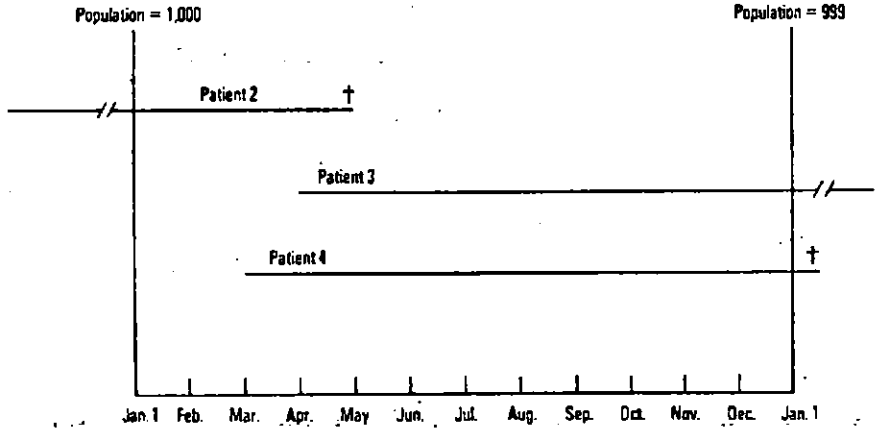


FIG. 2. Temporal profile of the 3 patients from Hometown with primary intracranial neoplasms.

TABLE 6. *Definition and calculation of mortality rate*

Mortality measures the frequency of deaths within a specific population and is calculated for a given time interval and given place. It is often expressed as a mortality rate.

$$\text{Mortality rate} = \frac{\text{Total number of deaths during given time interval in a specified population}}{\text{Average number of individuals at risk of dying during that same time interval in a specified population}}$$

$$\text{Mortality rate (primary brain tumors)} = \frac{1 \text{ (death)}}{\frac{1000 + 999}{2}} = \frac{1}{999.5} = 0.00100 = 100 \text{ deaths (from primary brain tumors) per 100,000 population per year}$$

lation at risk by taking an average of the population at the beginning and end of the interval: $(1{,}000 + 999)/2 = 999.5$. The mortality rate as calculated in Table 6 is 0.00100. It is usually expressed as deaths per 100,000 persons at risk of dying per year. The rate then becomes 100 deaths (from primary brain tumors) per 100,000 population per year.

Morbidity, on the other hand, measures the frequency of illness within a specific population and is calculated for a given time interval and given place. There are two indicators of morbidity: prevalence and incidence.

Table 7 defines both period and point prevalence; it provides examples for the calculation of prevalence ratios for primary brain tumors based on the data from Hometown. The numerator for the calculation of the period prevalence ratio consists of the three cases existing in Hometown during the course of the year. An average population at risk is calculated in the same manner as was done for the mortality rate. The period prevalence is usually expressed in the form of cases per 100,000 population. Point prevalence is much more commonly used. In Table 7 it is calculated for two dates, the first and last days of the year. On the first day of the year there was one brain tumor patient living in Hometown, and the population was 1,000. On the last day of the year there were two brain tumor patients living in Hometown, and the population was 999. The point prevalence ratio is also usually expressed as cases per 100,000 population. Prevalence is particularly useful in indicating the need for medical and social care to cope with the current cases. This index is applied to chronic diseases.

Another measure of morbidity is incidence; it is defined in Table 8. In calculating the incidence rate for primary brain tumors in Hometown, the numerator consists of the two new cases that developed during the course of the year. The average population at risk (denominator) is somewhat more difficult to determine. At the beginning of the year, one person in the population already had a primary brain tumor and was not at risk for developing

TABLE 7. *Definition and calculation of prevalence ratio*

Prevalence measures the frequency of all current cases of disease (existing cases and new cases) within a specific population and is calculated for a given time and given place. There are two types of prevalence:

1. *Period* prevalence measures the frequency of all current cases of disease within a specified population for a given place for a given *time interval*, usually 1 year.
2. *Point* prevalence measures the frequency of all current cases of disease in a specified population for a given place at a *single point in time*, usually on a given day.

$$\text{Period prevalence ratio} * = \frac{\text{Number of persons with the condition during the given time interval in a specified population}}{\text{Average number of persons capable of having the condition during that same time interval in a specified population}}$$

$$\text{Period prevalence ratio (primary brain tumors)} = \frac{3 \text{ cases}}{\frac{1000 + 999}{2}} = \frac{3}{999.5} = 0.00300 = 300 \text{ cases (primary brain tumor) per 100,000 population}$$

$$\text{Point prevalence ratio} * = \frac{\text{Number of persons with the condition on the given day in a specified population}}{\text{Number of persons capable of having the condition on the given day in a specified population}}$$

First day of year:

$$\text{Point prevalence ratio} * \text{ (primary brain tumors)} = \frac{1 \text{ case}}{1000} = 0.00100 = 100 \text{ cases (primary brain tumor) per 100,000 population}$$

Last day of year:

$$\text{Point prevalence ratio} * \text{ (primary brain tumors)} = \frac{2 \text{ cases}}{999} = 0.00200 = 200 \text{ cases (primary brain tumor) per 100,000 population}$$

* CRUDE OR RAW (RATE), POINT QUICK RATIO ! (H.J.)

one for the first time. Thus the population at risk at the beginning of the year was 999. At year's end, 997 persons remained (1 person died and 2 developed primary brain tumors during the year) who were free of primary brain tumors and were at risk for developing them. The average population at risk was $(999 + 997)/2 = 998$. As with other epidemiologic indices, the incidence rate is usually expressed in terms of cases per 100,000 population per year.

Incidence is an important indicator of the need for preventive measures, and it can provide etiologic clues. It is potentially useful for both acute and chronic diseases. It should be noted that the calculation of incidence requires determining when a disease begins. This may be extremely difficult with many neurologic disorders that have an insidious onset (e.g., Alzheimer's disease). Incidence and prevalence, the two measures of morbidity discussed here, are related to each other as follows: prevalence approximately equals incidence

ulti by ; c d n o ase.

194

TABLE 8. *Definition and calculation of incidence rate*

Incidence measures the frequency of addition of *new* cases of a disease within a specific population; it is calculated for a given time interval and a given place.

$$\text{Incidence rate} = \frac{\text{Number of new cases of the disease during the given time interval in a specified population}}{\text{Average number of people at risk of developing the disease during the given time interval in a specified population}}$$

$$\text{Incidence rate (primary brain tumors)} = \frac{2\ (\text{new cases})}{\frac{999 + 997}{2}} = \frac{2}{998} = 0.00200 = 200 \text{ cases (primary brain tumor) per 100,000 population per year}$$

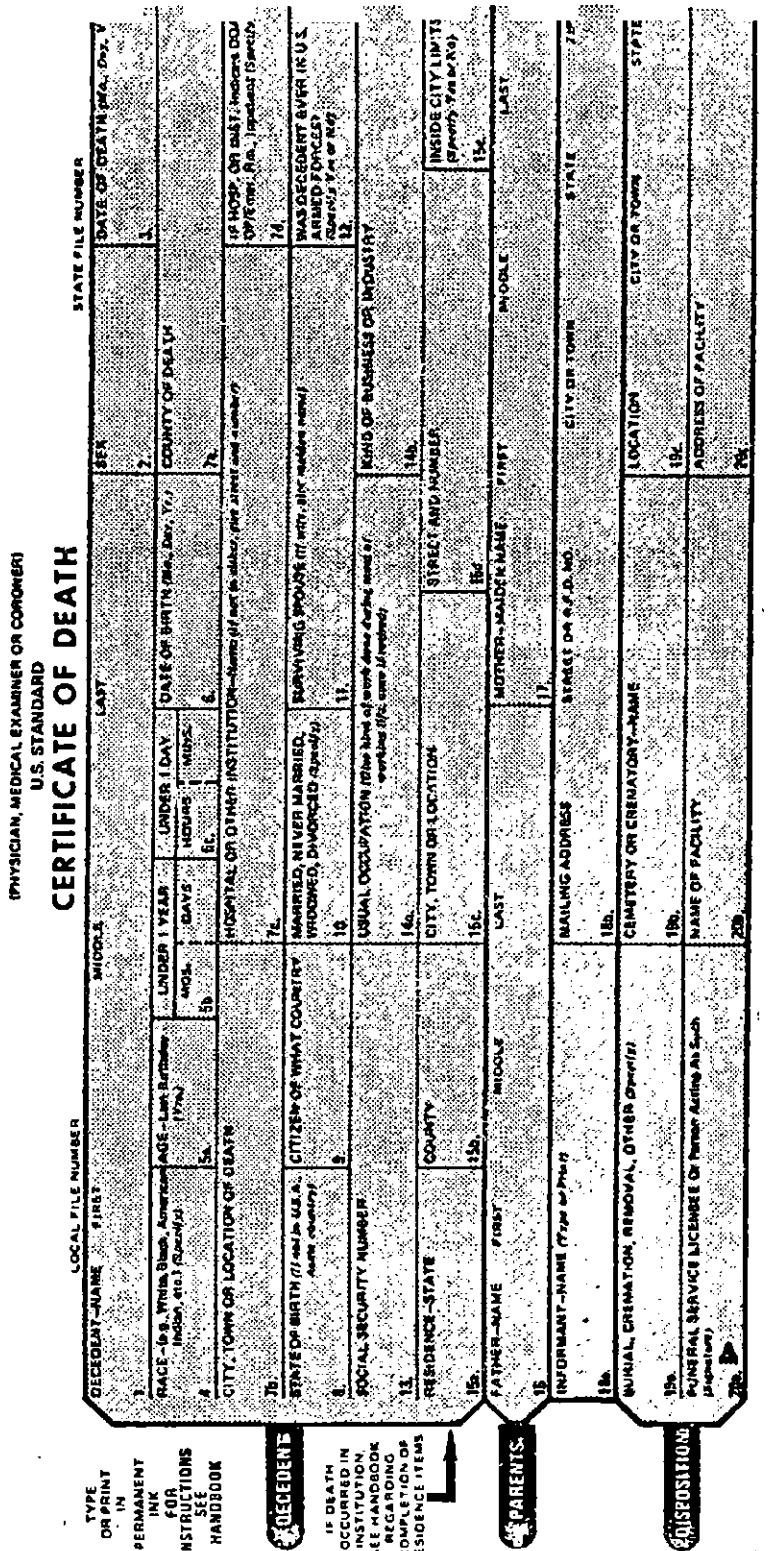
In addition, when calculating these epidemiologic indices, it is often helpful to further characterize the disease by person, place, and time. For example, which particular people in the community are most commonly affected by a certain disease? Are specific ages more commonly affected? Is the disease more common among men or a particular racial or occupational group? What is the geographic distribution of the disease? Is it more common in rural areas or in certain cities? Do the cases seem to cluster along particular streets or in certain homes? Is the disorder more common during certain seasons of the year? Is the disease frequency changing over time? If so, does the change show any periodicity? Are there long-term trends of an increase or decrease in the occurrence of the disease? Answers to questions such as these often provide valuable clues toward a better understanding of the etiology of a given disorder.

COMPARISON WITH KNOWN INDICES OF DISEASE

In order to determine if the results obtained from Hometown represent marked disparities from what might be expected, it is essential to compare the data from Hometown with tabulations from other sources. Before attempting such comparisons, however, it is important to explore the origins of such tabulations.

Mortality data are relatively easy to obtain. Many countries have routine systems for collecting such information in a centralized registry. In the United States, all deaths must be reported to a central agency within each state. Data accumulated in each such state agency are transferred to the National Center for Health Statistics. Each year this information is analyzed, and the resulting tabulations are made available in a standard series of reports.

We must bear in mind that these tabulations are based on information contained on death certificates. Figure 3A illustrates a prototype death certificate.



195

CERTIFIER

To be Completed by CERTIFYING PHYSICIAN Only

21a. To the best of my knowledge, death occurred at the time, date and place and due to the cause(s) stated.

(Signature and Title) ▶

21b. DATE SIGNED (Mo., Day, Yr.)	21c. HOUR OF DEATH M

21d. NAME OF ATTENDING PHYSICIAN IF OTHER THAN CERTIFIER (Type or Print)

To be Completed by MEDICAL EXAMINER or CORONER Only

22a. On the basis of examination and/or investigation, in my opinion death occurred at the time, date and place and due to the cause(s) stated.

(Signature and Title) ▶

22b. DATE SIGNED (Mo., Day, Yr.)	22c. HOUR OF DEATH
22d. PRONOUNCED DEAD (Mo., Day, Yr.)	22e. PRONOUNCED DEAD (Hour) AM M

23. NAME AND ADDRESS OF CERTIFIER (PHYSICIAN, MEDICAL EXAMINER OR CORONER) (Type or Print)

REGISTRAR

24a. (Signature) ▶	24b. DATE RECEIVED BY REGISTRAR (Mo., Day, Yr.)

CONDITIONS IF ANY WHICH GAVE RISE TO IMMEDIATE CAUSE STATING THE UNDERLYING CAUSE LAST

CAUSE OF DEATH

PART I

25. IMMEDIATE CAUSE (ENTER ONLY ONE CAUSE PER LINE FOR (a), (b), AND (c).)

(a) _____ Interval between onset and d

DUE TO, OR AS A CONSEQUENCE OF

(b) _____ Interval between onset and death

DUE TO, OR AS A CONSEQUENCE OF

(c) _____ Interval between onset and d

PART II OTHER SIGNIFICANT CONDITIONS – Conditions contributing to death but not related to cause given in PART I (a)

AUTOPSY (Specify Yes or No) 26.

WAS CASE REFERRED TO MEDICAL EXAMINER OR CORONER (Specify Yes or No) 27.

28a. ACC., SUICIDE, HOM., UNDET., OR PENDING INVEST. (Specify)	28b. DATE OF INJURY (Mo., Day, Yr.)	28c. HOUR OF INJURY M	28d. DESCRIBE HOW INJURY OCCURRED
28e. INJURY AT WORK (Specify Yes or No)	28f. PLACE OF INJURY – At home, farm, street, factory, office building, etc. (Specify)		28g. LOCATION STREET OR R.F.D. No. CITY OR TOWN STA

CONDITIONS IF ANY WHICH GAVE RISE TO IMMEDIATE CAUSE STATING THE UNDERLYING CAUSE LAST

CAUSE OF DEATH

PART I

25. IMMEDIATE CAUSE (ENTER ONLY ONE CAUSE PER LINE FOR (a), (b), AND (c).)

(a) Pneumonia — Interval between onset and death: 2 days

DUE TO, OR AS A CONSEQUENCE OF:

(b) Glioblastoma Multiforme — Interval between onset and de: 3 mos.

DUE TO, OR AS A CONSEQUENCE OF.

(c) _____ Interval between onset and death

PART II OTHER SIGNIFICANT CONDITIONS – Conditions contributing to death but not related to cause given in PART I (a)

Diabetes Mellitus

AUTOPSY (Specify Yes or No) 26.

WAS CASE REFERRED TO MEDICAL EXAMINER OR CORONER (Specify Yes or No) 27.

28a. ACC., SUICIDE, HOM., UNDET., OR PENDING INVEST. (Specify)	28b. DATE OF INJURY (Mo., Day, Yr.)	28c. HOUR OF INJURY M	28d. DESCRIBE HOW INJURY OCCURRED
28e. INJURY AT WORK (Specify Yes or No)	28f. PLACE OF INJURY – At home, farm, street, factory, office building, etc. (Specify)		28g. LOCATION STREET OR R.F.D. No. CITY OR TOWN STATE

We shall concentrate our attention on that area labeled "cause," where the physician lists the various conditions responsible for the patient's demise. The first line is reserved for the immediate cause of death. The second and third lines are for conditions, if any, that led to the immediate cause of death. The instructions indicate that the underlying cause of death is to be listed last, within these first three lines. The fourth line is intended for other significant conditions that may contribute to death but that are not directly related to the disorders appearing on the preceding lines.

Let us now examine how the death certificate is filled out. Suppose the decedent was an 80-year-old white male with a 20-year history of diabetes mellitus. For the preceding 3 months he had had severe headaches, associated with nausea and vomiting. He had also noted loss of strength in his right arm and hand, from which he often dropped objects. He was hospitalized. Tests revealed the presence of a primary intracranial neoplasm—glioblastoma multiforme. He underwent a craniotomy, after which he developed pneumonia; he died 2 days after the operation.

The death certificate would be completed as noted in Fig. 3B. Pneumonia, present for 2 days, would be listed as the immediate cause of death. Glioblastoma multiforme, known to have been present and causing symptoms for 3 months, is entered as the underlying cause of death. Diabetes mellitus could be listed under other significant factors. There seems to be little problem in this instance. But suppose the patient had had a long history of coronary artery disease characterized by angina pectoris. Further assume that he had suffered a myocardial infarction after the operation, then developed pneumonia, and died. Is the glioblastoma the underlying cause of death? Or is it the coronary artery disease manifested by the myocardial infarction? It is only the underlying cause of death that appears in the routine mortality tabulations of most countries. Obviously, the practice of abstracting a single cause of death, when several diseases are present and contribute to the patient's demise, poses many problems for the use of death certificate information. Another difficulty is the accuracy of the diagnoses appearing on the death certificate. Often a physician not previously involved in the care of a patient may be called on to pronounce that patient dead. Lacking adequate medical records, he may get this information from relatives and use this unconfirmed data in filling out the certificate.

Morbidity data are much more difficult to obtain since most countries do not have established systems for collecting this information. Table 9 outlines several potential sources of morbidity data. Many of these categories are reviewed in considerably more detail later in this volume.

Tax-financed plans include medical records from the Veterans Administration medical system, the armed forces medical system, the Medicare and Medicaid programs, etc. Although in some of these record systems it is possible to obtain relatively accurate and complete medical information, the patients in these programs are not representative of the national population.

TABLE 9. *Sources of morbidity data for neurological diseases*

I. Tax-financed plans
II. Ad hoc studies of geographic isolates
III. Special surveys
A. Cross-sectional
B. Longitudinal
IV. Record-linkage systems
V. Prepaid medical plans
VI. Surveillance systems
A. Diseases demanding prompt attention
B. Population-based registries

Ad hoc studies of small geographic isolates have been extremely valuable in neuroepidemiology. Examples are the foci of motor neuron disease in Guam and kuru in New Guinea. Studies of such isolates yield genetic or ecological associations that may not be apparent in larger population groups.

Special surveys, although not conducted on a routine basis, are carried out from time to time to collect information on specific diseases. Most of these are cross-sectional studies, which attempt to collect data on a population at one point in time or a relatively brief period of time. The National Institute of Neurological and Communicative Disorders and Stroke is conducting several cross-sectional surveys on intracranial neoplasms, cerebrovascular disease, multiple sclerosis, and head and spinal cord injuries. These studies have the advantage of employing appropriate statistical techniques in choosing a representative sample of the population. The disadvantages are the very high costs and the difficulty in ascertaining the validity of the diagnoses for the large number of cases included in the sample. A type of special survey that uses the longitudinal approach is represented by the Framingham study of cardiovascular disease. Ongoing since 1950, it consisted initially of careful examination of a group of approximately 5,000 men and women ages 30 to 62. This group has been carefully followed since that time. It has made possible the identification of stroke precursors and has clarified the relationships between vascular disease and hypertension. Senile dementia is now being examined in this group as it has been aging since 1950. As with cross-sectional surveys, the longitudinal approach is also very costly.

Another data resource for the neuroepidemiologist is the record-linkage system in a defined population. Over the course of a lifetime a single individual generates a plethora of medical records. These include prenatal records, delivery records, birth certificate, pediatric records (inpatient, outpatient, house calls, emergency room reports), etc. The fact that all of these records are usually stored in different places makes it extremely difficult to trace this information in order to obtain a complete medical profile of any given individual.

The prepaid medical plans represent another important potential source

of morbidity information. Examples of such medical care systems are the Hospital Insurance Plan of New York and the Kaiser-Permanente system. In such plans, individuals pay fixed fees and obtain their medical care as needed without extra cost. Data from such sources may present difficulties because the participants are not representative of the general population. Patients with chronic conditions or those who are particularly concerned about their health have a much greater incentive to join such programs.

Numerous programs for the routine surveillance of several specific diseases exist within the United States. These surveillance systems are usually for rare or serious diseases requiring immediate attention such as rabies or smallpox. Additionally, a number of registries based on well-defined populations have been established. Most of these deal with one disease or a group of closely related disorders (e.g., cancer or motor neuron disease). Israel has established a registry for all residents of the country who are discharged from an Israeli hospital with a diagnosis of any neurological disorder.

We have explored in some detail the problem of representativeness and diagnostic accuracy in mortality and morbidity data concerning neurological diseases. Another very important consideration is the completeness of case ascertainment for each of the various procedures commonly employed in collecting information for descriptive epidemiologic investigations. Figure 4A provides a schematic diagram of all individuals in a well-defined population who have a specific disease of interest. A certain proportion of these individuals will have symptoms, and the remainder will not. Some percentage of symptomatic patients may not have clinical or laboratory signs of disease. Thus, symptomatic patients with migraine may have no signs of an abnormality. A proportion of the symptomatic as well as a proportion of those without symptoms will have clinical or laboratory evidence of disease. Finally, there will be some asymptomatic individuals who do, in fact, have a particular disease, but who have no signs detectable by routine clinical or laboratory investigations. Therefore an individual without any history of neurologic dysfunction may on autopsy examination have evidence of scattered plaques characteristic of multiple sclerosis. The relative number of patients in any of these categories will vary from one disease to another. In summary, the diseased individuals can be classified into one of four groups:

1. symptomatic with clinical and/or laboratory evidence of disease
2. symptomatic without clinical and/or laboratory evidence of disease
3. asymptomatic with clinical and/or laboratory evidence of disease
4. asymptomatic without clinical and/or laboratory evidence of disease

Figure 4B presents the same type of diagrammatic analysis for the routine sources of mortality statistics. Obviously only those diseases that significantly ibut[illegible] the [illegible]t's [illegible]e a[illegible] [illegible]clude[illegible] ... mo[illegible] ta[illegible]ons. O[illegible]

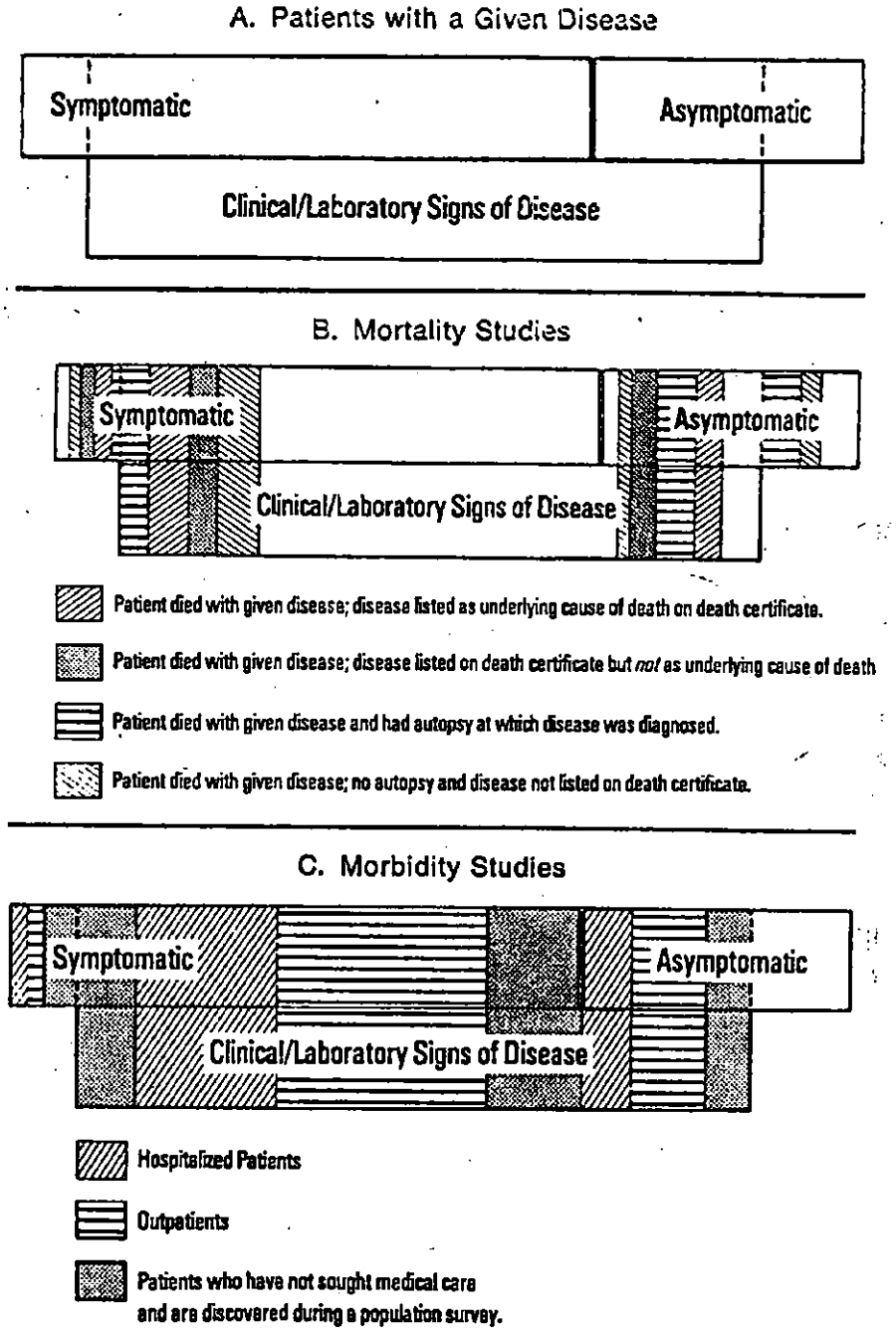


FIG. 4A. Hypothetical distribution of presence or absence of signs and symptoms in all individuals in a well-defined population who have a specific disease of interest. 4B. Hypothetical distribution of deaths in same population of patients. 4C. Hypothetical distribution of type of medical care (hospi[illegible] in ver[illegible] o[illegible] t ca[illegible] lined [illegible] vidu[illegible] me p[illegible] n of [illegible] s.

diseases present at death, those that are acute and rapidly fatal are most likely to be listed on the death certificate. The proportion of patients who die with a given disease of interest are noted by the shaded areas of Fig. 4B. Of these patients who die with the particular disease present, only those in whom the disease is listed as the underlying cause of death will appear in the routine mortality tabulations of most countries. For those instances in which the disease appears on the death certificate in some status other than the underlying cause of death, special tabulations of all death certificate diagnoses are required to reveal the magnitude of these disorders as a force of mortality.

Cases that are first discovered at autopsy are also indicated on the chart. Autopsy studies provide one of the few available methods for detecting the presence of disease in category 4—the asymptomatic patient without clinical or laboratory evidence of disease. Without the ideal situation of performing necropsy examinations on all people who die in a community, the exact magnitude of this fourth category will remain unknown. Note that autopsy studies represent a varied sample from each of the four classes of patients listed above. The results of such necropsy investigations may not be at all representative of all those individuals with the specific disease of interest. Only if the results of autopsy examinations are taken into account in certifying the cause of death can cases first diagnosed at autopsy be reflected in the routine or special mortality tabulations. Finally, any patient dying with a given disease, for whom (a) there is no autopsy and (b) the disorder is not listed on the death certificate, will not be included in any mortality statistics for that disease. The relative number of cases in each of the several categories in Fig. 4B will vary according to the disorder and the medical and death certification practices of the community.

Figure 4C presents the same type of analysis for morbidity studies. By surveying hospitals, the investigator will detect only those patients admitted to such institutions. If the study also includes outpatient facilities a certain additional number of affected individuals will be brought to the attention of the neuroepidemiologist. It is only with a complete medical survey of the entire community (not relying on existing medical records) that one can detect affected individuals who have not sought medical care. However, such individuals must have either symptoms or clinical/laboratory evidence of the disease. Diseased persons who are asymptomatic and have no clinical or laboratory evidence of disease will be missed. Ideally, the complete medical survey (screening each member of the population for the presence of disease) would pick up all affected individuals with symptoms or signs of the disorder. Because of the high cost of this approach and the nature of current screening techniques, this theoretical ideal is never achieved. Surveys of hospital facilities are usually least costly, but they may miss a significant proportion of affected individuals. With prior knowledge of the medical care practices for patients with a given disease, the neuroepidemiologist attempts

to design a study that will reveal at the lowest possible cost a very high percentage of the diseased individuals in the population.

Now that we have carefully considered other sources of mortality and morbidity statistics, we shall turn our attention back to the problem in Hometown. We may wish to compare the mortality experience in Hometown with death data for primary brain tumors in the United States and other countries worldwide (3). Most of the mortality rates for primary intracranial neoplasms fall within the range of 5 to 10 deaths per 100,000 population per year. These rates appear to be considerably smaller than the rate of 100 deaths per 100,000 population per year in Hometown. Before forming any conclusions with regard to this apparent discrepancy, there are a number of other important considerations that must be taken into account. We must first make sure that the definition of primary intracranial neoplasms used in tabulating the data from Hometown is equivalent to that used for the other sources of information. This is difficult to do for mortality data, since the diagnoses on the death certificate are accepted regardless of the evidence supporting such diagnoses. Unfortunately, a similar problem frequently exists with morbidity statistics as well. The majority of investigators do not state their precise diagnostic criteria. This makes it extremely difficult to make comparisons between studies. If all potential patients in one community were seen by a neurologist and had CT scans or cerebral angiograms, this would tend to increase the accuracy of the diagnosis. Patients in another community may not have the benefit of such medical expertise or advanced technology. But just as we are concerned about comparability of the numerators (criteria for identification of cases), we must also be alert to comparability of the denominators. Is the population structure of Hometown similar to the population structure of other areas yielding mortality statistics? Why is this so important?

If the mortality rates for primary brain tumors are analyzed according to the age of the patient, it becomes obvious that the risk of dying from a primary brain tumor varies greatly as a function of the individual's age. Death rates for primary brain tumors are much higher in the older age groups than among the young. Rates are also higher for males than for females and higher for whites than for nonwhites. If the population of Hometown consisted primarily of nonwhite females in their early twenties, the mortality rate for primary brain tumors would be expected to be much lower than if elderly white males accounted for the majority of the inhabitants.

To correct for possible differences, the epidemiologist uses specific rates. The mortality rate we generated for Hometown was a cause-specific rate, since it was limited to one cause of death, i.e., primary brain tumors. We may make the rates even more specific by restricting them to a given race, sex, and age group. Thus we can have the number of primary brain tumor deaths per 100,000 population per year for white males for the ages 0 through 4. Both the numerator and denominator are confined to a particular

661

age, sex, and race, and the resulting figure is an age-, sex-, and race-specific mortality rate. When dealing with relatively small populations, it becomes impractical to subcategorize the rates into many smaller groups, since one is left with only 0 or 1 case in each small category.

Rates that are not restricted to particular subgroups and include all deaths are known as crude mortality rates. Such crude rates have two major disadvantages: (a) they oversummarize complex patterns of rates; (b) they lack comparability for communities with populations that differ by age, sex, race, etc.

It is possible to adjust this crude rate to account for important differences in the population. The result is an adjusted or standardized rate, and the adjustment can be made for age, sex, race, etc. Thus one can generate an age-, sex-, and race-adjusted mortality rate. Similar considerations also apply to morbidity data.

The mortality rate calculated for Hometown is an exact rate. To the limits of our diagnostic abilities we identified all deaths in Hometown due to primary brain tumor. If we confine our comments concerning this rate to the experience of Hometown, there is no need to consider the problem of sampling variability. On the other hand, one might choose to regard Hometown as a random sample of some larger unknown population. If one chooses to accept this assumption, then one might use the mortality rate for Hometown as an estimate of the mortality rate in the larger population. Since the Hometown rate is based on only one death, the confidence interval associated with using this figure as an estimate for the rate in the larger population is quite large. Thus the rate for Hometown is 100 primary brain tumor deaths per 100,000 population per year; the 95% confidence interval for the rate in the larger population is 3 to 557 primary brain tumor deaths per 100,000 population per year. In general, the smaller the sample size, the greater will be the size of the confidence interval surrounding the sample estimate of the rate in the larger population. These comments are applicable to morbidity statistics as well.

COHORT ANALYSIS IN DESCRIPTIVE EPIDEMIOLOGY

Thus far in our discussions of descriptive epidemiology we have examined health status in discrete populations over rather limited specified periods of time (1 day in the case of point prevalence, and 1 year in the case of mortality or incidence rates). Such studies are often referred to as cross-sectional, since they provide a picture of disease in a population at one relatively brief moment in time. For investigations aimed at determining the natural history of disease and survival, it is necessary to use a longitudinal approach. One identifies a group of patients after the onset of disease and follows them over a period of time in order to observe the outcome.

Theoretically one could apply a similar approach to all individuals born

at the same time in a given population. The group of individuals born in the same year or narrow group of years is referred to as a birth cohort. The word "cohort" derives from a division within a Roman legion. Just as members of such a legion marched together, so members of a particular birth cohort grow old together and are subjected to different diseases or environmental factors existing in different calendar years. Thus one can observe these cohorts over a period of time to observe their lifelong experience with a particular disease. This strategy is quite difficult to implement in practice, and it involves an inordinate expenditure of time and money.

If one carries out repeated cross-sectional surveys of the same population at regular intervals, one can reconstruct such data to approximate the birth cohort pattern of analysis. In using cross-sectional data collected over extended periods of time, one must exercise caution to be certain that diagnostic methods and disease classification have not changed. In any given calendar year members of different age groups represent members of different birth cohorts. Table 10 illustrates the assignment of a mortality or incidence rate for a certain age group (20–24 years of age) and period (1930–1934) to a particular birth cohort (1905–1914).

Thus the mortality rate in 1930 for those 21 to 22 years of age belongs to the 1908 birth cohort. As the cohort grows older, it moves along the scale of both calendar time and age. It should be emphasized that this strategy only approximates a birth cohort analysis because of migration into and out of the population. For example, someone who is 21 years old in 1930 may have just moved into the community 1 year earlier and may not have been subjected to the same risks as a 21-year-old who has lived in that area throughout his lifetime.

TABLE 10. *Assignment of a mortality or incidence rate for individuals of a certain age (20–24 years) and period (1930–1934) to a particular birth cohort (1905–1914)*[a]

Birth cohort (year of birth)	Year of observation or death				
	1930	1931	1932	1933	1934
1905	24–				
1906	23–24	24–			
1907	22–23	23–24	24–		
1908	21–22	22–23	23–24	24–	
1909	20–21	21–22	22–23	23–24	24–
1910	–20	20–21	21–22	22–23	23–24
1911		–20	20–21	21–22	22–23
1912			–20	20–21	21–22
1913				–20	20–21
1914					–20

[a] For a given year of birth and year of observation (or death), exact age at observation (or death) depends on whether the day of observation (or death) occurred before or after the birthday during that year. Thus persons born in 1908 and dying in 1930 would either [illegible] 24 y[illegible] at d[illegible]

The birth cohort study design is useful in that factors responsible for the occurrence of disease in a given calendar year may have different effects on each age group, depending on that group's past experience. Thus certain age groups may be immune to an infectious agent present during a given year because of prior exposure and resulting immunity to this same agent.

The reasoning behind the birth cohort approach was used informally in an analysis of Parkinson's disease patients admitted to a given medical institution between 1875 and 1961 (4). The data were tabulated by age at onset of Parkinson's disease. Since the late 1920s the mean age of onset of new cases of Parkinson's disease admitted to this given hospital over a 35-year period increased by almost 30 years. This increasing age at onset corresponded remarkably to the cohort who were 5 to 59 years of age in 1920. This is the same cohort in which clinical von Economo's encephalitis occurred in the epidemics between 1918 and 1926. The investigators concluded that most cases of Parkinson's disease have a single etiology, possibly related to subclinical infection with von Economo's encephalitis and resulting in the occurrence of Parkinson's disease up to 40 years later. If this hypothesis is correct, the incidence of Parkinson's disease should decline as this cohort ages and dies. Although another study based on a well-defined population followed over 30 years has failed to document an increasing age at onset for Parkinson's disease (2), this example demonstrates the usefulness of the birth cohort type of analysis.

The birth cohort approach was formally applied to an analysis of incidence rates for amyotrophic lateral sclerosis on Guam (5). Figures 5A and 5B present these results and illustrate the two procedures for graphing data analyzed according to birth cohort. In Fig. 5A, incidence rates for each

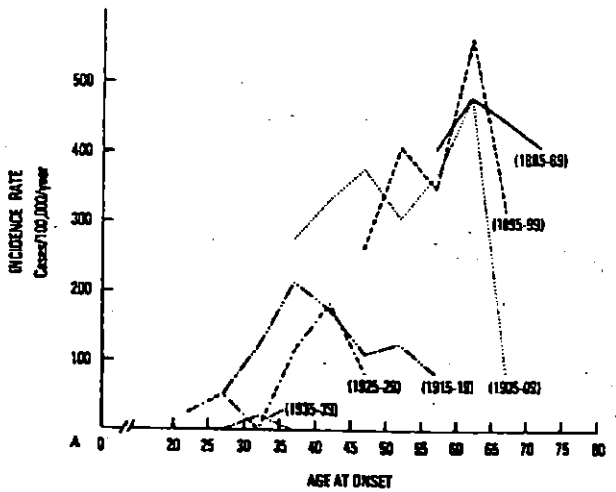


FIG. 5A. Average annual incidence rates for males in separate birth cohorts as a function of age. For clarity, graph does not show all birth cohorts for which data are available. Tabulations are from Guam (5) for amyotrophic lateral sclerosis.

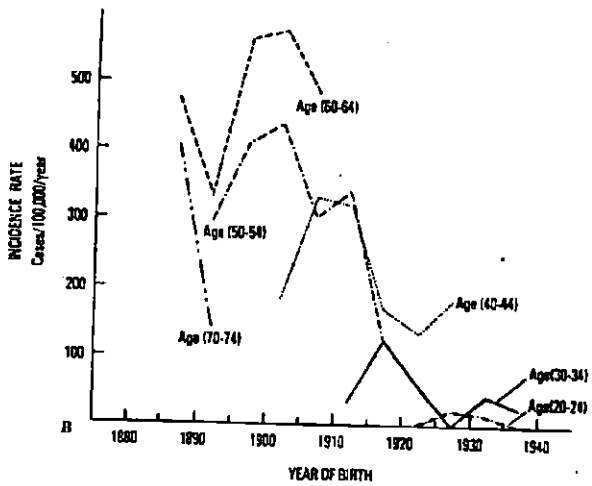


FIG. 5B. Average annual incidence rates for males in separate age groups as a function of year of birth. For clarity, graph does not show all age groups for which data are available. Tabulations are from Guam (5) for amyotrophic lateral sclerosis.

separate birth cohort are plotted as a function of age. In Fig. 5B, incidence rates for each separate age group are plotted as a function of birth cohort. In either instance, a general trend is observed: for males of the same age, those born in more recent years have a lower rate of amyotrophic lateral sclerosis. The data available for Guam are extremely valuable because it is unusual to find information on neurological disease in a well-defined population over a large number of years.

On the basis of observations of a well-defined population, as outlined in this chapter, it is possible to formulate hypotheses that can be tested using the techniques of analytic epidemiology.

REFERENCES

1. Galen, R. S. (1977): *J.A.M.A.*, 238:2466.
2. Kurland, L. T., Hauser, W. A., Okazaki, H., and Nobrega, F. T. (1969): In: *Proceedings 3rd Symposium on Parkinsonism*, pp. 12–15. Livingstone, Edinburgh.
3. Massey, E. W., Schoenberg, D. G., and Schoenberg, B. S. (1978): (*submitted for publication*).
4. Poskanzer, D. C., Schwab, R. S., and Fraser, D. W. (1969): In: *Progress in Neurogenetics*. Excerpta Medica International Congress Series. 175:497–505.
5. Reed, D. M., and Brody, J. A. (1975): *Am. J. Epidemiol.*, 101:287–301.

METH. INFORM. MED., VOL 17, N°1, 1978.

202

roblems of Medical Diagnosis

A Bibliography —

UDC 616—07 : 01

G. Wagner, P. Tautu, U. Wolber

The present bibliography aims at informing and stimulating researchers when studying the multiple facets of medical diagnosis. The complexity of investigation in this domain can be judged by the fact that papers dealing with medical diagnosis do not appear in medical journals only. Clearly, medical diagnosis is no longer a strictly medical problem. Mathematical and statistical models and analyses must be created and improved before computer programs for diagnostics can be elaborated.

A bibliography in such a complex and fast growing field of science cannot be exhaustive. Moreover, we have deliberately not considered short notes, abstracts and editorials. The bibliography given here starts with the few papers published about twenty years ago and ends in November 1977. It contains 827 titles and is divided up into two parts: an alphabetically arranged list of papers and a subject classification. The latter is arranged according to the following headings:

A. General Aspects

A1. Theories of medical diagnosis, surveys, analyses
A2. Computer approaches (e.g. computer programs, diagnosis simulation, artificial intelligence, computer linguistics)

B. Logical and Mathematical Aspects

B1. Logical problems (e.g. 2-or many-valued logic)
B2. Inductive logic (personals probabilities, utilities)
B3. Algebraic aspects (e.g. graph theory, pattern recognition, formal languages)
B4. Fuzzy sets
B5. Game theory
B6. Information theory and questionnaires
B7. Stochastic representations

C. Statistical Aspects

C1. Bayes theorem
C2. Discrimination and classification
C3. Stepwise algorithms
C4. Statistical decision analysis
C5. Error rates
C6. Other statistical approaches (e.g. likelihood ratio, Neyman's critical quotient, evaluation of computer programs)

D. Medical Applications

D1. Infectious and communicable diseases
D2. Diseases of the cardiovascular system
D3. Diseases of the respiratory system
D4. Diseases of the digestive tract
D5. Metabolic and nutritional diseases
D6. Diseases of the endocrine and reproductive system
D7. Diseases of the urinary tract
D8. Diseases of the hematopoietic and immune system
D9. Diseases of the nervous and muscular system
D10. Diseases of the skeletal system, skin and connective tissues
D11. Diseases of the eye, ear, nose, and throat
D12. Behavioral and mental disorders
D13. Injuries, occupational diseases, poisoning
D14. Neonatal diseases and abnormalities
D15. Multiple diseases studies

Some collateral branches were deliberately left out, as, for example, normal values in medicine, data acquisition and reduction, reliability of laboratory methods, etc. Although the domain of this bibliography is the clinical diagnosis, some hints about chemical and cell diagnosis were also included. For reasons of simplification, the subject classification is not finely detailed. However, its study will give a general idea not only about the diagnostic process as a multifaceted domain but also on the lacunae still existing in this field.

HEALTH SCIENCES LIBRARY

(à paraître dans la collection Mercure du Centre Canadien d'Etudes sur la Culture Traditionnelle, Ottawa).

LE DISCOURS SUR LA SANTE ET LA MALADIE DANS DEUX POPULATIONS DE QUEBEC

Par: GILLES BIBEAU
LOUISE PELLETIER

Note: Les auteurs remercient Brigitte St-Pierre et Donald Monty, étudiants terminant leur baccalauréat en Sciences infirmières, pour leur participation à toutes les étapes de la recherche.

Le 23 septembre 1980.

Au début de son étude sur les croyances actuelles des paysans du Bocage français dans la sorcellerie, J. Favret-Saada (1977) écrit que le paysan a été considéré jusqu'ici par les scientifiques comme "l'image inversée du savant ". La croyance persistante en la sorcellerie serait le signe d'une certaine incohérence intellectuelle et d'une mauvaise manipulation des relations de causalité: tel aurait été, d'après Favret-Saada, le crédo des chercheurs spécialistes de la mentalité populaire depuis Van Gennep (Belmont, 1974). Quant à elle, Fravet-Saada refuse d'emblée de considérer le paysan traditionnel comme un 'ignorant' et préfère chercher la logique interne qui organise ses conceptions et pratiques liées à la sorcellerie.

Nous avons inscrit nos propres recherches sur le savoir médical de deux populations du Québec Métropolitain dans la même orientation théorique et nous présentons dans la première partie de cet article les principales caractéristiques de ce savoir en nous inspirant principalement des travaux actuels des théoriciens des sciences. Cette orientation prolonge également la réflexion présentée par un des auteurs dans sa thèse doctorale (Bibeau 1979).

Il y a un lien direct entre la conception que le chercheur se fait du contenu et de l'organisation propre au savoir populaire et la méthode qu'il emploie pour rejoindre ce savoir. Nous pensons qu'on ne peut véritablement atteindre l'ensemble des connaissances détenues par un groupe qu'à travers le concours de plusieurs instruments complémentaires de collecte d'informations. Nous présentons justement dans notre deuxième partie les instruments que nous avons utilisés.

La troisième partie discutera les résultats préliminaires de notre enquête en s'attardant aux éléments qui entrent dans la construction bio-culturelle de la santé et de la maladie, telle que cette construction se présente dans deux catégories sociales de Québec. Nous montrerons qu'il est impossible de réfléchir sur l'organisation de cette construction sans présenter

les théories étiologiques, les conceptions du fonctionnement du corps et la catégorisation des maladies telles qu'elles ont cours dans ces groupes. Dans cet article, nous ferons cependant davantage porter l'analyse sur la délimitation de l'espace occupé par la santé et la maladie.

1. Le statut du savoir médical populaire

Nous n'entendons nullement le savoir médical de la population dans le sens de Rosenblatt et Suchman (1965) qui ont mesuré la quantité et la qualité des informations médicales détenues par les cols bleus américains. Dans ce genre de recherches, les auteurs s'intéressent principalement au transfert du savoir des experts médicaux vers les profanes dont le savoir antérieur n'est le plus souvent pas pris en considération. Bien qu'une telle approche n'implique pas nécessairement que les chercheurs ignorent l'existence d'une organisation spécifique du savoir médical dans les diverses catégories sociales, leurs considérations analytiques se limitent toutefois à mettre en relief les carences et les déficits de ce savoir populaire en le comparant à celui des spécialistes.

L'effort de reconstitution des savoirs médicaux détenus par la population des profanes s'est fait dans plusieurs directions depuis une vingtaine d'années. Rosenstock (1966; 1978) par exemple, qui s'inscrit de plain-pied dans la tradition nord-américaine de recherche sur les 'beliefs' (Fishbein et Ajzen, 1975) a proposé d'organiser l'ensemble des connaissances médicales d'un groupe social sous le nom de 'health belief model'. Cette tradition américaine a également donné une place importante aux recherches sur les attitudes différentielles des groupes face à la santé et à la maladie; dans le même contexte, plusieurs auteurs ont utilisé la notion de valeur comme cadre conceptuel. Cette approche par les 'croyances, attitudes et valeurs' est encore des plus

actives et elle a été récemment utilisée chez nous par le docteur Saucier dans sa recherche sur les conceptions de santé chez les adolescents de Montréal (1979). Les instruments généralement utilisés dans cette approche sont des questionnaires fermés complétés par des entrevues pour les 'beliefs', des échelles d'attitudes et le 'values clarification'.

Chez les chercheurs européens, le concept heuristique de 'représentation sociale' proposé par Durkheim a été préféré aux concepts de croyances de conceptions ou de système de pensée. Moscovici (1961) fut à l'origine du renouveau de popularité de ce concept par son étude sur la représentation que les Français se faisaient de la psychanalyse; il définit la représentation sociale de la manière suivante: "c'est un système cognitif qui a une logique et un langage particulier, une structure d'implication qui portent autant sur des valeurs que sur des concepts, un style de discours qui lui est propre" (Preface Herzlich, page 10, 1969). Ce concept de représentation veut mettre en évidence l'entreprise de reconstitution, de remodelage ou de construction que tout groupe fait subir aux éléments de son environnement; de plus, comme le dit l'expression,ce travail d'interprétation se fait de manière collective ou sociale. Cette approche explore le niveau cognitif et tout en insistant sur le fait que le conceptuel s'organise de manière cohérente dans les différents groupes sociaux, elle ne fournit cependant pas aux chercheurs les instruments susceptibles d'organiser ce savoir selon le point de vue des groupes étudiés.

C'est d'ailleurs sur le plan de la mise en ordre du savoir médical détenu par deux catégories sociales de la ville de Paris que l'excellent ouvrage de Herzlich (1969) trouve sa limite: les conceptions populaires y sont restituées mais elles ne sont pas insérées dans le cadre général de la science médicale propre à chacun des groupes. Pour étudier la représentation sociale de la santé et de la maladie, les chercheurs emploient

le plus souvent un guide d'entrevue qui explore successivement une série de thèmes particuliers. Les questionnaires fermés ainsi que les échelles de type Lykert sont considérés comme inadéquats pour ce genre d'études; certains chercheurs utilisent des tests projectifs lesquels sont susceptibles de faire émerger la configuration conceptuelle et symbolique.

De nombreux auteurs américains, spécialement des sociologues médicaux, se situent depuis quelques années dans la perspective française: c'est le cas de Freidson (1971) qui a étudié la construction sociale de la maladie, et de Mechanic (1978) qui a établi un contraste entre les maladies 'self-defined' et 'other-defined'. Ces deux auteurs situent le savoir médical d'une catégorie sociale par rapport à celui des spécialistes médicaux; leurs études sont nécessaires et dignes de valeur mais elles n'ont pas pu conduire à la mise en évidence du caractère original de la science médicale partagée par les différents groupes sociaux et par la population en général (1). A travers l'utilisation d'expressions comme construction sociale, définition sociale ou représentation sociale de la santé et de la maladie, il ne fait pas de doute que les sociologues médicaux veulent implicitement signifier que ce corpus de connaissances établit un ordre dans les faits, les met en interrelation et leur donne un sens. C'est là un net progrès mais ces chercheurs ne se sont pas donné les instruments qui leur auraient permis de dégager cet ordonnancement interne des connaissances médicales dans les différents groupes sociaux. Plusieurs sociologues médicaux se refusent à continuer à mesurer le savoir médical populaire à l'aune de celui des professionnels de la santé et ils s'interrogent sur la voie à suivre dans leurs recherches qualitatives. Le sociologue britannique Dingwall (1976) n'a pas craint d'affirmer que l'impasse actuelle ne peut être évitée que par le recours à l'ethnoscience et à l'ethnomédecine.

En ce qui concerne l'organisation interne des différents éléments qui entrent dans la construction culturelle ou la représentation sociale de la

maladie, nous sommes en face de deux grandes approches. Il y a tout d'abord le courant de l'ethnoscience qui présuppose que l'organisation des connaissances médicales de tout groupe social se présente sous la forme d'une science qui a des caractéristiques spécifiques. Face à cette perspective théorique qui accorde d'emblée le statut scientifique aux systèmes traditionnels d'interprétation des réalités physiques et sociales se dresse une seconde approche que l'on peut globalement qualifier de symbolique: les partisans de cette approche ont une orientation linguistique explicite et le vocabulaire qu'ils manipulent est fortement inspiré de la linguistique comme le démontrent les termes suivants: signe, symbole, signifiant, signifié, métaphore, métonymie, etc.

Ces deux perspectives ne doivent pas être considérées comme antithétiques puisqu'elles saisissent la santé et la maladie sous deux angles différents: les utilisateurs du symbolisme, de la sémantique et de la sémiotique considèrent généralement les réalités de santé et de maladie comme un langage et un discours dont il importe de découvrir les codes de fonctionnement alors que les partisans de l'ethnoscience ont davantage tendance à rechercher l'organisation interne d'un ensemble conceptuel qui est d'emblée appréhendé comme une science. Les deux perspectives sont fondamentalement orientées vers la compréhension des principes organisationnels qui intègrent les représentations en un ensemble cohérent lequel donne sens à chacun des éléments. Nous sommes d'avis qu'il faille nécessairement employer cette double perspective pour restituer la totalité de la construction bio-culturelle de la santé et de la maladie dans une société.

Afin de mieux préciser d'une part le statut du savoir médical populaire et d'autre part la méthode que nous avons utilisée dans l'enquête exploratoire menée dans deux populations du Québec métropolitain au cours de l'été 1980, nous croyons utile d'envisager le problème du savoir populaire de trois points de vue différents: celui des théoriciens des

sciences, celui des anthropologues médicaux et enfin celui des épidémiologues. Cette présentation sommaire des problèmatiques qui ont cours dans chacun de ces domaines nous permettra de mieux préciser le cadre théorique à l'intérieur duquel nous situons le savoir médical populaire ainsi que les caractéristiques de notre propre méthode d'investigation.

Le point de vue qui nous semble le plus déterminant dans la réorientation méthodologique actuelle est celui des théoriciens, historiens et philosophes des sciences. Dans son étude de la structure des révolutions scientifiques, Kuhn a clairement démontré, entre autres choses, que la science fonctionne comme un système social: il a identifié les processus sociaux de conservation, de transmission et d'innovation à l'oeuvre dans la science académique ou plus précisement encore dans les communautés scientifiques (1970). Campbell soutient de son côté que les systèmes de perpétuation et de maintien du savoir dans une communauté scientifique présentent des similitudes fondamentales avec les 'self-perpetuating maintenance systems' à l'oeuvre dans le domaine des croyances et de la superstition (1979). Tout groupe dépositaire d'un savoir doit donc rencontrer, selon Campbell, des exigences tribales fondamentales de cohésion de groupe, d'identité collective et d'adhésion aux systèmes traditionnels de validation, de préservation et de transmission de ce savoir.

Si nous appliquons cette conception des processus sociaux de la science non plus à la science des communautés académiques comme l'ont fait Kuhn et Campbell, mais à celle qui est détenue par la population en général ou par des sous-groupes de population, nous sommes conduits à deux constatations qu'on ne peut continuer d'ignorer. Tout d'abord, il est évident qu'on ne peut pas envisager la société globale contemporaine dans nos pays comme partageant les caractéristiques tribales présentées plus haut, comme cela se fait encore trop souvent dans des travaux du genre " Croyances et pratiques populaires au Canada Français" (Des Ruisseaux 1973).

Sur le plan méthodologique, il apparaît fondamental de se limiter à recencer la science des sous-groupes dans lesquels les 'social maintenance systems' sont les plus actifs: parmi ceux-ci, on peut penser à une catégorie professionnelle comme les forgerons ou les sages-femmes, à un groupe bien délimité dans un territoire comme les Iles de la Madeleine, à un groupement idéologique comme le sont les Bérêts Blancs, à une classe d'âge comme les adolescents d'une école secondaire, ou encore à un village de type traditionnel. Nous plaidons donc en faveur de la rédaction de monographies portant sur des groupements clairement identifiables sur le plan des processus sociaux de transmission des croyances et du savoir, bien que nous soyions conscients du fait qu'il n'existe actuellement aucun regroupement qui soit à l'abri d'une contamination de ses valeurs internes par celles de la société au sens large.

De ce point de vue, il faut donc viser à écrire des monographies limitées. Il ne faut cependant pas oublier de présenter la science médicale telle que partagée par les différentes catégories sociales. Le découpage le plus fréquent que l'on opère dans nos sociétés, se fait en effet par le biais du statut socio-économique et les classes sociales constituent pour les chercheurs le principal lieu d'études des différences qui surgissent au sein de la société globale. Spécialement en milieu urbain, les investigateurs n'ont pas le choix et ils doivent nécessairement sélectionner des échantillons de populations qui représentent les divisions sociales. Afin de minimiser les dangers liés à une telle division de la société, il faut prendre au moins les deux précautions suivantes: dresser d'abord des échantillons qui se situe aux extrêmes des groupes que l'on compare, et limiter chacun des échantillons à un quartier ou périmètre dans lequel fonctionnent ou sont susceptibles de fonctionner des réseaux de solidarité de classe.

La seconde conclusion des théoriciens de la science dont il nous faut tenir compte concerne la rapport entre croyance et science à l'intérieur

de ce que nous avons appelé la construction socio-culturelle de la santé et de la maladie. Schutz (1962) a été l'un des premiers philosophes des sciences à prendre au sérieux ce que l'on appelait le sens commun et une bonne partie de son oeuvre est consacrée à l'étude des structures de la science de la vie quotidienne. Horton (1971) a poussé encore plus loin cette réflexion en comparant la pensée traditionnelle africaine et la science occidentale. Ce que plusieurs spécialistes des sciences humaines appellent les 'beliefs', terme qu'ils utilisent pour exprimer la dimension cognitive dans un groupe, s'est vu promu par les théoriciens de la science au rang de savoir. Tout en acceptant de reconnaître à leur suite un statut scientifique particulier au savoir populaire, il ne faut cependant pas oublier l'avertissement de Barnes: "It would be characteristic of the exciting and provoking new sociology and history of science, which takes as a duty the working hypothesis that science is no truer than other forms of tribal magic, to interpret the scientist's laboratory experiment as just another divination ritual" (1974 :). Si "la science n'est pas plus vraie que les autres formes de magie tribale", nous sommes donc de nouveau confronté au vieux débat entre l'authentique savoir né de l'expérimentation et les croyances qui sont considérées comme proches parentes de la magie. Bien sûr, les termes du débat sont anciens mais la perspective est radicalement différente puisque les théoriciens de la science ont introduit une inversion dans les rapports entre science et croyance: nous ne sommes plus en face d'un ensemble de croyances qui revendique le statut scientifique mais nous sommes plutôt en face d'une science populaire qui comme toutes les autres sciences partage certaines caractéristiques de la magie et des croyances (2).

Sur le plan méthodologique, il va sans dire que les conséquences de ce tête-à-queue sont de première importance. Si le corpus de représentations et de conceptions qu'un groupe institue dans sa réponse à la maladie et à la santé existe sous la forme d'une science, il faudra nécessairement

utiliser des instruments de recherche aptes à recencer les données qui composent ce savoir ainsi que ses axes de structuration. C'est justement à ce carrefour que les tenants de l'ethnoscience rejoignent les théoriciens de la science et les dépassent en proposant aux chercheurs des instruments qui permettent d'éliciter l'information pertinente et de l'organiser dans un ensemble cohérent (voir Tyler 1969 pour une présentation de cette méthode).

Le second point de vue qui nous a guidés dans notre recherche est celui des anthropologues médicaux à orientation culturelle. Plusieurs parmi ceux-ci ont tendance à évacuer la dimension biologique dans l'interprétation de la maladie et de la santé au profit de la seule construction culturelle qu'une population met en place à partir du donné biologique de base. Fabrega (1974 ; 1979) a bien mis en évidence les limites de l'approche socio-culturelle des études ethnomédicales qui se sont polarisées sur les conceptions des causes et mécanismes de la maladie, du diagnostic et des noms assignés aux maladies, de même que du contexte social des problèmes de santé et de la thérapie. Toute l'ethnomédecine n'est pas réductible à ces thèmes de recherche mais on peut sans trop se tromper affirmer que l'ethnomédecine est fondamentalement parcourue par une assomption fondamentale: les systèmes médicaux populaires ne partagent pas la conception bio-médicale de la maladie et des soins qui caractérise la médecine occidentale officielle. Le récent ouvrage de Foster et Anderson (1978) est particulièrement paradigmatique à cet égard par son insistance sur les dimensions socio-culturelles de la santé et de la maladie et l'évacuation concomittante de la dimension biologique.

Chez certains auteurs, la maladie et la santé ont été finalement réduites soit à un système d'interprétation (théories étiologiques, système de classification des maladies, techniques thérapeutiques) qui donne un sens au surgissement du pathologique dans un groupe donné, soit à une expérience subjective du mal dont le contenu est analysé en termes psychologiques.

Ces deux perspectives sont des plus valables mais elles ne peuvent jamais, selon nous, être dissociées à la fois de la dimension biologique qui sert de point d'ancrage à la santé comme à la maladie, et du savoir technique que tout groupe développe dans son effort pour maintenir la santé et lutter contre la maladie.

Les ethnologues médicaux ont certainement raison de considérer la santé et la maladie comme un langage ou un discours dont ils s'efforcent de décripter le code et le sens lorsqu'ils approchent par exemple ces phénomènes dans le cadre de la sémiologie, de la sémantique ou du symbolisme. Le meilleur de l'anthropologie médicale réside peut-être d'ailleurs dans cet effort de lecture symbolique; cependant tout anthropologue qui dispose d'une formation dans les sciences biologiques ou cliniques se sent mal à l'aise devant ce réductionnisme qui évacue trop facilement le biologique et le technique. Si l'anthropologue veut vraiment restituer la totalité du contenu de la santé et de la maladie, il devra nécessairement articuler son discours symbolique sur un discours biologique: ces deux registres existent évidemment comme des réalités relativement autonomes que des chercheurs pourraient appréhender indépendamment mais l'originalité du travail de l'anthropologue vise justement à saisir l'articulation, l'interdépendance, l'interpénétration des deux registres.

Sur le plan méthodologique, cela signifie que le chercheur doit utiliser des instruments lui permettant de saisir la dialectique entre d'une part l'univers biologique et technique, et d'autre part l'univers symbolique et culturel. La méthode propre à l'ethnoscience donne accès au premier versant alors que l'approche par la sémiologie permet d'interpréter le monde culturel. Plusieurs chercheurs ont récemment validés divers instruments comme les tests projectifs, les associations libres (Maranda, 1972), la sémantique différentielle (Murphy, 1974), qui visent à faire émerger l'organisation conceptuelle et symbolique de la santé et de la maladie.

Ces instruments permettent de saisir partiellement comment les changement physiques intervenant dans le corps sont ressentis, percus, nommés et interprétés par différents groupes sociaux et culturels.

Le troisième point de vue qui nous a permis de mettre au point une approche originale face au savoir médical populaire nous vient du courant subjectif qui se développe actuellement en épidémiologie. L'épidémiologie classique a toujours visé à mesurer le niveau de santé objective d'une population ou d'une classe d'âge comme les personnes âgées en utilisant des index de santé dits objectifs; depuis une dizaine d'années, les épidémiologues prètent une attention de plus en plus grande aux paramètres socio-culturels (les habitudes culturelles, les conditions de vie,...) dans leur interprétation de la distribution des problèmes de santé dans les différents groupes. Cette épidémiologie socio-culturelle est complétée depuis peu par une épidémiologie de type subjectif qui s'intéresse à l'évaluation et à la perception que les membres d'un groupe font de leur propre état de santé. (Maddox 1964; Kennedy 1973, par exemple).

L'auto-diagnostic et l'auto-évaluation doivent être considérés comme le lieu d'actualisation de la totalité du savoir médical d'un groupe: la décision qu'une personne prend face à sa propre santé (en se considérant comme malade, en nommant sa maladie, en choisissant de consulter tel spécialiste) repose en effet sur ses théories étiologiques, sa classification des signes et symptômes, ses conceptions des processus physiologiques et psychologiques, etc. En d'autres mots, les études portant sur la perception subjective de l'état de santé atteignent le système médical populaire dans son lieu primordial de visibilité c'est à dire là où il affleure à l'occasion des épisodes concrets de maladie (Alonzo 1979; Garrity et autres 1978). Les recherches sur la distribution sociale des perceptions de santé se transforment en quelque sorte en des études sur la distribution différentielle du savoir médical dans la population.

Sur le plan méthodologique, cette nouvelle épidémiologie nous apprend à

fonder toute étude des savoirs populaires sur des épisodes de maladie, lesquels actualisent ce qui a été évoqué plus haut sous l'appellation de construction bio-culturelle de la santé et de la maladie. Nous reviendrons sur ce point dans la section qui explicite notre méthode d'investigation.

Nous avons examiné le statut du savoir médical populaire par le recours aux travaux de trois catégories de chercheurs qui ont chacun pour leur compte mis en évidence un ou des aspects spécifiques de ce savoir médical. Il n'est cependant pas facile, sur la base de ces travaux de déterminer comment se fait la conceptualisation et quel est l'ordonnancement théorique du corpus de connaissances médicales que détiennent les différents groupes sociaux dans nos populations; peu d'auteurs ont pris le risque de préciser les propriétés du paradigme médical populaire.

Construisant sur les comparaisons entre le savoir de l'homme de la rue et celui de l'expert établies par Schutz, auquel nous nous referrions précédemment, Hayes-Bautista (1978) a proposé de caractériser les savoirs médicaux populaires à partir de six critères: 1) informalité: le savoir n'est ni codifié ni écrit; 2) variété de sources: le savoir se constitue à partir de plusieurs sources; 3) son acquisition se fait au hasard des rencontres et des contrats; 4) ce savoir tend vers un accroissement progressif par l'addition des expériences; 5) ce savoir peut accepter des contradictions; et 6) ce savoir n'est que partiellement réflexe et il est agi. On pourrait encore caractériser le savoir médical populaire à partir d'autres critères mais ceux déjà cités nous semblent suffisants pour introduire à nos propres travaux.

2. Présentation de la méthode de recherche

La phase exploratoire du projet de recherche "Savoir médical populaire et démarches de quête de thérapie" s'est déroulée durant l'été 1980. Cette recherche visait à recencer les connaissances médicales de deux populations du Québec métropolitain à travers des méthodes sociologiques et anthropologiques. Nous considérons cette enquête comme exploratoire parce qu'elle nous a servi à tester plusieurs instruments de collecte d'informations que nous pourrons éventuellement utiliser dans une enquête plus large sous une forme améliorée. Le but ultime de la recherche est de jeter les bases d'une éducation à la santé qui tienne compte du savoir médical des populations et de la logique qui organise leurs comportements (help-seeking behaviors) dans les situations de maladie. En vue de cette application pratique, les enquêtes se sont déroulées à l'intérieur du territoire d'un Département de santé communautaire de Québec.

Deux populations clairement définies sur le plan socio-économique ont été sélectionnées pour les enquêtes: leur profil est semblable en ce qui concerne l'âge, le sexe et la structure familiale et dissemblable par rapport à la scolarité et à la profession. Une population ouvrière habitant un quartier clairement délimité dans la périphérie de Québec, quartier qui est le développement d'un vieux village, a été sélectionnée; la plupart des unités familiales interrogées résident dans le quartier depuis quelques générations. Les hommes et les femmes de cette population ont, en moyenne, entrepris des études secondaires mais ne les ont pas complétées. Les métiers sont dans tous les cas sauf un des métiers manuels. La seconde population sélectionnée vit dans un quartier huppé de Québec qui est habité par des membres des professions libérales, des employés cadres, des professeurs d'université,... Cette population est venue s'installer plus ou moins récemment dans le quartier mais elle manifeste néanmoins une grande solidarité de classe. Les hommes ont, en moyenne, complété des études universitaires et les femmes ont entrepris

des études collégiales sans toutefois les avoir complétées. Les enquêtes se sont déroulées en trois étapes comprenant chacune une visite dans l'unité familiale et l'application d'un des instruments de recherche. Nous avons d'abord procédé à une enquête extensive dans chaque milieu auprès de 50 unités familiales choisies sur la base des listes électorales et répondant à nos critères, à qui nous avons appliqué un questionnaire fermé visant à obtenir de l'information sur les problèmes de santé expérimentés par l'une ou l'autre des personnes de la maison durant les six mois précédant l'entretien. Cette enquête préliminaire de morbidité avait pour but de nous permettre d'isoler les unités familiales qui avaient vécu un épisode de maladie dans un passé récent; il nous était ensuite possible de travailler avec ces unités familiales dans une double perspective: l'investigation de leurs connaissances médicales et des différentes démarches entreprises pour contrôler le problème de santé. Nous nous limitons à présenter ici les instruments qui ont été employés pour investiguer le savoir médical des deux populations.

Nous avons essayé d'avoir accès à ce savoir médical populaire par le biais de trois instruments: une entrevue ouverte structurée (01), un diagramme de sémantique différentielle (02) et un questionnaire sur la perception face à certains problèmes limités (03). L'entrevue ouverte structurée (01) s'est faite auprès d'un échantillon de 20 personnes dans le milieu 1 et de 24 dans le milieu 2, personnes qui appartenaient toutes à une unité familiale vivant ou ayant vécu une situation problématique. L'organisation de l'instrument (01) peut être caractérisée comme suit: 1. Les questions ne sont jamais théoriques mais elles partent toujours de la situation concrète de maladie: nous n'interrogeons pas par exemple sur le fonctionnement du corps en général ou sur les causes générales des maladies mais nous

appliquons toujours nos questions à la maladie vécue dans la famille; 2. Les questions s'inscrivent dans un ensemble systémique: nous abordons successivement mais dans la continuité les problèmes de la définition de la santé et de la maladie, des signes et symptômes des maladies, de leurs causes, de leurs effets sur le corps, etc. Nous pensons en effet qu'il est impossible d'accéder à la construction bio-culturelles de la santé et de la maladie sans obtenir des informations sur les théories étiologiques, sur le fonctionnement du corps, sur les moyens à utiliser pour se maintenir en santé, etc. 3. L'instrument 01 veut enfin être émique en ce sens qu'il ne structure pas les réponses à l'avance et qu'il laisse aux personnes la possibilité d'exprimer leurs propres conceptions sur la santé et la maladie.

Les instruments 02 et 03 se situent directement dans le prolongement de la dimension émique de l'enquête. Le tableau de sémantique différentielle (02) vise à obtenir une information plus structurée susceptible de nous permettre d'organiser l'espace occupé par le savoir médical populaire en distinguant l'écart sémantique entre les concepts, leur recouvrement éventuel, etc. Cet instrument comme l'instrument 03 d'ailleurs avait aussi pour but d'opérer un certain contrôle de qualité sur la validité des informations obtenues dans l'entrevue (01) et c'est pourquoi nous les avons appliqués à l'occasion d'une troisième visite auprès d'un sous-échantillon des unités familiales sélectionnées.

En ce qui concerne l'instrument 03 qui porte sur les perceptions et qui devrait normalement faire émerger le domaine symbolique, notre questionnaire s'est avéré inadéquat; nous pensons qu'un test projectif ou un test d'associations libres ou d'autres tests du même genre seraient probablement plus adaptés.

Notre méthode se caractérise donc de quatre manières: 1. *elle reconstitue le savoir à partir d'une pratique* qui actualise ce savoir ou plus précisément sur la base d'épisodes concrets de maladie qui mettent ce savoir en forme; 2. *elle est émique* en ce sens qu'elle laisse aux personnes la possibilité d'exprimer leurs propres connaissances sans les structurer à l'avance dans des questionnaires pré-codifiés; 3. *elle est systémique* parce qu'elle resitue le savoir sur la maladie et la santé dans l'ensemble du savoir médical populaire; 4. enfin *elle est sémantique* car elle vise à découvrir l'ordonnancement interne de la construction bio-culturelle de la santé et de la maladie dans les deux populations.

Nous avons considéré cette construction bio-culturelle à la fois comme une science et comme un discours qui sont complémentaires tout en étant isolables au moment de la collecte des informations et lors de l'interprétation. Les données recueillies par l'entrevue ouverte (01) ont été codifiées et retransposées de façon à ce que nous puissions dégager les profils individuels et de groupe; cette codification nous a également permis d'établir un certain nombre de corrélations entre les différentes parties de l'ensemble du savoir médical, corrélations qui sont indispensables dans une approche systémique. Une fois notre espace bien délimité, l'analyse a consisté à saisir l'ordre interne de ce domaine à partir des données recueillies à la fois par la méthode de la sémantique différentielle (02), du questionnaire de perception (03) et de l'entrevue structurée (01).

3. Construction bio-culturelle de la santé et de la maladie. Données préliminaires.

Notre point de départ a consisté à demander aux personnes comment elles évaluaient leur propre état de santé. Lorsqu'il s'agissait d'une auto-évaluation positive, nous les interrogions immédiatement sur leur conception de la santé; dans les cas d'évaluation négative, nous nous centrions dès le début de l'entrevue sur la conception de la maladie. Pour avoir accès à l'autre pôle, nous demandions ensuite à la personne en santé si elle connaissait quelqu'un de son entourage immédiat qui était malade, et à la personne malade si elle pouvait nous parler de la santé en se référant à un proche qu'elle considérait comme sain. Cette procédure de référence à des situations expérientielles avait pour but de fournir un support constant à la réflexion du répondant.

Auto-évaluation de l'état de santé par les 44 répondants

	milieu ouvrier (20)	milieu bourgeois (24)	Total (44)
. parfait	2 (10%)	3 (13%)	5 (11%)
. très bon	5 (25%)	6 (25%)	11 (25%)
. bon	7 (35%)	13 (54%)	20 (46%)
. passable	5 (25%)	2 (8%)	7 (16%)
. mauvais	1 (5%)	0	1 (2%)

Les personnes interrogées sont majoritairement en bonne, très bonne ou excellente santé (36/44); seulement 8/44 perçoivent leur santé comme passable ou mauvaise. Dans ces 8 cas, il s'agit de personnes qui vivaient effectivement au moment de l'enquête un épisode de maladie.

Nous nous sommes intéressés aux raisons qui justifiaient l'évaluation de l'état de santé. Dans 8 cas, les personnes ont dit fonder leur perception sur des examens médicaux récents qui s'étaient révélés bons. Dans les autres cas, les répondants se sont auto-évalués en référence à trois catégories générales de critères (4):

	Ouvrier	Bourgeois	Total
. absence de problème spécifique	4 (20%)	15 (63%)	19 (43%)
. éléments physiques	19 (95%)	16 (67%)	35 (80%)
. éléments psychologiques	16 (80%)	4 (17%)	20 (45%)

Bien que le petit nombre de répondants ne permettent pas de tirer des conclusions valables pour l'ensemble des deux catégories sociales, il est intéressant de signaler que l'identification de problème précis ne semble pas servir de mesure aux personnes du milieu ouvrier sauf dans le cas où ces problèmes entraînent des perturbations psychologiques importantes. Le milieu ouvrier manifeste en effet une nette tendance à évaluer son état de santé à partir de connotations psychologiques (16/20 en parlent explicitement) alors que le milieu bourgeois semble préférer mesurer son niveau de santé à partir de problèmes identifiables (15/24 en parlent). 35 répondants sur 44 ont inclus dans leur évaluation des justifications d'ordre physique. En résumé, les répondants du milieu ouvrier associent dans presque tous les cas des éléments physiques et psychologiques dans leur auto-évaluation alors que les répondants de l'autre milieu se réfèrent plutôt au physique et à la présence ou à l'absence de problèmes précis.

Définition de la santé

	Ouvrier	Bourgeois	Total
. par la négative (absence de fatigue, de plaintes, de douleurs de problèmes, de maladie,...)	9 (45%)	6 (25%)	15 (34%)
. par la présence de problèmes mineurs qui sont sous contrôle	1 (5%)	5 (21%)	6 (14%)
. par la référence à la capacité fonctionnelle (pouvoir faire ce qu'on a à faire: travail, occupations journalières,...)	13 (65%)	18 (75%)	31 (70%)
. par la résistance (bon fond de santé, robustesse)	3 (15%)	2 (8%)	5 (11%)
. par l'apparence physique (bon teint, poids et yeux)	1 (5%)	0	1 (2%)

Définition de la santé (suite)

	Ouvrier	Bourgeois	Total
. par certaines fonctions physiologiques (bon appétit, sommeil facile,...)	6 (30%)	8 (33%)	14 (32%)
. par un bon état mental (de bonne humeur, joie de vivre, bon moral, bien dans sa peau, pep, entrain)	18 (90%)	11 (46%)	29 (66%)

Les deux éléments les plus fréquents auxquels les personnes se réfèrent pour définir la santé sont la capacité fonctionnelle (31/44) et l'état mental (29/44), le dernier étant surtout caractéristique du milieu ouvrier (18/20) et le premier du milieu bourgeois (18/24). La santé est associée chez les répondants du milieu ouvrier à la joie de vivre, au bon moral, à la bonne humeur; on peut donc aisément affirmer que la définition populaire de la santé inclut primordialement des éléments liés à la santé mentale. Un second marqueur du bon état de santé se trouve dans la capacité d'accomplir ce que l'on a à faire; cet indicateur rejoint en positif la définition de 'rôle de malade' interprété dès 1951 par T. Parsons comme le droit à se retirer des activités habituelles. Cette définition sociologique semble donc s'enraciner profondément dans la pensée populaire qui fait de la suspension des activités le marqueur de la détérioration de l'état de santé. L'idée généralement admise dans les milieux médicaux suivant laquelle la population appréhende son état de santé et de maladie à partir des seuls processus physiologiques trouve cependant ici un démenti total puisque la très grande majorité des répondants (29/44) prennent d'abord en considération la qualité de leur vie psychique.

De fait, plusieurs répondants (29/44) se réfèrent explicitement à des considérations physiologiques en plus des deux marqueurs déjà cités. Cet indicateur organique de santé est exprimé positivement par le bon appétit, le sommeil facile, la bonne digestion ..., ou négativement

par l'absence de douleur, ou de fatigue. Il nous semble important de signaler que cette attention au corps ut sic vient en troisième position derrière les indicateurs sociaux et psychologiques.

Cette définition de la santé par les processus biologiques mérite cependant d'être examinée en détail non seulement parce que la rencontre avec le système bio-médical professionnel va se faire à ce niveau au moment des consultations mais également parce que le corps constitue un espace privilégié de visualisation des signes de santé et de maladie. Au terme de son étude sur la représentation sociale de la santé à Paris, C. Herzlich (1969) établit un parallèle entre la santé et l'individu qui serait par nature sain dans son corps et son esprit: "La genèse de la santé et de la maladie doit se concevoir comme une lutte entre l'individu-santé et le mode de vie-maladie. Cette lutte oppose un élément actif, le mode de vie, qui déclenche, par ses attaques, la maladie, et un élément passif (passivité étant ici synonyme de résistance), la santé toute entière de l'individu (ou l'individu tout entier santé)" (p. 42).

Cette conception suivant laquelle la santé serait fondamentalement attribuable à l'individu à cause de sa robustesse, de sa bonne hérédité, ou de son fond de santé et que la maladie serait à placer totalement du côté de l'environnement et du style de vie ne se retrouve que partiellement dans les réponses de nos deux échantillons de populations. La grande rigidité du schéma de Herzlich peut possiblement s'expliquer à partir des caractéristiques de la vie parisienne mais nous comprenons néanmoins assez mal la connotation essentiellement positive attachée au corps. Si la maladie est totalement interprétée dans les termes d'une agression qui vient de l'extérieur, de l'environnement, cela implique que sans agression, il y aurait permanence de la santé ou, selon la pensée de Herzlich, que le corps est par nature sain. Les québécois sont certainement moins affirmatifs que les parisiens sur ce point et préférent utiliser un vocabulaire de type énergétique que nous discuterons plus loin.

Dans notre enquête, trois indicateurs sont donc simultanément utilisés par la très grande majorité des répondants pour définir la santé: un indicateur social, un indicateur psychologique et un indicateur physique. Cependant tous affirment qu'il y a plusieurs manières d'être en santé et que celle-ci peut à la limite varier d'une personne à une autre. Sans entrer dans les détails statistiques, nous signalons que les répondants semblent distinguer sur une base bi-polaire entre ceux qui font attention à leur santé et ceux qui n'en prennent pas soin tout en étant parfois aussi bien portants que les premiers, entre ceux qui ont une grosse santé de base et ceux qui ont hérité d'une petite santé, entre ceux qui ont une santé qui ne connait aucune défaillance et ceux qui sont en santé tout en vivant des problèmes qui restent sous contrôle, entre ceux enfin qui présentent peu de sensibilité au mal et ceux qui sont très sensibles à la moindre douleur. Il serait certainement possible de préciser davantage les modalités de la santé sur la base de ces quatre axes mais ce travail est encore en chantier et nous ne pouvons en dire davantage pour le moment.

Une chose est certaine: les modalités de la santé varient avec l'âge. L'enfant est considéré par 22 répondants (50%) comme ayant plus de résistance et d'énergie que l'adulte; 9 signalent que sa santé se mesure à ses activités (mouvements, jeux,...) signifiant ainsi que toute diminution de dynamisme traduit une détérioration de la santé. Cette interprétation de la santé chez l'enfant en termes énergétiques prend un sens spécial lorsque l'on constate que la personne âgée est à l'opposé, considérée comme disposant de moins d'énergie et de résistance. Les répondants des deux milieux sont presque unanimes à affirmer que l'énergie diminue avec l'âge. Tant chez l'enfant que chez la personne âgée, certaines maladies bénignes sont considérées comme normales et elles n'entamment pour ainsi dire pas le fond énergétique de santé. La différence entre ces deux groupes d'âge est située par

les répondants dans le fait que les enfants semblent ne pas se préoccuper de leur santé, négligence qui s'explique par leur dépendance vis-à-vis des adultes, de l'avis de plusieurs répondants, alors que les personnes âgées sont attentives aux moindres problèmes.

Dans cette approche énergétique, l'adulte est considéré comme en perte progressive de résistance, perte qui devient sensible au niveau d'un seuil critique que certains situent à 40 ans et d'autres à 50. Plus qu'aux autres âges de la vie, l'adulte a la pleine responsabilité de sa santé mais c'est pour en abuser le plus souvent, affirment plusieurs répondants. La différence qualitative dans la santé aux différents âges de la vie est donc attribuée dans les deux milieux à la diminution progressive de la résistance et à l'augmentation concomittante des maladies.

La santé est définie comme un donné dans la mesure où elle est comprise comme un potentiel de résistance mais elle est en même temps considérée comme une réalité relativement fragile qu'il faut protéger par différents moyens:

Moyens de conserver la santé

	Ouvrier	Bourgeois	Total
. alimentation	14 (70%)	24 (100%)	38 (86%)
. exercices physiques, loisirs	9 (45%)	18 (75%)	27 (61%)
. travail	1 (5%)	1 (4%)	2 (4.5%)
. environnement sain	2 (10%)	5 (21%)	7 (16%)
. éviter les stress, les tensions	5 (25%)	7 (29%)	12 (27%)
. bon moral	4 (20%)	3 (13%)	7 (16%)
. contrôle médical régulier	2 (10%)	3 (13%)	5 (11%)
. contrôle sur les habitudes de vie	18 (90%)	18 (75%)	36 (82%)

La responsabilité à l'égard du maintien de la santé s'exerce suivant des lignes qui sont incontestablement culturelles et circonstantielles. Il faudrait cependant pouvoir distinguer parmi les réponses recueillies, celles qui sont spontanées et qui ne sont peut-être que le simple écho affaibli des campagnes gouvernementales d'éducation à la santé (bien manger, faire de l'exercice "va jouer dehors"...) des véritables réponses réflexes qui intégrent à la fois le système personnel de valeurs et les comportements concrets. Cette discrimination n'est pas faite dans le tableau précédent qui est construit à partir des réponses spontanées de la population; une enquête parallèle qui vise les réponses réflexes et les comportements sera bientôt mise en place.

Quatre catégories de moyens s'inscrivent dans le domaine de l'organisation du style personnel de vie sont mises en avant dans les deux milieux: ce sont l'alimentation, le contrôledes habitudes de vie, les exercices physiques et l'évitement des stress et tensions. La responsabilité du maintien de la santé est entièrement placée du côté de l'individu, lequel sera tenu éventuellement coupable s'il tombe malade. Nous sommes en pleine société libérale qui remet la totalité du contrôle entre les mains des personnes individuelles. Il n'y a que 7 personnes qui signalent que l'environnement peut avoir des effets sur la qualité de la santé, et uniquement 2 s'intéressent aux conditions de travail alors que plusieurs personnes du milieu ouvrier ont par ailleurs signalé qu'elles travaillent dans une ambiance pénible. Toujours dans ce refus de céder aux autres le contrôle sur sa santé, il n'y a que 5 personnes qui suggèrent que la supervision médicale intervienne dans le maintien de la santé.

Si notre interprétation des données préliminaires concernant les moyens de se maintenir en santé est juste, nous nous trouvons en présence d'une étrange situation: d'une part, il y aurait dans la population une faible prise de conscience de l'importance de la qualité de l'en-

vironnement et des responsabilités des différentes institutions publiques dans la création de bonnes conditions de santé; d'autre part, l'insistance sur la responsabilité individuelle, laquelle peut très bien être une traduction des campagnes "Au Québec, il faut se tenir en santé", pave la voie au courant de culpabilisation des malades, qui se développe de plus en plus chez les planificateurs. Les utilisateurs des services de santé seront un jour appelés à payer un 'ticket modérateur' (comme cela se fait dans plusieurs pays européens) puisqu'ils seront tenus personnellement responsables de leur état et qu'ils n'auront plus le droit de faire endosser les conséquences de leur mauvais style de vie par toute la communauté. Nous pensons pouvoir conclure de nos données sommaires que la mentalité écologique n'organise encore la pensée et les comportements qu'à l'intérieur du domaine privé individuel; il nous apparait aussi que la remise totale de la responsabilité entre les mains de chaque individu risque de conduire à une culpabilisation progressive des malades.

Nous arrivons maintenant au second versant de la construction bio-culturelle de la santé et de la maladie, celui qui touche justement aux problèmes de la maladie. Notre point de départ a consisté à demander, comme dans le cas de l'étude sur la santé, aux différentes personnes si elles étaient actuellement malades ou si elles connaissaient quelqu'un de leur entourage qui l'était. Le déroulement de l'enquête s'est ensuite poursuivi en référence à ce cas de maladie. Le tableau suivant présente les principaux indicateurs de la présence de la maladie:

Indicateurs de la maladie

	Ouvrier	Bourgeois	Total
. diminution ou perte de capacité physique	19 (95%)	17 (71%)	36 (82%)
. dépendance accrue	3 (15%)	4 (17%)	7 (16%)
. états psychologiques spécifiques (agressivité, diminution du moral,...)	18 (90%)	13 (54%)	31 (70%)
. douleur et souffrance	11 (55%)	6 (25%)	17 (38%)

Etre malade signifie d'abord pour les répondants une diminution ou une perte des capacités physiques; dans les deux milieux, on note également que la maladie s'accompagne de conséquences psychologiques particulières comme l'agressivité ou la diminution du moral chez le malade. Ces deux indicateurs de maladie sont explicitement signalés par 90% (18/20) des répondants du milieu ouvrier et 54% (13/24) dans le second milieu. Si nous établissons une comparaison avec les marqueurs de l'état de santé, nous constatons qu'il y a une nette consistance. La dépendance accrue à laquelle se réfèrent 7 personnes peut être interprétée comme l'envers de l'indicateur social déjà présenté: non seulement le malade suspend ses activités comme l'a montré Parsons mais il a de plus besoin que les autres s'occupent de lui. Il faut peut-être signaler que 55% des répondants du milieu ouvrier associent la douleur et la souffrance à l'expérience de la maladie alors que seulement 25% établissent une telle association dans le milieu bourgeois.

Quels sont plus précisément les signes, les manifestations, les indices qui caractérisent la présence de la maladie? Nous avons regroupé les réponses en distinguant entre le signe et le symptôme, donnant à ces mots aux sens multiples la signification que leur attribue Barthès dans son étude sur la sémiologie médicale (1972): le signe est le symptôme soumis à une lecture positionnelle de sens et le symptôme est davantage un matériau brut non encore décripté par l'analyse.

Signes et symptômes de la maladie

Signes:	Ouvrier	Bourgeois	Total
. changement dans l'apparence physique	14 (70%)	6 (25%)	20 (45%)
. modification des fonctions physiologiques	10 (50%)	8 (33%)	18 (41%)
. modification du degré d'activités	15 (75%)	14 (58%)	29 (66%)
. modification de l'état psychologique	11 (55%)	16 (66%)	27 (71%)

Symptômes:	Ouvrier	Bourgeois	Total
. fièvre	8 (40%)	6 (25%)	14 (32%)
. faiblesse, fatigue	14 (70%)	13 (54%)	27 (61%)
. douleur	9 (45%)	8 (33%)	17 (39%)
. yeux, teint changés	13 (65%)	6 (25%)	19 (43%)
. perte de poids	6 (30%)	2 (8%)	8 (18%)

Le tableau des signes n'appelle pas beaucoup de commentaires car les résultats concordent parfaitement avec ce que nous avons antérieurement discuté concernant les marqueurs spécifiques de la santé et de la maladie. Lorsque nous jetons un coup d'oeil du côté des symptômes, nous sommes en un premier temps surpris de constater que les manifestations typiques de la présence de la maladie se réduisent soudainement à des modifications physiques. Cette polarisation de type organiciste se comprend plus difficilement en fonction d'une hypothétique intériorisation du modèle bio-médical par la population qu'en référence au langage qu'il convient de parler lorsqu'on est malade si l'on veut être compris des professionnels. Les symptômes signalés sont suffisamment généraux pour être considérés comme des indices aisément identifiables de modification des états corporels et on peut penser que ces marqueurs très empiriques servent effectivement à la population pour discriminer entre ceux qui sont malades et ceux qui sont en santé.

A ce stade-ci de la description de la construction bio-culturelle que deux populations du Québec métropolitain font de la santé et de la maladie, la présentation des données et leur interprétation devraient normalement s'orienter dans une double direction. Tout d'abord, il conviendrait de situer les informations précédentes dans le cadre plus large de la pensée médicale populaire telle qu'elle s'exprime à travers les théories de l'étiologie des maladies, les explications des effets des maladies sur le corps et la personne, de même que par le biais de l'organisation nosologique populaire du domaine des maladies (5).

Nos enquêtes encore fort limitées nous ont cependant permis d'obtenir des informations systématiques sur chacun de ces domaines d'une part à partir de cas concrets de maladies et d'autre part en référence à des maladies courantes (cancer, diabète, hypertension, rhume,...). (6) Cette première direction dans le prolongement de notre réflexion viserait donc directement la présentation de la construction bio-culturelle en tant que science médicale populaire.

La seconde direction devrait envisager cette même construction comme un discours ou un langage qui va plus loin que la simple science médicale descriptive pour se hausser au niveau des symboles et des valeurs culturelles. Les chercheurs se heurtent généralement à cette épaisseur qualitative parce que leurs instruments ne sont pas suffisamment adéquats pour la traverser. Ce fut également notre cas: le test de sémantique différentielle que nous avons utilisé avait pour but de mesurer le poids des différents prédicats par rapport à la santé et à la maladie; ce premier test était complété par un questionnaire qui visait à faire délimiter l'écart sémantique entre différents termes communément utilisés pour se référer au pathologique. Cette enquête qui cherche non seulement à organiser le champ du conceptuel sur la base de prédicats mais également à remonter jusqu'au niveau de la vie symbolique, à celui de la charge affective dont son investies certains états ou organes physiques, est actuellement en cours et se révèle prometteuse. Nous sommes cependant conscients du fait que la méthode de la sémantique différentielle a des limites principalement par rapport au surgissement du symbolique et nous pensons qu'elle devra être complétée par l'utilisation d'instruments projectifs ou de tests d'association libre et d'association narrative semblables à ceux employés par Maranda (voir note 3).

En conclusion de ces réflexions sur le statut du savoir populaire, la méthodologie complexe proposée pour l'atteindre et la présentation de quelques résultats préliminaires, il nous parait utile de nous arrêter sur le difficile problème des lieux de 'captage' de la construction bio-culturelle de la santé et de la maladie. Toutes les informations présentées ci-dessus ont été recueillies dans le cadre d'enquêtes qui se sont déroulées dans des maisons privées auprès de personnes qui avaient plus ou moins récemment vécu un épisode de maladie. Nous pensons que ce lieu de surgissement de l'information, bien que très important, ne suffise pas à donner accès à toutes les dimensions de la construction bio-culturelle.

Nous comptons maintenant saisir l'information à partir de deux lieux complémentaires de 'captage'. Tout d'abord, il faudrait pouvoir transformer en informateurs privilégiés certaines personnes qui sont actuellement engagées dans un processus de quête de thérapie. Ces personnes devraient idéalement être rejointes à différents lieux de la quête (une clinique médicale, une consultation auprès d'un acuponcteur, d'un chiropraticien, d'un homéopathe ou d'un guérisseur, une visite à un lieu de pélerinage, etc) car on peut à bon droit faire l'hypothèse suivante: la diversité dans les lieux de consultation s'explique par les différences dans la demande et l'attente.

Le troisième lieu qui nous apparait comme fort propice à l'élucidation complète du discours populaire sur la santé et la maladie est celui des malades qui expérimentent une maladie chronique. Ces personnes vivent souvent l'aventure de la maladie depuis plusieurs années et c'est probablement chez eux que le fonds symbolique a des chances de s'exprimer avec le plus de visibilité. De nombreux problèmes méthodologiques sont liés à un tel genre d'enquêtes mais il faut cependant essayer de les surmonter.

NOTES

(1) En conclusion de son étude comparative des travaux des sociologues médicaux sur la définition de la santé et de la maladie, Idler (1979) écrit: "Neither Parsons, Mechanic or Freidson, or even their Marxist critics have attempted to analyse illness (much less health) independant of the medical profession. While a functional definition of illness provides social data for analysis that the biomedical lacks, limiting the definition of illness to that which is medically validated severely restricts the possible area of sociological inquiry both theoretically and practically" (1979:724).

(2) Comaroff écrit à ce propos que l'étude comparée des systèmes de connaissances démontre clairement que toutes les sciences peuvent être considérées comme des systèmes symboliques et qu'il faut nécessairement, sous cet angle, rapprocher le système médical officiel, en tant qu'ensemble normatif et système défini d'action, de l'ethnoscience populaire qui s'est développée en parallèle dans les sociétés occidentales (1978:250).

(3) Les travaux en cours dans l'équipe dirigée par P. Maranda méritent d'être signalés explicitement pour une double raison: tout d'abord à cause de leur importance théorique et ensuite parce qu'ils se déroulent principalement dans la ville de Québec. P. Maranda vise à faire la sémantographie de différents domaines comme le travail (1979), la maladie (Blouin 1980), la sexualité,... en utilisant deux tests complémentaires (le TAL ou Test d'Associations Libres et le TAN ou Test d'Associations Narratives) qui sont construits sur ce qu'il appelle les 'symboles lourds' dont l'existence est transculturelle. Afin de

limiter les informations et de rendre possible l'analyse, Maranda se limite, pour le moment, à l'utilisation de trois symboles (femme, homme et serpent) qui sont fortement marqués dans la culture occidentale. On peut se demander si d'autres symboles comme le sang, la mort par exemple ne seraient pas plus adaptés à la clarification du domaine sémantique de la santé et de la maladie. La méthode initiée par Maranda et actuellement appliquée dans plusieurs recherches mérite certainement d'être étendue à l'utilisation d'autres symboles.

(4) Dans la plupart des tableaux qui suivent, les différents codes sont généralement exclusifs du point de vue formel mais il se peut très bien qu'une même personne se réfère simultanément à plusieurs d'entre eux pour qualifier un phénomène comme la santé ou la maladie. Chaque répondant a un profil spécifique qui s'organise autour de plusieurs codes que nous avons hiérarchisés à partir de ce qu'il nous disait de leur importance relative. Cette référence de chaque personne à plusieurs codes explique donc pourquoi il n'y a pas adéquation entre le nombre de réponses et le nombre de répondants.

(5) Dans ce dernier domaine, un travail exploratoire a été mené par G. Brunel (1978), travail qui s'inscrit dans la ligne théorique de l'ethnoscience.

(6) Une présentation même sommaire de ces résultats dépasserait considérablement l'espace qui nous est imparti et nous nous voyons donc dans l'obligation de référer le lecteur au texte du rapport complet de la recherche.

BIBLIOGRAPHIE

ALONZO, A.A. Everyday Illness Behavior: A Situational Approach to Health Status Deviations, in Social Science and Medicine, Vol 13 A, No 4 (1979) 397-404.

BARTHES, R. Sémiologie et médecine, in R. Bastide (ed). Les Sciences de la folie, Paris: Mouton, 1972.

BELMONT, N. Van Gennep, le créateur de l'ethnographie française, Paris: Payot, 1974.

BIBEAU, G. De la maladie à la guérison. Essai d'analyse systémique de la médecine des Angbandi du Zaïre, Thèse doctorale, Université Laval, 1979.

BIBEAU, G., PELLETIER, L. La santé et la maladie: science et discours dans deux populations du Québec métropolitain (Rapport de recherche), Ecole des sciences infirmières, Université Laval, 1980.

BLOUIN, M., Culture santé et maladie. Mémoire de maîtrise, Département d'anthropologie, Université Laval, 1980, 115p.

BRUNEL, G., MORISSETTE, L. Guérison et ethno-étiologie populaire, Anthropologica (1979).

CAMPBELL, D.T. A Tribal Model of the Social System Vehicle Carrying Scientific Knowledge, in Knowledge: Creation, Diffusion, Utilization, Vol 1, No 2 (1979) 181-201.

COMAROFF, J. Medicine and Culture: Some Anthropological Perspective, in Social Science and Medicine, Vol 12 B (1978) 247-254.

CUMMINGS, K.M., JETTE, A.M., ROBERTSON, I.M. Construct Validation of the Health Belief Model, in Health Education Monographs, Vol 6, No 4 (Winter 1978) 394-405.

DES RUISSEAUX, P. Croyances et Pratiques Populaires au Canada Français, Montréal: Ed. du Jour, 1973.

DINGWALL, R. Aspects of Illness, New York: St Martin's Press, 1976.

FABREGA, H. Disease and Social Behavior: An Interdisciplinary Perspective, Cambridge: The MIT Press, 1974.

FABREGA, H. The Ethnography of Illness, in Social Science and Medicine Vol 13 A, (1979) 565-576.

FAVRET-SAADA, J. Les mots, la mort, les sorts. La sorcellerie dans le Bocage, Paris: Gallimard, 1977.

FISHBEIN, M., AJZEN, I. Beliefs, Attitudes, Intention and Behavior: An Introduction to Theory and Research, Reading, Mass.: Addison-Wesley, 1975.

FOSTER, G.M., ANDERSON, B.G. Medical Anthropology, New York: John Wiley and Sons, 1978.

FREIDSON, E. Profession of Medicine: A Study of the Sociology of Applied Knowledge, New York: Dodd, Mead, 1970.

GARRITY, T.F., SOMES, G.W., MARX, M.B. Factors Influencing Self-Assessment of Health, in Social Science and Medicine, Vol 12, No 2 (1978) 77-81.

HAYES-BAUTISTA, D.E. Chicano Patients and Medical Practioners: A Sociology of Knowledges Paradigm of Lay - Professional Interaction, in Social Science and Medicine, Vol 12, No 2 (1978) 83-90.

HERZLICH, C. Santé et maladie. Analyse d'une représentation sociale, Paris: Mouton, 1969.

HORTON, R. African Traditional Thought and Western Science, in M.F.D. YOUNG (ed), Knowledge and Control, London: Collier-Macmillan, 1971.

IDLER, E.L. Definitions of Health and Illness and Medical Sociology, in Social Science and Medicine, Vol 13 A (1979) 723-731.

IGUN, U.Z. Stages in Health-Seeking: A Descriptive Model, in Social Sciences and Medicine, Vol 13 A (1979) 445-456.

KENNEDY, D.A. Perceptions of Illness and Healing, in Social Sciences and Medicine, Vol 7 (1973) 787-805.

KUHN, T.S. La structure des révolutions scientifiques, Paris: Flammarion, 1972.

MADDOX, G.L. Self-Assessment of Health Status, in Journal of Chronic Diseases, 17, 449, (1964).

MARANDA, P. Introduction to Anthropology: A Self-Guide, Englewood Cliffs: Prentice Hall, 1972.

MARANDA, P. Sémantographie du domaine "travail" dans la Haute-Ville et dans la Basse-Ville de Québec, in Anthropologica, Vol. XXI, 1979, (no spécial sur l'Appropriation sociale de la logique).

MECHANIC, D. Medical Sociology, New York: Free Press, 1978.

MOSCOVICI, S. La phychanalyse, son image et son public. Etude sur la représentation sociale de la psychanalyse, Paris: Presses Universitaires de France, 1961.

MURPHY, H.B.M. Differences of Mental Disorders Between French Canadians and British Canadians, Canadian Psychiatric Associations Journal, Vol 19 (1974) 247-257.

PARSONS, T. The Social System, London: Routledge and Kegan, Paul, 1951.

ROSENSTOCK, I.M. Why People Use Health Services, in The Milbank Memorial Fund Quarterly, Vol XLIV (1966) 94-127.

SAUCIER, J.F., STEINBERG, M. Adolescents et Prévention. Rapport Préliminaire, Montréal: Université de Montréal, 1979.

SCHUTZ, A. Collected Papers Vol 1: The Problem of Social Reality, The Hague: Martinus Nijhoff, 1962.

TYLER, S.A. Cognitive Anthropology, New York: Holt, Rinehart and Winston, 1969.

23 septembre 1980
/lm /dt

HERZLICH, Claudine. La conception psychosociale de la maladie. Anthropologie et psychosociologie médicale (introduction), pp. 13-29, in HERZLICH, Claudine. Santé et maladie; analyse d'une représentation sociale. Paris. Mouton, 1969.

INTRODUCTION

I. — LA CONCEPTION PSYCHOSOCIALE DE LA MALADIE. ANTHROPOLOGIE ET PSYCHOSOCIOLOGIE MÉDICALE

Que signifient pour nous la santé et la maladie ? Que sont-elles pour un membre de notre société ? Par l'expérience personnelle de chacun, mais aussi à travers l'information qui circule dans la société, modelés par ses valeurs, leur image se structure : être malade, être en bonne santé, c'est, semble-t-il, d'abord et seulement, une expérience individuelle qui ne se peut partager. Pourtant, nous savons que l'incommunicabilité est déjà rapport aux autres ; on est malade ou bien portant pour soi, mais aussi pour la société et en fonction de la société.

Ainsi, la maladie d'abord vécue, subie sans intermédiaires, est aussi apprise : l'enfant, pensons-nous, ne comprend pas ce qu'est la maladie et ce qu'elle va signifier dans sa vie ; il doit apprendre à insérer son expérience dans un réseau d'explications et de règles sociales — le jeu du « malade et du docteur », en particulier, semble avoir ce rôle [1]. L'adulte lui-même apprend de la société à être malade ; cet apprentissage débute avec le nom donné par le médecin à la maladie, puis viennent les prescriptions qui sont des règles d'action, la rencontre avec des institutions : l'hôpital, la Sécurité Sociale. Pour le malade, comme pour le médecin, pour l'assuré social en face de l'administration, des rôles se différencient et des normes de conduite s'imposent : au regard de la société, le malade n'est pas équivalent au bien portant.

Étudier la représentation sociale de la santé et de la maladie, c'est observer comment cet ensemble de valeurs, de normes sociales et modèles culturels est pensé et vécu par des individus de notre société, étudier comment s'élabore, se structure, logiquement et psycholo-

1. Cf., par exemple, l'analyse qu'en fait J. P. Valabrega (1962), chapitres VI et VII.

giquement, l'image de ces objets sociaux que sont la santé et la maladie. Nous approchons ainsi l'image que chacun se fait du réel, son organisation et son sens. Mais, puisque l'image contribue à modeler les comportements, est source de comportements, son étude sera aussi celle d'une des « sources » du réel.

*

Une telle tentative nous semble cependant inséparable de la conception caractéristique de notre culture qui, autant que comme faits organiques, envisage aujourd'hui santé et maladie *sous l'angle psychosocial, comme faits psychosociaux*. Que l'on tende d'ailleurs à inclure dans la maladie des états d'inadaptation — personnelle et sociale — qui, naguère, ne relevaient que d'une approche morale et légale, montre aussi que, pour nous, la maladie a cessé d'être exclusivement un fait du corps, objectivable en seuls termes organiques. Cette conception prévaut chez certains médecins ; sous diverses formes, elle s'exprime dans les manuels médicaux, mais aussi dans la presse quotidienne [1]. Elle commence à s'inscrire dans les institutions par le biais de réformes des études médicales ou de la formation du personnel hospitalier par exemple. Elle est enfin à la source de différents modes d'approches du phénomène maladie.

Une telle conception est moins neuve qu'il n'y paraît : du moins affirmerons-nous qu'elle peut, sous sa forme la plus générale, *être reliée à toute l'histoire des idées médicales*. On peut distinguer les théories centrées sur l'objectivation de la maladie dans le corps de celles qui envisagent santé et maladie davantage comme mode de relation — équilibre et déséquilibre — de l'homme avec son milieu ; où interviennent donc facteurs humains, conditions écologiques et structures sociales. En un sens, le développement de la médecine scientifique nous apparaît comme le triomphe de l'approche centrée sur le corps ; maints historiens de la médecine voient en Hippocrate le fondateur de la médecine parce qu'il a inventé la méthode clinique, l'observation du malade, de ses symptômes, de son corps. Mais l'on sait aussi qu'il notait l'influence du climat, de l'air, de l'eau, du sol, sur la fréquence de certaines maladies.

Au cours des siècles, l'orientation « vers le corps » s'est imposée et c'est le développement, à partir du XVI^e^ siècle, de l'anatomie, puis de la physiologie. Plus tard viendra l'utilisation des concepts de la physique et de la chimie. Cependant, l'autre courant subsiste ; il se manifestera par exemple dans l'intérêt pour les « maladies étran-

1. On y lit, par exemple : « Pour le philosophe comme pour le médecin, il n'y a pas de définition satisfaisante de la maladie... elle est pourtant un fait social... », Professeur A. Lemaire, *Le Monde* du 24-XII-1960.

gères », tel qu'on le trouve chez Paracelse et renforcé, dit Sigerist (1951) [1], par la découverte de « terres nouvelles » où les colonisateurs rencontreront des maladies jusque-là inconnues. Plus tard, aux XVIII^e^ et XIX^e^ siècles, on verra l'éclosion d'une véritable « géographie médicale », les auteurs traçant, comme par exemple August Hirsch [2], une carte de la répartition des maladies dans l'espace et dans le temps. Aux mêmes époques, philosophes et médecins se pencheront sur le problème des liens entre maladies et conditions sociales. Les campagnes pour l'hygiène et l'amélioration des conditions de vie seront nombreuses [3]. Que ces approches des faits de maladie soient bien liées à une certaine vision de la nature du pathologique, un fait l'atteste : au cours du XIX^e^ siècle, avec le triomphe des idées pastoriennes, les géographies médicales sont abandonnées en même temps que la conception du pathologique change. Bien que l'existence de facteurs géographiques soit indéniable — il y a une géographie des maladies — ils apparaissent comme négligeables et la recherche s'absorbe dans l'étude de l'agent microbien lui-même. De même, les théories sur l'hygiène changeront de sens — l'accent n'est plus mis sur l'importance intrinsèque des conditions de vie — tandis qu'on se rattache aux théories pastoriennes du germe.

De nos jours se développe une réaction contre une conception exclusivement organiciste et « ontologique » de la maladie et elle prend la forme de *l'étude systématique des facteurs psychosociaux impliqués dans la maladie et dans la pratique médicale*. On peut, nous semble-t-il, y distinguer *trois courants* :

— Le premier, par son intérêt pour la nature même des phénomènes, *par ses buts préventifs et curatifs*, fait partie de la médecine : *la médecine psychosomatique*, quelles que soient parfois les ambiguïtés de sa définition, s'intéresse à la nature, à la genèse psychologique autant qu'organique de la maladie ; la médecine et la psychiatrie sociales tendent à prévenir et à guérir.

— Le second a pour thème : *la relativité culturelle* des conceptions et des comportements dans le domaine de la maladie et de la santé. C'est, par excellence, le domaine de l'*anthropologie* qui a su mettre en évidence, dans son étude des sociétés plus élémentaires que la nôtre, les répertoires culturels de mythes, de croyances et de pratiques médicales ainsi que leurs liaisons avec l'ensemble des valeurs de chaque culture.

— Le troisième courant affirme l'importance des *comportements sociaux* concernant la santé et la maladie, qu'il étudie dans notre

1. Les noms suivis d'une date renvoient à la bibliographie.
2. Cité par Sigerist (1951).
3. Cf., à ce sujet, R. Dubos (1961), chapitre v.

738

société. C'est là, nous semble-t-il, le point de départ de la *psychosociologie médicale*.

C'est à l'intersection des deux derniers courants, participant à la fois de l'optique anthropologique et de l'optique psychosociologique, que nous voudrions situer notre étude.

L'Anthropologie

Les études de la *médecine primitive* sont anciennes et nombreuses. Les motifs en sont divers : on a souvent pensé qu'il s'agissait d'un bon moyen d'approche des sociétés et de la pensée primitive elle-même. Leur but fut parfois pratique : faire accepter à des peuples qui l'ignorent la médecine et l'hygiène occidentales. Sigerist (1951) signale encore d'autres perspectives : intérêt historique ; la médecine des sociétés primitives encore existantes peut nous renseigner sur les origines de la pensée et des pratiques médicales modernes. On postule alors un parallélisme entre les sociétés primitives actuelles et anciennes. On s'est intéressé également aux croyances et pratiques médicales primitives dans la mesure où on les retrouve dans la médecine populaire et les superstitions des sociétés occidentales ; ce qui suppose une identité des mécanismes mentaux. Dans tous les cas cependant, les auteurs ont abordé, comme problème essentiel, l'étude des représentations de la maladie dans les sociétés qu'ils étudiaient.

Si la psychosociologie médicale se caractérise, comme nous le verrons, par un découpage des problèmes — et donc par une série d'apports réels mais — partiels entre lesquels les liaisons manquent parfois, les études anthropologiques se caractérisent au contraire par leur caractère *intégrateur* : face aux cultures qu'ils étudient, les auteurs ne dissocient pas les pratiques médicales des représentations qui contribuent à les orienter. En outre, ils relient celles-ci à d'autres conceptions et valeurs plus générales : conceptions de l'univers, des dieux, des rapports humains, du bien et du mal.

Le contenu des études est orienté par deux idées fondamentales. Premièrement, les représentations de la maladie, comme d'ailleurs les comportements des malades et de leur entourage ou les pratiques médicales, sont variables selon les cultures. Mais cette diversité peut être ordonnée selon quelques directions essentielles. Ainsi, on distingue classiquement des conceptions endogènes et exogènes de la maladie. Dans les conceptions exogènes — que l'on a appelées aussi centripètes et additives [1] —, la maladie est causée par l'intrusion, réelle ou symbolique, d'un objet dans le corps du malade. Dans les

1. J. P. Valabrega (1962).

conceptions endogènes, ou encore centrifuges et soustractives, la maladie est causée par le vol magique de l'âme de l'individu. Il semble que les théories exogènes soient davantage employées pour expliquer les maladies douloureuses et qu'on recoure aux théories endogènes pour des troubles accompagnés de perte de conscience. Les théories exogènes se rencontrent surtout en Amérique, en Australie et en Asie du Sud-Est ; le berceau des théories endogènes est la Sibérie. On a également fait des hypothèses [1] sur des époques d'apparition différentes, paléolithique et néolithique, des deux conceptions.

La seconde idée fondamentale est celle-ci : les représentations et pratiques sont liées au système global de valeurs de la société. La maladie peut être, a-t-on dit, d'origine magique : elle est causée par l'action d'un autre homme ou d'un sorcier. Ou bien elle est d'origine religieuse : elle provient d'un dieu ou d'un esprit. Mais le fait essentiel est l'absence de différences nettes entre médecine, magie et religion : les pratiques se rejoignent comme les systèmes de valeurs et de représentations se confondent. La maladie, en dernière analyse, provient d'une faute envers les dieux, les morts ou la société, de la rupture d'un tabou ou d'un manquement aux valeurs du groupe ; et, pour le malade lui-même, la maladie est une sanction. D'ailleurs, le malade, même s'il est rejeté, a souvent un caractère sacré et, de même, l'on sait que le « medecine man » ou le « chaman » joue, dans maintes sociétés, le rôle de prêtre ou de sorcier.

Plus récemment sont apparues des études un peu différentes, touchant aux problèmes d'acculturation et de résistance au changement, centrées sur l'introduction de l'hygiène et de la médecine occidentale au sein de cultures différentes. Les auteurs [2] nous montrent qu'il est souvent impossible de faire adopter une technique qui va à l'encontre d'une conception solidement établie dans une société, ou encore ils étudient comment s'enchevêtrent sans se détruire, dans les conceptions et pratiques médicales d'une population, celles héritées de la culture traditionnelle et celles de la médecine scientifique moderne. Dans tous ces cas, le problème de la liaison entre représentation et conduite s'impose.

On a souligné que certaines analyses des anthropologues pourraient s'appliquer avec profit à notre société. Ainsi, J. Stoetzel (1959) remarque qu'on peut classer historiquement les principales théories médicales occidentales en conceptions endogènes et exogènes de la maladie. Les conceptions exogènes prévalent dans certaines théories antiques : dans l'explication hippocratique de l'épilepsie (celle-ci a

1. Par exemple, F. Clements (1932).
2. Cf., par exemple, M. Mead (1953), L. Saunders (1954), C. M. Carstairs (1955), H. A. Gould (1957).

239

pour origine un phénomène externe, les vents), dans les théories sur les épidémies et la contagion qui se développent au Moyen Age et au XVII[e] siècle ; enfin, elles trouveront leur apogée avec Pasteur. Les conceptions endogènes sont centrées autour de notions de terrain, de disposition, d'hérédité. La maladie provient alors de l'homme lui-même. Canguilhem (1943) qualifie par exemple la médecine grecque de dynamique et totalisante. « La maladie n'est pas quelque part dans l'homme, elle est en tout l'homme et elle est toute entière de lui. »

D'ailleurs, certains travaux historiques incluent dans leur analyse les conceptions du normal et du pathologique. Parfois, sur le seul plan de l'histoire des sciences, comme l'œuvre décisive de G. Canguilhem (1943), où l'inspiration est tout à la fois historique et épistémologique. Mais d'autres auteurs — nous pensons à l'œuvre de Sigerist (1943, 1951) ou, en France, à celle de Foucault (1961, 1963) — dépassent le cadre d'une histoire des seules théories et découvertes médicales pour envisager les interrelations entre l'évolution des idées scientifiques et les conceptions des sociétés elles-mêmes, voire leurs pratiques (ainsi chez Foucault, l'internement des malades mentaux). Dans de tels travaux, l'histoire des sciences, par sa volonté de ne pas séparer la science des conditions sociales où elle se crée et de ses conséquences concrètes dans la société, inclut une anthropologie historique.

Pour l'époque actuelle, l'intérêt d'une « anthropologie moderne des faits de santé et de maladie » [1] paraît certain si l'on songe au pouvoir d'expression symbolique que ces sujets ont pour nous. Une simple réflexion sur le nombre d'œuvres romanesques qu'ils ont inspirées nous le confirmera ; de même le fait que la plupart d'entre elles entendent exprimer, par le truchement de la maladie, la condition humaine toute entière. On connaît aussi l'ampleur de la littérature populaire, le nombre de films centrés sur ces thèmes, en bref, l'existence de tout un « folklore ». Qu'une telle entreprise n'ait guère été tentée montre, une fois de plus, la difficulté qu'il y a à aborder l'étude de notre société d'un point de vue anthropologique.

La Psychosociologie médicale

Les premiers écrits de Psychosociologie médicale apparaissent aux États-Unis durant les années quarante, mais le domaine a connu depuis un développement considérable. Une enquête d'A. R. Strauss, en 1957, pouvait déjà dénombrer les travaux de 110 chercheurs se réclamant de ce domaine. Le but des premiers ouvrages apparaît le

1. Pour reprendre le terme de J. P. Valabrega (1962).

plus souvent comme pragmatique : contribuer à une « humanisation » de la médecine, par exemple en faisant acquérir au personnel hospitalier les rudiments d'une formation en sciences humaines [1] ; de là le caractère de généralité ou de pétition de principes propres à certains de ces ouvrages. Bien que les auteurs aient rapidement élargi leur perspective, on peut dire qu'ainsi qu'il est normal dans un tel domaine, la double orientation — vers la pratique ou centrée sur la recherche — subsiste encore aujourd'hui. C'est elle que recouvre, nous semble-t-il, la distinction par A. R. Strauss, en 1957, entre une sociologie *de* la médecine et une sociologie *dans* la médecine.

Dans de nombreux cas, les démarches et les concepts de la sociologie ont été transposés à l'étude de la médecine : ainsi, les *recherches épidémiologiques* s'attachent à préciser les rapports entre santé, maladie, conditions écologiques et facteurs sociaux [2]. De son côté, la sociologie des organisations trouve en *l'hôpital* un objet de choix ; l'étude des *professions médicales*, enfin, apparaît comme une branche de la sociologie des professions. Dans tous les cas, les apports sont loin d'être négligeables ; le besoin pourtant s'est fait sentir d'une articulation entre facteurs sociaux et aspects psychologiques ou psychosociaux. Ainsi en est-il pour les études épidémiologiques — dont on a souligné que, si elles peuvent établir des correspondances, elles ne peuvent en mettre en évidence les processus — et, en 1953, Ruesch tentait déjà d'étudier non seulement le rôle de l'appartenance à une classe, mais aussi ses rapports avec le type de technique sociale (dominance ou soumission) habituelle à l'individu. Des préoccupations du même type se font jour dans les travaux plus récents de Myers et Roberts [3] (1962) ou encore dans ceux de French (1962, 1966) : ce dernier cherche à déterminer les chaînons reliant environnement social objectif et santé, il souligne l'importance de la « réponse » de l'individu aux conditions écologiques et, par là, des variables psychosociales d'identité et d'estime de soi. Par ces affinements de l'analyse et des outils conceptuels, on peut espérer parvenir à élucider les mécanismes de la sociogenèse des maladies.

En revanche, l'étude de la *réponse à la maladie* — des comportements et des relations interindividuelles dans la santé et dans la maladie — a, d'emblée, été abordée dans des perspectives plus différenciées. Certes, nombres d'études de l'information, des attitudes ou des comportements des malades — ou des bien portants envers

1. Cf. l'ouvrage de Simmons & Wolf (1954), par exemple.
2. Ainsi ont fait, pour les maladies mentales, Faris et Dunham (1934), dans un travail désormais classique et, plus récemment, Hollingshead et Redlich (1958). En France, on connaît la synthèse de R. Bastide (1965) sur ces problèmes. Pour les maladies physiques, citons, entre autres, les études de Downes (1948), Logan et Brooke (1957).
3. Travaux qui font suite à ceux de Hollingshead et Redlich.

malades et maladies — restent purement descriptives : celles notamment qui tentent de recenser les variables, culturelles ou socioéconomiques, de telle connaissance, attitude ou comportement ; celles encore qui, utilisant des schèmes temporels, décrivent les diverses phases de la situation de maladie [1]. Mais les analyses de *l'organisation* de la réponse à la maladie, qui tentent de trouver un schème directeur aux attitudes et comportements divers qui la composent, peuvent se réclamer d'une double orientation théorique : les idées freudiennes d'une part, l'analyse faite par Parsons des rôles du malade et du médecin dans notre société, d'autre part.

La psychanalyse a introduit une conception de la maladie et du malade commandée par les idées d'évasion dans la maladie et de bénéfices, primaires et secondaires, qu'elle engendre. Elle transpose sur le plan individuel les problèmes de l'étiologie : s'évader dans la maladie, c'est s'évader des problèmes personnels et des contraintes sociales. La psychanalyse met donc l'accent sur la psychogenèse possible des troubles ; la détermination des maladies fait place à la motivation à la maladie. Ce sont de telles notions qui inspirent, par exemple, Balint (1957). Partant du rôle de la personne même du médecin, de la relation avec lui dans la guérison — le médecin, dit-il, est le médicament le plus couramment utilisé dans la pratique médicale — il a analysé brillamment divers mécanismes des relations qui se nouent au cours de la consultation médicale : les « offres » du malade, expressions diverses de son conflit, de ses tentatives pour le fuir, et la « fonction apostolique » du médecin qui leur répond. Sur le plan des comportements, celui du malade est dominé par l'introversion et la régression ; en font partie la dépendance vis-à-vis de l'entourage et l'exaltation du moi. Certains auteurs non psychanalystes — Barker (1946), Stoetzel (1960) — s'accordent d'ailleurs sur ces analyses.

A l'opposé de cette conception, Parsons (1952, 1953, 1958) définit la maladie comme « déviance » que la société doit contrôler. L'optique est apparemment antithétique de celle issue des idées psychanalytiques. Le malade est caractérisé non par son destin individuel mais par son statut dans la société ; la maladie n'est pas vue dans son déterminisme mais dans ses incidences, n'est plus une issue mais un problème à résoudre.

Cependant, les deux perspectives sont conciliables, voire complémentaires. C'est parce que l'évasion dans la maladie est une tentation pour chacun que la société doit exercer un contrôle sur les malades et leur entourage. Elle crée donc des rôles à tenir : celui du médecin a pour fonction de canaliser, chez le malade, la déviance que représente la maladie. Celui du malade combine les aspects régressifs, bénéfices secondaires autorisés (« l'exemption des responsabilités » et « le droit à l'aide ») avec des comportements plus « positifs » (le malade a des devoirs, il doit « désirer guérir » et coopérer avec le médecin en vue de la guérison). Pour Parsons, la maladie n'est donc pas seulement l'aboutissement d'un comportement régressif mais aussi une conduite d'ajustement. Le malade ne se définit pas seulement par ses manques par rapport à l'homme sain ; comme lui, il garde un rôle social.

L'articulation des concepts psychanalytiques et des concepts parsoniens paraît nécessaire. Il est difficile de réduire le comportement du malade à des conduites régressives et passives sans envisager, comme le fait Parsons, leur fonction positive dans un processus plus général d'ajustement à la situation : par son rôle défini au sein de la société, le patient s'adapte à la fois à la maladie et aux besoins de la société.

Mais on a également marqué [1] les limites de l'analyse de Parsons : valable surtout pour nos sociétés occidentales, limitée à un certain type et à une certaine étape de la maladie, celle du recours au soin médical. Surtout, bien qu'ayant inspiré maintes études empiriques [2], cette analyse semble confondre le plan de l'analyse sociologique des normes et des institutions et celui du vécu psychologique. Parfaitement pertinente quant au « point de vue de la société » sur la maladie, elle paraît moins adéquate quant à l'individu lui-même, malade dans la société, et échoue alors à rendre compte de certains de ses comportements. Ainsi, l'individu malade n'accepte pas toujours ce rôle qu'on veut lui imposer, il le refuse parfois et se fait, en quelque sorte, déviant dans sa déviance. Ce sont les « malades difficiles » qui ne se conforment pas aux anticipations du médecin, de l'entourage, de la société.

Ainsi, et dans l'analyse de Parsons et dans nombre d'études empiriques, *la relation même de l'individu à la santé et à la maladie ne nous paraît pas élucidée*. Dans l'enquête de D. Mechanic (1961), par exemple le concept de « rôle de malade » n'éclaire pas véritablement ce qu'il prétend atteindre : la relation spécifique de l'individu à son état. Il

1. Ainsi Lederer (1952) décrit différentes « phases » de la maladie (installation, maladie déclarée, convalescence), Balint (1957) distingue le stade inorganisé et le stade organisé dans la maladie, Barker (1946) utilise les concepts topologiques lewiniens, Goldstein et Dommermuth (1961) conçoivent le rôle du malade en termes de « cycle », etc.

1. Par exemple E. Freidson (1961-62).
2. Nous pensons par exemple au livre de R. C. Fox (1959) décrivant le comportement de malades graves soumis à des thérapeutiques expérimentales, aux travaux de D. Mechanic (1960, 1961) sur le rôle et le comportement de maladie; ou encore à différentes études de l'interaction du malade et de son entourage, par exemple Yarrow *et al.* (1955), Freeman et Simmons (1958).

se borne à établir une corrélation : adopte plus facilement le rôle de malade (c'est-à-dire consulte un médecin) l'individu plus enclin à se percevoir comme tel. D'une part, dans cette étude comme dans d'autres, le recours aux soins médicaux constitue une définition peut-être trop étroite du comportement de maladie ; d'autre part, l'étude des comportements n'est pas clairement reliée à la façon dont l'individu perçoit la situation et lui donne sens.

Ce problème est pourtant au cœur de certaines recherches. Freidson (1961), par exemple, étudie différents types de pratiques médicales à partir de la perception qu'en élaborent les malades. De même, Levinson et Gallagher (1964), dans une étude dont ils soulignent l'orientation sociopsychologique, se donnent pour but essentiel la conception que le malade mental a de son rôle et de sa situation à l'hôpital. Complémentaire de la description des comportements eux-mêmes, et pouvant en faire progresser l'analyse, une nouvelle question se pose donc : *si des attitudes se forment, si des rôles s'institutionnalisent, à quelles conceptions de la situation de « malade dans la société » répondent-ils pour l'individu ?*

II. — LA REPRÉSENTATION SOCIALE DE LA SANTÉ ET DE LA MALADIE : PROBLÉMATIQUE ET ORIENTATION MÉTHODOLOGIQUE

Dans l'étude de la réponse à la maladie, comme dans celle des problèmes épidémiologiques, le besoin d'une orientation psychosociale — centrée sur l'articulation de la personne et du système socioculturel — se fait jour. Autant que les facteurs sociaux dont dépendent santé et maladie, autant que les institutions et normes sociales les concernant, il nous faut étudier la perception que l'individu en élabore et le sens qu'ils prennent pour lui. Dans cette perspective, par l'étude de la représentation sociale de la santé et de la maladie, on tentera de comprendre les attitudes et comportements qu'elles engendrent, le savoir qui circule à leur propos, *dans la relation même qui se crée entre l'individu, la santé et la maladie.*

Depuis quelques années, les études de représentation sociale, longtemps abandonnées, connaissent en France un nouvel essor. Qu'il s'agisse de la représentation d'une théorie scientifique comme moyen de connaissance et d'action (la psychanalyse) [1], de celle d'un rôle

1. S. Moscovici (1901).

social (la femme) [1], ou d'un secteur d'activité particulièrement valorisé (la culture) [2], leur étude correspond à la nécessité de mieux comprendre la façon dont l'individu « construit » la réalité sociale et s'y oriente. Chaque auteur souligne, en effet, que la représentation d'un objet particulier fait partie de celle, plus vaste, de la société toute entière. A travers elle, c'est toujours la relation de l'individu à la société qui est en cause.

On doit noter d'abord le caractère complexe de la réalité psychologique que désigne le terme de représentation : concept et perception se fondent, images individuelles et normes sociales se rencontrent dans ce mode d'appréhension d'un objet ou de la réalité sociale. Dans le cas présent, par représentation sociale de la santé et de la maladie, nous entendrons l'élaboration psychologique complexe où s'intègrent, en une image signifiante, l'expérience de chacun, les valeurs et les informations circulant dans la société. Son champ n'est pas moins vaste : elle porte tout à la fois sur des états (ceux de santé et de maladie) et leurs critères, sur des personnes (le malade et le bien portant), leurs comportements et leurs rôles.

Cette image que l'on se fait du réel peut cependant s'étudier à un double niveau : elle est *saisie et évaluation d'une expérience* — immédiatement informée par les normes sociales — et, par là, elle oriente attitudes et comportements. Elle est aussi *construction d'une notion* : la conceptualisation de l'expérience, dont rendent compte les notions mêmes de santé et de maladie, représente la contre-partie, sur le plan cognitif, des modèles de comportement. C'est à une mise en correspondance de ces deux séries d'éléments, cognitifs et dynamiques, que l'on peut tenter de parvenir.

Il faut aussi marquer les liens étroits d'une telle représentation avec un mode spécifique de pensée sociale : *la science*. On connaît l'importance de la vulgarisation médicale dans notre société et l'avidité du public à son égard. Le savoir se transmet aussi par d'autres voies dont l'une des plus importantes est la consultation médicale. Nous traiterons pourtant la représentation sociale de la santé et de la maladie comme phénomène *sui generis*. De même que l'on a pu étudier la pénétration sociale d'une théorie scientifique — la psychanalyse [3] — comme phénomène autonome, dans son rôle de transformation du réel et non comme simple distorsion ou appauvrissement d'un modèle scientifique, de même, plutôt que de recenser les connaissances médicales, leur source ou l'exactitude de leur contenu, nous tenterons de mettre au jour l'image même du réel chez l'individu

1. P. H. Chombart de Lauwe *et al.* (1963).
2. R. Kaës (1966).
3. S. Moscovici (1961).

et, par là, le cadre de référence où l'information s'ordonne, les lignes de force selon lesquelles elle s'organise et prend sens pour les membres d'un groupe social.

Cependant, la représentation nous intéresse d'abord par son rôle dans la construction de la réalité sociale. Si l'expérience individuelle prend sens de par sa confrontation avec les valeurs et modèles culturels, si elle se coule dans les catégories d'une conceptualisation commune, par là même nous assistons à sa fusion en réalité sociale, ou plutôt à leur interpénétration en un phénomène unique à propos duquel communication, consensus et normes collectives peuvent s'établir. *L'étude devient alors celle de la santé et de la maladie, pour l'individu, dans la société.* Dans la représentation sociale se cristallise la double relation de l'individu à la maladie — ou à la santé — et à la société.

*

Notre étude porte donc sur l'élaboration même d'un objet social pour des membres de notre société. « L'approche du monde tel qu'il apparaît à l'autre »[1] est, dans cette mesure, au centre de ses préoccupations. Si donc s'y incluent des processus particuliers (ayant trait, d'une part à la genèse des conduites envers un objet social, d'autre part à son organisation sur le plan cognitif), qui sont du domaine de la *psychologie sociale*, elle souhaite cependant emprunter à *l'anthropologie* sa perspective intégratrice. De même, si l'observation se centre d'abord sur *l'observation d'un contenu* (celui que l'on attribue à la santé et à la maladie), elle se tourne aussi vers *l'analyse d'un phénomène*, vers les mécanismes par lesquels la représentation s'édifie et sa fonction d'orientation des comportements.

Une telle étude se veut d'abord suggestive, exploratoire. Cela ne signifie pas que l'on doive renoncer — du moins sur le plan du contenu même de l'image — à une problématique de départ qui, dans notre cas, s'est formulée au contact de la littérature psychosociologique et psychanalytique, ethnographique et historique sur la santé, la maladie, la médecine. L'expérience personnelle et la littérature romanesque y jouent aussi leur rôle. Ainsi se sont dégagées les principales directions de recherche :

— Les premières questions ont trait aux *critères mêmes des notions de santé et de maladie* et à leur organisation : quels sont les liens réciproques des deux concepts ? Sont-ils des concepts symétriques ? S'excluent-ils réciproquement ?

1. R. B. Mac Leod (1951).

— Nous nous demanderons aussi quelle est la *conception causale de la maladie, de la santé ?* Peut-on définir dans notre société une conception endogène ou exogène de la maladie ? Quelle nature attribue-t-on aux phénomènes du normal et du pathologique ?

— En troisième lieu, se pose le problème des rapports entre *les concepts de santé et de maladie et les valeurs.* La santé se confond-elle avec le bien ? La maladie est-elle le mal ? Provient-elle, de quelque façon, d'une faute ? Y a-t-il, dans notre société, responsabilité et culpabilité face à ces problèmes, et quel retentissement cela peut-il avoir sur les conduites du malade et du bien portant ?

— Comment conçoit-on l'incidence de la santé et de la maladie sur la participation à la société ? La santé est-elle une *norme requise par la société ?* La maladie est-elle perçue comme une *conduite déviante ?* La société apparaît-elle comme protectrice ou rejetante vis-à-vis des malades ? Quelles seront, dans cette perspective, les normes de conduite du malade et du bien portant ?

— Enfin, quelles relations perçoit-on entre *la santé, la maladie et la mort ?*

L'orientation même de l'étude obligeait à accorder au sujet une certaine liberté, dans la mesure où sa vision propre constitue l'objet même de l'observation : cela interdisait tout recours à des méthodes d'observation trop fermées et suggérait de choisir l'entretien libre individuel comme seule technique adéquate de collecte des données. Le matériel ainsi recueilli n'est pas celui d'une enquête classique, centrée sur le recensement exhaustif des opinions ou des comportements. Il ne peut pas non plus rendre compte du plus profond, du plus intime de l'expérience ; il n'atteint pas l'implicite, l'inconscient. Il nous apparaît comme le discours où l'individu explicite sous diverses formes (opinions et informations, sentiments, récits d'expériences et de comportements) son image de la santé et de la maladie. L'interprétation veut atteindre, dans ce discours, l'intégration de chaque élément en une image organisée, non un sens caché à l'individu mais le sens qu'elle a pour l'individu.

III. — PRÉSENTATION DE L'ÉTUDE

Il était hors de nos moyens de procéder à une étude extensive sur un échantillon représentatif de la population française. Là n'était pas non plus notre but : plus qu'à la description sociologiquement

exacte, plus qu'au dénombrement des phénomènes, nous voulons prétendre à l'analyse de contenus et de mécanismes psychologiques généraux. Dans un second temps, une enquête extensive pourrait permettre de valider le tableau que nous présentons et d'en apprécier la variation selon des critères pertinents.

La population interrogée se compose de 80 personnes, pour moitié *cadres, membres de professions libérales et « intellectuelles »* (sujets ayant tous fait des études supérieures), et pour l'autre moitié, membres des *classes moyennes* c'est-à-dire employés, petits fonctionnaires, artisans, petits commerçants ou femmes dont les maris appartiennent à ces catégories professionnelles (ces sujets ayant fait des études primaires ou secondaires). A l'intérieur de ces deux groupes, les sujets se répartissent en nombre sensiblement égal selon des variables d'âge et de sexe, c'est-à-dire hommes et femmes de 25 à 40 ans et plus de 40 ans [1].

Notre choix s'est porté sur ces deux groupes — classes moyennes et intellectuelles — en fonction du rôle dynamique qu'on peut leur attribuer, nous semble-t-il, dans la genèse, la transformation, le renouvellement des idéologies dans notre société. La présence d'un petit groupe de personnes habitant la campagne [2] constitue une tentative, certes très limitée, pour introduire dans l'observation une variable de « genre de vie » que l'on peut supposer importante dans les problèmes de santé et de maladie.

Nous n'avons pas choisi notre population en fonction d'une variable « malade-bien-portant », dans la mesure où nous ne voulions pas définir *a priori* la situation de malade et de bien-portant. Il nous est cependant possible de classer les sujets selon la nature et le degré de leurs expériences [3] : ainsi, 24 % des sujets nous ont dit « être actuellement malades » ; en outre, 13 % pensent avoir été gravement malades » dans le passé ; 20 % ont eu un ou plusieurs accidents.

Préalablement à la *collecte du matériel*, une série d'entretiens de technique non directive nous a permis de sélectionner les thèmes d'entretien pertinents par rapport à la population étudiée. Il s'agit de traduire, avec ajustements, les problèmes posés par le chercheur en guide d'entretien pour les enquêtés.

L'élaboration de ce guide s'est faite sous forme de thèmes à explorer, et non sous forme de questions à poser dans un ordre ou dans une formulation stricte. L'effort principal consistant à mettre au jour la problématique propre à chaque sujet, les différents thèmes d'entretien n'étaient pas imposés aux sujets mais abordés différentiellement quant à leur formulation et à leur succession [1].

1. Cf. Annexe I.
2. Le groupe des « classes moyennes » comprend un petit échantillon de 12 personnes habitant un petit village normand.
3. Cf. Annexe I.

Nous ne posions qu'une seule question de façon directe et standardisée à la fin de l'entretien [2] : « Pensez-vous possible qu'il existe un jour un monde où la maladie n'existerait pas, *un monde sans maladie ?* » Cette question « provocatrice » obligeait chaque sujet à se placer dans une optique « futuriste » et « déréelle » qui, pensons-nous, constitue une des optiques propres à éclairer les problèmes de la santé et de la maladie. Il suffit de penser à un certain nombre d'utopies, d'anticipations scientifiques où la santé et la maladie jouent un rôle important. Il semble que le matériel ainsi recueilli ait répondu à nos attentes.

Avec 20 personnes (c'est-à-dire le quart de la population interrogée), nous avons eu un second entretien au cours duquel nous leur demandions de commenter une retranscription du premier. Il s'agissait là de dépasser les aléas possibles de la situation d'entretien unique et d'étudier les retentissements d'une éventuelle « maturation » du problème chez les personnes interrogées [3].

Au total, notre matériel se compose donc de cent entretiens, chacun ayant eu une durée d'une heure et demie à deux heures environ. Si donc on ne peut, en aucun cas, assimiler notre travail à une étude extensive, nous croyons, en revanche, pouvoir insister sur le caractère assez intensif de l'observation et du matériel recueilli.

Le matériel recueilli pose des problèmes à l'analyse de par son étendue d'une part, de par son hétérogénéité, d'autre part. Celle-ci doit être acceptée comme nécessaire durant la phase de collecte du matériel. Lors de la phase d'analyse et d'interprétation, il faut cependant la réduire, trouver les unités d'analyse commune. Il faut déterminer d'une part les *objets* de l'analyse et d'autre part, *ses outils.*

Une première idée s'impose à nous : *l'analyse d'un tel matériel ne peut prétendre à l'exhaustivité, elle se doit donc d'être sélective.* Par là même s'opère la structuration et l'homogénéisation indispensables. Cette sélection, en fonction de la problématique définie et aussi d'une première approche clinique du matériel, s'est faite dans trois directions, et a défini trois principaux « objets » à l'analyse sur le plan du contenu lui-même.

1. On en trouvera la liste en annexe.
2. Cependant, l'idée d'explorer ce thème nous étant venue en cours d'enquête, la question n'a pas été posée à l'ensemble de la population, mais à 37 personnes seulement.
3. La technique nous a été en partie suggérée par les réflexions de certaines personnes à la fin de l'entretien, du type : « Je sens que je vais continuer à penser à ces problèmes ».

1. Les thèmes, concepts et constructions qui rendent compte de la *genèse* de la santé et de la maladie.

2. Ceux qui rendent compte des *définitions*, limitations, classifications des états de santé et de maladie.

3. Ceux qui rendent compte des *conduites* du bien portant et du malade, et envers la santé et la maladie.

A partir de là, les problèmes qui se sont posés à nous sont des problèmes classiques d'analyse de contenu. Ce sont :

1° *Le choix des catégories d'analyses et leur organisation.* — Deux types d'analyses nous ont paru utiles :

— L'analyse de l'apparition (« occurrence ») [1] de certaines notions ou thèmes permet d'obtenir une première image du matériel ainsi que les grandes lignes de différenciation du contenu.

— L'analyse de l'apparition simultanée (« cooccurrence ») de deux ou plusieurs éléments. Dans ce cas, on s'intéresse à l'existence ou à l'absence de *relation* entre deux ou plusieurs éléments ainsi qu'à leur nature. Par là, on peut espérer dépasser la simple description du contenu et aborder son organisation.

2° *Le choix des unités d'analyse.* — L'unité d'analyse la plus employée a été « le cas » lui-même, l'entretien dans son ensemble. En particulier, lors de l'analyse de la seule apparition de notions ou thèmes simples. Mais une unité plus fine s'est parfois révélée nécessaire : d'une part, dans le cas où il s'agit de comprendre la nature des relations entre deux éléments, d'autre part lorsque la fréquence relative d'un terme, d'une notion par rapport à une autre, le nombre de références qui y sont faites, peuvent paraître indicateurs de son importance dans la représentation.

En outre, nous recourrons à des analyses purement qualitatives sans données numériques. De même, la réflexion porte parfois sur l'analyse d'un cas individuel où l'intuition clinique joue alors pleinement son rôle. C'est dans la convergence de ces données, de ces différents indices que réside la possibilité de leur unification. C'est elle seule aussi qui peut assurer à notre observation quelque validité.

En effet, au stade actuel de la psychologie sociale où n'existe aucune méthode d'observation pleinement rigoureuse et satisfaisante, aucune unité méthodologique stricte ne nous paraît possible ni réellement souhaitable ; c'est à partir, en effet, de différents modes d'approche que la technique pourra progresser. Ils n'ont, au demeurant, qu'un seul but : nous faire réfléchir à l'image sociale de la santé et de la maladie et tenter de la comprendre.

1. Osgood C. E. (1959).

Nous remercions Monsieur le professeur J. Stoetzel qui a bien voulu diriger cette étude. Elle a été réalisée au Laboratoire de psychologie sociale de la Sorbonne, en liaison constante avec M. S. Moscovici, directeur à l'École Pratique des Hautes Études. Sans ses conseils et l'intérêt qu'il nous a manifesté, elle n'aurait pu être menée à bien.

Nous remercions également Mademoiselle M. Gluge et Madame M. Fichelet qui nous ont aidé à recueillir le matériel ainsi que Madame Sella qui a participé à son analyse. Enfin, nous sommes reconnaissants à Monsieur Fouilhé de nous avoir procuré les moyens financiers nécessaires à cette étude.

L'UNION MEDICALE DU CANADA. Vol. 109, Mai 1980, No 5.

Tribune éditoriale

Gilles Bibeau
Jacques Tremblay

Résistance et soumission, face à l'émergence d'une autre conception de la santé.

Deux textes, qui précèdent, parlent de promotion de la santé, de prévention de la maladie et de la modification de comportements. Le premier, celui de M. O'Neill, se présente sous la forme d'une revue critique de la littérature, qui vise à familiariser les professionnels de la santé avec les principales "théories du changement de comportement"; le second est plus limité : il présente les stratégies généralement utilisées dans les programmes d'éducation physique, stratégies qui sont incitatives, correctrices et essentiellement dirigées vers les individus, comme le démontrent les auteurs. Globalement, on semble en être resté dans ces programmes à une conception classique de la pédagogie Individuelle, qui gratifie pour les bonnes actions, corrige les inadéquates et pénalise pour les mauvaises.

M. O'Neill signale dans sa contribution les limites inhérentes aux diverses conceptions du changement de comportement en matière de santé; notre réflexion veut justement prolonger cette dimension critique sur trois points particuliers, celui du rapport entre comportement individuel et norme culturelle, celui du niveau souhaité de changement, en surface ou en profondeur et, enfin, celui des rapports entre gestion autoritaire des comportements personnels par les professionnels et appropriation du savoir par la population.

Le comportement humain ne peut être valablement compris, dans sa dynamique, que s'il est simultanément saisi dans son enracinement individuel immédiat et, plus profondément, sur l'horizon du système de valeurs culturelles du groupe auquel appartient la personne. L'accent mis par de nombreux programmes sur la responsabilité individuelle, postulant que l'addition de nombreux changements individuels va provoquer une mutation de la structure, oublie que les interventions de santé publique sont, d'abord et avant tout, un exercice de changement culturel. Bien sûr, l'adoption de nouveaux comportements de santé implique pour la personne individuelle un rapport nouveau à certaines valeurs mais, l'adoption généralisée des nouvelles valeurs pose en fait, "plus fondamentalement", un problème au niveau culturel, c'est-à-dire, celui de la concurrence des "innovations" avec les caractéristiques centrales de la civilisation à laquelle appartient la personne. C'est, en effet, dans les "valeurs pivots" de notre civilisation capitaliste avancée que les habitudes néfastes à la santé, la suralimentation, le tabagisme, le bruit, la vitesse, les paradis de l'alcool et de la drogue, les stress structurels de la vie quotidienne, la pollution des lieux de travail et de l'environnement, vont puiser leurs racines les plus profondes. Ce serait une erreur grave d'oublier que nous sommes ici en face d'épiphénomènes qui tirent leur existence de la structure culturelle, qui leur sert de support et qui les génère. Bien des auteurs, par exemple, ont montré les liens intimes qui relient la surconsommation de médicaments et, une certaine conception de la technologie et des énergies fortes.

Un exemple permettra de mieux comprendre pourquoi il importe de poser le problème de changement sur l'horizon de la culture en direction de laquelle il faudra redéployer nos efforts d'éducation. Nous empruntons cet exemple au livre de V.F. Fuchs, *Who shall live*[1] qui contraste les états de santé des individus habitant les deux états voisins du Utah et du Nevada dans l'Ouest des Etats-Unis. Les statistiques de morbidité situent ces deux états aux deux extrémités: les habitants de l'Utah sont parmi les Américains les plus florissants, alors que leurs voisins du Nevada sont parmi les moins bien portants: un exemple entre mille, la mortalité des enfants blancs du Nevada est de 40% plus élevée dans cet état que dans l'Utah. Pourtant les similitudes ne manquent pas entre ces deux états: sensiblement le même climat, la même scolarisation, la même urbanisation mais aussi, la même médicalisation et, à peu près un budget identique engagé dans la promotion de la santé. En fait, un épidémiologiste ne peut rien comprendre à cette différence de morbidité, aussi longtemps qu'il n'a pas comparé les deux états du point de vue culturel. S'il le fait, il constatera la chose suivante: l'Utah est une région fortement influencée sur le plan culturel par la religion mormonne, qui y domine, maintenant un fort puritanisme, rejetant le tabac et l'alcool et, favorisant la structure familiale. Le Nevada, son voisin, se présente au contraire comme le paradigme d'une certaine "american way of life" avec Las Vegas, la capitale du jeu et Reno, dite la capitale du divorce. Notre intention n'est pas de porter un jugement de valeur sur les styles de vie caractéristiques des deux états, mais uniquement de faire la démonstration de la corrélation directe existant entre les valeurs culturelles d'un groupe et les comportements individuels reliés à la santé.

L'éducation à une autre santé ne peut donc se faire pratique dans nos sociétés occidentales avancées que dans le changement socio-politique. L'éducation à la santé pose par conséquent le problème transformation de notre culture, dans sa conception du travail et des milieux de production, dans celle de l'habitat et des groupements humains, dans le développement d'une autre relation à l'environnement et au corps, nous semble se situer dans nos sociétés davantage du côté des mobilisations populaires nées

de la base, que dans des projets gouvernementaux réformistes. Non seulement il ne nous semble pas possible d'isoler les nouveaux comportements de santé du développement d'une nouvelle culture, dont il faut mettre en place les fondements, mais il nous semble de plus que le travail des éducateurs de santé doit s'inscrire de manière privilégiée dans les mouvements sociaux et les nouvelles mobilisations qui travaillent à la construction d'une alternative à la vie actuelle. Nous sommes conscients du fait que cette réorientation de la culture occidentale vers des valeurs fondamentales de vie porte incohativement en soi le risque d'une certaine oppression des libertés individuelles mais, nous n'en sommes pas encore là, et ce problème, qui deviendra central pour les systèmes sanitaires des prochaines décennies, aura alors amplement l'occasion d'être discuté.

Dans le prolongement de la réflexion sur les rapports entre comportements individuels et normes culturelles, nous signalons que la culture existe dans notre société sous la forme de sous-cultures relativement autonomes. Les épidémiologistes ont mis clairement en évidence la nature des corrélations existant entre les formes de morbidité et les catégories socio-économiques, en montrant spécialement la prévalence de nombreuses maladies dans les secteurs les plus défavorisés de la population. Spontanément, les programmes d'éducation à la santé ont été dirigés vers ces catégories sociales mais, ici encore, les animateurs de santé ont souvent oublié que ces milieux possèdent une sous-culture qui ne reproduit pas nécessairement toutes les valeurs de la culture de la classe moyenne. L'éducation à la santé en tant qu'exercice de changement culturel devient un véritable défi, lorsque les éducateurs n'appartiennent pas à cette sous-culture ou, qu'ils l'ignorent le code.

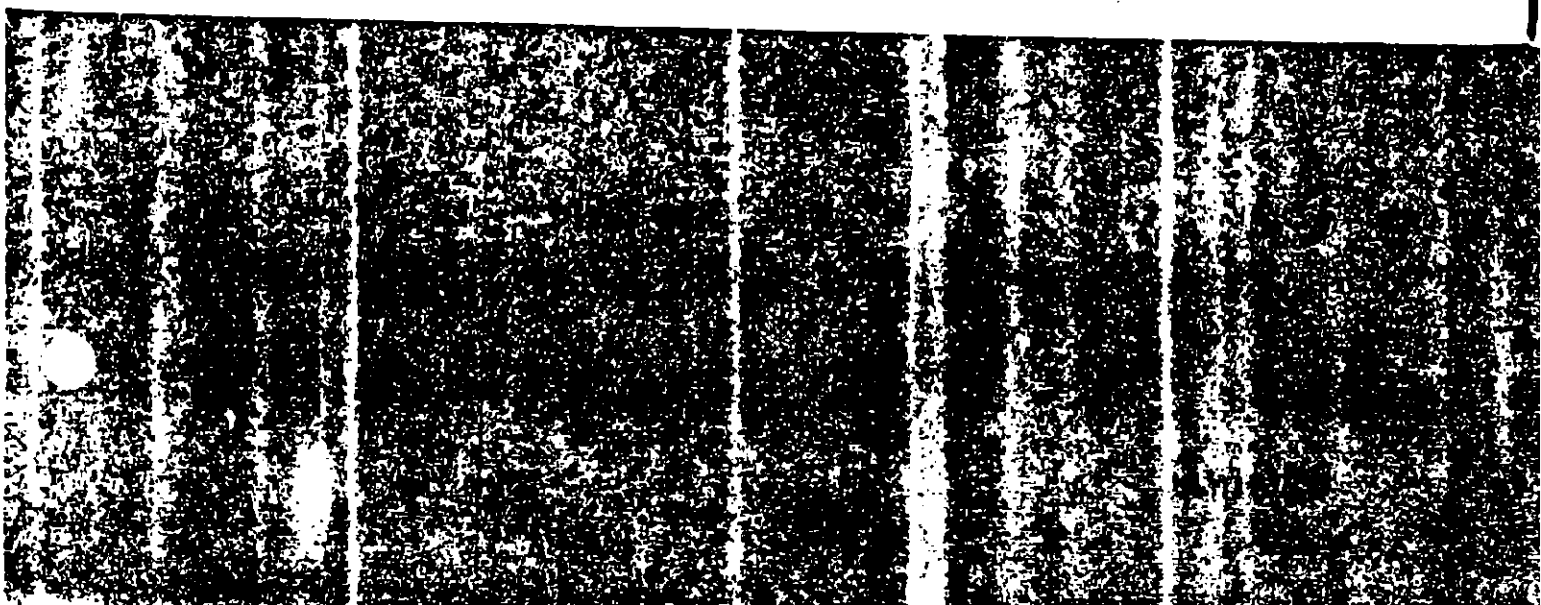
Notre discussion du second point, celui des rapports entre changement superficiel et changement profond, sera courte. Bien des changements, que l'on avait crus permanents se sont révélés éphémères et sans lendemain. Plus clairement que dans bien des travaux qui inspirent communément les éducateurs de santé, le groupe de Palo Alto, structuré à ses débuts autour de l'anthropologue G. Bateson, a réexaminé fondamentalement le problème du changement, ses paradoxes et ses contradictions et, a proposé de distinguer deux niveaux dans ce phénomène: le changement 1 et le changement 2.[2]

Le changement 1 est celui qui s'effectue totalement à l'intérieur de la structure existante, permettant à la personne de composer avec la situation en propos ou un nouveau comportement, qui est accommodation. Dans ce cas, on peut dire sans erreur : "Plus ça change, plus c'est pareil". Le changement 2 implique une sortie du système et, un réaménagement du contexte dans lequel s'inscrit la structure. Il s'agit, dans le langage des théoriciens de la communication, d'un méta-changement. Le changement 2 implique donc le changement de la structure dans laquelle se situe une personne en tant que totalité et, cette réflexion rejoint ce que nous avons dit plus haut concernant la nécessité d'enraciner le changement de comportements individuels dans celui d'une modification de l'ensemble des valeurs culturelles.

Notre dernière réflexion concerne le problème de l'appropriation du savoir par les non-spécialistes. Les échecs nombreux des programmes d'éducation à la santé et, l'absence d'impact d'actions correctrices, qui étaient par ailleurs scientifiquement valables et présentées sous la forme d'une excellente vulgarisation, s'expliquerait par le peu de considération accordée à ce que l'on appelle, de plus en plus, les "savoirs populaires". Il existe à côté des savoirs scientifiques que les éducateurs s'efforcent de traduire dans des termes compréhensibles, un autre savoir, dont il faut comprendre l'articulation et la logique interne, celui des non-spécialistes. Il ne suffit pas de populariser et, de rendre accessible la science qui explique pourquoi tel comportement est meilleur pour la santé, mais il faut que le travail des éducateurs de santé s'intéresse davantage à la manière dont les personnes expliquent leurs habitudes dans le domaine du manger, du boire, de l'exercice, etc. Une autre stratégie d'éducation à la santé devrait alors naître : celle qui se construit sur l'articulation des savoirs scientifique et populaire, à l'intérieur d'une nouvelle conception du changement socio-culturel.

Bibliographie

1. Fuchs, V.F.: Who Shall Live? Health Economic and Social Choice, New York : Basic Books Establishers, 1975.
2. Watzlawick, P., Weakland, J., Fisch R. : Changements. Paradoxes et Psychothérapie, Paris: Seuil, 1975.



LECTURE POUR JEUDI LE 9 JUIN 1983

A Michel
en bien amical hommage

249

TOME XI N° 3 1982

L'Espace Géographique, n° 3, 1982, 161-175.
Doin, 8, place de l'Odéon, Paris-VI°.

Nature et société

GÉOGRAPHIE MÉDICALE GÉOGRAPHIE DES MALADIES GÉOGRAPHIE DE LA SANTÉ

Henri PICHERAL

Université Paul Valéry, Montpellier

ENVIRONNEMENT
ESPACE (ANALYSE DE L')
GÉOGR. MÉDICALE
GÉOGR. DES MALADIES
GÉOGR. DE LA SANTÉ

RESUME. — Du fait du glissement du concept de maladie à celui de santé, la géographie médicale s'est orientée dans deux directions : la géographie des maladies et la géographie des soins de santé. La géographie de la santé peut rassembler ces deux tendances à travers l'analyse régionale dans une recherche spatiale et interdisciplinaire sur la qualité des soins en relation avec toutes les variables de l'environnement.

ENVIRONMENT
GEOGR. OF DISEASES
GEOGR. OF HEALTH
MEDICAL GEOGR.
SPATIAL ANALYSIS

ABSTRACT. — *Medical geography, geography of diseases, geography of health.* — Discrepancies between definitions of illness and health are reflected in a bifurcation of research directions within medical geography : the geography of disease on the one hand, and the geography of health care on the other. A geography of health could re-integrate these two tendencies through an interdisciplinary spatial approach to regional analysis on the quality of health care in relation to all environmental variables.

La naissance officielle de la géographie médicale date de 1949, au Congrès de l'U.G.I. à Lisbonne. Mais c'est à New-Delhi en 1968 qu'elle reçoit sa vraie consécration. Huit ans plus tard, la Commission changeait d'appellation au Congrès de Moscou et devenait un « groupe de travail de géographie de la santé ». Pourquoi un nouveau baptême ? pouvaient se demander les géographes encore mal convaincus de la nature « géographique » de cette nouvelle branche ou plutôt de ce nouveau greffon. Etait-ce bien de la « géographie » ? En quoi et pourquoi un géographe aurait-il une quelconque compétence en matière de maladie ou de santé ?

L'idée n'est pas neuve pourtant. Elle a même plus de trois mille ans ! Mais les géographes ne s'y intéressent que depuis une cinquantaine d'années. Guère plus. Dans ce décalage réside sans doute la première interrogation d'une discipline qualifiée récemment d'« enfant adoptif » [4]. Fille de la géographie ou de la médecine ?

I. MALADIE ET SANTÉ, MÉDECIN ET GÉOGRAPHE.

1. *Pathologie géographique ou anthropologie médicale ?*

Force est de reconnaître la légitimité des médecins quand ils en revendiquent la paternité. Car on leur doit les premières descriptions des relations entre la présence ou la fréquence des maladies et les éléments du milieu. Hippocrate et ses disciples sont du nombre, bien sûr, mais déjà bien avant eux les médecins de l'école Ayurvédique de l'Inde (I-II° millénaire av. J.C.) [97]. Les médecins arabes, comme Avicenne, apportent leur lot d'observations empiriques avant qu'une multitude discrète de « topographes médicaux » ne brosse avec une précision d'arpenteur le tableau des conditions sanitaires des villes et des

campagnes de l'Europe du XVIII[e] et du XIX[e] siècles. Et parmi eux, un médecin allemand, L.L. Finke aborde le premier en 1792-95 les rapports entre géographie et maladie [26]. La première mention d'une « géographie médicale » et sa première formulation reviennent pourtant à un médecin français, J. Ch. Boudin qui publia en 1843 un ouvrage sous le titre : *Essai de géographie médicale, ou étude des lois qui président à la distribution géographique des maladies ainsi qu'à leurs rapports topographiques entre elles, lois de coïncidence et d'antagonisme* [12]. Le terme plaît et connaît un succès foudroyant si l'on en juge par la floraison de traités et de manuels de *medical geography, medizinische geographie, geomedizin* et autres *nosogéographie, pathogéographie* et *medicinskaja geografia*. Sans doute cette vogue répond-elle à l'air du temps, au goût de l'encyclopédisme et de l'exotisme renouvelé par les conquêtes coloniales. Mais elle se situe surtout dans le sillage des découvertes pastoriennes, les succès de l'hygiène, la foi dans le Progrès et la Science. Il ne se passe plus d'année jusqu'en 1914 sans la sortie d'un nouvel ouvrage des presses de Londres, Paris, Berlin, Stuttgart, Leipzig ou Saint-Petersbourg. Le plus important, le plus solide aussi, s'avère le *Manuel de pathologie géographique et historique* de A. Hirsch (1860-64) [32].

Le milieu du XIX[e] siècle marque une étape importante dans la démarche des médecins. Jusque-là ils se sont attachés à rechercher des lois régissant les relations entre la maladie et le lieu par la juxtaposition d'observations locales et localisées sur les effets du climat ou de la topographie. Désormais ils se demandent « quelle est la production, la marche et la répartition des maladies » tant est manifeste « le besoin de comparer les phénomènes qui se produisent dans notre *milieu* avec ceux du même ordre qui se produisent dans les autres pays » [11]. Le Dr Bordier sait de quoi il parle : il est alors en 1884 professeur de géographie médicale à l'Ecole d'anthropologie de Paris. Lui et ses correspondants à l'étranger cherchent une vision globale de la distribution des maladies à différentes échelles : le monde, l'Etat, la région (la province), la ville... Et en toute logique ils font appel à la carte.

La première carte « médicale » date sans doute de 1832 (Shapter) suivie de près par celle de J. Snow en 1855. L'une et l'autre sont des cartes par points, localisant avec précision le domicile des malades lors des grandes épidémies de choléra à Exeter et Londres. Mais Snow va plus loin en situant sur une carte la fameuse fontaine publique de Broad Street responsable de la diffusion de la maladie. Il s'agit sans nul doute de la première carte étiologique. Mais bientôt de véritables atlas apparaissent... sans aucune participation de géographes. La seule exception concerne Peterman qui profite d'un séjour en Angleterre pour publier en 1852 une carte de la répartition du choléra en 1831-33 dans les Iles Britanniques [98,28].

Car à la veille de la première guerre mondiale, la géographie médicale demeure une exclusivité médicale. Etablie et mise en forme par des médecins à l'usage des médecins. Finke, Schnurrer, Boudin, Fuchs, Mühry, Haviland, Lombard, Hirsch, Bordier ou Clemow, tous, sans exception, sont médecins [4, 51]. Dans le débat qui anime de multiples académies de médecines et d'innombrables sociétés savantes dans toute l'Europe, les géographes sont absents et curieusement muets. Ce serait faire injure à Ritter, Ratzel, Reclus ou Vidal de la Blache que de croire à leur ignorance de la question. Ils s'en désintéressent pourtant avec superbe. Leur silence surprend car les faits avancés par les médecins auraient dû servir d'arguments aux tenants du déterminisme et de l'environnementalisme, ou de l'anthropogéographie [18]. L'on s'étonne d'ailleurs que Miss Semple n'ait pas émis quelque jugement catégorique à propos du rôle de la marée sur le destin de la peste ou du choléra ! Les pères de la géographie humaine du XIX[e] siècle se seraient-ils rangés à l'avis péremptoire et sans appel du Dr. Navarre, parus dans un article des *Annales de Géographie* de 1904 : « La géographie médicale, un bien grand mot pour une petite chose... (elle) n'est pas une science. On ne peut donner ce nom à une collection de faits instables et contingents d'où ne jaillit aucune idée générale » ? (1).

2. *La géographie médicale selon M. Sorre.*

Mis à part l'intérêt mesuré accordé par J. Brunhes à cette géographie-là en 1925, c'est bien Max. Sorre qui établit les véritables fondements de la géographie des maladies, même s'il conserve le terme de géographie médicale [99, 100]. Un géographe trace enfin avec rigueur le champ et les limites de la discipline. Il forge un nouveau concept, le « complexe pathogène », compréhensible par ses pairs et encore d'actualité cinquante ans plus tard. Du moins auprès des spécialistes des maladies transmissibles, de la pathologie exotique notamment, et d'une façon plus générale, des éco-pathologistes et des épidémiologistes.

Car à la suite des travaux des « hygiénistes » et des apôtres de la « médecine sociale » du XIX[e] siècle, une sorte de clivage s'instaure entre médecine curative et médecine préventive, entre fondamentalistes, cliniciens et épidémiologistes. Les uns s'adressent à l'individu, les autres à la collectivité. Les uns ne parlent que du malade quand les autres s'interrogent sur la santé publique, sa surveillance, son contrôle. Ces attitudes divergent et divergeront longtemps, rendant de plus en plus inconfortable, sinon ambiguë, la situation du géographe.

La démarche de Sorre était en effet tributaire de son temps. L'objet prioritaire de l'épidémiologie, comme le souligne l'étymologie, était bien alors l'étude des maladies épidémiques, infectieuses, parasitaires et tropicales en particulier. Aussi M. Sorre puise-t-il abondamment dans la littérature médicale de l'époque et oriente-t-il délibérément la géographie médicale dans ce sens. Parce que la répartition des

(1) Voir, sur tout ce qui précède, la préface de L. et G. Masse, in Jenicek [44].

maladies transmissibles sert son propos et sa démonstration et que, seule à ses yeux, cette pathologie entretient des liens étroits avec le milieu. C'est-à-dire le milieu naturel, physique et biologique. Dans cette optique son système apparaît très cohérent et fort novateur. Les aires pathogènes se moulent en effet dans un cadre spatial rigoureux, défini par les limites mêmes du « complexe ». En revanche, Sorre se refuse à toute analyse géographique de la pathologie non infectieuse. Ne déclare-t-il pas en 1943 : « du cancer nous ne dirons rien parce que sa nature infectieuse est sujette à controverse et surtout parce qu'il ne nous paraît pas possible d'admettre une relation entre les conditions géographiques et la distribution du cancer. » (2). Voilà qui est clair, net et tranché... bien que Sorre ait par la suite tempéré son jugement (3).

Cette prise de position catégorique s'avère déterminante pour l'essor de la géographie médicale. Ce choix dicté pour l'essentiel par les connaissances médicales de l'époque, a bien failli compromettre son avenir.

D'abord aux yeux des géographes eux-mêmes qui taxeront cette géographie-là d'un déterminisme excessif, par trop tourné vers les facteurs naturels de la maladie et négligeant ou niant ses composantes sociales (A. Demangeon, F. Braudel, P. George...). On dénoncera alors sans peine l'incompétence du géographe en matière médicale et sa propension à l'encyclopédisme et à l'expansionnisme. Ce qui justifiera le jugement récent d'A. Reynaud : « voilà la médecine devenue une des sciences annexes de la géographie » (4). Quant aux médecins, leur attitude est plutôt réservée, du moins pour la plupart d'entre eux. Car la démarche *géographique* « n'expliquerait » finalement que la partie visible de l'iceberg. L'épidémiologie moderne n'embrasse-t-elle pas toute la pathologie, y compris les maladies chroniques, sociales, mentales, voire héréditaires ? Dans la mesure aussi où la méthode descriptive et analytique apparaît désormais insuffisante, voire décevante et de portée limitée. Et puis, à quoi bon s'attarder sur l'étiologie des maladies quand se profile l'avenir triomphant des sulfamides et des antibiotiques ? En définitive, le résultat le plus tangible demeure l'opinion de la plupart des médecins à l'égard de l'objet et de la nature de la géographie. Elle se résume à des noms de lieux, à l'analyse des longitudes et des latitudes, des températures et des tranches d'eau, à la rigueur du sol et du sous-sol. Mais avant tout, le géographe n'est à leur avis qu'un faiseur de cartes.

En schématisant à l'extrême, disons que deux voies s'offrent alors aux géographes. Dans la lignée de Sorre, les uns développent, seuls ou en collaboration avec des médecins, des recherches d'écologie médicale sur les maladies transmissibles. Outre de multiples monographies, on leur doit plusieurs travaux d'une qualité exceptionnelle. L'école de médecine allemande se distingue tout particulièrement avant, pendant et après la guerre avec W. Rimpau et H. Zeiss et surtout la prestigieuse équipe réunie autour d'E. Rodenwaldt et H. J. Jusatz. Naissent alors de splendides atlas (*Seuchen Atlas* en 1943-45 et *Welt Seuchen Atlas* en 1952-60) [88, 89]. Aux Etats-Unis, après la somme de McKinley [60] sortent les 17 livraisons du *World Atlas of Diseases* publiées par J. May, de 1950 à 1955 dans la *Geographical Review*. Ce médecin français, émigré aux Etats-Unis, a des conceptions très proches de celles de Sorre, qu'il va développer et élargir dans deux ouvrages fondamentaux : *Ecology of Human Diseases* et *Studies in Disease Ecology* [63, 65]. Leurs titres ne laissent pas de doute sur l'orientation choisie. Une tendance très voisine se manifeste en Union Soviétique avec Pavloski, Sosin et Ignatev dont le laboratoire de géographie médicale dépend directement, notons-le, du département de biogéographie de l'Université de Moscou.

Dans le même temps, d'autres géographes, surtout anglo-saxons, s'engagent derrière les « néo-épidémiologistes » et adoptent leur démarche, fondée sur des méthodes statistiques et mathématiques de plus en plus sophistiquées. Leur champ d'étude couvre toute la pathologie, de la grippe à la schizophrénie, du choléra à l'infarctus du myocarde ou du cancer du foie aux accidents de la route. Leur démarche ne diffère guère de celle des médecins, sinon que les uns appellent géographie « médicale » ou « pathologique » ce que les autres nomment « pathologie géographique » (l'ordre des termes n'est pas neutre...). Mais le géographe s'efface derrière l'épidémiologiste à force de vouloir s'identifier à lui.

De fait, la géographie médicale n'est toujours pas émancipée. Elle reste dans le sillage des grandes orientations de la recherche étiologique dont elle n'est qu'une des approches. Une parmi d'autres. Pourtant, en une vingtaine d'années, elle bénéficie d'une triple opportunité et d'un changement des mentalités à son égard.

3. Le tournant des années 50-60.

Peu à peu est admise et se répand une nouvelle conception de la « santé » qui bouscule les habitudes. Selon le préambule de la constitution de l'Organisation Mondiale de la Santé, en 1946, la santé est un « état de bien-être complet, physique, mental et social et non pas simplement l'absence de maladie ou d'infirmité ». L'événement est de taille. Aussi ambitieux et sans doute utopique soit-il, ce postulat a le mérite de sortir la « santé » de l'orbite strictement médicale. De négatif, le concept devient positif. Il acquiert surtout une dimension culturelle et sociale toute nouvelle.

Cette attitude se trouve renforcée au début des années 60 par l'émergence de la notion d'environnement et de son corollaire, la qualité de la vie. La santé en constitue un champ d'application privilégié tant elle a valeur de symbole. On s'interroge sur l'inter-

(2) P. 405-406 de Sorre [100].

(3) P. 1089 à 1106 de Journaux [45], plus précisément p. 1091.

(4) P. 18 de Reynaud [86].

face santé/environnement. Notamment R. Dubos qui publie coup sur coup *Man, Medicine and Environment* et *L'homme et l'adaptation au milieu* [20, 21]. Eminent biologiste et humaniste, R. Dubos est sans doute un géographe qui s'ignore ! A ses reflexions qui justifieraient à elles seules le bien-fondé de la géographie « médicale », répondent celles d'un géographe : *Man, Environment and Diseases* et *Environmental Medicine* de M. Howe [37, 38]. Sans que le mot soit explicitement prononcé (ferait-il encore peur ?), le déterminisme retrouve une nouvelle jeunesse. Mais il ne s'agit plus désormais du seul déterminisme naturel mais de déterminisme social, économique, culturel... La santé est maintenant comprise comme un tout, comme un système et non comme les relations du couple antithétique santé/maladie.

Par ailleurs, les pouvoirs publics réalisent que la santé a un prix. La montée vertigineuse des dépenses force l'Etat à considérer la santé en termes de coûts, d'investissements, de choix. Et comme il devient clair que chacun n'est pas égal devant la maladie et la mort et que l'accès aux soins demeure bien mal partagé, la nécessité de réduire les disparités spatiales et les inégalités sociales devient elle-même évidente. Dès lors la santé devient l'objet d'une véritable politique.

Ainsi, en quelques décennies, le champ d'étude s'est tellement élargi que le médecin n'est plus capable de le maîtriser à lui seul. Son monopole millénaire en matière médicale en est ébranlé, sinon menacé. La scène s'est enrichie de nouveaux acteurs car à ses côtés figurent désormais le sociologue et l'économiste, le psychologue et le démographe, le juriste et l'hydro-géologue, l'historien et l'anthropologue, le bio-physicien et l'ethnologue, le généticien et le... géographe. Tous jouent en fait la même pièce : l'état de santé de la population ou plutôt des populations. Chacun conçoit sa contribution à sa manière, selon son tempérament même si le médecin tient à conserver le premier rôle, ce qui paraît légitime compte tenu du sujet !

Le géographe « médical » a tout à gagner de ces nouvelles conditions. Sans savoir au juste ce qu'est la géographie, chacun sent confusément la nécessité d'intégrer des variables géographiques dans l'analyse des phénomènes de santé. Par ailleurs ces préocupations rejoignent les questions et les courants qui animent à partir des années 60 la communauté géographique [18]. Le domaine de la santé offre un champ d'application remarquable à la « nouvelle géographie », aux méthodes quantitatives, à la cartographie automatique, aux fécondes approches d'analyse de l'espace, à l'étude des processus de diffusion et des comportements... (P. Haggett, J. Hunter, N. McGlashan, M. Howe, G. Pyle...).

Dès lors, il apparaît que toute référence spatiale en matière de santé est du ressort du géographe. Mais sa contribution ne peut plus se limiter à la géographie des maladies, à leur répartition et à leurs causes (*geography of diseases*) : celle-ci demeure un sujet privilégié mais qui doit de plus en plus à l'épidémiologie. Son approche devient cependant plus originale quand il s'attache à évaluer l'impact d'une situation morbide sur les activités et le développement d'une population dans un espace donné. Elle s'avère beaucoup plus spécifique encore quand elle cherche à mesurer l'ajustement de l'offre et de la demande de soins, que celle-ci soit effectivement déclarée ou ressentie ou simplement perçue. L'analyse spatiale du système de distribution de soins et des services de santé, de l'accès aux soins et de leur « consommation », entrent de plain-pied dans le champ d'étude du géographe. Enfin la recherche d'indicateurs capables de cerner le niveau de santé d'une population urbaine ou rurale, riche ou pauvre, nationale ou régionale, relève aussi, à l'évidence, des préoccupations du géographe.

Aussi semble-t-il préférable de parler de géographie de la santé (*geography of health*) plutôt que de géographie médicale [113, 59, 43, 81]. S'il faut à tout prix proposer une énième définition, peut-être pourrions-nous dire simplement qu'il s'agit de l'étude spatiale de la qualité de la santé dans ses relations avec l'environnement physique, biologique, socio-économique, comportemental et culturel. Nul doute dans cette optique qu'il s'agisse d'une science de l'homme. Nul doute non plus qu'elle participe largement de la géographie sociale et qu'elle doive beaucoup aux autres sciences humaines, la sociologie médicale en particulier (et qu'il faudrait sans doute appeler aussi sociologie de la santé) [103, 72]. Nul doute enfin sur le caractère volontariste de la géographie de la santé qui se situe dans le droit fil d'une géographie appliquée et applicable. N'est-ce pas le sens du dernier ouvrage de G. Pyle *Applied Medical Geography* [82]... même s'il conserve la terminologie chère aux anglo-saxons ? Car, en dernière analyse, la « santé » n'est-elle pas la condition première de tout épanouissemnt de l'homme ?

Vaste programme ! Utopie, prétention ou nouvelle manifestation de la volonté d'encyclopédisme du géographe ? Non, s'il admet une fois pour toutes qu'il n'est ni médecin, ni épidémiologiste, ni sociologue, ni... Non, si ces derniers acceptent le dialogue et ne se prétendent pas géographes dès qu'ils citent le nom d'un pays ou d'une région et en dressent une carte approximative. Non, si chacun comprend l'obligation d'une démarche pluridisciplinaire dans les faits et non dans les discours.

La simple énumération de ces préalables suffit à montrer le chemin encore à parcourir. Car si certains points sont acquis, d'autres demeurent flous.

II. GÉOGRAPHIE DES MALADIES, GÉOGRAPHIE ET MALADIES.

1. *De la description à la cartographie.*

Constater l'inégale distribution de telle ou telle maladie dans l'espace a été le premier résultat

tangible de la géographie pathologique. En mesurer la fréquence représente la deuxième étape qui, au-delà des observations empiriques, implique la collecte de données fiables et comparables.

Encore faut-il d'abord parler le même langage. L'on doit à l'Office International d'Hygiène Publique de la Sociétés des Nations puis à l'OMS, la mise en forme d'une classification internationale des maladies et causes de décès (C.I.M.). Périodiquement remise à jour et riche à présent d'un millier de rubriques, elle sert de référence nosologique officielle. Il s'en faut pourtant qu'elle soit utilisée partout et de façon rigoureuse. L'information sanitaire reste encore bien souvent embryonnaire, faute d'infrastructures administratives suffisantes et d'éducation du personnel médical. A en croire certains, la « densité » des maladies ne serait que le reflet de la densité des médecins ! Ce n'est pas tout à fait une boutade et il convient de garder un sens critique aiguisé vis-à-vis des statistiques de mortalité et, plus encore, de morbidité. Il paraît souvent tout à fait vain de livrer à l'ordinateur des masses de données dont on connaît les tares congénitales. On ne peut comparer que ce qui est comparable.

Une autre difficulté réside dans le choix des unités spatiales de référence. Même en optant pour des circonscriptions administratives, commodes mais dont on sait les limites, se pose un problème d'échelle. Par nature le géographe penche pour l'analyse la plus fine possible mais il s'expose alors à des risques statistiques évidents. Un nombre trop réduit de malades ou de morts dans un espace trop restreint rend toute interprétation peu ou non significative. Il existe donc un seuil, variable selon la fréquence de la maladie, au-delà duquel toute analyse devient impossible. Rappelons à ce propos que les taux de mortalité et de morbidité s'expriment pour 100 000 habitants...

Une fois résolus ces problèmes méthodologiques, le géographe établit des cartes de localisation et de fréquence. Il est là en terrain connu et familier. De nombreux médecins estiment d'ailleurs qu'il s'agit là de la seule contribution spécifique du géographe... La cartographie des maladies ne diffère pas de la cartographie tout court : les techniques sont les mêmes [35, 49]. Et des essais de cartographie automatique et de géocodage ont donné des résultats tout à fait probants [1, 5] (5).

Aussi dispose-t-on désormais d'une très vaste panoplie de cartes, tant sur la pathologie infectieuse et parasitaire que sur les maladies chroniques ou mentales, et ce, à toutes les échelles. D'une vision planétaire chère à l'OMS au canton ou au quartier urbain, l'on débusque la maladie au point de pouvoir affirmer que le cancer de l'œsophage atteint une fréquence record dans tel « pays » breton, l'onchocercose dans telle section de la Volta et l'infarctus du myocarde dans tel îlot de Chicago. Certains pays comme les Etats-Unis, le Canada, disposent d'atlas de mortalité suivant les différentes causes de décès, le plus expressif et le plus « géographique » dans sa conception étant le *National Atlas of Disease Mortality in the United Kingdom* de M. Howe [34, 50, 62] (6). D'autres ont publié des atlas nationaux d'une maladie ou d'une cause spécifique de décès, les plus nombreux concernant les maladies cardio-vasculaires et les diverses localisations cancéreuses (Etats-Unis, Canada, Suisse, Japon, Chine, Italie...). Mais la France n'est pas encore du nombre... Il reste certes des terres vierges, inconnues, faute d'informations fiables, surtout dans les pays sous-développés. L'on procède alors à des enquêtes directes sur le terrain, monographies locales riches en cartes et dans le droit fil des « topographies médicales » d'antan.

Mais si l'on dispose de séries chronologiques de données, il devient possible de dresser des cartes à des dates différentes, soulignant le sens et le rythme de l'évolution (progression/régression) de telle affection, sa diffusion ou sa rétraction dans l'espace [31, 41, 83, 104]. Spectaculaire, le procédé a beaucoup servi en épidémiologie historique pour l'étude des pandémies (choléra, peste, paludisme, variole...) et l'analyse du « destin des maladies infectieuses » (Ch. Nicolle). Mais il demeure encore largement pratiqué par l'OMS ou le Centre d'Atlanta pour la surveillance mondiale de la grippe.

Quelle que soit la nature de la maladie, son extension et sa diffusion appellent des questions sur les causes de cette répartition. Surtout si la superposition de plusieurs cartes nosologiques révèle dans un même espace des situations analogues ou voisines. Car alors l'association pathologique se double d'une association géographique et renforce l'idée de conditions spécifiques propres à un espace donné, à son « milieu », à son « environnement ». Qu'une fréquence élevée de tuberculose et de silicose apparaisse autour d'une exploitation minière en même temps qu'une fréquence tout aussi forte d'alcoolisme et de cancers bronchopulmonaires, n'est pas une constatation anodine. Pas plus que la superposition de la distribution de l'hépatite B et du cancer du foie au Sénégal, ou du lymphome de Burkitt et du paludisme en Afrique intertropicale.

2. *L'inné et l'acquis.*

La mise en évidence de différences spatiales de fréquence s'apparente à la photographie. Elle n'explique rien mais sert de révélateur. Le rôle de la carte est provocateur : elle pose implicitement des questions et appelle, si possible, des réponses.

La première est d'ordre démographique. En effet et compte tenu des variations des risques selon l'âge, les différences géographiques de la mortalité ou de

(5) Voir aussi dans PYLE [82], le chap. 3 « Meaning and method in disease mapping » et surtout le chap. 8 « Applications of geocoding systems in medical geography ».

(6) Citons à titre d'exemples et pour le cancer [93, 108, 112, 115, 2].

la morbidité peuvent seulement traduire le degré de vieillissement ou de jeunesse des populations concernées. En d'autres termes les inégalités spatiales seraient le plus souvent fonction de la structure par âge (mais aussi par sexe). Dès lors l'interprétation de taux bruts déforme la réalité des risques encourus par une population, soit en les exagérant, soit en les minimisant. A juste raison, démographes et épidémiologistes recommandent d'utiliser des taux standardisés pour l'âge et le sexe en fonction d'une population de référence [44, 90].

Pourtant, même après avoir opéré cette correction et annulé le biais démographique, subsistent de grandes disparités spatiales. Toutes les cartes récentes établies selon cette technique en témoignent. C'est alors, et alors seulement, que se pose vraiment la question en termes d'étiologie.

L'origine des maladies est au cœur du débat passionné sur l'inné et l'acquis. Le fatalisme est-il de rigueur ? L'homme est-il condamné à subir les effets conjugués de facteurs héréditaires et inévitables, et de facteurs externes dépendants de son environnement global ?

Sans doute les géographes étaient-ils allés trop vite en ne retenant de l'étiologie que ce qui était « exogène » et donc, a priori, de leur champ d'étude ? Car les recherches les plus récentes en matière génétique modifient les données du problème. On savait déjà qu'il existait une géographie du sang [7, 6]. On découvre à présent une géographie des gènes [13, 25, 92]. L'étude du système H.L.A. (7) vient de révéler des fréquences très variables selon les différents groupes ethniques de certains gènes. Or leur association avec plusieurs maladies est aujourd'hui incontestable. Il en est ainsi, entre autres affections, du diabète juvénile, de la spondylarthrite ankylosante, de la sclérose en plaques, peut-être des cancers du colon et du sein... On suspecte aussi le rôle de certains antigènes dans la physiopathologie du paludisme. Ces résultats encore provisoires, sont d'une importance capitale dans la mesure où ils permettent de déterminer des processus de sélection naturelle, d'immunisation ou de prédisposition, responsables, seuls ou en association avec d'autres facteurs, d'une inégale distribution spatiale des maladies. On en devine les conséquences d'intérêt géographique évident, ne serait-ce que par le biais des migrations qui modifient la répartition des gènes et, de fait, des risques.

Il va de soi néanmoins que ce sont les facteurs de l'environnement qui concernent le plus le géographe. L'OMS leur a d'ailleurs consacré un gros ouvrage [75] qui confirme combien la pathologie s'inscrit à la fois dans les éco-systèmes et les géo-systèmes [63, 20, 21, 39, 76]. N'en déplaise à certains, il serait aujourd'hui absurde, voire ridicule, de nier le rôle direct ou indirect des éléments du milieu naturel.

Il n'y a pas lieu ici de recenser tous les facteurs et leurs effets pathologiques. Mais il n'est plus scientifiquement défendable de négliger le rôle des propriétés naturelles de l'air, de l'eau, du sol ou du sous-sol. On a fait quelques progrès depuis Hippocrate en la matière et il convient de faire appel à la météoropathologie, à la géo-chimie, à la pédologie... pour mieux comprendre certains processus morbides [63, 20, 21, 37, 38, 39, 76]. Qu'il s'agisse du rôle du rayonnement solaire dans la fréquence des cancers de la peau ou de la dureté de l'eau dans l'étiologie des maladies de l'appareil circulatoire. Ou qu'il s'agisse encore de la responsabilité des turbulences atmosphériques, des coups de froid ou de chaleur et des coulées de foehn sur l'apparition d'accidents cérébro-vasculaires ou de troubles de l'appareil respiratoire. Il n'est même plus tout à fait exclu que la mauvaise réputation des « maisons à cancer » soit une légende au vu des travaux en cours sur les effets des radiations ionisantes issues de tel ou tel « terrain » géologique. De même le rôle des micro-organismes, virus, bactéries, parasites et autres champignons, n'est plus à démontrer. La présence ou le maintien de la bilharziose ou de la malaria tiennent aux conditions naturelles rigoureuses des biotopes des agents pathogènes, de leurs réservoirs et de leurs vecteurs, inertes, passifs ou vivants.

Inversement, combien de maladies tirent leur origine de facteurs anthropiques, voire artificiels, en tout cas sociaux ? Combien sont tributaires des modes et conditions de vie ou de travail, du niveau des revenus, de la technologie, des comportements individuels et collectifs... ? La liste en serait presque infinie. Pourtant le rôle de l'alimentation mérite une attention toute particulière tant ses incidences sont multiples. L'on se doit à ce propos d'évoquer l'œuvre malheureusement inachevée de J. May [64, 66], consacrée aux effets pathogènes de la malnutrition et aux maladies de carence. La géographie de la faim (J. de Castro) participe directement de la géographie de la santé. Mais les excès alimentaires et le nouveau modèle de nutrition des sociétés riches les exposent à des risques élevés de maladies cardio-vasculaires, de diabète ou de certains cancers. L'étude des conséquences pathologiques des régimes alimentaires ouvre un immense champ de recherche aux géographes : en tant que telles et à travers leur évolution mais surtout comprises dans leur environnement global, agricole, économique, social et culturel [51, 73]. De même, chiquer du bétel ou du coca pour oublier sa faim n'est pas une habitude sans risque, pas plus que le fait de fumer ou de boire de l'alcool. On le sait suffisamment par ceux qui, justement, n'en font pas usage comme les Mormons ou les Adventistes du 7e Jour. Les traits culturels d'une communauté, d'une ethnie, d'une religion ne peuvent plus être niés. La circoncision de leurs partenaires ne « protège »-t-elle pas les femmes juives du risque de cancer du col utérin ? Quant aux effets pathogènes du développement, de l'urbanisation, de l'industrialisation, on ne sait plus au juste par où commencer tant s'accumulent les procès à leur encontre. Du stress au cancer du poumon, des accidents de la route à l'ulcère de l'estomac, de l'alcoolisme à l'infarctus du myocarde...

(7) HLA : Human Leucocyte Antigen.

Pêle-mêle on découvre les effets délétères des additifs alimentaires, de l'amiante, des pesticides et des engrais chimiques, des polluants de l'eau, de l'air, du sol en passant par les risques occasionnés par la sédentarité, le travail à la chaîne, l'usage de la pilule contraceptive, des adoucisseurs d'eau ou des vide-ordures... Ultime paradoxe, la médecine « fabrique des malades » (J.C. Sournia). L'essor des maladies iatrogènes en témoigne (I. Illich) : une consommation excessive de médicaments devient elle-même pathogène, une transfusion sanguine peut provoquer une hépatite virale ou un paludisme, une irradiation trop prolongée au cours d'un traitement peut entrainer un cancer des os, de la thyroïde ou une leucémie... Haro sur la « société pathogène » (E. Barral) ! Sus au Progrès ? Se complaire dans la condamnation aveugle et morbide des sociétés industrielles et urbaines, c'est oublier un peu vite et faire fi de leur extraordinaire bilan positif depuis cent ans. Que l'on songe seulement à l'allongement de l'espérance de vie. Jeter un voile pudique sur leurs excès, c'est nier l'évidence et oublier les très grandes inégalités sociales qui subsistent ou même s'aggravent, devant la maladie et la mort. Suivant les sociétés, les milieux sociaux et professionnels, l'accessibilité aux soins, le niveau d'éducation... Le géographe ne saurait les ignorer.

En fait, tout ce qui altère, compromet ou menace la vie, de façon directe ou indirecte, entre désormais dans le champ étiologique. Bien sûr se pose la question de la norme, du seuil entre le normal et le pathologique (C. Canguilhem, M. Foucault) variable selon l'époque, les individus, la société. La relativité de la notion même de santé s'en trouve renforcée. Ce qui justifie sans aucun doute que ne soit plus prise en compte la seule morbidité déclarée, officielle au sens pathologique du terme, mais aussi la morbidité ressentie et perçue. La sociologie médicale s'y emploie.

3. *Systèmes pathogènes et corrélations spatiales.*

De toute évidence, les facteurs de risque ne sont pas de même nature et n'agissent pas de la même manière. Les relations de causalité semblent de moins en moins simples à démêler à mesure que s'affine la recherche épidémiologique.

L'enseignement majeur des découvertes pastoriennes réside dans la démonstration d'une relation directe de cause à effet de type linéaire : un agent pathogène provoque une maladie selon un enchaînement logique, cartésien, de phénomènes. De cette conception est née la notion de *chaîne épidémiologique*. Sa simplicité et sa cohérence ont fait son succès. Surtout d'un point de vue prophylactique, car il suffit de supprimer n'importe quel maillon de la chaîne pour rompre tout le cycle et faire disparaître tout risque. Ce concept s'appliquait bien à la pathologie infectieuse et parasitaire. Mais il apparaît vite mal adapté à d'autres situations qui aboutissent visiblement à d'autres processus. Si bien que l'on est amené aujourd'hui à considérer la maladie non plus comme l'aboutissement d'un enchaînement, mais comme le produit de multiples facteurs. Facteurs provoquants que l'on peut appeler « agents directs », facteurs conditionnants ou « agents indirects et associés » et facteurs prédisposants ou « circonstances ». Dès lors, force est de reconnaître dans la maladie le résultat d'un faisceau de « causes », inscrites dans un *système plurifactoriel*. Les processus sont cumulatifs mais aussi synergiques : leurs effets s'additionnent et se démultiplient (ou s'annulent dans certains cas).

M. Sorre avait parfaitement senti la nécessité d'intégrer plusieurs variables dans son concept de « complexe pathogène ». Mais en limitant son champ d'étude aux maladies transmissibles, il reprenait implicitement le cheminement classique de la chaîne épidémiologique. Et, en toute logique, il concluait à l'existence d'aires pathogènes spécifiques et aux limites strictes dictées par les conditions du milieu, sous-entendu naturel. Dans ces conditions mieux vaudrait parler de *complexes éco-pathogènes*, notion admise et mise en pratique en écologie médicale par de nombreux parasitologistes. Son application s'avère d'une grande efficacité et prend souvent des aspects spectaculaires. Ne peut-on circonscrire le foyer d'une maladie à partir de la seule carte de la végétation ? Il en est ainsi de la géographie de la leishmaniose viscérale dont les insectes vecteurs ne peuvent subsister dans le Midi méditerranéen de la France que dans les étages du pin d'Alep et de la chênaie d'yeuses [87].

Mais tout n'est pas aussi « simple ». On ignore encore tout de l'agent pathogène de bon nombre de maladies. Par contre, on cerne peu à peu leur « climat épidémiologique » par la découverte de multiples facteurs qui leur sont associés. Or la plupart sont d'essence sociale et dépendent des modes et niveaux de vie, des conditions de travail, des comportements... La combinaison de ces facteurs constitue alors des *complexes socio-pathogènes* [H. Picheral, 78] dans la mesure où la maladie est bien un produit social. Leur analyse réclame des moyens considérables de traitement des données afin d'évaluer le poids relatif de chacune des composantes étiologiques et de mesurer leurs interactions. On sait ainsi qu'il existe des risques graves de cardiopathies ischémiques dans des populations présentant une prédisposition génétique et une propension à l'obésité, consommant une alimentation trop riche en graisses saturées d'origine animale génératrices de cholestérol et d'hypertension artérielle, buvant de l'eau de boisson trop douce, pauvre en calcium et magnésium, et fumant plus de 20 cigarettes par jour dans l'exercice de professions citadines et sédentaires réclamant une tension nerveuse importante... On conçoit qu'il ne soit pas facile de démêler l'écheveau !

La situation se complique encore quand interviennent des facteurs découlant de l'usage de techniques ou de technologies propres à telle ou telle société [27]. L'exposition professionnelle à un agent toxique précis (mercure, fibre d'amiante, chlorure de vinyle, dioxyde, solvants...) comme des conditions de travail

spécifiques (bruit, vibrations, lumière, chaleur...) ou des conditions de vie particulières dans un environnement pollué, contribuent à composer les *complexes techno-pathogènes* des maladies « de civilisation ». A tel point que certains n'hésitent pas à classer dans cette catégorie des concepts aussi flous que la « ville », la « vie urbaine », l'« urbanisation » ou l'«industrialisation ». Ce qui, à la réflexion, n'est pas faux tant ces notions synthétiques incluent une somme de variables socio-économiques. C'est dire aussi combien le souci didactique des typologies peut cacher la complexité des complexes...

Le cas de la sclérose en plaques en fournit une excellente illustration [68]. Rares sont les maladies aussi « géographiques » ! De l'équateur aux plus hautes latitudes, sa fréquence ne cesse d'augmenter : nulle en Ethiopie, la prévalence atteint 80 p. 100 000 à Oslo et 153 dans les Iles Shetland. S'agit-il d'une question de climat, de rayonnement solaire ? Ou d'une prédisposition génétique (car les populations des Iles Faroe et du Japon sont fort peu menacées) ? Ou d'une exposition avant l'âge de 5 ans à un agent pathogène inconnu (car, à l'occasion de migrations d'un pays de haute incidence à un pays de faible incidence, les adultes conservent à l'arrivée les risques de leur pays d'origine alors que les jeunes enfants présentent les risques de leur pays de destination). Pour l'heure la question reste entière faute de réponse formelle.

Mais voilà une maladie dont la fréquence semble être fonction d'au moins quatre variables géographiques : zonale, ethnique, démographique et environnementale. Chacune peut être cartographiée et confrontée avec succès à la carte de morbidité, sans apporter pour autant de « preuve » définitive. C'est tout le problème des corrélations spatiales.

Problème qui se complique d'abord par la question du temps de latence. On sait en effet que dans le cas des affections chroniques et dégénératives, un délai plus ou moins long, de 5 à 40 ans environ, sépare le déclenchement de la maladie de l'exposition supposée du malade à tel ou tel facteur pathogène. Le fait, déjà capital pour le médecin, a des incidences essentielles pour le géographe. Car il pose la question de la mobilité de la population.

Les épidémiologistes se sont abondamment servis des migrants comme témoins du rôle respectif de l'inné et de l'acquis, et comme indicateurs de l'impact de facteurs d'environnements différents [67]. Aussi ce thème revient-il très souvent dans les études comparées des risques encourus par les différentes communautés ou groupes ethniques des Etats-Unis, du Canada, d'Israël ou d'Afrique du Sud, ou dans l'analyse des grandes migrations internationales des Juifs ou des Japonais. L'épidémiologie des maladies cardio-vasculaires et des cancers doit beaucoup à cette méthode d'une grande fécondité. Mais l'on ne sait toujours pas exactement comment appréhender le cas des migrations intérieures. Il s'agit pourtant de la situation la plus répandue, au lendemain d'une vague sans précédent d'exode rural et en période d'intense mobilité de la population. A quoi sert d'observer la fréquence du cancer broncho-pulmonaire dans l'agglomération parisienne en 1981 si l'on ignore ou si l'on ne prend pas en compte la consommation de tabac, le mode de vie, le métier, le type de chauffage et d'énergie utilisés... des malades 20 ou 30 ans plus tôt alors qu'ils vivaient ailleurs, à la campagne, dans une autre ville, dans une autre région ? A quoi sert de mesurer l'incidence actuelle des cardiopathies dans l'agglomération lyonnaise sans savoir quoi que ce soit de l'alimentation passée des néo-citadins venus de Bresse, du Jura, de Savoie, du Forez ou du Dauphiné ? Le décalage temporo-spatial est pourtant une des clefs de l'étiologie de ces maladies. Mais l'inadéquation est totale. Car faute de mieux et par la force des choses, l'on confronte le plus souvent la morbidité, ou pire encore la mortalité du moment à des données étiologiques du même moment. En fait, on ne pourra mesurer les éventuels effets de l'environnement actuel que dans quelques décennies.

Le biais est trop important pour ne pas mériter une approche approfondie et pluridisciplinaire. Mais la solution passe par la disposition d'un dossier individuel très précis, rétrospectif et retraçant en fait toute la vie du malade. Ce qui paraît à la fois utopique et tout à fait contraire à l'éthique. On imagine en effet les risques de telles pratiques pour les libertés individuelles.

Cette question en appelle pourtant une autre. A considérer la pathologie d'une population dans un espace donné, ne s'expose-t-on pas à une généralisation excessive ? Un taux régional, cantonal ou urbain ne signifie pas que chaque habitant de cette région, de ce canton ou que chaque citadin court effectivement ce risque. Le groupe n'est pas forcément la somme des individus, si tant est qu'il n'y ait que des malades et non des maladies. L'interprétation des différences géographiques doit en tenir compte. Cette critique de la géographie pathologique rejoint le procès fait à l'épidémiologie descriptive, utile certes mais insuffisante. On s'oriente désormais vers des études longitudinales et prospectives en suivant des échantillons de populations et des populations témoins tout au long de leur existence. Mais le prospectif est-il du champ de la géographie, sinon à très court terme ?

En revanche subsiste une ultime question, la plus importante peut-être. Etablir une corrélation, aussi significative soit-elle, n'autorise pas à conclure à une relation de causalité. Les épidémiologistes, forts de méthodes statistiques de traitement des données de plus en plus raffinées, en conviennent. Seules des présomptions plus ou moins lourdes peuvent être relevées à l'encontre de tel ou tel facteur. Il revient aux médecins et aux biologistes d'en démontrer la responsabilité objective. Une fois encore, il ne peut y avoir substitution ou confusion des rôles. Le géographe doit se garder des conclusions précipitées et de la griserie de la superposition de cartes trop parfaitement corrélées. La prudence, le sens critique et l'humilité sont de rigueur en la matière. La carte est un indicateur des risques courus par une popu-

lation en fonction de ses caractères, de ses activités, de sa dynamique et de son environnement. Elle alerte, elle focalise la recherche en resituant le problème pathologique dans son contexte, tout son contexte, mais sans prétendre en donner la clef. C'est bien la voie suivie depuis une quinzaine d'années par les géographes (Hunter, McGlashan, Mead, Pyle, Stamp..) [59, 82, 78, 42, 69, 102] et notamment M. Howe dans son ouvrage *A World Geography of Human Diseases* [40].

. En définitive, il n'est pas inutile de citer le jugement, réservé mais mesuré de deux médecins [111] sur la géographie des maladies : « *la pathologie géographique a le mérite de nous donner l'image probablement assez fidèle de la distribution des cancers... Si les progrès qu'elle nous a permis de faire... n'ont pas été très nombreux, ni très spectaculaires, c'est moins à la technique qu'il faut s'en prendre qu'à notre propre impuissance de concevoir des hypothèses valables qui rendent compte de telle ou telle distribution particulière. A cet égard, la pathologie géographique a joué parfois un rôle utilement négatif en réduisant à néant des hypothèses trop hâtives ou trop audacieuses qui se trouvaient contredites par les faits bruts mais précis qu'elle apportait* ».

III. ESPACES, SOCIÉTÉS ET SANTÉ.

Un des objectifs de l'épidémiologie est d'identifier, à des fins préventives, des *populations à risque*. On tente ansi de cerner le profil des groupes particulièrement exposés à telle maladie en fonction de son étiologie. Les résultats sont déjà impressionnants et confirment le postulat « à chaque maladie sa géographie ». En termes de géographe, il s'agit de géographie « générale ». Mais ne peut-on inverser les termes et s'interroger sur l'état de santé d'une population dans un espace donné ? Autrement dit, peut-on établir une géographie « régionale » et affecter à des populations, donc à des espaces, des « niveaux de santé » ? On le voit, la problématique change et s'élargit tout à la fois.

Le bon sens populaire et l'expérience voudraient que la santé aille de pair avec la richesse, le progrès, le développement... Historiens et démographes ont montré ce que la transition démographique devait à la *transition épidémiologique* [A. R. Omran, 74]. De fait, le monde s'est fendu en deux. Ne serait-ce pas d'ailleurs un des caractères des disparités Nord/Sud ? D'un côté, les populations encore soumises à la pathologie infectieuse et parasitaire et aux maladies de carence qui hypothèquent leur avenir. De l'autre, des sociétés dont l'espérance de vie s'allonge, dont les modes de vie se modifient très vite entrainant l'essor des maladies chroniques, dégénératives et « de civilisation ». L'on sait aussi qu'à de rares exceptions près, l'évolution se fait de la première à la seconde situation. Un tel transfert, une telle « révolution », avec toutes leurs implications, ne peuvent laisser indifférent le géographe.

La démonstration est presque trop facile. Entre les populations qui disposent d'un médecin pour 30 ou 50 000 habitants et celles qui en ont un pour 500 ou 1 000, il y a tout de même une marge. Et l'on sait bien que l'absence d'hôpitaux ou de médecins figurent parmi les critères du sous-développement (Y. Lacoste). C'est si vrai que tous les annuaires de grandes institutions internationales, de l'ONU à la Banque Mondiale, utilisent ces données comme indicateurs, au même titre que la production de blé ou d'acier.

A cette échelle point d'hésitations, point de doutes en dehors des interpellations, fondées d'ailleurs, sur les excès de la médecine et des médecins. Mais la relation entre ressources sanitaires et niveau de santé se vérifie-t-elle à n'importe quelle échelle ? Le sous-développement est une notion relative qui dépasse de loin les grands clivages macroscopiques et planétaires. En outre, tout Etat a désormais une politique de santé, nationale d'abord mais aussi, et de plus en plus, régionale, pour respecter les besoins spécifiques des populations.

1. *La distribution et la consommation de soins.*

La place et le rôle du système de soins dans le complexe pathogène ne sont guère contestables car de son efficacité dépend, quoi qu'on en dise, la poursuite ou la rupture de continuum morbide. L'institution médicale participe de plain-pied à toute évaluation des problèmes de santé, comme facteur correcteur mais aussi comme facteur promoteur des disparités spatiales et sociales, ce qui va souvent de pair.

La géographie des services de santé s'apparente à la géographie de la distribution. Elle en fait même partie intégrante. Se posent alors des questions classiques de densité de l'appareil de soins, en termes de personnel et de matériel, mais aussi des problèmes de desserte de la population, ou si l'on préfère, d'accessibilité aux soins.

Innombrables sont les travaux à propos des professions médicales et des équipements sanitaires, et de leur répartition dans l'espace. D'un point de vue quantitatif et qualitatif. Mais tous aboutissent à la même constatation : l'inégalité est manifeste, elle est la règle partout, quel que soit le système politique et économique, quelle que soit l'échelle. Ainsi la « densité médicale » est très variable entre nations, entre pays riches et pauvres, entre villes et campagnes, d'une ville à l'autre, entre quartiers urbains... De même pour les dentistes, sage-femmes, infirmiers, pharmaciens... De même pour les hôpitaux, maternités, hospices, dispensaires et autres maisons de cure... Cette approche est utile, intéressante mais finalement de portée limitée. Elle s'affine dès que l'on distingue le statut (public/privé), la qualification

et la spécialisation du corps médical, la nature et l'affectation des équipements. A cette analyse très répandue de la localisation des ressources sanitaires s'ajoutent des réflexions sur les motifs et les choix d'implantation d'un cabinet ou d'un hôpital dans telle ou telle région, tel ou tel quartier. On dresse alors des « cartes sanitaires », on établit des typologies, des hiérarchies qui confirment ou renforcent les caractères des réseaux urbains et les poids des contraintes administratives. Ces tableaux de la situation, d'une situation, gagnent pourtant à être vus en perspective : prendre un certain recul permet de saisir les tendances et de savoir si l'on s'engage ou non dans la voie d'une réduction des inégalités. De fait, toutes ces recherches de nature économique, démographique et sociologique, voire historique (8), concernent aussi directement la géographie car elles ont toutes une assiette spatiale.

Mais la plupart d'entre elles se réfèrent à un seul modèle : la médecine « moderne » et le système « occidental » de soins. Pourtant, là où il a été importé, imposé et appliqué, l'on découvre (ou redécouvre) qu'il cohabite ou se trouve en concurrence avec d'autres systèmes, « traditionnels » certes, mais peut-être trop vite condamnés ou ridiculisés. D'où le regain d'intérêt pour l'ethno-médecine et la médecine dite populaire (*folk-medicine*) (9). Le choix entre plusieurs systèmes de soins débouche sur la question plus large de la pratique médicale, celle des praticiens comme celle des populations.

Outre les apports considérables de la sociologie médicale en matière de comportements, il apparaît que l'accès aux soins est largement tributaire du couple distance/temps. Ainsi se multiplient les études des aires de clientèle, des zones d'attraction selon la proximité, la durée et les difficultés du trajet... Ces recherches connaissent un grand développement, en particulier dans les pays scandinaves et anglo-saxons et doivent beaucoup, on s'en doute, à Lösch, Hägerstrand et Christaller [13, 47, 55, 77, 94, 29, 95]. Inversement se mesure le poids du sous-équipement et de l'isolement dans de nombreux espaces défavorisés, angles morts ou « périphéries ». Dès lors et suivant son tempérament, le géographe emboîte le pas de l'économiste ou celui du sociologue selon qu'il privilégie la construction de modèles ou l'enquête directe sur le terrain.

A ce stade, le champ de la géographie des soins de santé (*geography of health care*) est suffisamment vaste pour que l'on admette son entière autonomie vis-à-vis de la géographie des maladies. Mais peut-on vraiment séparer l'organe de la fonction ?

Car la mission des services de santé est bien de distribuer des soins à des consommateurs. En ce sens, il s'agit ni plus ni moins d'un marché. Cette vision économique de la santé ne manque pas d'intérêt pour le géographe : en tant que telle et pour l'observation d'un type particulier de circuit économique. Mais l'analyse de la « consommation médicale » constitue surtout une approche différente des besoins de santé. Une somme considérable d'états morbides échappe en effet complètement aux mailles de la collecte classique des données épidémiologiques, fondée sur la morbidité diagnostiquée, officielle, répertoriée. En revanche, tous ces soins ont un coût dont on peut évaluer le volume. Du moins dans les sociétés qui disposent de systèmes d'assurances sociales.

On peut alors connaître le type des dépenses de santé et leur évolution, suivant leur nature, suivant le sexe et l'âge des consommateurs, voire leur catégorie socio-professionnelle, mais surtout selon leur lieu de résidence. Car une véritable géographie de la consommation de soins se dessine et les disparités spatiales, observées à n'importe quelle échelle, en disent long sur les comportements des populations. Mais l'on aperçoit aussi à quel point l'importance des services médicaux, leur densité et leur qualité, pèsent sur l'accès aux soins et, de fait, sur leur consommation, au même titre que la distance (matérielle, socio-psychologique). Ce champ de la recherche est déjà largement prospecté par les économistes et les sociologues, ne serait-ce que par le CREDOC en France (10). Il n'a jusqu'à maintenant et assez curieusement guère attiré les géographes [78]. La consommation médicale traduit pourtant des besoins effectifs (mais pas forcément réels, dit-on...). Elle exprime des comportements propres à tel milieu social, à telles mentalités, à telle région. Elle figure enfin en bonne place parmi la panoplie des indicateurs de santé.

2. *Indicateurs et indices de santé.*

Porter un diagnostic sur l'état de santé d'une population, d'une région ou d'un pays reste l'objectif ultime, idéal et convoité. Depuis les années 30 et les travaux de Pittaluga et de Stouman et Falk cités par M. Sorre [99, 100] jusqu'aux volumineux rapports de l'O.M.S. [33], prévaut le même souci : disposer d'un indice unique et synthétique. Une récente mise au point donne une bonne analyse critique de ces multiples tentatives [30].

Les premiers essais étaient fondés sur l'utilisation d'indicateurs de mortalité puis de morbidité. Qu'il s'agisse de taux spécifiques, surtout démographiques

(8) Parmi une littérature pléthorique, citons Bridgman [13], Labasse [47], Phillips [77], Shannon & Dever [94]. En France, on trouvera de très riches analyses dans les *Cahiers de Sociologie et de Démographie médicale* (Paris) ; voir notamment Lévy [55].

(9) Voir à ce propos l'analyse critique de Retel-Laurentin & Epelboin [85] et Elling [23].

(10) Outre la somme impressionnante des travaux du CREDOC, on peut citer quelques ouvrages d'économistes français : Barral [3], Brunet-Jailly [14], Lévy [56].

(mortalité infantile, espérance de vie...), ou d'indices pondérés suivant de nombreuses variantes. La démarche s'est ensuite affinée avec l'introduction de nouvelles variables, comme les handicaps ou les degrés d'incapacité... Du même coup s'élargissait la notion de morbidité : aux maladies diagnostiquées, officielles en quelque sorte, s'ajoutaient leurs séquelles, fonctionnelles, psychologiques et sociales, mais aussi toute la pathologie ressentie et perçue.

Toutes ces approches font uniquement référence à la maladie : c'est leur principal défaut. Car tous les indices proposés ne mesurent que l'absence de santé ou la mauvaise santé, mais pas la « santé ». On cherche alors, dans une quête toujours actuelle et apparemment sans fin, la définition d'un indice de santé « positive » ou de santé « globale ». On construit des tests de santé « totale » et des modèles théoriques intégrant de multiples indicateurs dont la pondération elle-même pose problème. Car en fait, chaque variable retenue possède un poids variable dans le temps et dans l'espace, selon le milieu social et culturel, suivant les sociétés. D'un point de vue de géographe, il apparaît vain et utopique d'attendre un indice applicable partout, par tous et à tous.

Faut-il alors renoncer pour autant ? En l'état actuel des choses, mieux vaut une approche aussi imparfaite soit-elle, que l'absence de mesure et d'évaluation. Même si la « sanométrie » (Goldberg) se cherche encore !

3. *Géographie et politique de santé.*

Car on ne peut plus attendre. Le « laisser-faire » est générateur d'inégalités. Aussi l'OMS proclame-t-elle « la santé pour tous » et les médias établissent-ils des palmarès de la « santé » des villes et des régions, pendant que la « santé » fait vibrer les tribunes électorales. Chacun affirme la nécessité d'une politique et pour certains, d'une planification de la santé. Doivent-elles se faire au jugé ? Peut-on appliquer partout un modèle unique ? Peut-on opérer des choix indispensables sans aucun critère ? De telles attitudes paraîtraient inconcevables et irresponsables. Et pourtant...

Toute politique suppose une assise spatiale. En toute logique, l'OMS conçoit son action à l'échelle du monde, encore qu'elle ait jugé indispensable de créer six bureaux régionaux (Amérique, Europe, Méditerranée orientale, Afrique, Sud-Est asiatique et Pacifique occidental). Chaque Etat a évidemment sa politique nationale de santé mais même parmi les plus centralisés, se manifeste le besoin de régionaliser choix et investissements. Et dans les pays à système fédéral, il est fréquent que chaque province ou « état » ait le souci d'adapter sa politique à une échelle spatiale encore inférieure. Convaincu de la discontinuité de l'espace et de la spécificité des espaces, le géographe ne peut que s'en féliciter. Mais au-delà des intentions et des institutions, quelle connaissance a-t-on de la santé « régionale » ?

A dire vrai, on n'en a pour l'instant qu'une vision très partielle et sectorielle, bien souvent éclatée et parcellisée. La plupart des études, au demeurant remarquables, répondent encore à une interrogation d'ordre essentiellement étiologique. Il s'agit en général d'une géographie régionale d'une maladie et, de fait, d'*épidémiologie spatiale* (Y. Verhasselt). Le souci du géographe est bien ici de localiser et de resituer dans leur environnement global les « populations à risque » identifiées par l'épidémiologiste. Ce type d'analyse demeure beaucoup plus fréquent que les tentatives d'évaluation de niveaux de santé.

Un grand nombre de recherches s'inspirent de ce fait de l'écologie médicale et de la détermination d'« aires pathogènes », à la manière de Sorre et de ses applications au monde méditerranéen et aux massifs forestiers équatoriaux. Mais il s'agit alors davantage de géographie zonale que de géographie régionale. Dans le même esprit, E. Pavlovski et des équipes de médecins et de géographes soviétiques ont défini des « paysages épidémiologiques » (*landscape epidemiology*) qui sont en fait des foyers naturels de la maladie [114]. Cette conception vient d'être récemment reprise et élargie par G. Rémy en Afrique occidentale : « *les nuances de la morbidité propres à une région et l'ensemble des facteurs qui en rendent compte — ce que l'on peut convenir d'appeler son paysage épidémiologique — sont une composante de son paysage géographique. Toute évolution de l'une retentit sur l'autre* » [84]. On dépasse là le stade de l'épidémiologie descriptive pour chercher à mesurer le retentissement de la maladie sur le peuplement, les activités, la maîtrise et la mise en valeur de l'espace. C'est un peu dans la même optique qu'ont été élaborés, sous la direction d'H. Jusatz à Heidelberg, des atlas régionaux en Libye, Ethiopie, Afghanistan et Koweit [46].

Beaucoup plus rares sont les tentatives de dresser le tableau pathologique complet d'une région, en particulier dans les sociétés développées. Citons les études sur le Québec, le Rio Grande do Sul, et en France, sur plusieurs régions, notamment le Limousin et le Midi méditerranéen [78, 10, 96]. Par contre, l'on ne compte plus les recherches sur la pathologie comparée des différents quartiers ou des « aires sociales » des grandes villes : Chigago, Montréal, Londres, Glasgow, Bruxelles... [57, 58, 36, 80]. L'intérêt de cette série d'études est de pousser l'analyse à un niveau de plus en plus fin. Mais toutes demeurent fondées sur le critère médical, sur des données de mortalité et (ou) de morbidité diagnostiquées.

Le souci d'intégrer d'autres variables a conduit plusieurs chercheurs à formuler des indices plus synthétiques. Sont pris alors en compte les services de santé ou la consommation médicale pour s'interroger sur l'efficacité du système de soins, les effets de la sous-médicalisation ou encore sur les relations entre état de santé et ressources sanitaires. Soit à l'échelle

(11) Ainsi, en France : LEBART, SANDIER & TONNELIER [53], DUPUY [22], FAURE, SANDIER & TONNELIER [24].

des régions (11), soit, ce qui est rarissime, à l'échelle d'une région et à l'intérieur de cette région comme au Québec (L. Loslier, J.-P. Thouez) ou dans le Midi de la France (H. Picheral). Le bilan ne manque pas d'intérêt sans pour autant être encore pleinement convaincant. Certes, les résultats contribuent à souligner les disparités régionales et intra-régionales, d'ordre économique ou social. Ils les confirment même, mais l'adéquation entre l'offre et de la demande de soins et les niveaux de santé est loin d'être toujours significative à cette échelle. Sans doute, faute de données suffisantes et surtout, d'indicateurs adéquats.

Il est temps de réfléchir à une méthode rigoureuse qui permette de saisir l'interface santé/espace/environnement aux niveaux régional et infra-régional. Ne nous leurrons pas : cette démarche est certainement la plus difficile parce que la plus ambitieuse. C'est pourtant la plus « géographique ». Il est temps désormais d'imaginer et de construire les outils de « topographies de santé », d'esprit bien plus large que les topographies médicales d'autrefois. Ce qui implique la collaboration de multiples disciplines car le géographe ne saurait y suffire à lui seul. Les médecins non plus d'ailleurs. Et peut-être même serait-il nécessaire de se détacher quelque peu de la démarche des épidémiologistes.

Car de deux choses l'une. Ou l'on veut localiser à des fins préventives des foyers, des zones, des aires « à risque ». Et alors seule la méthode épidémiologique est capable d'y parvenir. Inutile de chercher d'autres voies, il n'en existe qu'une, à quelques variantes près. Ou bien l'on tente de guider l'élu, l'aménageur, le « décideur » dans ses choix en lui indiquant où et quelles sont les priorités. Les problématiques sont différentes. Elles ne sont ni concurrentes, ni exclusives : elles se complètent et devraient aller de pair [79].

Dans les faits cela signifie deux approches parallèles et complémentaires. D'un côté s'il s'agit d'évaluer les risques respectifs dans des espaces différents, on ne peut en aucun cas éviter le préalable de la standardisation. Car comment pourrait-on comparer des populations dotées de caractères propres, démographiques et socio-professionnels ? Chaque fois que les données le permettent (d'où la nécessité de recensements fiables et détaillés à cette échelle), il convient d'annuler les biais de structure. C'est l'optique de la géographie des maladies et de l'épidémiologie spatiale.

Par contre, il n'est plus concevable d'évaluer l'état de santé de populations « fictives » et abstraites vivant dans des espaces devenus neutres par l'artifice de la statistique. Et avec des normes ou des populations de référence nationales et internationales ! La modélisation va à l'encontre du réel, du tangible. C'est-à-dire l'appartenance à un milieu spécifique, montagneux ou deltaïque, rural ou urbain, avec des populations jeunes ou vieilles, riches ou pauvres dont les comportements sociaux et culturels pèsent directement sur leur accès aux soins et sur leur santé. Les besoins sont spécifiques : ici à un quartier d'immigrés ou à une haute vallée montagnarde, là à une banlieue industrielle ou à un littoral prisé des retraités... Existe-t-il ici et là assez de médecins, de couveuses, de maisons de retraite, de dispensaires ou de centres de planning familial ? Ou à la limite, y en aurait-il « trop » ? Où faut-il concrètement mener des actions préventives, amener l'eau, la traiter, installer des centres de soins, raser des logements insalubres, bâtir, reconvertir... ? En un mot comment mener une politique sans savoir où se situent les besoins effectifs, réels. Très bien, dira-t-on, mais en vertu de quelles normes ? Précisément on se rend compte aujourd'hui de la fatuité de vouloir imposer partout un modèle unique, universel, imaginé à Washington, Genève ou Paris. La gestion de la santé suppose une connaissance du milieu tel qu'il est et non tel qu'il pourrait être si... Le géographe ne dérogerait pas en s'engageant sur cette voie. Ne serait-ce pas aussi et dans un certain sens, un retour aux sources ?

*
* *

En définitive, si l'on en juge par l'impressionnante production scientifique de ces deux dernières décennies, la géographie de la santé fait preuve d'une vitalité remarquable. Elle est présente sur tous les fronts, à l'affût des dernières techniques de cartographie ou de traitement des données, attentive aux nouveaux courants de la pensée géographique, vigilante et critique à l'égard des excès des sociétés industrielles, sensible aux effets du dénuement et de la pauvreté. Pourtant, en dépit de cette santé de bon aloi, cette géographie semble mal comprise, mal perçue, voire ignorée. Davantage encore dans les pays latins et en France jusqu'à une date récente, que dans les pays anglo-saxons. Cela tient probablement aux mentalités mais aussi au cloisonnement des systèmes universitaires et de la recherche. Cela tient sûrement à certaines ambiguïtés manifestes; le vocabulaire, la terminologie, la méthode tardent à se fixer. Enfin et surtout, cela tient au double visage de ce qui fut longtemps écartelé entre deux spécialités, deux types de géographes venus d'horizons bien distincts : la géographie des maladies (*geography of diseases*) et la géographie des soins de santé (*geography of health care*). Leurs origines différentes expliquent leurs cheminements respectifs et leurs voies longtemps parallèles. Ecartant le risque d'une dichotomie artificielle, elles tendent aujourd'hui, semble-t-il, à se rejoindre dans la géographie de la santé (terme préférable, encore une fois à la formule équivoque des Anglo-saxons : *medical geography*). Leur meilleur terrain de convergence paraît bien être l'analyse régionale des phénomènes de santé, de loin la plus opérationnelle. En ce domaine, beaucoup reste à faire, sinon tout, dans les pays sous-développés mais aussi dans les sociétés industrielles.

N'oublions pas qu'il s'agit d'une branche encore jeune de la géographie et qui cherche à s'émanciper sans aucun ostracisme ni volonté d'isolationnisme ou d'impérialisme. Au contraire. La géographie de la

santé est une discipline ouverte aux autres démarches, aux autres langages. Elle n'a d'ailleurs pas le choix. Car que serait la géographie des maladies sans l'épidémiologie ? Elle ne peut ni renier cette filiation directe, ni s'en passer. On pourrait multiplier les exemples de carrefours du même ordre. Quant à l'incompétence supposée du géographe en matière de santé, n'est-ce pas finalement un mauvais prétexte ? Car enfin combien de thèmes consacrés et familiers cette géographie ne recoupe-t-elle pas ? Le géographe ne peut être dépaysé ou désarmé quand il s'agit de géographie générale ou régionale, physique ou humaine, urbaine ou rurale, zonale, sociale, industrielle, commerciale, appliquée, culturelle, administrative... En réalité, est-ce si curieux, si choquant, si paradoxal de la part de géographes qui traitent de l'homme ?

Manuscrit prêt en juin 1981.

RÉFÉRENCES BIBLIOGRAPHIQUES

[1] ARMSTRONG (R.W.), 1972, « Computers and mapping in medical geography », in MCGLASHAN, 1972.

[2] *Atlas of cancer mortality in the People's Republic of China.* Pékin, 1980.

[3] BARRAL (P.E.), 1978, *Economie de la santé : faits et chiffres.* Paris.

[4] BARRETT (F.), 1980, « Medical geography as a foster child », in MEADE (M.), ed., *Medical Geography,* Chapel Hill, p. 1-15.

[5] BERLIE (J.), 1978, « La répartition géographique de la mortalité par cancer de l'œsophage en France. Méthodologie de traitement des données ». *Bull. de la Section de Géogr. du Comité des Trav. histor. et scient.*, n° 2, p. 147-168.

[6] BERNARD (J.), 1965, « Esquisse d'une géographie des maladies du sang ». *Annales de Géographie,* p. 271-280.

[7] BERNARD (J.), RUFFIÉ (J.), 1965, *Hématologie géographique.* Paris.

[8] BESANCENOT (J.P.), 1974, « Premières données sur les stress bioclimatiques en France ». *Annales de Géographie,* t. 83, n° 459, p. 497-530.

[9] BESANCENOT (J.P.), 1978, « Le contexte climatique des accidents cardio-vasculaires en façade méditerranéenne de l'Europe ». *Bull. de la Section de Géogr. du Comité des Trav. histor. et Scient.*, t. 2, p. 75-89.

[10] BLANCHET (M.), et al., 1975, *La géographie de la mortalité au Québec, 1969-1973.* Québec.

[11] BORDIER (A.), 1884, *La géographie médicale.* Paris.

[12] BOUDIN (J.C.), 1843, *Essai de géographie médicale ou études sur les lois qui président à la distribution géographique des maladies ainsi qu'à leurs rapports topographiques entre elles, lois de coïncidence et d'antagonisme.* Paris.

[13] BRIDGMAN (R.F.), 1963, *L'hôpital et la cité.* Paris.

[14] BRUNET-JAILLY (J.), 1972, *Essai sur l'économie générale de la santé.* Paris.

[15] *Cahiers de Sociologie et de Démographie Médicale.* Paris.

[16] CARTER (C.O.), PEEL (J.) (ed.), 1976, *Equalities and inequalities in health.* Londres.

[17] CHARLOT (M.), 1976, *Vivre avec la mort.* Paris.

[18] CLAVAL (P.), 1976, *Essai sur l'évolution de la géographie humaine.* Paris.

[19] DAUSSET (J.), DECOS (L.), 1979, « Les marqueurs génétiques dans la susceptibilité aux maladies selon les populations ». *Revue Epidém. Santé Publique,* t. 27, n° 5-6, p. 369-379.

[20] DUBOS (R.), 1968, *Man, medicine and environment.* New York.

[21] DUBOS (R.), 1973, *L'homme et l'adaptation au milieu.* Paris.

[22] DUPUY (J.-P.), 1973, *Relations entre dépenses de santé, mortalité et morbidité.* Paris, CEREBE, rapport.

[23] ELLING (R.) (ed.), 1981, « Traditional and modern medical systems ». *Soc. Sc. Med.* v. 15 A, n° 2 (special issue).

[24] FAURE (H.), SANDIER (S.), TONNELIER (F.), 1977, *Analyse régionale des relations entre l'offre et la demande de soins médicaux.* Paris, CREDOC, rapport.

[25] FEINGOLD (N.), 1979, « Pathologie géographique et marqueurs génétiques : détection des associations ». *Revue Epidém. Santé Publique,* t. 27, n° 5-6, p. 381-388.

[26] FINKE (L.L.), 1792-1795, *Versuch einer allgemeine medizinisch-praktischen Geographie.* Leipzig, 3 vol.

[27] GEORGE (P.), 1978, « Perspectives de recherche pour la géographie des maladies ». *Annales de Géographie,* t. 87, n° 484, p. 641-650.

[28] GILBERT (E.W.), 1958, « Pioneer maps of health and disease in England ». *Geographical Journal,* n° 124, p. 172-183.

[29] GODLUND (S.), 1961, « Population, regional hospitals, transport facilities and regions ». *Human Geography,* n° 21.

[30] GOLDBERG (M.) et al., 1979, « Indicateurs de santé et « sanométrie » : les aspects conceptuels des recherches récentes sur la mesure de l'état de santé d'une population ». *Revue Epidém. Santé Publ.,* t. 27, n° 1, p. 51-68; n° 2, p. 133-152.

[31] HAGGETT (P.), 1972, « Contagious processes in a planar graph : an epidemiological application », in MCGLASHAN, 1972.

[32] HIRSCH (A.), 1860-64, *Handbuch der historisch-geographischen Pathologie,* Stuttgart, 2 vol.

[33] HOLLAND (W.), IPSEN (J.), KOSTRZEWSKI (J.), 1979, « Measurements of levels of health ». *OMS Public. Reg.,* série Europe, n° 7, 456 p.

[34] HOWE (G.M.), 1970, *National atlas of disease mortality in United Kingdom.* Londres.

[35] HOWE (G.M.), 1970, « Some recent developments in disease mapping ». *Royal Society of Health Journal,* n° 90, p. 16-20.

[36] HOWE (G.M.), 1972, « London and Glasgow : a comparative study of mortality patterns », in *International Geography.* Toronto.

[37] HOWE (G.M.), 1972, *Man, environment and disease : a medical geography of Britain through the ages,* Newton Abbot.

[38] HOWE (G.M.), LORAINE (J.A.), 1973, *Environmental medicine.* Londres.

[39] Howe (G. M.), 1976, « Environmental factors in diseases », in Lenihan & Fletcher, 1976.

[40] Howe (G. M.), 1977, *A world geography of human disease*. Londres.

[41] Hunter (J. M.), 1971, « Diffusion of influenza in England and Walles ». *Ann. of Assoc. of Amer. Geogr.*, n° 61, p. 637-653.

[42] Hunter (J. M.), 1974, *The geography of health and disease*. Univ. North Carolina, Chapel Hill.

[43] Hunter (J.M.), 1974, « The challenge of medical geography », in Hunter, 1974.

[44] Jenicek (M.), 1976, *Introduction à l'épidémiologie*. Saint-Hyacinthe, Paris.

[45] Journaux (A.) (sous la direction de), 1966, *Géographie générale*. Paris, La Pléiade.

[46] Jusatz (H. J.), (éd.), 1967-1971, *Geomedical Monographs series*. Heidelberg.

[47] Labasse (J.), 1980, *L'hôpital et la ville, géographie hospitalière*. Paris.

[48] Lannoy (W. de), 1978, *Atlas socio-géographique de Bruxelles*. Bruxelles.

[49] Learmonth (A.T.A.), 1965, « Maps of some standardized mortality ratios for Australia ». Dep. of Geogr., Univ. of Camberra, *Occasional Paper* n° 3.

[50] Learmonth (A.T.A.), 1972, « Atlases in medical geography 1950-70 : a review », in McGlashan, 1972.

[51] Learmonth (A.T.A.), 1978, *Patterns of disease and hunger*. Londres.

[52] Learmonth (A.T.A.), 1978, « Medicine and medical geography before the second world war : a personal perspective ». *Geographia Medica*, n° 8, p. 67-102.

[53] Lebart (L.), Sandier (S.), Tonnelier (F.), 1974, « Aspects géographiques du système de soins médicaux ». *Consommation*, n° 4, p. 5-50.

[54] Lenihan (J.), Fletcher (W.), (ed.), 1976, *Health and the environment*. Londres & Glasgow.

[55] Lévy (D. R.), 1980, *Les médecins dans les zones rurales*. Paris 1, thèse.

[56] Lévy (E.) et al., 1975, *Economie du système de santé*. Paris.

[57] Loslier (L.), 1976, *La mortalité dans les aires sociales de la région métropolitaine de Montréal*. Québec.

[58] Loslier (L.), 1978, « Modes de vie, milieux de vie et disparités de santé au Québec ». *Bull. de la Section de Géogr. du Comité des Trav. histor. et scient.* n° 2, p. 169-184.

[59] McGlashan (N. D.), 1972, « Medical geography : an introduction », in McGlashan (ed.), *Medical geography. Techniques and field studies*. Londres.

[60] McKinley (E. B.), 1935, *A geography of disease*. Washington.

[61] Masironi (R.), 1979, « Géochimie et maladies cardio-vasculaires ». *Revue Epidém. Santé publ.*, 27, n° 5-6, p. 455-463.

[62] Mason (T. J.), McKay (F.), 1975, *Atlas of cancer mortality for U.S. counties 1950-69*. Washington.

[63] May (J.), 1958, *The etiology of human disease*. New York.

[64] May (J.), 1961, *The ecology of malnutrition in the far and near East : food resources, habits and deficiency*. New York.

[65] May (J.), 1961, *Studies in disease ecology*. New York.

[66] May (J.), 1963, *The ecology of malnutrition in five countries of East and Central Europe*. New York.

[67] Mayer (J. D.), 1980, « Migrant studies and medical geography : conceptual problems and methodological issues », in Meade, 1980.

[68] Mayer (J.D.), 1981, « Geographical clues about multiple sclerosis ». *Ann. Assoc. Americ. Geogr.*, v. 71, n° 1, p. 28-39.

[69] Meade (M. S.), 1980, *Conceptual and methodological issues in medical geography*. Univ. North Carolina, Chapel Hill.

[70] Meade (M. S.), (ed.), 1980, *Medical geography*, Chapel Hill.

[71] Meade (M. S.), 1980, « An interactive framework for geochemistry and cardiovascular disease », in Meade, 1980.

[72] Mechanic (D.), 1978, *Medical sociology*. New York.

[73] Newman (J.-L.), 1980, « Dietary behavior and protein-energy-malnutrition in Africa South of the Sahara : some themes for medical geography », in Meade, 1980.

[74] Omran (A. R.), 1971, « The epidemiologic transition : a theory of the epidemiology of population change ». *Milbank Mem. Fund. Qu.*, t. 49, n° 4, p. 509-538.

[75] Organisation Mondiale de la Santé, 1972, *Risques pour la santé du fait de l'environnement*. Genève.

[76] Pampalon (R.), 1980, *Environnement et santé*. Québec.

[77] Phillips (D. R.), 1981, *Contemporary issues in the geography of health care*. Norwich.

[78] Picheral (H.), 1976, *Espace et santé. Géographie médicale du Midi de la France*. Montpellier.

[79] Picheral (H.), 1978, « Santé et régions françaises ». *Bull. de la Sect. de Géogr. du Comité des Trav. histor. et scient.*, n° 2, p. 225-251.

[80] Pyle (G. F.), 1971, *Heart disease, cancer and stroke in Chicago*. Chicago.

[81] Pyle (G. F.), 1976, « Foundations to medical geography ». *Economic Geography*, t. 52, n° 2, p. 95-102.

[82] Pyle (G. F.), 1979, *Applied medical geography*. New York.

[83] Pyle (G. F.), 1980, « Geographical perspectives on influenza diffusion : the United States in 1940's », in Meade, 1980.

[84] Rémy (G.), 1980, *Paysages épidémiologiques d'Afrique occidentale*. Avant-projet EHESS.

[85] Retel-Laurentin (A.), Epelboin (A.), 1980, « Bilan et perspectives des recherches médicales en sciences humaines », in *La recherche en sciences humaines 1979-1980*, Paris, rapport CNRS.

[86] Reynaud (A.), 1974, « La géographie entre le mythe et la science : essai d'épistémologie ». Reims, *Travaux de l'Institut de Géographie de Reims*, n° 18-19.

[87] Rioux (J.A.), Golvan (Y.J.) et al., 1969, *Epidémiologie des leishmanioses dans le sud de la France*. Paris.

[88] Rodenwaldt (E.), Jusatz (H. J.), 1942-45, *Seuchen Atlas*. Gotha.

[89] Rodenwaldt (E.), Jusatz (H. J.), 1952-1961, *Welt-Seuchen Atlas*. Hambourg.

[90] Rouquette (C.), Schwartz (D.), 1970, *Méthodes en épidémiologie*. Paris.

[91] Sakamoto-Momiyama (M.), 1977, *Seasonality in human mortality : a medico-geographical*. Tokyo.

[92] Salmon (C.), 1979, « Groupes sanguins érythrocytaires et pathologie géographique ». *Revue Epidém. Santé Publique*, t. 27, n° 5-6, p. 389-397.

[93] Segi (M.) et al., 1976, « Cancer mortality by site and by city and county in Japan ». *Kosei no Shihyo*, t. 23, n° 8.

[94] Shannon (G. W.), Dever (G. E. A.), 1974, *Health care delivery : spatial perspectives*. New York.

[95] Shannon (G. W.), 1980, « Space and time in medical care geography », in Meade, 1980.

[96] Shimada (A.), et al., 1981, « Regional differences of death from chronic disease in Rio Grande do Sul from 1970 to 1976 ». *Soc. Sc. Med.*, v. 15 D, n° 1, p. 187-198.

[97] Singh (R. L.), Singh (R. P.), Singh (B. P.), 1976-77, « Approaches towards geography of health : a synoptic review », *Geographia Medica*, n° 6-7, p. 1-21.

[98] Snow (J.), 1885, *On the mode of communication of cholera*. Londres, 2e ed.

[99] Sorre (M.), 1933, « Complexe pathogène et géographie médicale ». *Annales de Géographie*, n° 235, p. 1-9.

[100] Sorre (M.), 1943, *Les fondements biologiques de la géographie humaine*. Paris.

[101] Sorre (M.), 1966, « Géographie de l'état sanitaire et des maladies », p. 1091, *in* Journaux, 1966.

[102] Stamp (L. D.), 1964, *Some aspects of medical geography*. Londres.

[103] Steudler (F.), 1972, *Sociologie médicale*. Paris.

[104] Stock (R.), 1976, *Cholera in Africa*. Londres.

[105] Thouez (J.-P.), 1978, *Espace régional et santé*. Sherbrooke.

[106] Thouez (J.-P.), 1978, « La dureté de l'eau potable et la mortalité vasculaire cérébrale dans l'est du Québec ». *Bull. de l'Assoc. Géogr. Franç.*, n° 451, p. 115-123.

[107] Thouez (J.-P.), et al., 1978, *Atlas médical des cantons de l'Est*. Sherbrooke.

[108] Thouez (J.-P.) et al., 1980, « Specific primary sites of malignant neoplasms, their incidence and their etiology in the province of Quebec 1971-75 ». Montréal, Départ. de Géographie, *Notes et documents* n° 2.

[109] Tromp (S. W.), 1963, *Medical biometeorology*. Amsterdam.

[110] Tromp (S. W.), 1973, « The relationship of weather and climate to health and disease », in Howe & Loraine, 1973.

[111] Tuyns (A. J.), Repetto (F.), 1979, « Pathologie géographique et cancers digestifs. Situation présente et perspectives d'avenir ». *Revue Epidém. Santé Publ.*, t. 27, n° 5-6, p. 465-477.

[112] Verhasselt (Y.), 1975, *Maps of cancer distribution*. Bruxelles, Geo. Inst., Vrije Univ.

[113] Verhasselt (Y.), 1977, « Quelques réflexions sur la géographie médicale ». *Bulletin de la Société belge d'Etudes géographiques*, t. 46, n° 2, p. 185-193.

[114] Voronov (A. G.), 1977, « The geographical environment and human health ». *Soviet Geography*, n° 4, p. 230-237.

[115] Wigle (D. T.), 1980, *Répartition géographique de la mortalité au Canada*. Hull. vol. 1 : cancer.

Une vue sur les sciences humaines de la santé

Signalons la parution d'un nouveau volume dans la collection « Synthèse et Documentation » du Centre de documentation Sciences humaines du Centre national de la Recherche scientifique (1). Dans le domaine des sciences humaines de la santé, ce volume se présente comme une mise au point et un ouvrage technique. Il comporte trois parties.

La première est la synthèse des travaux peu diffusés, réalisés dans les années 1975 à 1981 et qui ont été enregistrés et analysés dans la base bibliographique RESHUS (REseau documentaire en Sciences HUmaines de la Santé). L'intérêt essentiel de cette tentative est qu'elle permet de préciser quels ont été les thèmes, les méthodes, les orientations disciplinaires, parfois les motivations ou les présupposés de ces travaux. Le lecteur n'est donc pas seulement informé de ce qui a été fait, il est introduit à une réflexion sur la nature réelle et les limites d'un effort d'investigation d'une grande ampleur et d'une extraordinaire diversité. Cette synthèse est réalisée selon les trois thèmes suivants : les inégalités devant la maladie, la dynamique du système de soins, l'évaluation et le contrôle.

La seconde est la présentation de RESHUS, réseau documentaire mis en place par le Laboratoire d'Economie et de Sociologie du Travail et le Centre de Documentation Sciences humaines du CNRS. La troisième partie est constituée par la présentation des « recherches en cours » dans ce domaine en France. Par « recherches en cours », on entend : recherches actuellement en cours financées sur fonds publics et n'ayant pas encore donné lieu à publication même restreinte. Pour chaque équipe on donne : la ville de localisation, l'intitulé et l'adresse du maître d'œuvre, le titre de la recherche, un résumé détaillé, la durée de la recherche, les sources de financement, les noms des responsables de la recherche et leur rattachement administratif et/ou scientifique. En bref, un instrument de travail certainement utile dans un domaine tout neuf, qui devrait faire progresser l'idée de transparence et de publicité des contrats de recherches financés sur fonds publics. — Régine VANDUICK, CNRS.

(1) Mossé (Philippe), 1982, *Une vue sur les Sciences humaines de la santé (1975-1981)*. Paris, CNRS, CDSH, coll. Synthèse et Documentation, 225 p.

Prévalence de l'usage de la cigarette en milieu scolaire: Québec, 1978

Paul-Marie Bernard[1], Jacques Tremblay[2], Nicole Rousseau[3] et Clermont Bégin[4]

Résumé

Une étude de prévalence, menée auprès de 6000 écoliers de la région de Québec, a permis d'établir l'importance du tabagisme en milieu scolaire, tant par la précocité de l'apprentissage, que par la croissance rapide du phénomène, lors du passage du niveau primaire au secondaire. Les indicateurs de performance pour la mesure d'impact et le moment d'intervention privilégié sont présentés en rapport avec les programmes visant à la réduction du comportement tabagique en milieu scolaire.

Dans un document qui fait état de l'ampleur du problème du tabagisme au Québec[1], la Direction de la planification des services de santé du Ministère des Affaires sociales du Québec énonce un objectif important de lutte contre ce problème: la formation graduelle de générations de non-fumeurs. Cet objectif se situe dans le cadre d'un programme d'intervention à long terme dont le groupe-cible serait vraisemblablement la population écolière du Québec.

Par ailleurs, les statistiques officielles n'établissent l'usage du tabac que pour les 15 ans et plus. Une étude canadienne menée en 1971-72 par "The Canadian Home and School and Parent-teacher Federation", a considéré le problème sur des groupes d'âge plus jeune (de 8 à 15 ans), mais n'a pas distingué les Québécois de l'ensemble[2]. Une autre étude plus récente (1975)[3], menée par le Département des sciences de la santé communautaire de l'Université de Sherbrooke, visait l'identification des facteurs déterminant le phénomène tabagique chez les jeunes et la détermination du degré d'exposition de ceux-ci au tabagisme. Dans cette étude, le faible échantillon partiel de 351 écoliers de secondaire I ne permet pas de faire une analyse détaillée de la prévalence du tabagisme.

Objectifs et population visée

Une étude descriptive, commandée par le ministère des Affaires sociales du Québec, a été réalisée à l'automne 1978 par une équipe interfacultaire de chercheurs de l'Université Laval, auprès des écoliers francophones de la région métropolitaine de Québec, fréquentant les niveaux scolaires primaire (4, 5 et 6) secondaire (I, II et III). Cette étude avait comme objectifs 1) d'établir la prévalence de l'usage de la cigarette dans la population visée; 2) de déterminer l'âge d'apprentissage du tabagisme chez les écoliers et 3) de décrire l'évolution temporelle du phénomène de l'usage du tabac dans la population scolaire.

Matériel et méthode

Un questionnaire prétesté a été appliqué à un échantillon de 6211 écoliers francophones de la région métropolitaine de Québec. Les niveaux scolaires visés étaient le primaire (4, 5 et 6), secondaire (I, II et III).

Echantillonnage

L'échantillonnage, qui visait comme unité la classe, s'est déroulé en deux étapes. Dans un premier temps, un échantillon aléatoire simple avec remise a été tiré de l'ensemble des écoles, la probabilité pour une école d'être choisie étant proportionnelle à sa taille. Soixante-dix écoles ont ainsi été choisies. Dans un deuxième temps, on a prélevé de chaque école tirée un échantillon de classes stratifié par niveau scolaire. Le nombre de classes tirées de chaque école, pour chacun des niveaux offerts,

1) M.Sc. S.M., professeur agrégé, Département de médecine sociale et préventive, Université Laval.

2) M.D., M.Sc., F.S.A.M.S., chercheur santé nationale, Ecole de pharmacie, Université Laval.

3) M.N., professeur assistant, Ecole des sciences infirmières, Université Laval.

4) Ph.D., professeur agrégé, Faculté des sciences de l'administration, Université Laval.

Tirés à part:
Paul-Marie Bernard, professeur agrégé, Département de médecine sociale et préventive, Faculté de médecine, Université Laval, Ste-Foy.

Le présent article est issu d'une étude de prévalence sur l'usage de la cigarette en milieu scolaire, effectuée dans les écoles de la région de Québec en 1978. Cette étude, qui a fait l'objet d'un rapport publié chez l'Editeur officiel, a été menée grâce à une subvention du MAS et, en partie, à une subvention nominative de Santé et Bien-être social No 6605-1370-48.

était fixé par le nombre de remises subies par l'école. Le choix des classes dans les écoles s'est fait de façon non aléatoire en accord avec l'équipe de recherche, par le principal ou directeur d'école à qui avait été donnée la consigne de ne choisir que des classes typiques de son école. Au total 232 classes ont été choisies pour composer un échantillon global de 6211 écoliers, soit environ 10% de la population totale.

Stratégie et instrument de cueillette des données

La cueillette de l'information a été réalisée par l'application d'un questionnaire prétesté et adapté. Ce questionnaire se compose de deux parties: une première partie comporte une lettre d'information et une série de questions destinées à tous les répondants; une seconde partie correspond au statut du répondant (non-fumeur, ex-fumeur ou fumeur). Ce statut était déterminé par la réponse donnée à la question "Fumes-tu la cigarette?", apparaissant dans la première partie du questionnaire.

Pour l'application de ce questionnaire, dix monitrices ont été recrutées et entraînées par une simulation de situation. Elles ont suivi une procédure prédéterminée.

Saisie et validation des données

Les données furent saisies directement à partir des questionnaires par les préposés à la perforation. Une validation fut entreprise, identifiant les cas de codes inexistants et les défauts de consistance interne. La consistance des déclarations d'âge et de niveau scolaire fut établie par tabulation SPSS. Les erreurs furent corrigées à partir des questionnaires identifiés à la validation. Les cas de non-réponses furent exclus aux fins d'analyse.

Résultats

Des 6211 questionnaires passés dans les écoles, 95 furent exclus de l'analyse parce qu'ils marquaient des lacunes importantes sur les informations de base telles l'âge, le sexe, le niveau scolaire ou le statut. L'analyse a donc porté sur 6116 questionnaires convenablement remplis, ce qui correspond à 98,5% des enfants interrogés.

Les résultats ont été analysés dans le cadre des objectifs poursuivis par l'étude. Les données ont permis d'établir des résultats de prévalence, d'identifier les âges d'apprentissage et de tracer l'évolution temporelle du phénomène de l'usage du tabac.

Les répondants ont été classifiés suivant les cinq catégories décrites:

Non-fumeur (NF): qui dit n'avoir jamais fumé.

Fumeur régulier (FR): qui dit fumer actuellement une cigarette ou plus par semaine.

Fumeur occasionnel (FO): qui dit fumer, mais moins d'une cigarette par semaine.

Ex-fumeur régulier (EFR): qui dit ne plus fumer, mais avoir déjà fumé une cigarette ou plus par semaine.

Ex-fumeur occasionnel (EFO): qui dit ne plus fumer, mais avoir déjà fumé moins d'une cigarette par semaine. Cette catégorie peut regrouper les écoliers qui n'ont fait qu'essayer une cigarette, parfois même seulement quelques "bouffées".

Les résultats sont présentés en distinguant les niveaux scolaires, les sexes et les catégories de fumeurs. Ces dernières peuvent être regroupées en catégories plus larges comme non-fumeur (NF), fumeur (F = FR + FO) et ex-fumeur (EF = EFR + EFO).

Prévalence de l'usage du tabac

Le tableau I montre les données de prévalence suivant les différents niveaux scolaires et les sexes. Il permet d'identifier une décroissance rapide du pourcentage de non-fumeurs (NF) sur les niveaux scolaires. Chez les garçons, ce pourcentage de 60% qu'il est en primaire 4 passe à 28% en secondaire III. Chez les filles, les pourcentages correspondants sont de 67% et 18,9%.

Tableau I

Les habitudes de fumer décrites en pourcentage

Niveau scolaire	Sexe	NF (%)	FR (%)	FO (%)	EFR (%)	EFO (%)	1 + hier* (%)	Nombre total
4	M	60,2	2,0	2,0	5,0	30,8	2,3	603
	F	67,0	0,5	1,1	3,3	28,1	0,5	570
5	M	51,1	1,2	1,9	5,8	40,0	1,9	573
	F	57,4	1,2	0,7	4,9	35,8	0,5	589
6	M	42,7	2,2	2,9	7,8	44,4	2,4	588
	F	47,7	1,2	2,1	5,4	43,5	1,4	570
I	M	32,9	4,0	8,7	8,3	46,1	9,5	423
	F	33,7	5,7	10,1	11,1	39,4	10,7	495
II	M	29,2	6,1	12,7	9,7	42,2	13,0	424
	F	25,6	7,8	17,5	12,5	36,6	19,1	503
III	M	28,0	3,8	18,9	8,6	40,7	20,0	339
	F	18,9	8,0	29,4	8,4	35,3	32,6	439
Total	M	42,9	3,0	6,6	7,3	40,3	6,8	2950
	F	43,3	3,8	9,1	7,4	36,4	9,7	3166

*Pourcentage des écoliers qui ont déclaré avoir fumé une cigarette ou plus la veille du questionnaire.

La décroissance rapide du pourcentage de NF sur les niveaux scolaires s'accompagne chez les deux sexes d'une croissance marquée du pourcentage des EFO et des EFR au primaire et d'une croissance relayée au secondaire du pourcentage des FO.

On retrouve d'ailleurs, pour les deux sexes, un profil semblable dans l'évolution des pourcentages de fumeurs F (FR+FO) et d'ex-fumeurs EF (EFR+EFO) sur les niveaux scolaires. Le pourcentage d'EF, déjà élevé en primaire 4, croît rapidement jusqu'en secondaire I où un maximum est atteint (54,4 pour les garçons et 50,5% pour les filles). Le pourcentage des F est bas et stable au niveau primaire, mais amorce dès le niveau secondaire I une croissance rapide. De primaire 6 à secondaire I ce pourcentage fait plus que doubler chez les garçons (de 5,1% à 12,7%) et presque quintupler chez les filles (de 3,3% à 15,8%).

Une comparaison établie pour chaque niveau scolaire laisse voir une différence significative dans les habitudes de fumer entre les garçons et les filles en primaire 4, primaire 5 et en secondaire III. En primaire 4 et en primaire 5, il y a significativement plus de garçons qui ont expérimenté la cigarette que les filles ($\chi^2 = 5,89$, $p < 0,02$ pour le primaire 4; $\chi^2 = 4,57$, $p < 0,04$ pour le primaire 5). Pour les niveaux subséquents cette différence s'amenuise et même s'inverse. En secondaire III, il y a significativement plus de filles qui ont expérimenté la cigarette que de garçons ($\chi^2 = 8,97$, $p < 0,01$). Pour ce même niveau, il y a significativement plus de F chez les filles; par ailleurs on retrouve chez les garçons un pourcentage plus élevé d'EF.

Les réponses à la question "Combien de cigarettes as-tu fumées hier?" ont été compilées et ont permis de calculer les pourcentages rapportés dans la colonne "Une ou plus hier" du tableau I. Au primaire, moins de 2% des écoliers ont rapporté avoir fumé une cigarette ou plus la veille du questionnaire. Par ailleurs dès le secondaire I, ces pourcentages quintuplent presque et continuent de croître de façon géométrique sur les niveaux suivants. Pour les garçons, ces pourcentages sont de 9,5% en secondaire I et de 20,0% en secondaire III. Chez les filles, une augmentation encore plus marquée: 10,7% en secondaire I et 32,6% en secondaire III. Les filles au secondaire sont donc plus nombreuses à consommer la cigarette.

Quant à la quantité fumée, elle croît au secondaire avec le niveau scolaire et la consommation s'avère plus importante chez les garçons (Tableau II, figure 1).

L'âge à l'apprentissage

L'étude visait comme second objectif l'identification de l'âge de la première cigarette et l'âge à la régularité, des comparaisons devant être établies entre les sexes. Aucune différence dans les moyennes d'âge

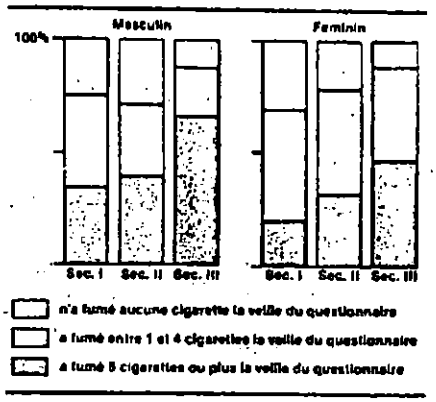


Fig. 1 — Répartition en pourcentage des écoliers fumeurs au secondaire suivant le nombre de cigarettes fumées la veille du questionnaire, par sexe et par niveau scolaire.

Tableau II

Pourcentage des écoliers fumeurs suivant leur consommation de cigarettes la veille du questionnaire

Niveau scolaire	Sexe	Consommation 0	1–4	5 +	Total	(effectif)
Primaire* 4, 5 et 6	M	44	44	12	100	(70)
	F	62	24	14	100	(37)
Sec. I	M	25	41	34	100	(53)
	F	31	49	20	100	(77)
Sec II	M	29	31	40	100	(78)
	F	22	46	32	100	(123)
Sec. III	M	12	22	66	100	(76)
	F	20	44	36	100	(162)
Total Sec.	M	22	30	48	100	(207)
	F	20	44	36	100	(362)

*Les trois niveaux du primaire ont été regroupés en raison des petits effectifs.

Tableau III

L'âge moyen à la première cigarette chez les écoliers qui ont déclaré avoir déjà fumé, suivant le sexe et le niveau scolaire

Sexe	Masculin			Féminin		
Niveau scolaire	x	σ	n	x	σ	n
4	7,98	1,47	223	7,94	1,57	179
5	8,50	1,84	266	8,91	1,58	243
6	9,23	1,92	331	9,77	1,78	291
Total Primaire	**8,65**	**1,85**	**820**	**9,01**	**1,81**	**713**
I	9,97	2,14	274	10,72	1,85	320
II	10,76	2,31	293	11,36	1,85	369
III	11,25	2,27	238	11,81	1,83	347
Total Secondaire	**10,63**	**2,30**	**805**	**11,31**	**1,90**	**1036**

à la première cigarette ne fut trouvée entre les différentes catégories de fumeurs ou ex-fumeurs, quelque soit le sexe et le niveau scolaire considéré. Le tableau III décrit pour chaque sexe et chaque niveau scolaire l'âge moyen à la première cigarette chez les écoliers qui ont déclaré avoir déjà fumé.

Ce tableau indique une croissance de l'âge moyen à la première cigarette avec le niveau scolaire. Pour tous les niveaux scolaires à l'exception du primaire 4, les garçons ont déclaré avoir fumé leur première cigarette à un âge moyen inférieur à celui déclaré par les filles. Ce délai dans l'expérimentation de la première cigarette entre les garçons et les filles est de 5 à 9 mois.

Par ailleurs la figure 2 illustre le comportement des deux sexes face à l'essai. Jusqu'à l'âge de 12 ans, proportionnellement moins de filles ont fait l'essai de la cigarette. Cependant, avant cet âge, soit à partir de 10 ans, la croissance des pourcentages se fait plus rapide chez les filles et conduit à une proportion plus grande d'essayeuses chez celles-ci dès l'âge de 12 ans.

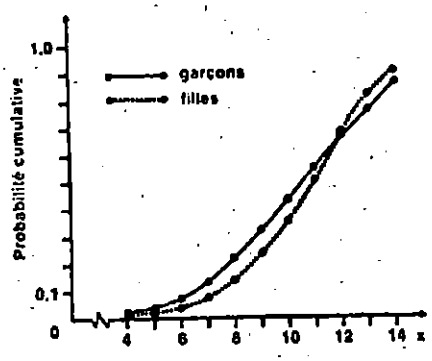


Fig. 2 — Probabilité cumulative de l'essai avant l'âge x selon le sexe.

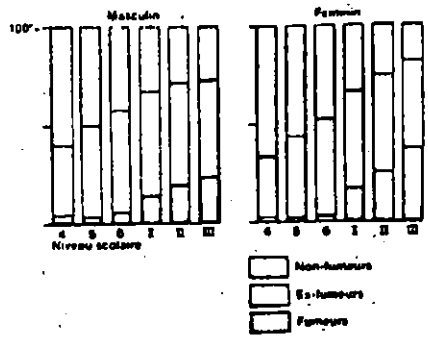


Fig. 3 — Distribution de fréquences en % des habitués face à la cigarette, suivant le sexe et le niveau scolaire. Ces graphiques sont construits à partir des données du tableau.

Evolution temporelle

La figure 3, tirée des données du tableau I, illustre bien le fait que le pourcentage des fumeurs croît avec le niveau scolaire. Alors que la proportion des ex-fumeurs marque une certaine stabilité sur les niveaux scolaires, la proportion des fumeurs, elle, croît rapidement à partir du secondaire et surtout chez les filles. La figure 4 montre par ailleurs un déplacement vers la gauche, avec l'âge actuel, des courbes décrivant la probabilité cumulative de l'essai avant l'âge X. Ce comportement des courbes indique que la proportion des essayeurs avant l'âge X est plus forte chez les plus jeunes. Le même phénomène avait été observé par Hanley et Robinson[3].

Discussion

Echantillonnage

Le mode d'échantillonnage utilisé ne peut pas être considéré comme tout à fait aléatoire. Il a dû être adapté au contexte scolaire, c'est-à-dire à l'organisation des cours et au fonctionnement de l'école. Le choix des classes par les principaux d'écoles a pu conduire à un échantillon biaisé.

Si un tel biais existe, la cohérence des résultats supporte l'hypothèse qu'il soit systématique. Par exemple, la plupart, sinon la totalité des 47 principaux, auraient pu retenir, pour l'échantillon, des classes où l'on fume moins. Mais il apparaît peu probable que les 47 principaux concernés aient biaisé leur choix dans le même sens.

Profil du comportement

Les principaux indicateurs du phénomène tabagique utilisés, soit l'essai et l'usage, subissent une transformation maximale entre le primaire 6 et le secondaire I. C'est dans cette période que l'on retrouve le plus fort pourcentage d'essayeurs (EFO), soit 46,0% chez les garçons et 44,0% chez les filles. C'est aussi dans cette période que l'on observe le plus fort taux de croissance dans les pourcentages de fumeurs (FO et FR). Parallèlement, on y observe la décroissance la plus rapide de la proportion des non-fumeurs. Il est donc possible de conclure que la transition du primaire au secondaire implique chez l'écolier une modification radicale des habitudes face à la cigarette. L'écolier essaie au primaire et devient fumeur au secondaire.

La catégorisation des écoliers en fumeurs (réguliers ou occasionnels), en ex-fumeurs, etc . . . , est basée sur les réponses de ceux-ci à des questions qui leur demandent un effort de mémoire et/ou qui les forcent à interpréter leur propre comportement: par exemple, les questions: "Fumes-tu à toutes les semaines?" ou "Fumes-tu à tous les jours?". Ce contexte peut limiter grandement la validité des réponses.

La question "Combien as-tu fumé de cigarettes hier?" permet de définir une classe d'indicateurs plus factuels que les précédents, non soumis aux facteurs qui en limitent la validité. Le seul désavantage est que la définition repose sur une courte période d'observation. Il serait tentant d'identifier chaque niveau de consommation à une catégorie de fumeurs; ainsi "0" à fumeurs occasionnels, "1 cigarette ou plus" à fumeurs réguliers et "5 cigarettes ou plus" à fumeurs sérieux.

Les pourcentages des écoliers qui déclarent avoir fumé une cigarette ou plus la veille du questionnaire, obéissent à un même profil

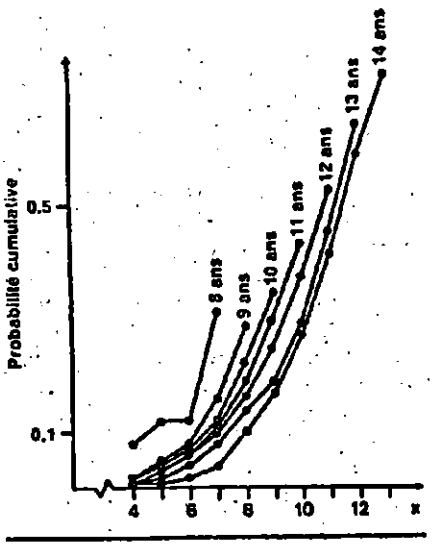


Fig. 4 —Probabilité cumulative de l'essai avant l'âge x suivant l'âge actuel.

que celui observé pour les pourcentages de fumeurs, avec une forte croissance de primaire 6 à secondaire I. Qui plus est, ces pourcentages sont similaires à ceux observés pour la catégorie des fumeurs occasionnels.

Cette similitude dans les deux séries de pourcentages marque peut-être l'existence d'un biais dans la classification des fumeurs, si les pourcentages de "1+hier" sont considérés comme "standards". Les écoliers fumeurs auraient eu tendance à sous-estimer leur propre consommation de cigarettes, biais qui favorise la classe des FO au détriment de la classe des FR.

L'âge à la première cigarette

Le premier essai de la cigarette se fait très tôt et davantage chez les garçons. Ainsi, 32% des garçons et 24% des filles en ont déjà fait l'expérience avant l'âge de 9 ans, soit le primaire 4 (figure 3). La croissance des moyennes d'âge avec le niveau scolaire peut s'expliquer par les effets non mutuellement exclusifs:

1) d'un biais de "mémoire"*;
2) d'une tendance réelle à fumer de plus en plus tôt chez les écoliers;
3) d'un facteur de vieillissement des cohortes considérées. La distribution d'âge à l'essai ou à la régularité subit avec le vieillissement de la cohorte un déplacement vers des valeurs supérieures, ce qui se traduit nécessairement par une croissance de la moyenne.

L'âge à la première cigarette n'apparaît pas comme un indicateur valide de l'acquisition de l'habitude de fumer. Le premier apprentissage pourrait se faire de façon diffuse. Toute intervention systématique pour prévenir ce type d'apprentissage serait donc difficile à établir et coûteuse à évaluer.

A l'appui de ce commentaire, il faut évoquer deux constatations. 1) Il n'y a pas de différence entre les moyennes d'âge à la première cigarette chez les fumeurs et les exfumeurs, quelque soit le sexe et le niveau scolaire. Cette absence d'effet différentiel de l'âge au premier essai suggère que cette variable ne soit pas déterminante dans le processus de l'adoption de l'habitude de fumer. 2) L'âge au premier essai constitue un idicateur qui présente certains problèmes de validation. Le déplacement des courbes de probabilités cumulatives (figure 4) en est une illustration. Ce déplacement peut s'expliquer tant par une tendance réelle à fumer de plus en plus tôt que par l'existence du biais de "mémoire". Par ailleurs, ce biais ne saurait être contrôlé sans multiplier les études de cohortes, longues et coûteuses.

Comparaison avec les études antérieures

Il pourrait être tentant d'établir une comparaison au niveau des résultats avec certaines études canadiennes antérieures dont celle de Hanley et Robinson[2] et celle de Laforest[3]. Cependant, une telle comparaison ne peut être établie que si les études à comparer présentent des similitudes au plan des objectifs, de la population visée, de la méthode et de l'analyse. Or les deux études citées marquent avec la présente étude des dissimilitudes qui peuvent interdire toute comparaison sérieuse.

Conclusion

La présente étude a permis de constater, en octobre et novembre 78, l'importance du tabagisme dans la population scolaire francophone de la région métropolitaine de Québec et d'établir:

1) qu'environ 32% des garçons et 24% des filles ont déjà expérimenté la cigarette avant l'âge de 9 ans, soit le primaire 4;
2) que les pourcentages d'essayeurs croissent rapidement au primaire, pour ensuite se stabiliser au secondaire, tant chez les garçons que chez les filles;
3) que les pourcentages de fumeurs (occasionnels et réguliers), d'abord stables au primaire, croissent rapidement au passage de primaire 6 à secondaire I;
4) que les filles commencent à fumer généralement plus tard que les garçons, mais sont plus nombreuses à le faire.

La présente étude s'inscrit dans le cadre d'une démarche d'intervention face au problème du tabagisme dans les écoles. Elle devait permettre d'identifier le temps fort du phénomène. L'étude de Laforest proposait en conclusion d'entreprendre une action préventive avant le secondaire I. La présente étude soutient dans ses résultats cette proposition et identifie le primaire 6 comme groupe-cible d'une intervention préventive, soit juste avant le passage au secondaire I.

Summary

A prevalence study, among 6000 schoolchildren of the Quebec Region, has established the importance of tobacco smoking. An early smoking exposition and a rapid increase in smoking behavior during transition from primary to secondary school are stressed. Performance indicators and optimal intervention period are identified, with respect to health education programs aimed at smoking by schoolchildren.

Bibliographie

1. Pigeon, J. et Montambault, P.: Le tabagisme: un problème de santé prioritaire au Québec. Toxicomanies. 10: 35-44, 1977.
2. Hanley, J.A. et Robinson, J.C.: Cigarette Smoking and the Young: a National Survey. CMAJ., 114: 511-517, 1976.
3. Laforest, L.: Exploration des facteurs déterminant le degré d'exposition de la population étudiante au tabagisme (cigarette). Bureau de la recherche. Direction de l'usage non médical des drogues, sept. 1975.
4. Bynner, J.M.: Government Social Survey. The Young Smoker. A Study of smoking schoolboys carried out for the Ministry of Health. London, Her Majesty's Stationery Office, 1969.
5. Kelson, S.R. et coll.: The Growing Epidemic. Amer. J. Public Health. 65: 923-938 1975
6. Le tabac et la santé au Canada. Notes de recherche. Direction générale de la planification à long terme (santé) et Direction de l'usage non médical des drogues. Santé et bien-être social Canada, mars 1977.
7. Politique de lutte contre le tabagisme. Document préliminaire. Planification des services de santé. Ministère des Affaires sociales, août 1978.
8. Rapport d'un comité d'experts de l'O.M.S.: Les effets du tabac sur la santé. Série de rapports techniques no 568, O.M.S. Genève, 1975.
9. Russell, M.A.H.: Realistic Goals for Smoking and Health. A Case for Safer Smoking. Lancet. 1: 254-258, 1974.
10. Usage du tabac au Canada de 1965 à 1974. Rapports techniques no 1. Bureau de la recherche. Direction de l'usage non médical des drogues. Protection de la Santé. Santé et bien-être social Canada, janvier 1976.
11. Weiss, Carol H.: Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, New Jersey, 1972, p. 34.

*Les répondants font mention de l'âge où ils croient avoir essayé la cigarette pour la première fois. Les répondants d'un âge plus avancé ont tendance à situer l'événement du premier essai à un âge plus avancé en raison d'un mauvais rappel mnémonique.

LECTURE POUR VENDREDI LE 10 JUIN 1983

AMERICAN
Journal of Epidemiology
Formerly AMERICAN JOURNAL OF HYGIENE

© 1978 by The Johns Hopkins University School of Hygiene and Public Health

VOL. 108 OCTOBER, 1978 NO. 4

Reviews and Commentary

CAUSATION AND DISEASE: A CHRONOLOGICAL JOURNEY

THE THOMAS PARRAN LECTURE[1,2]

ALFRED S. EVANS[3]

The invitation to give the Thomas Parran Lecture is both a pleasure and an honor for me. I wish first to indicate something of the life and contributions of Thomas Parran. Dr. Parran was born in 1892, and received the Master of Arts degree at the University of Maryland and his medical degree from the Georgetown University, both in 1915. He joined the Public Health Service shortly after graduation, became Chief of the Division of Venereal Disease, and then was appointed Health Commissioner of the State of New York by Franklin Delano Roosevelt when he became Governor of that state. When Roosevelt was elected to the Presidency he appointed Dr. Parran to Surgeon General of the Public Health Service. In addition to these important posts in the practice of public health Dr. Parran was a founder of the World Health Organization and both the founder and first Dean of the Graduate School of Public Health at the University of Pittsburgh. On retirement from Pittsburgh he became President of the Avalon Foundation. Some of Dr. Parran's contributions to public health included his courageous attack on venereal diseases, his promotion of clinical research at the National Institutes of Health, his emphasis on the right of individuals to good health and medical care, his promotion of improvements in diet, particularly in good milk supplies, and his participation in international health movements, such as the WHO and the Pan American Health Organization. Dr. Parran's book on syphilis, entitled *Shadow on the Land: Syphilis* (published by Reynal and Hitchcock, New York, 1937), in which he presented his program for the control of venereal disease in the United States, was a pioneering publication that dared to bring this problem to the open attention of the public. It should still be of interest to those concerned with the control of infectious diseases in this country.

[1] Delivered at the Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, December 15, 1977.

[2] Adapted, by permission, from "Causation and disease: The Henle-Koch postulates revisited," by A. S. Evans. *Yale Journal of Biology and Medicine*, Volume 49, pages 175–195, 1976.

[3] Professor of Epidemiology and Director of WHO Serum Reference Bank, Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT 06510.

Henle-Koch Postulates

In my presentation on the chronology of causation and disease, I wish to emphasize three points: first, that the Koch postulates should really be termed the "Henle-Koch postulates"; secondly, that those postulates

were not regarded as rigid criteria by Koch himself at the time they were presented and should not be regarded as such today; and thirdly, that all of our concepts of causation are limited by the technology available to prove them and our understanding of the pathogenesis and epidemiology of the disease at the time of the investigation. A fuller exposition of this chronology has previously been published (1). Let us first turn to the Henle-Koch postulates. Jakob Henle was a German-born physician who went to Zürich at age 31 as Professor of Anatomy where in addition to his brilliant contributions to our understanding of the histology of the retina and the kidney and many other organs he wrote in 1840 an essay, "On Miasmata and Contagie" (2), in which the concept was advanced that "Before microscopic forms can be regarded as the cause of contagion in man they must be constantly found in contagious material. They must be isolated from it and their strength tested." This essay was written some 42 years before the first bacillus, that of tuberculosis, was discovered by Henle's pupil, Robert Koch, in 1882. Henle left Zürich, went to Germany and finally became Professor at the University of Göttingen where Koch studied under him and adopted the concepts of his teacher. Koch's contribution to causation was made originally in 1884 and more formally presented in a lecture given in 1890 at the International Congress in Berlin (3). These postulates have become our classical point of reference in relating causative agents to disease. The major features were that the parasite must be present in every case of the disease under appropriate circumstances, that it should occur in no other disease as a fortuitous and non-pathogenic parasite, and finally that it must be isolated from the body in pure culture, repeatedly passed, and induce the disease anew. These criteria of causation have been slavishly held up as the only true basis for establishing causation. As recently as the October 29, 1977 issue of Lancet (4) some putative bacterial causes for Crohn's disease were discarded because they did not fulfill these postulates. Yet it is clear that failure to fulfill them fully was not regarded by Koch himself at the time of presentation as being grounds to exclude the possibility of a causal relationship and they certainly should not be so regarded today. Koch recognized that the cause of anthrax, tuberculosis, erysipelas, tetanus and many animal diseases fully fulfilled his criteria. At the same time he recognized that there were many causes of illness which did not fulfill the criteria, such as typhoid fever, diphtheria, leprosy, relapsing fever and Asiatic cholera. Cholera, in particular, should be noted because Koch isolated the organism of this disease in Egypt and was well aware of its capacity to produce epidemic disease in man. So, even at the time of presentation, Koch emphasized that not all the criteria were necessary for proof and that just the first two were sufficient. The second criteria of the postulates became a problem in fulfillment when the carrier state was recognized for diphtheria by Park and Beebe (5) and by the work of Chapin (6) and by Koch himself for typhoid fever (7) in the period 1895–1902.

Causation in virus diseases

The virus era opened in the 1900s, and by 1930 many viruses had been identified in mice and tissue culture. In 1937, Thomas Rivers, in his Presidential Address to the American Society of Immunology (8), emphasized the following limitations of the Koch postulates for virus diseases: 1) more than one agent might be needed to produce a given disease; 2) viruses could not be grown on lifeless media but required living cells in contrast to bacteria; and 3) asymptomatic carriers existed. Rivers proposed a set of criteria quite similar in essence to those of the Henle-Koch postulates but included two additional concepts: 1) the reproduction of the disease experimentally should exclude the possibility that the laboratory animal was infected with some laboratory virus and include appropriate con-

trols, and 2) that antibody to the disease should appear during the course of the illness.

HUEBNER: GUIDELINES FOR ESTABLISHING A VIRUS AS CAUSE OF DISEASE

These concepts of Rivers on the etiology of virus diseases became untenable when many viruses were sometimes found to be simultaneously present in ill persons and even in healthy people. In 1957 Dr. Robert Huebner who is now Chief of the Laboratory of Tumor Virology, National Institutes of Health, discussed the virologist's dilemma (9). He indicated that the presence of many viruses—even in normal people—made the identification of the presence of a virus of low order in establishing causation and one could derive spurious causative associations if you based the conclusions on this fact alone. He recognized that some infections were due to multiple viruses, and that sometimes viruses could produce chronic diseases, or carrier states, or that viral reactivation occurred. Huebner's contributions included what he called a bill of rights for prevalent viruses. They emphasized that the mere presence of the virus should not be regarded solely as the basis for etiology and he introduced epidemiologic principles by longitudinal and cross-sectional studies as an element of proof in causation. He also emphasized that the disease should be prevented by a specific vaccine to the suspected agent. This concept is of great importance in establishing the etiologic relationship of a suspected infectious agent to that disease. Huebner also emphasized that one needed money in order to accomplish the establishment of proof and included this as a ninth criteria for etiologic studies.

BROADENING OF THE CONCEPTS OF CAUSATION FOR VIRUS DISEASES

These concepts of causation for virus diseases were then broadened over the next 10 years by the recognition that the same clinical picture or syndrome could be produced by a variety of different agents, that different agents predominated under different epidemiologic circumstances, and that the host response to a given virus would vary from one setting to another. These ideas were expressed by Evans in 1960 (10) and 1967 (11) as they related to acute infectious diseases. It also became clear that the agent alone was a necessary but not sufficient factor needed to cause most diseases. Cofactors and the susceptibility of the host were of key importance in the occurrence of *clinical* illness.

IMMUNOLOGIC PROOF OF CAUSATION

The next problem arose with the isolation of a virus now known as the Epstein-Barr virus (12) and with the identification of an antigen known as the Australia antigen by Blumberg and his associates (13). These two agents were established as causally related to infectious mononucleosis and to viral hepatitis B, respectively, mainly by serologic studies. The agents were a particular challenge in terms of the Henle-Koch postulates because they could not be grown in pure culture nor did their injection reproduce the disease in experimental animals. This resulted in the establishment of an immunologic proof of causation. The criteria required that antibody to the agent is regularly absent prior to the disease and that it regularly appears during illness. The absence of the antibody should indicate susceptibility to infection and its presence indicate immunity. Antibody to no other agent should be similarly associated with the disease unless it is a co-factor. These criteria of immunologic proof were the basis for the association of Epstein-Barr virus and infectious mononucleosis (14). The discovery of this relationship was made in 1968 by Dr. Werner Henle, his wife, Dr. Gertrude Henle, and a young associate, Dr. Volker Diehl, in Philadelphia (15). It is a particular irony that Dr. Henle, who is the grandson of Jakob Henle, established the causative relationship of EBV to infectious mononucleosis without fulfilling a single

one of the postulates set up by his grandfather and by Robert Koch. This serves to emphasize that we must change our criteria with our technology.

Slow viral infections and chronic neurologic disease

The next challenge to the Henle-Koch postulates arose with the recognition of agents called slow viruses. These caused kuru, a disease occurring in the Fore tribe in New Guinea, and Jakob-Creutzfeldt disease, a pre-senile dementia. The work on these agents carried out by Dr. Gajdusek and his associates at the National Institutes of Health resulted in the award of the Nobel prize to Dr. Gajdusek in 1977 (16). The particular difficulties posed by these agents in relation to causation were their long incubation period, their association with chronic neurologic diseases and, most importantly from the standpoint of the Henle-Koch postulates, the fact that the agents could not be isolated in tissue culture in the laboratory. Furthermore, these unique agents did not produce an immune response and were highly resistant to a great variety of physical and chemical agents. A new set of criteria for causation was therefore needed for this group.

Dr. Richard Johnson of The Johns Hopkins University and Dr. Clarence Gibbs, an associate of Dr. Gajdusek's at the NIH, proposed guidelines (17) for establishing causal relationship which were based on 1) the consistency in animal transmission of the agent, 2) the fact that the agent should be serially transmissible in experimental animals with filtered material and 3) that similar results should not be obtained from normal tissues. In light of our knowledge of the presence of many viruses in the intestines and throat of healthy persons, it would not be surprising to me if normal brain tissue may not sometimes contain latent agents. It is well recognized that herpes zoster and simplex and papova viruses can remain latent in many tissues, including nerve tissue. Thus, their third criteria may produce difficulties in interpretation in the future. There are also some "conventional viruses" associated with chronic neurologic diseases, namely measles virus in relation to subacute sclerosing panencephalitis (SSPE) (18) and papova virus in relation to progressive multifocal leukoencephalopathy (PML) (19). In both of these conditions reactivation of a virus in which the primary exposure was years before is responsible for the chronic disease observed. In SSPE the main epidemiologic event appears to be infection with measles virus very early in life at a time of inadequate host response (20). With PML, alteration in our immune system by immunosuppression either naturally induced, as in Hodgkin's disease, or artifically induced by immunosuppressive drugs is responsible for the reactivation of this agent (21). Thus the criteria for causation in these diseases depends on the identification of the agent in brain tissue, the presence of high antibody titers to the virus and, at least for SSPE, the experimental reproduction of the disease in a laboratory animal (22).

Evidence relating viruses to cancer

The possibility that viruses might cause cancer required new criteria of causation. The problems in establishing this relationship include the long incubation period between exposure to the putative agent and the cancer and the probability that disease results from reactivation of the virus rather than from a primary infection. The agents most commonly recognized as the best oncogenic candidates are the herpes viruses. But proof of causation is difficult because they are common and ubiquitous viruses, probably require co-factors, and there are difficulties in reproduction of the cancer in animals. In addition, human volunteer studies are not possible. There is also the probability that the cancer may have different causes in different geographic areas or under different epidemiologic settings (23). The most likely oncogenic candidate is Epstein-Barr virus in relationship to Af-

rican Burkitt lymphoma and to nasopharyngeal cancer. The elements of causation consist immunologically of the presence of higher antibody titers in cases than in healthy, matched controls and the demonstration of high levels of antibody prior to the cancer; and virologically of the presence of the viral genome in tumor cells and the ability of the virus to transform cells; and thirdly, (at least for Burkitt lymphoma), experimental evidence that the virus can induce a malignant lymphoma in animals. Similar virologic and serologic proof exists for EBV and nasopharyngeal cancer, except this cancer has not been reproduced in experimental animals. It is important to emphasize that co-factors are involved in the appearance of both of these cancers. In Burkitt lymphoma the essential co-factor appears to be malaria. Infection both by EBV and by the malarial parasite in the first year of life may be requirements for a malignancy to develop. This infection occurring early in life is reminiscent of the relation of SSPE and measles. The presence of high antibody titers prior to the development of Burkitt lymphoma has recently been demonstrated in a massive prospective study of 45,000 children in West Africa (24). It is important to recognize that not all tumors fulfilling the histologic criteria of Burkitt lymphoma are associated with EBV. American Burkitt lymphoma, for example, is not usually associated with high EBV titers, the genome has been absent from the tumor tissue and in some cases EBV antibody has been entirely absent from the blood, indicating that EBV infection had never occurred (25). In nasopharyngeal cancer genetic factors seem to be an important risk factor since HLA antigen 2 and a new antigen called Singapore 2 have both been associated in Chinese with a fivefold higher incidence of their susceptibility to nasopharyngeal cancer than in Chinese lacking these characteristics (26). Another virus of the herpes group associated with possible malignancy is herpes type 2 in cervical cancer. Here the virologic and serologic evidence is much less convincing than in African Burkitt lymphoma (27). It is clear that at least 40 per cent of cancer of the cervix may occur in the absence of antibody to HSV 2. My own view of the relationship of viruses to cancer is that the host plays a critical role in the development of the malignancy. Such factors as the age at the time of infection, the immunologic state of the host, the presence of co-factors such as concurrent infection, and the genetic attributes as expressed by human leukocyte antigens, all play a role in the pathogenesis of the malignant process.

Causation in chronic disease

Let us now turn to the issue of causation in chronic disease. Interest in this began around 1957 and was based on the Henle-Koch postulates. Important in this early work was the contribution of the late Dr. Yerushalmy and the late Dr. Palmer (28) who proposed a series of relationships that include the concepts that 1) the suspected characteristic must be found more frequently in persons with the disease than in persons without, 2) that persons possessing the characteristic should develop the disease more frequently than those not possessing it, and 3) that this association must be tested for its validity by investigating the relationship of the characteristic to other diseases in order to establish the specificity of the characteristic and the disease. These guideposts for implication of an etiologic factor in a chronic disease were amplified by Dr. Abraham Lilienfeld, who proposed five additions: 1) the incidence of the disease should increase in relation to the duration and intensity of the suspected factor, 2) its distribution should parallel that of the disease in all relevant aspects, 3) a spectrum of illness should be related to the exposure to this suspected factor, 4) reduction or removal of the factor should reduce or eliminate the disease and 5) human populations exposed to the factor in controlled or even in natural experiments should de-

velop the disease more commonly than those not exposed (29).

The contributions of the committee appointed by the Surgeon General to delineate the relationship between smoking and health should also be recognized (30). They included five features of association between a putative cause and the disease; namely, consistency, strength, specificity, temporal relationship, and the coherence of the association.

Criteria for causation: a unified concept

A review of these various concepts of causation in infectious diseases and in chronic diseases has led me to try and formulate a unified concept (1). It combines the elements of many of the guidelines previously mentioned. The main features are 1) the prevalence of the disease should be higher in those exposed than in those not exposed, 2) that exposure to the putative cause should be present more commonly in those with the disease than in those without the disease, 3) that the incidence should be higher in persons who are so exposed than in those not exposed as shown in prospective studies, 4) that exposure to the suspected factor should precede the disease, 5) that there should be a measurable biologic spectrum of host responses, 6) that experimental reproduction of the disease should be demonstrated, 7) that elimination of the putative cause should decrease the incidence of the disease, and 8) that prevention or modification of the host response should decrease or eliminate the expression of the disease. These concepts are not original with me but put together a concept of causation applicable to both infectious and non-infectious diseases. Just as the Henle-Koch postulates cannot be regarded with any finality, so too, these concepts should be taken only as guidelines, subject to our changing knowledge of technology and causation.

In summary, the original Henle-Koch postulates had many limitations, some recognized then, and others later (table 1). Fulfillment of the postulates is certainly reasonable grounds for accepting a causal role of the putative agent but lack of fulfillment of the postulate should not exclude such a relationship. This chronologic view has focused primarily on the relation of a single cause to a disease. This is a simplistic view because most infectious agents are a necessary but not sufficient cause of disease; indeed many viral infections are inapparent. Causation in both infectious and non-infectious disease involves a complex interplay of agents, environmental, and host factors. The latter include the host's immunologic status, genetic background, socioeconomic level, hygienic practices, behavioral patterns, age at the time of exposure and the presence of co-existing disease. Different qualitative and quantitative mixes of the agent, environment, and host may result in the same clinical and pathological diseases under different circumstances. These more sophisticated approaches to causality are discussed in recent books (31–33) and articles (34–39).

Another view of the causal relationship of an agent to disease might be framed in legal terms. Table 2 presents some of these comparisons (35). In criminal law, the presence of the criminal at the scene of the

Table 1

Limitations of Koch's postulates of causation*

1) Not applicable to all pathogenic bacteria.
2) May not be applicable to viruses, fungi, parasites.
3) They do not include the following concepts:
   a) The asymptomatic carrier state.
   b) The biologic spectrum of disease.
   c) Epidemiologic elements of causation.
   d) Immunologic elements of causation.
   e) Prevention of disease by elimination of putative cause as element of causation.
   f) Multiple causation.
   g) One syndrome has different causes at different settings.
   h) Reactivation of latent agents as cause of disease.
   i) Immunologic processes as cause of disease.

* More properly termed the Henle-Koch postulates.

TABLE 2

Rules of evidence: criminality and causality

Mayhem or murder and criminal law	Morbidity, mortality and causality
1) Criminal present at the scene of the crime	Agent present in lesion of the disease.
2) Premeditation	Causal events precede onset of disease.
3) Accessories involved in the crime.	Co-factors and/or multiple causality involved.
4) Severity or death related to state of victim.	Susceptibility and host response determine severity.
5) Motivation—the crime must make sense in terms of gain to the criminal.	The role of the agent in the disease must make biologic and common sense.
6) No other suspect could have committed the crime in the circumstances given.	No other agent could have caused the disease under the circumstances given.
7) The proof of the guilt must be established beyond a reasonable doubt.	The proof of causation must be established beyond reasonable doubt or role of chance.

crime would be equivalent to the presence of the agent in a lesion of the disease. Premeditation would be similar to the requirement that the causal exposure should precede the onset of the disease. The presence of accessories at the scene of a crime might be compared to the presence of co-factors and/or multiple causes for human diseases. The severity of the crime or the consequence of death might be loosely equivalent to susceptibility and the host responses which determine the severity of the illness. The motivation involved in a crime should make sense in terms of reward to the criminal, just as the role of a causal agent should make biologic sense. The absence of other suspects and their elimination in a criminal trial would be similar to that of the exclusion of other putative causes in human illness. Finally, need that the proof of guilt must be established beyond a reasonable doubt would be true for both criminal justice and for disease causation.

What will the future hold in terms of new infectious diseases in which it may be necessary to look for causal relationships? First, we have those new diseases which might be produced by changes in the agent or its point of entry. Resistant organisms will continue to emerge which are resistant to antibiotics and other therapeutic drugs, or will appear as a result of plasmid transfer of genetic factors of resistance. Recently we have recognized the plasmid transfer of resistance of typhoid organisms to chloramphenical, of the gonococcus to penicillin and most recently of penicillin resistant forms of the pneumococcus. All these indicate that there is an epidemic potential of these highly resistant organisms. It is also frightening to think that this resistance might be transferred *in vivo* to other organisms. For example, is it possible that a resistant gonococcal infection of the mouth might lead to the transfer of that plasmid to a streptococcus or to a meningococcus present in the same environment? Then we have new agents in old portals. The recognition of new agents of infectious diseases —probably transmitted by the respiratory routes, such as Lassa fever and Marburg fever, are examples of this. We also have old agents in new portals, such as the transfer of herpes type 2 virus from its normal habitat beneath the diaphragm to a new habitat in the oropharyngeal cavity as a result of changing sexual practices. New agents will be identified for old diseases like the common cold of which we now only recognize about half of the causative agents or for encephalitis in which only 25–50 per cent have a recognized cause or for diarrheal diseases. There are important new advances in the discovery of rotavirus in infant diarrhea, of toxogenic *Escherichia coli* in the traveler, and of *Giardia lamblia* in the camper, but the majority of common

gastroenteritides remain unexplained. Finally, there are latent infections that are being reactivated which were formerly unrecognized as human pathogens. For example, reactivation of latent papova virus causes progressive multifoccal leukoencephalopathy but we have not recognized any disease syndrome associated with the common primary infection with this agent.

There are also environmental factors which may result in new diseases. The expanded tourist travel, research exploration, and military excursions to the remote corners of the globe may bring us in contact with new pathogens not previously identified. Our exploration of the ocean bottoms and of outer space may expose us to new microbial forms. Our own creation of special environments, such as intensive care units or kidney dialysis units may produce new and unusual host responses.

Changing economic factors are more likely to result in the re-emergence of old diseases than the creation of new ones. Funds which are now provided for the surveillance and immunization of important diseases, like smallpox, cholera, polio and measles may not be appropriated as these diseases are brought under control and lose their emotional and political clout. Increased economic development and better hygiene may delay exposure to common agents like polio, infectious mononucleosis, and hepatitis from early childhood to young adult life. This may result in the occurrence of more *clinical illnesses* due to these agents because the host response is more severe in older life.

There are also host factors which may play a role in the emergence of new diseases. These include those which are associated with our increasing use of new therapies, such as immunosuppressive drugs, new antibiotics, organ transplants and tissues from man and perhaps other primates, and the use of synthetic organs and valves. With a large proportion of our population now living into the sixth, seventh and eighth decade, latent infections may reactivate and cancer increase as our normal immunologic controls diminish in efficacy. The control of umbrella diseases like malaria and schistosomiasis in tropical settings may permit recognition of host responses which were previously hidden by the coexistence of these infections. Changing host susceptibility may occur as a consequence of new environmental stresses, of immunologic manipulations, or possibly even of genetic engineering. Changing social and cultural habits may result in new patterns of illness. So we may be faced in the future with a variety of new and unrecognized infectious diseases whose etiology must be established. Proof of this relationship must be based on common sense, good guidelines of causation appropriate to existing technology and a keen sense of the biologic basis of disease.

Causation of Health

This presentation has reviewed our changing concepts of the causation of disease. Perhaps it is time we also look at the causation of health (table 3). Let us direct our attention not only to those factors that produce disease but also to those that produce health. Let us change our "don'ts" for "do's" in medical practice and public

TABLE 3

Postulates for the causation of health

1) The preventive factor must be consistently present in persons of good health or free of a particular disease.
2) The factor must be isolatable in a pure form, (i.e., can be identified as causal).
3) The extent to which the factor is effectively applied must parallel an increase in good health and/or freedom from that disease.
4) Experimental application of the factor to one segment of a population should significantly increase their good health as compared with matched controls.
5) Withdrawal of the preventive factor should be associated with an increase of disease associated with that factor.
6) The effect of the factor shall be measured in terms of lower morbidity and mortality, longer life, and lower medical costs.

health. Let us approach the proof of causation of health with a keen scientific, statistical, and biologic orientation.

Let us do studies of what makes us survive and not only of what makes us die. Let us do case/control studies of healthy persons with ill persons as the control; let our prospective studies of populations focus on those who remain in good health rather than only those who become ill. Let us recognize that there will be multiple causes for health just as there are multiple causes for disease and that health like disease is a biologic spectrum. Let us recognize that establishing the proof of the causation of health will be more difficult and challenging than the causation of disease. But let us begin.

References

1. Evans AS: Causation and disease: The Henle-Koch postulates revisited. Yale J Biol Med 49:175–195, 1976
2. Henle J: On Miasmata and Contagie. (Translated from German by G Rosen). Johns Hopkins Press, Baltimore, 1938
3. Koch R: Ueber bakteriologische Forschung. *In* Verh. X. International Medical Congress, Berlin 1890. p 35, 1892
4. Ward M: The pathogenesis of Crohn's disease. Lancet 2:903–990, 1977
5. Park WH, Beebe AL: Diphtheria and pseudodiphtheria. Med Rec 46:385–401, 1894
6. Chapin CV: The Sources and Modes of Infection. New York, Wiley, 1910
7. Koch R: Die Bekämpfung des Typhus. Vortrag gehalten in der Sitzung des wissenschaftlichen Senats bei der Kaiser Wilhelms Akademie am 28 November, 1902
8. Rivers TM: Viruses and Koch's postulates. J Bacteriol 33:1–12, 1937
9. Huebner RJ: The virologist's dilemma. Ann NY Acad Sci 67:430–445, 1957
10. Evans AS: Common clinical syndromes of infectious disease. I. Introduction and common respiratory diseases. II. Common infectious exanthem. III. Common infections of the central nervous system. IV. Common diarrheal diseases. Wis Med J 59:508–512, 553–556, 656–660, 700–702, 1960
11. Evans AS: Clinical syndromes in adults caused by respiratory infection. Med Clin N Am 51:803–815, 1967
12. Epstein MA, Achong BG, Barr YM: Virus particles in cultured lymphoblasts from Burkitt lymphoma. Lancet 1:702–703, 1964
13. Blumberg BS, Sutnick AI, London WT, et al: Australia antigen and hepatitis. N Engl J Med 282:349–354, 1970
14. Evans AS: New discoveries in infectious mononucleosis. Mod Med 42:18–24, 1974
15. Henle G, Henle W, Diehl V: Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis. Proc Natl Acad Sci USA 59:94–101, 1968
16. Gajdusek DC: Kuru and Creutzfeldt-Jakob disease. Ann Clin Res 5:254–261, 1973
17. Johnson RT, Gibbs CJ Jr: Editorial. Koch's postulates and slow infections of the nervous system. Arch Neurol 30:36–38, 1974
18. Sever JL, Zeman W (Editors): Conference on measles virus and subacute sclerosing panencephalitis. Bethesda, Maryland, September 19, 1967. Neurology 19:30, 1968
19. Zu Rhein GM: Association of papova-virions with a human demyelinating disease (progressive multifocal leukoencephalopathy). Prog Med Virol 11:185–247, 1969
20. Detels R, Brody JA, McNem J, et al: Further epidemiologic studies of subacute sclerosing panencephalitis. Lancet 2:11–14, 1973
21. Narayan O, Penney JB, Johnson RT, et al: Etiology of progressive multifocal leukoencephalopathy. N Engl J Med 289:1278–1282, 1970
22. Byington DP, Johnson KP: Experimental subacute sclerosing panencephalitis in the hamster. Correlation of age with chronic inclusion-cell encephalitis. J Infect Dis 126:18–26, 1972
23. Evans AS (Editorial): The spectrum of infections with Epstein-Barr virus: A hypothesis. J Infect Dis 124:330–337, 1971
24. de Thé G, Geser A: A prospective seroepidemiological study to investigate the role of EBV in Burkitt's lymphoma. *In* Unifying Concepts of Leukemia. Bull Haematol No 39 (edited by RM Durcher, L Chieco-Bianchi), Karger, Basel, 1973, p 448
25. Pagano JS, Huang CH, Levine P: Absence of Epstein-Barr viral DNA in American Burkitt's lymphoma. N Engl J Med 289:1395–1399, 1973
26. Simons MJ, Wee GB, Day NE, et al: Immunogenetic aspects of nasopharyngeal carcinoma. IV. Probable identification of HLA second locus antigen associated with high risk of nasopharyngeal carcinoma. Lancet 1:142–143, 1975
27. Melnick JL, Adam E, Rawls WE: The causative role of herpesvirus type 2 in cervical cancer. Cancer 34:1375–1385, 1974
28. Yerushalmy J, Palmer CE: On the methodology of investigations of etiologic factors in chronic diseases. J Chronic Dis 10:27–40, 1959
29. Lilienfeld AM: On the methodology of investigations of etiologic factors in chronic disease—Some comments. J Chronic Dis 10:41–46, 1959
30. Surgeon General, Advisory Committee of the USPHS: Smoking and Health. PHS Publication No 1103. Washington DC, US GPO, 1964
31. Lilienfeld A: Fundamentals of Epidemiology. New York, Oxford University Press, 1976
32. MacMahon B, Pugh TF, Ipsen J: Epidemiologic Methods. Boston/Toronto, Little, Brown, 1960
33. Susser MW: Causal Thinking in the Health Sciences; Concepts and Strategies of Epidemiology. New York, Oxford University Press, 1973, p 181

34. Bollet AJ: On seeking the cause of disease. Clin Res 12:305, 1964
35. Rothman KJ: Causes. Am J Epidemiol 104:587–592, 1976
36. Koopman JS: Causal models and sources of interaction. Am J Epidemiol 106:439–444, 1977
37. Cassel J: The contribution of the social environment to host resistance. Am J Epidemiol 104:107–123, 1976
38. Miettinen OS: Proportion of disease caused or prevented by a given exposure, trait, or intervention. Am J Epidemiol 99:325–332, 1974
39. Stewart GT: Limitations of the germ theory. Lancet 1:1077–1088, 1968

280

les études rétrospectives en épidémiologie clinique

P. Fournier *, J.P. Bocquet **

Les enquêtes de type rétrospectif tendent à être de plus en plus utilisées en épidémiologie clinique où elles constituent un net progrès par rapport au simple report de cas. Ces études peuvent être rapidement mises en œuvre et à un moindre coût. Toutefois, leur élaboration doit être rigoureuse car elles sont sujettes à de nombreuses erreurs systématiques. Dans ce domaine, l'établissement de standards méthodologiques est en cours et ce fait doit permettre leur extension. Il convient toutefois de garder à l'esprit qu'il s'agit d'observations structurées et que les conclusions que l'on peut en tirer ne sont en rien comparables à celles découlant de démarches plus purement expérimentales.

LE but de cet article est d'exposer, de la manière la plus explicite possible, les avantages et inconvénients des enquêtes de type rétrospectif [1], afin d'aboutir à des considérations pratiques qui permettront aux chercheurs cliniciens d'entreprendre ou de faire entreprendre ce type de recherche.

D'une manière générale, l'objet de la démonstration du chercheur est l'approche de relations de type cause à effet. Il peut s'agir de l'étude de nouvelles relations de ce type ou de la confrontation de relations établies avec des observations nouvelles.

Cette approche peut s'effectuer de deux façons :

- Soit en utilisant une méthode purement expérimentale. Un schéma expérimental très rigoureux, contrôlant effectivement toutes les variables susceptibles d'intervenir dans la relation étudiée, est alors mis sur pied, mais sa réalisation est délicate, voire souvent impossible.

- Soit en utilisant des méthodes d'observation. C'est généralement dans cette situation que se place le chercheur clinicien, et dans ce contexte les études rétrospectives lui sont extrêmement utiles.

L'accent sera mis sur les défauts inhérents à ce type d'étude, tant au niveau de schéma expérimental que de la limitation de l'interprétation des résultats mais aussi sur leurs avantages, ainsi que les raisons qui ont poussé à les voir se multiplier.

La recherche clinique s'intéresse essentiellement à l'évaluation des techniques diagnostiques, à la recherche des facteurs étiologiques des maladies, au pronostic de patients en fonction des différentes étapes du processus pathologique et de la présence ou l'absence de certains facteurs, à l'efficacité et aux effets iatrogènes des différentes thérapeutiques. Le domaine de l'expérimen-

MOTS CLÉS : *Etudes rétrospectives - Epidémiologie clinique.*

1. Tout au long de cet article on entendra par étude rétrospective les études cas-témoins (en anglais : case-control, case-referent, case-compeer, retrospective).

*Département de médecine sociale et préventive (Pr Jenicek), université de Montréal. ** Maître conf. agr. de médecine préventive, santé publique et hygiène, université de Nice, chef Serv. méd. int. et mal. infect., hôp. de l'Archet, Nice.

CM – 26-02-1983 – 105-09

tation thérapeutique a connu un développement important et a acquis, à l'aide de schémas expérimentaux adaptés, une individualité certaine au sein de la recherche clinique (1). Pour cette raison, il sera exclu des considérations de cet article.

Le report de cas est une forme usitée de présentation de facteurs étiologiques et pronostiques. Il peut ne concerner que des aspects ponctuels ou, dans sa forme la plus complète, relater l'histoire de la maladie.

Si cette démarche ne comporte pas de groupes témoins, elle peut être considérée comme purement descriptive, elle est alors utile pour attirer l'attention des cliniciens sur des affections ou des manifestations peu communes ou nouvelles. Les observations isolées doivent entrer dans une rubrique générale qui pourrait être intitulée « Faits cliniques ».

Si ces reports ont pour finalité l'approche d'une relation causale, ils doivent comporter un ou des groupes témoins. Quoique en situation d'observation, la démarche est alors plus analytique et doit se prévaloir d'un cadre méthodologique défini et rigoureux.

Toutes ces activités peuvent et doivent bénéficier des progrès récents de l'épidémiologie analytique. Dans ce domaine, les études prospectives et rétrospectives sont d'un grand recours pour le chercheur clinicien.

Les études prospectives et rétrospectives diffèrent fondamentalement en ce sens que, dans un cas (prospectif), *la cause sert à établir la distinction entre les groupes expérimentaux et témoins, l'effet est alors attendu ; alors que dans le cas des études rétrospectives, l'effet est déjà connu et permet d'établir les groupes à partir desquels les causes seront recherchées rétroactivement* (groupes de cas et de témoins).

L'enquête prospective s'inscrit plus dans la séquence temporelle naturelle observable dans le cadre d'une relation de cause à effet, une enquête rétrospective va à l'encontre de cette séquence, mais présente par ailleurs de nombreux avantages.

Quelques aspects historiques extension de ce type d'études

Si les premières études qui visaient à montrer rétroactivement des relations de cause à effet ont vu le jour au XIX[e] siècle, il n'en demeure pas moins qu'elles ont été peu développées jusqu'aux années 1920. Elles ont par la suite connu un essor lié au développement des statistiques qui leur ont permis de s'accompagner d'un cadre de référence méthodologique plus rigoureux. Ensuite, sous l'impulsion de biostatisticiens (Cornfield, Mantel, Haenszel, Miettinen), elles se sont considérablement développées.

Cette extension récente a été chiffrée (2) *(tableau 1).*

En vingt ans, la proportion des enquêtes rétrospectives parmi les études à caractère épidémiologique a été multipliée par près de 2,5. L'étude de la relation « tabac-cancer du poumon » (1950) a d'abord été abordée à l'aide d'études rétrospectives.

Tableau 1. – *Evolution de la proportion des enquêtes rétrospectives dans deux revues médicales.*

	Articles originaux à caractère épidémiologique parus dans deux revues *	*Part des enquêtes rétrospectives dans ces articles*
1956-1957	4 à 5 %	18 %
1976-1977	11 à 15 %	43 %

* Il s'agit de deux revues médicales non spécialisées à grande diffusion : *The Lancet, The New England Journal of Medicine.*

Les grandes étapes d'une enquête rétrospective

Au cours de ce chapitre et des suivants, on supposera que le schéma expérimental s'adresse à l'étude d'une relation de cause à effet.

Choix des cas

Il se déroule en quatre étapes.

1° ***Critère de diagnostic :*** il permet d'établir la distinction entre les cas et les témoins. Ce critère doit être rigoureux, s'il s'agit d'un ou d'une série de tests, ils doivent être spécifiques et sensibles.

2° ***Source des cas :*** elle peut être très variable. Dans le cadre de la recherche clinique, il s'agira le plus souvent d'une série de malades qui, au cours de leur affection, sont entrés en contact avec un ou des Services médicaux spécifiés ou parfois des sujets vus à l'occasion de consultations externes.

3° ***Mode de sélection des cas :*** parfois seront inclus dans l'étude tous les cas, s'ils sont disponibles et plus volontiers encore si l'affection est rare. Si l'on a recours à une fraction de ces cas, la sélection doit s'effectuer selon des critères probabilistes.

4° ***Cas incidents ou prévalents ?*** il est reconnu qu'il est préférable de ne choisir que les cas incidents (survenus au cours d'une période déterminée), plutôt que des cas prévalents (cas présents à un instant donné), afin d'éviter un biais lié aux caractéristiques qui ont permis aux cas prévalents de survivre.

Choix des témoins

C'est indiscutablement l'étape la plus importante et la plus délicate de ces enquêtes. Le bien-fondé des conclusions et la validité des observations reposent pour beaucoup sur le choix des témoins. En effet, ce choix est laissé à l'entière appréciation du chercheur, car il n'existe pas un groupe témoin indiqué comme cela est parfois le cas pour d'autres études, mais une multitude de groupes possibles parmi l'ensemble des personnes malades ou non malades, qui ne présentent pas l'affection étudiée.

1° ***Critère de diagnostic :*** en toute rigueur, les témoins devraient avoir subi la même démarche diagnostique que les cas et y avoir présenté une réponse négative. Pour des raisons évidentes d'éthique et de coût entre autres, il n'est pas convenable de recourir à cette procédure. Généralement, les témoins n'auront pas suivi l'ensemble des investigations appliquées aux cas.

Toutefois, le chercheur devra s'assurer que le groupe des témoins ne contient pas des sujets présentant un processus pathologique du même type que les cas mais à un stade infra-clinique.

2° ***Source de témoins :*** dans la pratique si les cas proviennent d'un hôpital, les témoins seront choisis parmi les personnes fréquentant le même hôpital. On considère en effet que l'ensemble de ces personnes présentent des caractéristiques communes qui les amènent à fréquenter cet hôpital.

Choisir ses groupes dans cette population augmente leur comparabilité. Toutefois, cette attitude pratique ne doit pas être érigée en loi car il est des situations dans lesquelles elle n'est pas fondée.

Si les témoins sont choisis dans le même hôpital, il faudra veiller à ce que leur motif d'hospitalisation ou de consultation ne soit pas lié à une pathologie trop proche de celle des cas. Par exemple, si l'on étudie la relation « tabac-cancer du poumon », les cas seront des sujets qui ont une néoplasie pulmonaire, les témoins ne doivent pas être choisis parmi des malades présentant une bronchite chronique car cette affection est également en partie due au tabac.

Dans le cas d'une étude où le facteur étiologique est l'alcool, le choix de témoins dans un Service de médecine générale n'est pas approprié car l'alcool peut être directement ou indirectement relié à la présence du témoin dans le Service.

3° ***Sélection :*** on aura parfois recours au

même mode de sélection que celui des cas, d'autres fois un mode différent sera utilisé.

L'appariement est actuellement très usité. Il consiste à choisir pour chaque cas un ou plusieurs témoins semblables pour certaines caractéristiques. Ces caractéristiques, équitablement réparties dans les deux groupes, seront dites contrôlées. Il s'agit généralement de l'âge, du sexe, de la catégorie socio-professionnelle, du groupe ethnique, de la date d'hospitalisation, etc. En pratique, il convient de ne pas apparier par plus de 4 ou 5 variables. Ces variables doivent être reliées à la fois à la cause et à l'effet étudié. C'est la seule situation dans laquelle l'appariement confère au schéma expérimental adopté une validité accrue (3).

Recueil des données

Toutes les différentes techniques de recueil des données peuvent être utilisées (interviews, questionnaires, dossiers médicaux...), l'important étant la comparabilité des données entre les différents groupes. Il est absolument nécessaire que la qualité de ces données ne varie pas d'un groupe à l'autre. Dans une étude rétrospective, cette qualité a tendance à ne pas être constante pour de nombreuses raisons.

Inconvénients des études rétrospectives

1° Le choix arbitraire de deux groupes dans lesquels la prévalence de la cause à étudier est 100 % (cas) et 0 % (témoins) fait que l'on ne peut déduire de ce type d'étude que des résultats relatifs et non absolus. La mesure qui résume mieux le degré d'association entre un facteur étiologique et une affection donnée est le risque relatif (RR). Il est commun de dire qu'il mesure la « force » d'une association. Exprimé par l'intermédiaire d'un chiffre absolu, sa valeur indique le surcroît de présence de l'affection étudiée chez les malades exposés par rapport aux sujets non exposés. Un risque relatif de 2 signifie que les sujets exposés présenteront 2 fois plus souvent la maladie étudiée que les non-exposés.

Il est par contre d'autres indicateurs qui ne peuvent pas directement être calculés à partir des études rétrospectives (risque attribuable, fraction étiologique du risque).

2° Quand la cause (exposition) est rare, le nombre de sujets nécessaires dans chaque groupe est élevé. Un exemple pratique infirme cette remarque. C'est ce type d'enquête qui a permis de mettre en évidence la relation « exposition in utero au diéthylstilbestrol et cancer vaginal » (4).

3° L'inconvénient majeur ou plutôt les inconvénients majeurs sont les très nombreuses sources de biais que comportent ces études. Plus que dans tout autre type d'étude, toutes les étapes des enquêtes rétrospectives sont susceptibles de conduire à des erreurs systématiques.

On peut grossièrement les regrouper en trois catégories :

a) ***Biais de sélection.*** Les plus importants à connaître sont :

• *Les biais dus aux cas prévalents* (biais de Neyman). Souvent pour des raisons de commodité, l'investigateur constituera ses groupes à partir de tous les malades qui présentent l'affection à étudier au moment où l'enquête débute. Ces groupes comprendront des malades dont les premiers symptômes et le diagnostic s'échelonnent sur de longues périodes antérieures. Cette fraction du groupe est constituée par les survivants et/ou ceux qui continuent à fréquenter l'établissement où se déroule l'étude. Ce fait amène à sur-représenter ces malades. Cela peut constituer une source de biais qui peut être évitée en ne sélectionnant que les cas (ou une fraction d'entre eux) qui ont été diagnostiqués pendant une période déterminée.

Par exemple, une étude portant sur le cancer verrait les formes mineures sur-représentées si l'on décidait de choisir tous les sujets qui fréquentent un Service hospitalier ou un Centre spécialisé à un moment donné.

Il conviendrait mieux de choisir l'ensemble des cas qui ont été diagnostiqués pendant une période donnée, ainsi toutes les différentes formes de cette maladie seraient représentées de façon satisfaisante.

• *Les biais dus aux taux d'hospitalisation différents dans les groupes et dépendant de la cause à étudier* (biais de Berkson).

Ainsi si l'on désire étudier la relation « alcool-cancer de l'œsophage », le choix de témoins, non atteints de cette affection, dans un Service de gastro-entérologie, amènera à sur-représenter la présence du facteur alcool dans ce groupe, car il peut constituer directement ou indirectement le motif d'hospitalisation. De ce fait, on ne décèlera pas de relation entre ce facteur et la maladie ou une relation inverse de celle qui existe.

b) ***Biais lié à la collecte des données***

– Selon la nature de l'information que l'on désire recueillir, il peut exister des différences de mémorisation ou d'attitudes chez les cas et chez les témoins ;

– La ou les personnes chargées de collecter les données peuvent avoir des attitudes différentes selon qu'elles s'adressent à des cas ou à des témoins, etc.

c) ***Biais dus au défaut de contrôle de variables intervenant dans la relation étudiée***

Ici, c'est généralement l'établissement du schéma expérimental qui est en cause.

Tous ces biais peuvent conduire à une distorsion de la relation étudiée. Selon les cas, cette relation apparaîtra renforcée ou minimisée, voire inversée. Certains ouvrages traitent de l'algèbre de ces erreurs de classification (5). Récemment, un article a établi un catalogue de tous ces biais, 57 sont actuellement recensés et la collaboration de tous les chercheurs est sollicitée afin de le compléter (8).

Avantages des études rétrospectives

1° Le premier des avantages est incontestablement *le coût réduit* de ces études et le fait qu'elles puissent être mises en œuvre rapidement. Ceci est particulièrement important dans le cas des maladies chroniques, où le laps de temps séparant l'exposition aux facteurs étiologiques et la survenue de ces maladies est généralement long. Une étude rétrospective, en plaçant l'observateur après l'effet, ne nécessite pas une surveillance qui, dans certaines situations, exigerait plusieurs années.

On peut remarquer que cet avantage est aussi un inconvénient dans la mesure où il rend plus difficile et aléatoire la collecte de données de bonne qualité.

2° Dans le cas de maladies rares ou peu fréquentes, c'est souvent *le seul moyen d'investigation* dont on dispose en recherche étiologique. On entend par seule méthode applicable le fait que des tests statistiques puissent être utilisés avec une puissance et une précision convenables.

Exemple

L'exemple suivant va permettre d'illustrer cet avantage. Il montre quel est le nombre de sujets nécessaires, dans le cas où l'on a un groupe de cas et un groupe témoin, pour des études rétrospectives et prospectives. Il s'applique à des affections et des expositions de fréquences différentes *(tableau 2)*.

Afin de pouvoir établir des comparaisons entre ces différents exemples, les erreurs de première espèce (α), de deuxième espèce (β) ainsi que le risque minimum à déceler (risque relatif (RR) = 2) sont identiques dans tous les cas.

Cela signifie que dans chaque exemple le risque de conclure à tort de l'existence d'une relation (erreur de première espèce) est de 5 %, le risque de conclure à tort de l'absence de relation (erreur de deuxième espèce) est de 20 %, le risque relatif minimum susceptible d'être décelé est 2.

On peut constater que dans le cas d'une maladie rare (tumeur hypophysaire), une étude rétrospective nécessitera un nombre

	Rétrospectif	*Prospectif*
Maladie fréquente		
● *Accidents de la route :* incidence 2 500/100 000		
● *Expositions :*		
– *Fréquente :* existence d'une affection chronique, 15/100	207	1 046
– *Rare :* état d'ivresse au volant, 1/100	2 394	904
Maladie peu fréquente		
● *Cancer du sein :* incidence 150/100 000		
● *Expositions :*		
– *Fréquente :* contraceptifs oraux 30/100	140	21 132
– *Rare :* nulliparité 5/100	515	18 780
Maladie rare		
● *Tumeur de l'hypophyse :* incidence 1/100 000		
● *Expositions :*		
– *Fréquente :* contraceptifs oraux 30/100	140	< 999 999
– *Rare :* nulliparité 5/100	515	< 999 999
Avec $\alpha = 5\%$; $\beta = 20\%$, RR = 2.		

Tableau 2. – *Nombre de sujets nécessaires par groupe dans le cas d'études prospectives et rétrospectives* (7).

raisonnable de sujets, 140 pour une exposition fréquente (contraceptifs oraux), 515 pour une exposition plus rare (nulliparité). Dans le cas d'une étude prospective, avec les mêmes exigences statistiques, le nombre de sujets nécessaires (supérieur à 1 million) rend ce type d'approche non envisageable.

Cet exemple illustre également la remarque émise précédemment au titre des inconvénients : plus l'exposition est rare moins l'enquête rétrospective semble justifiable.

Conclusion

Il est indéniable que la vogue que les études rétrospectives connaissent actuellement ira grandissant. Cela tient au coût réduit et à la rapidité avec laquelle ce type d'étude peut être entrepris. De plus, il est certaines situations, maladies ou effets rares où il n'est pas envisageable d'entreprendre un autre type d'étude. Toutefois, l'instigateur ou l'utilisateur de ces études devra garder présents à l'esprit deux points particulièrement importants :

1° ***La conception même de ces études entraîne la possibilité d'observer de très nombreux biais.*** Puisque ces biais sont conceptuels, il faut les connaître et utiliser un schéma expérimental qui les supprime ou les minimise. Dans ce domaine, l'établissement de standards méthodologiques est en cours, et les chercheurs cliniciens devraient en bénéficier.

2° ***Quelles conclusions tirer de ces études et quelles inférences peut-on faire en direction de populations mal définies, voire hypothétiques ?***

Il convient de considérer le résultat d'une enquête rétrospective comme une **observa-**

tion. La démarche empirique veut que ce soit la conjonction de nombreuses observations concordantes qui amène l'établissement de lois ou de relations. Il y a donc lieu d'examiner le résultat de chacune de ces études en le confrontant avec d'autres études portant sur le même sujet, tout en examinant le rôle que pourrait avoir chaque schéma expérimental sur l'observation d'une concordance ou d'une discordance des résultats.

Il est à remarquer que la tendance actuelle dans ces études est le choix de plusieurs (3 ou 4) groupes témoins qui appartiennent à des populations différentes. Dans le cas où tous les résultats concordent, cette situation conforte l'expérimentateur qui pourra ainsi plus facilement appliquer à la population générale les résultats concernant l'existence ou la non-existence de la relation qu'il étudiait.

Il est possible de conclure en affirmant que les études rétrospectives constituent un élément de promotion indiscutable de la recherche clinique mais que pour ce, il faut en connaître les aspects méthodologiques et ne pas tirer de l'observation de leurs résultats des conclusions hâtives et non fondées.

Le développement de systèmes de recueil systématique de données, tels que les fichiers « Cancer » mis en place dans certains départements ou certaines régions fournirait une abondante source de données que les études rétrospectives pourraient traiter aisément.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. **Schwartz D, Flamant R, Lellouch J.** L'essai thérapeutique chez l'homme. Paris, Flammarion Médecine Sciences éd., 1970.
2. **Cole P.** The Evolving Case-Control Study. *J Chronic Dis* 1979, 32, 15-27.
3. **Miettinen OS.** Matching and design efficiency in retrospectives studies. *Am J Epidemiol* 1970, 91, 111-118.
4. **Herbst AL et al.** Adenocarcinoma of the vagina : association of maternal stillbestrol therapy with tumor appareance in young women. *New Engl J Med* 1971, 284, 878-881.
5. **Fleiss J.** Statistics for rates and proportions. New York Wiley éd., 1973.
6. **Sackett D.** Biais in analytic research. *J Chronic Dis* 1979, 32, 51-63.
7. **Schlesselman JJ.** Sample size requirement in cohort and case control studies. *Am J Epidemiol* 1974, 99, 381-384.

l'essentiel

Les enquêtes de type rétrospectif sont de plus en plus utilisées en recherche clinique, particulièrement celles à visée étiologique. Elles permettent de mieux structurer la présentation d'observations et, par là, se différencient du simple report de cas.

Si ces études peuvent être aisément entreprises et à un moindre coût, il n'en demeure pas moins que leur mise en œuvre est sujette à de nombreuses erreurs systématiques. La principale difficulté réside dans le choix de groupes témoins, car ce choix conditionne les possibilités ultérieures d'inférence.

Les biais les plus fréquemment rencontrés concernent la sélection des malades et des témoins ainsi que la collecte des données.

L'établissement de standards méthodologiques est en cours, la collaboration des chercheurs cliniciens et des épidémiologistes doit permettre l'extension de ce type d'étude. Il n'en demeure pas moins que les conclusions à tirer des résultats obtenus à l'aide de cette approche doivent être prudentes.

SPECIAL ARTICLE

DOUBLE STANDARDS, SCIENTIFIC METHODS, AND EPIDEMIOLOGIC RESEARCH

ALVAN R. FEINSTEIN, M.D., AND RALPH I. HORWITZ, M.D.

Abstract Because suspected noxious agents can seldom be tested experimentally in people, epidemiologic investigation of the causes of disease depends on observational procedures, of which the most prominent types are cohort and case–control studies. In these studies the compared agents cannot be assigned experimentally, but many other basic experimental principles that could be employed are often overlooked, including verification of quality and accuracy in raw data, avoidance of major biases in comparison, vigilance in checking for methodologic errors, and maintenance of a careful distinction between research data that generate hypotheses and the new data needed to test these hypotheses. The problems that result when these principles are overlooked can be alleviated if etiologic research is performed and evaluated with the same standards used in other branches of science. (N Engl J Med. 1982; 307:1611-7.)

IN contemplating the health hazards of such diverse entities as pharmaceutical agents, nutritional ingredients, industrial products, air pollution, and personal habits, readers must evaluate evidence obtained in a manner that differs substantially from the customary scientific methods and materials of laboratory research. The evidence regarding health hazards often comes from epidemiologic studies in which the methods were not experimental and the material consisted of groups of people who could not be assembled or examined under laboratory conditions.

Because the results of such studies are often accompanied by elaborate statistical presentations, readers may be too awed or confused by the statistics to look closely at the scientific quality of the underlying evidence; or efforts to appraise the evidence may be thwarted by difficulties in understanding the unfamiliar investigative procedures used to obtain it. Suspending critical judgment and hoping, as an act of faith, that the epidemiologic methods were scientifically satisfactory, the reader may then concentrate on medical, sociopolitical, and other implications that arise from interpretations of the statistics.

The difficulties of performing observational epidemiologic research may sometimes lead to the use of investigative procedures that overtly violate the standards used in other branches of science. The violations can escape critical attention because they occur at such fundamental levels of research — the quality of data, fairness of comparison, and cogency of methods — that the maintenance of rigorous scientific principles may be taken for granted and seldom questioned. Unfamiliar with the basic strategies and tactics of the research, other scientists may not consider the possibility that defects can exist at such elementary levels.

Our purposes in this discussion are to outline two of the main investigative structures used in epidemiologic research (the observational cohort study and the retrospective case–control study), to discuss some of the basic methodologic challenges, and to suggest some easily attained improvements. To illustrate the problems, we compare the procedures used in observational cohort and case–control studies with procedures used in other forms of scientific investigation.

From the Yale University School of Medicine, New Haven, and the Cooperative Studies Program Support Center, Veterans Administration Medical Center, West Haven, Conn. Address reprint requests to Dr. Feinstein at Yale University School of Medicine, 333 Cedar St., P.O. Box 3333, New Haven, CT 06510.

Supported in part by grants from the Commonwealth Fund and the Robert Wood Johnson Foundation (No. 6309). Dr. Horwitz is a Henry J. Kaiser Family Foundation Faculty Scholar in General Internal Medicine.

THE OBSERVATIONAL STRUCTURES OF EPIDEMIOLOGIC RESEARCH

In traditional scientific circumstances a cause–effect relation between a suspected noxious agent and the disease allegedly produced by that agent would be investigated as an experiment. After being checked for eligibility and examined to verify the absence of the disease under investigation, the people who entered the experiment would be randomly assigned to receive either the potentially noxious agent or an alternative (or no) exposure. The two groups of exposed and nonexposed people would then be followed, with each person examined in the same way at regular intervals to detect the possible occurrence of the disease. At the end of a suitably long period of follow-up observation, the occurrence of the disease would be enumerated and expressed as the rate of risk for each group. If a statistically significant difference existed between the two rates, and if no other plausible explanation could be found, exposure to the agent would be held responsible for causing the difference in rates of risk.

Despite the obvious scientific advantages of such experiments they can almost never be used to investigate etiologic relations in epidemiologic research. Most investigators would be reluctant on ethical grounds to ask large numbers of people to expose themselves to a possibly noxious agent; even if asked, very few persons would volunteer to be randomly assigned to exposure or nonexposure; and those who did volunteer could not be coerced into maintaining the assigned regimen. Because of these ethical and logistic barriers to the performance of experiments, epidemiologists have relied instead on observational studies of people who received the noxious agents through mechanisms other than randomized assignment. The exposure may have been self-imposed (e.g., through personal habits or nutritional patterns), or it may have

occurred as a prescribed medication, an atmospheric pollutant, or an occupational toxin.

The epidemiologic research structure that most resembles an experiment is the observational cohort study. It has the same basic arrangement and forward direction as an experiment, but the suspected noxious agent and the control agent (or nonexposure) are not assigned by randomization. After the members of the exposed and unexposed groups are identified, they are followed forward in time to determine whether the disease occurs subsequently. The calculated rate of occurrence indicates the risk for each group, and the relative importance of the two risk rates is estimated by combining them into a single index. This index can be either a subtracted increment, which is called the "attributable risk," or a quotient, which is called the "risk ratio" or "relative risk."

For example, an investigator who suspects that tea drinking causes renal stones might assemble a group of 4000 tea drinkers and a group of 4000 people who did not drink tea. After arranging to follow the two groups for the next 10 years, the investigator might find the results shown in Table 1. From these data, the rates of

Table 1. Development of Renal Stones among 4000 Tea Drinkers and 4000 People Who Do Not Drink Tea.

Group	Renal Stones	No Renal Stones	Total
		no. of people	
Drink tea	200	3800	4000
Do not drink tea	50	3950	4000
Total	*250*	*7750*	*8000*

risk would be calculated as 200/4000, or 0.05, for the tea drinkers and 50/4000, or 0.0125, for those who do not drink tea. The risk ratio would be 0.05/0.0125, or 4.0, and the incremental difference (or attributable risk) between the two groups would be 0.05 − 0.0125, which equals 0.0375, or 37.5 per thousand. A risk ratio above 1 would suggest that tea drinking might cause renal stones, a relatively small incremental difference would suggest that the hazard was not frightful, and the discovery of a small but distinct risk in the unexposed group would indicate that tea drinking could not be the only cause of renal stones.

The logic of this type of forward-directed cohort research is readily recognized by scientific investigators, since the basic approach is analogous to the procedures used in laboratory work. An observational cohort study, however, can seldom be conducted to investigate the causes of most human diseases. Few investigators would be willing to spend 10 years observing the course of 8000 people. Even if the investigators were willing, it might be difficult to assemble a large number of people who would be willing to participate in a long-term study, and even if the people were willing, the cost of the research might be prohibitive. These problems in logistics and feasibility occur whenever the attack rate (or risk) of the disease in question is relatively low and whenever the duration of the allegedly pathogenetic exposure is relatively long. In these circumstances cohort research is extremely difficult to conduct, because too many people have to be assembled and followed for too long a period.

As a substitute for this customary type of forward-directed research, epidemiologists have devised the case–control study, which follows a reverse direction of scientific logic. The investigator begins at the end rather than at the beginning of the causal pathway by studying people who already have the disease. A control group is assembled from people in whom the disease has not been noted, and controls are often selected to match cases in such demographic features as age and gender. The two groups are then followed in a backward direction, with inquiries aimed at determining each member's antecedent exposure to the alleged etiologic agent.

For example, in a case–control study of the relation between tea drinking and renal stones the investigator might assemble 100 people with renal stones and 100 people without them. After inquiring about antecedent tea drinking, the investigator might tabulate the numbers shown in Table 2.

Unlike cohort data, these data cannot be used to determine the risk for exposed and unexposed people. If such calculations were attempted, they would show an occurrence rate of 79/128, or 0.62, for renal stones in tea drinkers and a corresponding value of 21/72, or 0.29, in those who did not drink tea. Both values are unrealistically high estimates of the risk of renal stones, and erroneous results would be obtained if these values were used to calculate the incremental risk or the risk ratio. To avoid this problem, investigators calculate the odds of exposure in the case and control groups by using the ratio of exposed to unexposed people in each group. Thus, the odds for previous tea drinking would be 79/21, or 3.76, in the patients with stones and 49/51, or 0.96, in the controls. These two ratios are then divided to form an "odds ratio," which has been shown mathematically[1] to approximate the value of the risk ratio that might be found in a forward-directed cohort study. In the cited example the odds ratio would be 3.76/0.96, or 3.92, which is close to the risk ratio (4.0) calculated for the corresponding cohorts described earlier.

Although the actual risks cannot be directly calculated, the other logistic and mathematical advantages of case–control studies have made them an appealing

Table 2. Antecedent Tea Drinking among 100 People with Renal Stones and 100 People without Them.

Antecedent Tea Drinking	With Stones	Without Stones	Total
		no. of people	
Drank tea	79	49	128
Did not drink tea	21	51	72
Total	*100*	*100*	*200*

epidemiologic mechanism for investigating the causes of disease. The case–control procedure avoids the large sample sizes and protracted follow-up periods needed in cohort research, and the results can be mathematically arranged to approximate the risk ratio. With this appeal, the case–control study has become so prominent that it is now the main source of most current epidemiologic research in the causes of disease.[2]

On the other hand, since neither observational cohort studies nor case–control studies are conducted under experimental circumstances, the results may not always accurately reflect what might be found with a randomized experiment, which serves as the scientific "gold standard" in cause–effect research. The rest of this discussion is concerned with the basic scientific principles of experimental research — quality of data, fairness of comparison, and cogency of methods — and with the way in which these principles are followed or neglected in observational epidemiologic studies.

Quality of Data

A fundamental principle of science is verification of the raw data used to identify and characterize the objects under study.

Identification of Basic Groups

Each member of the groups under comparison must be carefully checked to confirm the presence or absence of the characteristic (such as a particular exposure, disease, or treatment) that identifies the group. In a clinical experiment the compared groups are easily identified, since the fundamental act of the research is assignment of the active or control agent that defines the groups. In observational cohort research the investigator does not assign the compared agents, so the basic groups are identified on the basis of data indicating exposure or nonexposure to the agent in question.

This information is relatively easy to obtain and verify in a preplanned cohort study.[3] The people can be asked about their exposure as they enter the study, and other sources of data can be used to confirm their statements. If the groups are assembled in retrospect, years after the exposure or nonexposure, using data from medical records or other sources of information that were not collected to meet scientific investigative goals, the data may be less accurate and more difficult to check. For example, when information is obtained in retrospect for a study of the hazards of being employed in a particular occupation, the available data regarding the cohort of people who worked in that occupation 20 years earlier may not indicate the extent to which each person was actually exposed to the particular substance suspected of being toxic.

In a case–control study the groups are identified according to the presence or absence of the disease under investigation. Since all other analyses begin after the case and control groups have been chosen, this first step in the diagnostic identification of each participant is a crucial scientific activity. Unfortunately, in case–control research diagnostic status is not regularly verified for either the cases or the controls. Instead, the investigators usually accept the diagnoses that were entered in the patients' medical records. If the control group is assembled from a community population, diagnostic status may be determined only by statements made by the interviewed persons.

From time to time, case–control investigators may actually assemble and review the data, specimens, or other material used to diagnose the disease in question. This type of review is used to exclude patients with false-positive diagnoses of the disease, but it is almost never applied to the false-negative side of the diagnostic problem. The control group simply consists of people in whom the target disease has not been diagnosed; their true status as members of the control group is almost never checked. In many instances, their true status cannot be checked, because the appropriate diagnostic tests were not performed. Thus, in a study of patients with renal stones, the members of the control group may not have received the appropriate tests to demonstrate that their kidneys are indeed free of stones; the controls for a group of patients with breast cancer may not have had their breasts carefully examined; and the controls for patients with uterine cancer may not have received any of the special diagnostic tests needed to show that such a cancer is absent.

Confirmation of Subsequent Data

In experimental clinical science, after the basic groups are identified, vigorous efforts are made to ensure that the subsequent raw data are accurate and trustworthy. Fulfillment of this principle leads to such activities as developing and testing data-collection instruments and establishing quality control for chemical measurements.

Since the groups compared in cohort research are identified according to the presence or absence of exposure to the suspected noxious agent, the main raw data thereafter relate to the subsequent development of the disease under investigation. This information is commonly obtained from death certificates — a notoriously inaccurate source of data about specific diseases — and no further confirmation is attempted. When additional efforts are made, the work may be confined to checking only for false-positive diagnoses. Thus, the investigators may try to substantiate the reported diagnosis of disease, but efforts are seldom made to check for false-negative situations in which the disease was present but not recognized. The problem of a false-negative diagnosis is particularly important, as noted below, for the many infarctions, cancers, stones, and other noninfectious diseases that often escape detection during a patient's lifetime and are noted only on the increasingly infrequent occasions when autopsy is performed.

In case–control research the crucial subsequent information consists of data about each person's previous exposure to the suspected etiologic agent. This information is usually obtained from entries in medical records, from direct conversation with the patient, or

from questionnaires answered by the patient. The reliability of the data is seldom confirmed, however, by checking information obtained from the medical record against other sources, by repeating the interview or the questionnaire at a later date, or by other determinations of the accuracy with which the patient has recalled events that may have occurred many years earlier. Furthermore, the data are often obtained by excerpters or interviewers who are aware of the participant's status as a case or control and who also know the hypothesized causal relation under investigation.

The epidemiologic literature contains little discussion of these problems or the tactics to be used in resolving conflicts between data collected from two different sources. In two studies[4,5] in which the same information (about occupation and medication used during pregnancy) was checked in the same way on two different occasions, the results showed substantial disparities between the two sets of data. Our own research[6] has shown that the odds ratios can differ considerably according to the way in which the investigator combines or rejects discrepancies in data from different sources.

This neglect of quality control in basic data has been so pervasive in epidemiologic research that one of the current leaders[7] in the field recently decried the "extremely sketchy" way in which data-gathering methods are described in published reports of research; the failure to note "the type of training . . . the interviewers received"; and the infrequency with which basic epidemiologic data are replicated, validated, or confirmed. He concluded that epidemiologists have paid "inadequate attention to the quality of the data obtained in our investigations" and that the need for improving "the validity of the data is particularly urgent."

Fairness of Comparison

A scientifically fair comparison of the effects of two agents requires that all other pertinent elements of the comparison be arranged in an equitable manner. If substantial inequities result in a major bias, it may be impossible to determine whether subsequent differences in results are due to this bias or to differences in the impact of the compared agents.

In addition to verification of basic data, attempts to achieve a fair comparison are a hallmark of the scientific experiment. Among the more obvious safeguards against bias are careful selection of control groups, randomization, and double-blind procedures. A catalog of all the diverse forms of bias[8-10] that can arise when these safeguards are absent in nonexperimental research is beyond the scope of this discussion, but three of the most prominent scientific problems in both cohort and case–control studies will be mentioned briefly.

Susceptibility Bias

In an experimental trial the main purpose of randomization is to assign the compared etiologic agents in an equitable manner, thereby avoiding susceptibility bias. In the absence of randomized assignment, this bias can arise if the people who choose (or are chosen) to be exposed to the suspected etiologic agent are much more or less likely to contract the disease in question than are those who remain unexposed.

An excellent example of this problem in occupational research is the bias now known as the "healthy worker effect." Examining the effect of alleged industrial toxins in cohorts of exposed workers, investigators have often been surprised to find that the workers emerged with lower death rates than those found in a group of controls from the general population.[11] The paradox was explained after the discovery that preemployment requirements in the industry would delineate a group of workers who were initially much healthier before occupational exposure than were the unselected members of the general popuation. An example of an analogous problem in studying the hazards of pharmaceutical agents was provided by Stern et al.,[12] who showed that the prevalence of cervical dysplasia was much higher among a group of women who began to use oral contraceptives than it was among users of diaphragms or cervical caps. If this preexposure base-line difference were not known, oral contraceptives might later be erroneously identified as the cause of cervical carcinoma or other lesions preceded by cervical dysplasia.

If the investigators concluded that oral contraceptives caused the dysplasia itself, the false relation could be classified as protopathic bias[13] — a situation in which exposure to the alleged etiologic agent is begun, stopped, or otherwise altered because of early clinical manifestations of the same disease with which the altered exposure to the agent is later associated. For example, a falsely high relation would be created between tea drinking and renal stones if people with early symptoms of renal stones that had not yet been diagnosed began drinking tea to relieve those symptoms.

To reduce some of the problems of susceptibility bias in observational studies, investigators should analyze participants separately according to the presence or absence of clinical features that affect prognosis for the cited disease and that may act as indications or contraindications for exposure to the cited agent.[14,15] For example, if the suspected agent is a drug and the alleged disease is an adverse effect of that drug, the reason that the drug was taken is one of the most cogent clinical features to be considered in the analysis.

Unfortunately, such clinical analyses are rarely performed in observational epidemiologic studies. Instead, the search for susceptibility bias is usually restricted to an analysis of easily obtained demographic data, such as age, race, gender, and socioeconomic status; clinical data on symptoms, severity of illness, or comorbid ailments are seldom included. To avoid the problems of protopathic bias, investigators can examine the people admitted to a cohort study at base line to rule out the initial presence of the disease in question — such as cervical carcinoma or renal stones — before the exposure period begins. Such examinations

are impossible in case–control research, since the diseased cases are first encountered after exposure has already occurred. Compensating adjustments[13,15] can be made, however, if appropriate data are collected and analyzed.

Detection Bias

In an experimental trial investigators can make plans for equal and suitably "blinded" examination of the exposed and nonexposed groups to detect the occurrence of the disease that is the alleged effect of exposure to the suspected agent. The equality of examination must be maintained for both groups in the general-surveillance procedures, in the ordering of diagnostic tests, and in the interpretation of pertinent diagnostic evidence. If this equality is not maintained in an observational study — i.e., if examinations of the people receiving the two compared agents differ according to diagnostic intensity or criteria — detection bias can occur. The problem of detection bias is particularly important in the case of stones, infarctions, sclerotic arteries, cancers, and other chronic diseases that often remain undiagnosed during a patient's lifetime. Any phenomenon that leads to an increased diagnostic search for these diseases will result in the discovery of more "silent" occurrences and may thereby seem to increase the disease's true rate of occurrence.

Thus, if tea drinking provokes an abdominal discomfort that is not caused by renal stones, it may result in a more frequent diagnostic use of pyelography in tea drinkers, as compared with controls, revealing cases of renal stones that have hitherto been symptomless and that would otherwise be undiagnosed. This type of problem has only recently begun to receive serious attention in observational studies,[16] and no agreement has yet been reached about methods for reducing detection bias in case–control research. The problem was openly recognized in a cohort study[17] of the "increased morbidity" that followed use of oral-contraceptive agents, but no solutions were offered.

Migration Bias

This type of bias arises if an unequal migration has occurred in the compared groups between the onset of exposure (or nonexposure) and the subsequent collection of the members whose outcomes are enumerated. Inequalities can result from loss of group members in a cohort study or from decisions made in the arbitrary assembly of cases and controls in a case–control study.

In an experimental trial or observational cohort study the initial status of each member of the groups under comparison is known. Various forms of actuarial analysis can be used thereafter to avoid the bias, described by Neyman,[18] that can arise when members of the two compared groups are lost in unequal numbers or at unequal time intervals. In case–control research the investigator beginning at the end of the causal pathway and having no way of knowing who has been lost from the original exposed and unexposed groups can only hope that migration bias has not occurred.

A more identifiable form of migration bias in case–control research is produced by the investigator's arbitrary selection of cases and controls. In the original case–control investigation[19] of the association between reserpine and breast cancer, this type of "exclusion bias" probably occurred when the investigators decided to eliminate from the control group patients with cardiovascular disease. Since reserpine was particularly likely to be prescribed for hypertensive cardiovascular problems, the exclusion of such patients left the control group with a disproportionately small number of reserpine users, thus fallaciously elevating the odds ratio.[20]

Cogency of Methods

The cogency of scientific methods is maintained by two important standards: vigilance in seeking out and correcting methodologic errors and a firmly established principle that research hypotheses must not be tested with the same data that were used to generate the hypotheses. Neither standard has been well maintained in epidemiologic approaches to observational research.

Methodologic Vigilance

When a major error is suspected in a particular scientific method, the customary scientific response is to instigate research that will allow the suspicion to be either confirmed or refuted. For example, if a knowledgeable biochemist suggests that Benedict's solution may be yielding inaccurate measurements of urinary glucose, most laboratories that measure glucose in this manner will immediately check the accuracy of their mensuration and correct it if necessary.

A challenge that could evoke such methodologic reappraisals for observational studies occurred in 1946, when Joseph Berkson,[21] who was one of America's leading biostatisticians, described a major potential source of error in statistical associations. From theoretical considerations, Berkson developed a mathematical model showing that if patients with different diseases are hospitalized at different rates, substantial distortions may arise when data from a hospital population are used to draw conclusions about the concurrence of diseases and exposures. This same problem may also occur when rates of exposure to an antecedent agent are calculated for groups of hospitalized patients and controls.

Neither Berkson nor several clinical writers[22,23] who reiterated his suspicion had access to the kind of epidemiologic populations that would be needed to determine whether the conjecture was actually true or false. In the populations that might have been checked, however, no investigations were instituted. Without documentary evidence, the type of bias proposed by Berkson was then regarded merely as a theoretical hunch whose scientific reality had not been demonstrated.

No specific research was done to investigate the problem until more than 30 years later, when a group of clinical epidemiologists[24] working in Canada examined direct data for the concurrence of diseases and exposures in a community population and in the corresponding hospital population. The investigators found that the suspected bias occurred with distressing frequency and had important consequences. When noted in case–control arrangements of hospitalized patients, odds ratios for the associations between clinical conditions and possibly etiologic medications ranged from about one sixth of the true risk ratios to values that were 17 times higher than those for the community counterparts.

Retroactive Alteration of Research Hypotheses

After the "active" and control groups have been delineated for the comparison that tests the original hypothesis, scientific tradition calls for both groups to remain intact. They are not altered afterward if the investigator dislikes the original findings. The overt violation of this tradition has often been accepted in epidemiologic research in which "data-dredging" activities have changed the suspected agents or groups under analysis, or in which the fundamental control groups have been directly altered.

Changes in Suspected Agent

While collecting information about antecedent exposure to the agent, A, that is suspected of causing the disease in a case–control study, the investigator can also inquire about antecedent exposure to diverse other agents, B, C, D, E, F, and so forth. If the suspected hypothesis about agent A is not confirmed, the investigator may then calculate odds ratios for all the other agents about which information is available. Although a positive result can often be expected to occur by chance alone if enough agents are tested during these "hypothesis-generating" activities, the positive finding may be presented as though it confirmed a preconceived hypothesis.

Analysis of Multiple Subgroups

Without changing the agent suspected in the initial hypothesis, an investigator who does not find a positive result in the original analysis of data may then explore diverse subsets of the original groups until a positive finding emerges. For example, if an odds ratio of 1 is the unimpressive primary result of a case–control study of the relation of tea drinking to renal stones, the investigator may then check the odds ratios separately for women and men, whites and blacks, old and young, and for a variety of other characteristics examined singly or in combination. Such analyses — like the analogous activities described in the previous section — will often yield a positive result by chance alone if enough subgroups are checked, but the investigator may suggest that the agent has had a causal effect.

This type of subgroup analysis was performed in a large-scale case–control study of the relation between saccharin and bladder cancer.[25] The primary result of the study — an odds ratio with a narrow confidence interval around 1 — seemed to exonerate saccharin as a possible cause of bladder cancer. After multiple subgroup analyses, however, saccharin was left hanging under a cloud because of the elevated odds ratios found in nonsmoking white women who were not exposed to certain chemical substances and in white men who were heavy smokers.

Changes in Control Groups

Without changing the hypothesis of a case–control study and without resorting to analysis of multiple subgroups, the investigators can also attain the desired results by direct, retroactive alteration of the original control group. An example of this procedure occurred in a case–control study[26] that was instituted to confirm a previously noted association[19] between reserpine and breast cancer. The patients in the confirmatory study were drawn from "women with newly diagnosed neoplasms." "Patients with breast cancer" became the cases and "patients with other neoplasms" became the original controls. From the available "adequate drug-therapy data," the investigators calculated an odds ratio of 2.04 as the "estimated relative risk of breast cancer in patients exposed to rauwolfia derivatives," but the ratio was not statistically significant.

The investigators then formed a new control group by excluding patients with cancers of the "pancreas, skin, corpus uteri, kidney, [and] nervous system" from the original control group of "patients with other neoplasms." With this new control group as the focus of comparison, the odds ratio rose to 3.9 and achieved statistical significance. Rearranged in this manner, the data were then regarded as confirming the original hypothesis under investigation. Although the investigators were forthright in providing a detailed description of their changes, such retroactive alteration of the original control group is not an accepted standard in other branches of science.

Suggestions and Comments

Despite the extraordinary scientific difficulties of working with nonexperimental data, observational studies have been both persuasive and preeminent in demonstrating such phenomena as the hazards of thalidomide, the virtues of fluoridation, and the dangers associated with smoking and radiation. Since observational research is often the only feasible method for studying the causes of disease directly in people, our main concern in this essay is to acknowledge the importance of epidemiologic methods and to help improve their scientific quality.

Although the agents under study cannot be assigned by randomization, the other principles of experimental scientific research need not be abandoned in observational studies. With appropriate attention to methods, these principles can often be observed reasonably well.

To reduce the problems of migration and detection bias, the persons admitted to cohort or case–control studies can be selected according to the same eligibility criteria and subsequent surveillance procedures that would have been used in a randomized trial. To reduce the problem of susceptibility bias, efforts can be made to obtain such pertinent clinical information as psychological factors, parental longevity, health status, and reasons for exposure to etiologic or therapeutic agents; and such data can be suitably organized for stratifications or other compensatory adjustments.

Improvements in the collection and confirmation of basic data require only appropriate attention to the challenges and standards. The verification of diagnoses should include attention to false-negative as well as false-positive errors. Although seldom possible for all the participants in a large-scale study, confirmation of other basic data can always be attempted in suitably chosen smaller samples to provide an indication of accuracy. The methods used to acquire data can be described explicitly and reproducibly; and if "double-blind" procedures are unfeasible, information can be collected in a manner that reduces or eliminates bias.

Improvements in the cogency of basic methods also require only suitable attention and concern. The problems described by Berkson can be checked if large collections of suitable data are available; and new methods can be developed for appropriate sampling and analyses that will minimize distortions. Although the original control group should not be retroactively altered, investigation of additional agents and cogent subgroups is often a reasonable activity. When such investigation is undertaken, however, the choices should be made with suitable justification, and the results should be interpreted with suitable caution. Although agreement has not been reached on when and how to adjust the statistical levels of significance in a study in which multiple agents have been investigated without preceding hypotheses about them, the results of such studies should be viewed not as conclusions but as tentative hypotheses to be confirmed by further research. Similar precautions should be taken when multiple subgroups are analyzed without specific clinical justifications; and investigators should be alert to the "protection" that commonly occurs in the complementary subgroup when the original no-risk group is divided into two subgroups, one of which has an elevated risk.

Epidemiologic research has become increasingly important because it offers a substitute for the unattainable scientific gold standard of a randomized experimental trial. Even if the etiologic comparison cannot be conducted with randomized assignment of agents, however, the other scientific principles of an experimental trial can still be used to determine eligibility for admission to a study and to avoid bias in collecting data, forming groups, and making comparisons. An insistence on high scientific quality for etiologic as well as for therapeutic investigations and on adherence to the basic principles of experimental research, even when an experiment cannot be conducted, can help epidemiologic studies achieve the standards and rigor of research in other branches of modern science.

References

1. Cornfield J. A method of estimating comparative rates from clinical data: applications to cancer of the lung, breast, and cervix. J Natl Cancer Inst. 1951; 11:1269-75.
2. Cole P. The evolving case-control study. J Chronic Dis. 1979; 32:15-27.
3. Feinstein AR. Clinical biostatistics. St Louis: CV Mosby, 1977:305-19.
4. Chambers LW, Spitzer WO. A method of estimating risk for occupational factors using multiple data sources: the Newfoundland lip cancer study. Am J Public Health. 1977; 67:176-9.
5. Klemetti A, Saxén L. Prospective versus retrospective approach in the search for environmental causes of malformations. Am J Public Health. 1967; 57:2071-5.
6. Horwitz RI, Feinstein AR, Stremlau JR. Alternative data sources and discrepant results in case-control studies of estrogens and endometrial cancer. Am J Epidemiol. 1980; 111:389-94.
7. Gordis L. Assuring the quality of questionnaire data in epidemiologic research. Am J Epidemiol. 1979; 109:21-4.
8. Feinstein AR. Methodologic problems and standards in case-control research. J Chronic Dis. 1979; 32:35-41.
9. Sackett DL. Bias in analytic research. J Chronic Dis. 1979; 32:51-63.
10. Horwitz RI, Feinstein AR. Methodologic standards and contradictory results in case-control research. Am J Med. 1979; 66:556-64.
11. McMichael AJ, Haynes SG, Tyroler HA. Observations on the evaluation of occupational mortality data. J Occup Med. 1975; 17:128-31.
12. Stern E, Clark VA, Coffelt CF. Contraceptive methods: selective factors in a study of dysplasia of the cervix. Am J Public Health. 1971; 61:553-8.
13. Horwitz RI, Feinstein AR. The problem of "protopathic bias" in case-control studies. Am J Med. 1980; 68:255-8.
14. Feinstein AR. Clinical biostatistics. St Louis: CV Mosby, 1977:385-97.
15. Horwitz RI, Feinstein AR. Analysis of clinical susceptibility bias in case-control studies: analysis as illustrated by the menopausal syndrome and the risk of endometrial cancer. Arch Intern Med. 1979; 139:1111-3.
16. *Idem*. Alternative analytic methods for case-control studies of estrogens and endometrial cancer. N Engl J Med. 1978; 299:1089-94.
17. Royal College of General Practitioners. Oral contraceptives and health: an interim report. New York: Pitman, 1974.
18. Neyman J. Statistics — servant of all sciences. Science. 1955; 122:401-6.
19. Boston Collaborative Drug Surveillance Program, Boston University Medical Center. Reserpine and breast cancer. Lancet. 1974; 2:669-71.
20. Horwitz RI, Feinstein AR, Stewart KR. Exclusion bias and the false relationship of reserpine/breast cancer. Clin Res. 1981; 29:563A. abstract.
21. Berkson J. Limitations of the application of fourfold table analysis to hospital data. Biomet Bull. 1946; 2:47-53.
22. Mainland D. The risk of fallacious conclusions from autopsy data on the incidence of diseases with applications to heart disease. Am Heart J. 1953; 45:644-54.
23. Brown GW. Berkson fallacy revisited: spurious conclusions from patient surveys. Am J Dis Child. 1976; 130:56-60.
24. Roberts RS, Spitzer WO, Delmore T, Sackett DL. An empirical demonstration of Berkson's bias. J Chronic Dis. 1978; 31:119-28.
25. Hoover RN, Strasser PH. Artificial sweeteners and human bladder cancer: preliminary results. Lancet. 1980; 1:837-40.
26. Armstrong B, Stevens N, Doll R. Retrospective study of the association between use of rauwolfia derivatives and breast cancer in English women. Lancet. 1974; 2:672-5.

LECTURE POUR LUNDI LE 13 JUIN 1983

Chapter 8

Health Resource Analysis

Recognizing the importance of an expanded concept of epidemiology to include holism and wellness in determining community health status, we must also realize that, at present and also in the future, health must be lived and can only in small part be delivered. It becomes clear, therefore, that the various aspects of the health care delivery system must be analyzed in accordance with the epidemiological model for health planning and analysis.

Federal legislation has mandated the analysis of the health care delivery system in six areas: accessibility, availability, acceptability, quality assurance, cost containment, and continuity. At present, the federal government has placed the most emphasis on cost containment, that is, reduced or controlled hospital costs and physician fees. Within the framework of the federal mandates, the health care delivery system must be examined, with specific emphasis on patient-origin studies and on accessibility, resource analysis, location, and allocation analyses. These areas impose recurring requirements on health systems agencies (HSAs), state health planning and development agencies (SHPDAs), hospitals, and public health departments. Techniques used in each of the six areas will, when applied, facilitate plan development and lead to efficiency in service delivery.

DEFINING PROBLEM AREAS

Various techniques can be applied to the following specific operational problems involving resources, patient-origin, location, allocation, and accessibility:[1]

1. *Defining planning subregions.* What are the appropriate geographic units for developing a community health profile, for analyzing sub-

regional variations in health status and system performance, and for identifying other important characteristics such as demography and social indicators?

2. *Creating new regional capability.* This issue has gained importance in relation to burn-treatment centers, neonatal intensive care units, CAT (computer axial tomography) scanners, radiation therapy units, renal dialysis units, and other expensive services. The questions of best location (accessibility to population) and desired capacity (utilization levels) and the possibility of coordination between HSA and state are relevant here.
3. *Reorganizing services.* Reorganizing services, one of the goals of the federal planning legislation, poses complex questions regarding location and districting, particularly in the emergency medical service area. There planners have been concerned with determining which hospital should provide what level of services for what geographic areas, with assessing the problem of triage of emergency patients involving assignment to the most appropriate treatment site, and with minimizing distance in order to increase ambulance-response time in emergencies.
4. *Reducing duplication of services.* The problem of redundancy has been experienced most acutely in inpatient hospital services, particularly obstetric care because of the declining birth rate and the emphasis on reducing neonatal mortality. Planners have attempted to incorporate considerations of travel time, utilization levels, and consumer preference into decisions to keep facilities in service.
5. *Meeting increasing demands.* Projecting future demands for a health service is a recurring problem in community health planning. When the service has a seemingly unbounded growth potential, as in renal dialysis, the problem is particularly difficult. Facility location and allocation are dependent on forecasts of demand and severity of condition by subregion, as well as on travel time and network considerations.
6. *Reviewing certificate-of-need requests.* The agency decision in this case is a reaction to a proposal rather than a planning initiative. Often, the planning agency will be faced with provider-generated proof, possibly contradicting agency forecasts, of the "community need" for service initiation or expansion. The agency must be prepared to examine the assumption behind the provider's analysis, especially as it relates to demand considerations, subarea analysis, service area definitions, or market-penetration expectations.
7. *Adjudicating competing proposals for the same service or service area.* This is a special project review problem that occurs when a service need has been previously established and the planner must decide, among other things, which of the two or more proposed locations or services is preferable.

Each of these decision problems requires study, focusing on the definitions of accessibility, facility location and allocation, and delineation of service areas and utilizing patient-origin analysis. Planning for community health care delivery specifically requires the delineation of service areas to reflect either the market penetration of the health service or the consumer's preference for the facility.

DETERMINATION OF HEALTH SERVICE AREAS

The establishment of health service areas is important in determining the market share of patients for hospitals. In many cases, a hospital's market base is necessary to justify its needs in terms of beds, radiation therapy units, CAT scanners, or other special services. The hospital must show from its population base or utilization rate that it has sufficient usage to justify the new service. On the other hand, the health planning agency must determine service areas to meet its planning needs and to justify its decisions to support or reject hospital applications for new services. For these reasons, the health planner must recognize the potentials and limitations of the various methods for determining service areas. Among these methods are the following:

Consumer Surveys

Direct Method

In this method,[2] a consumer is asked, via questionnaire or interview to specify where the health care is obtained. In this type of survey, the health planner should consider the following:

1. Knowledge of survey design, sampling, and sampling errors is required.
2. The method is expensive.
3. Results can indicate potential needs and demands.
4. The method is very similar to the health interview survey.
5. A telephone survey may accomplish similar results.
6. This method's results are similar to those from relevance or commitment indexes.

296

Indirect Method

The following points should be noted with regard to this method.[3]

1. Hospital patient records can be reviewed to obtain patient-origin information.
2. Specific services (that is, discharge diagnoses) may be used with this method.
3. Specific knowledge of record sampling, time frames, and seasonal variations (adjustments), is required.
4. The method is relatively inexpensive, compared to the direct method.
5. Results reflect demand (morbidity and mortality).
6. Results yield fairly accurate locations.
7. Difficulties may arise when relating patient-origin data with social, economic, and demographic data.
8. For urban areas (Standard Metropolitan Statistical Areas), either census tracts, planning districts, or five-digit zip codes can be used.
9. For rural areas (where census data are available), the county or, in some instances, the enumeration district can be the geographic unit.

The use of the indirect method for analyzing patient-origin data, using five-digit zip codes, is illustrated in a study of 23 acute care, short-term hospitals in the Baltimore, Maryland, area.[4] The service areas were established by distributing the admissions across hospitals in direct proportion to the population of the zip code area. The demographic characteristics of each portion of a postal zone were assumed to be identical with those of the entire zone and were weighted according to the proportion of the zone's admissions to a given hospital's total admissions. For example, if 40 percent of the admissions generated in zip code area 31303 were admissions to Johns Hopkins Hospital, 40 percent of the zip code area's population ($50{,}000 \times .40 = 20{,}000$) would be allotted to the Johns Hopkins Hospital market service area. Likewise, if 20 percent of all the admissions to Johns Hopkins Hospital were from zip code area 31303, 20 percent of the characteristics of the hospital's market service area would be derived from the characteristics of zip code area 31303. This would be continued until all zip code areas were allocated. Mutually exclusive market service areas could be determined by assigning each zip code zone to the hospital receiving the highest percentage of its residents as patients. This method can be very useful for determining population-based planning results. It should be noted that the service area is not important in a geographic sense, but only in a population sense. A hospital using this method of apportionment and following stated, related assumptions can determine the total population which it serves and, in accordance with federal guidelines in Section 1501 of P.L. 93-641, can apply population-based standards.

Provider Surveys

In using a provider survey,[5] the following aspects should be considered:

1. The provider is asked to define the service area by listing the communities from which the patients originate.
2. The survey may be administered to the population of providers rather than to a sample of consumers.
3. The provider must estimate the percentage of patient visits coming from the communities listed on a questionnaire.
4. The list of communities should, according to federal guidelines, represent all incorporated places within a specified time distance (30 minutes). Other requirements may also apply.
5. Total patient visits would be recorded.
6. Such a survey involves primarily a rural state approach and would be most difficult to accomplish in an urban area.
7. The following calculations are possible: (1) the number of patient visits each community makes to each provider; (2) the number of patient visits each community makes, with each community having one or more providers; (3) the utilization rates for each community; and (4) overlapping service areas.
8. Complete enumeration is essential. For example, if only one rural provider fails to participate, inaccurate utilization rates will result.
9. Utilization rates can be calculated on the basis of the population of the community and the patient visits from that community.

Determination of Mental Health Catchment Areas

In mental health terminology, "catchment" area rather than "service" area is used to define market penetration. To determine catchment areas, the following assumptions apply:

1. Adoption of a community standard population ranging from 75,000 to 200,000.
2. Easy access to all facilities in the catchment area.
3. Omission of political boundaries.
4. Uniform distribution of mental health problems among all population groups in the catchment area.

5. Application of the catchment area concept can vary. For example, Illinois bases its catchment area on the distance from the mental health hospital, Pennsylvania adheres strictly to county boundaries, while Arizona bases its catchment area on geographic area and distance from the mental health center.

The following considerations relate to the determination of mental health catchment areas:

1. Because census data are based on political boundaries, boundaries of a population-based catchment area are difficult to determine.
2. Population shifts require boundary adjustments.
3. The problem of area definition in sparsely or densely populated places must be solved.

Birth and Death Record Analysis

When using birth and death records for service area determination, the following points should be considered:[6]

1. Information on birth and death certificates indicates place of residence, place of occurrence of the event, and hospital of occurrence, if applicable.
2. Such information identifies movement of patients to hospitals.
3. Service areas may be determined from birth and death records.
4. Birth records provide a most reliable means of justifying obstetrical services.
5. Birth records also provide crude estimates of demographic characteristics of the population (mothers).
6. Although they can provide estimates of patient movement, death certificates are obviously less reliable in providing a complete picture for service area analysis.
7. Information from birth and death records may have to be combined with a survey.
8. Given time constraints, birth and death record analysis may be the most appropriate method for the first year of the program.

Hierarchical Structuring (Population Threshold)

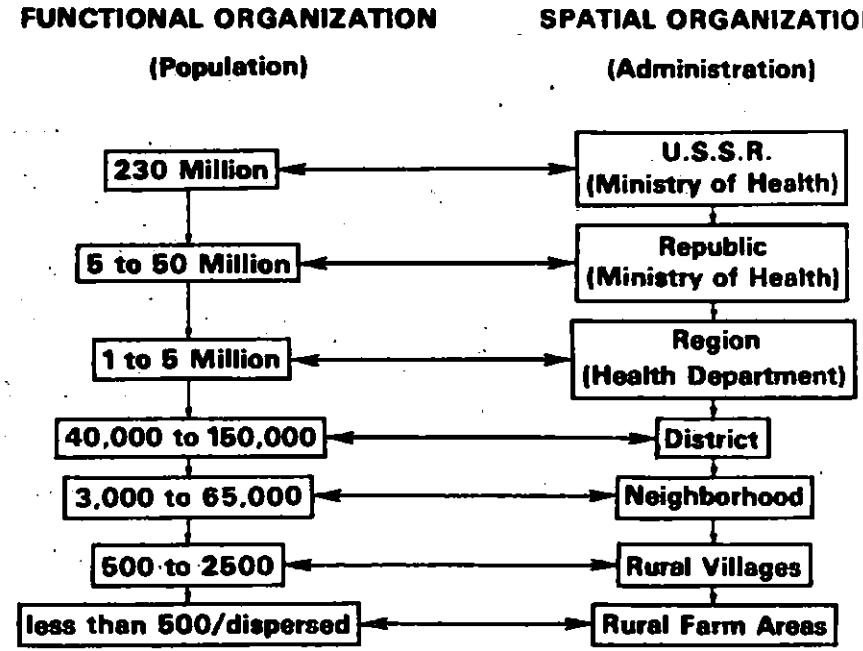
This method is based on population levels. The following aspects should be noted:

1. The minimum population required for providing categorical services to effectively support a specific resource must be determined.

2. The method provides services which are delivered within a hierarchical structure.
3. Primary, secondary, and tertiary services require a minimum population before they can be supplied.
4. An important variable is the minimum or the maximum distance people are willing to travel for the service.

Combining population and distance with administrative characteristics results in a functional/spatial organization of health services.[7] Examples include the functional/spatial organization of health care in the U.S.S.R. (Figure 8-1), in Sweden (Figure 8-2), and in England and Wales (Figure 8-3).[8] A domestic application for Rhode Island is shown in Table 8-1.[9] In addition, Table 8-2 provides a guide demonstrating the population base necessary to provide health services of various types.[10]

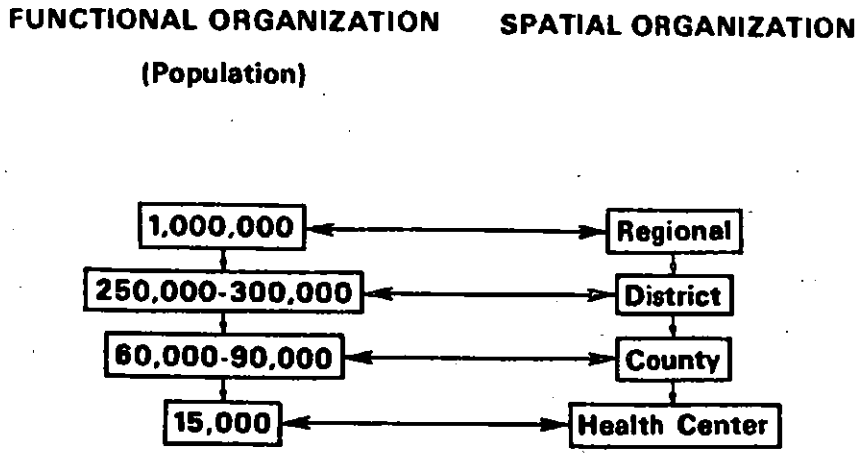
Figure 8-1 Spatial/Functional Organization of Health Care in the U.S.S.R.



Source: G. W. Shannon and G. E. Alan Dever, *Health Care Delivery—Spatial Perspectives* (New York: McGraw-Hill Book Co., 1974), p. 15.

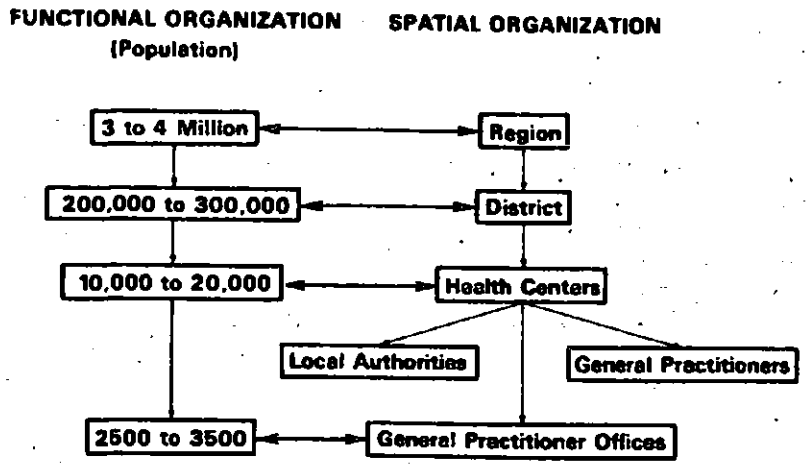
278

Figure 8-2 Spatial/Functional Organization of Health Care in Sweden



Source: G. W. Shannon and G. E. Alan Dever, *Health Care Delivery—Spatial Perspectives* (New York: McGraw-Hill Book Co., 1974), p. 20.

Figure 8-3 Spatial/Functional Organization of Health Care in England and Wales



Source: G. W. Shannon and G. E. Alan Dever, *Health Care Delivery—Spatial Perspectives* (New York: McGraw-Hill Book Co., 1974), p. 20.

Table 8-1 Criteria for Delineating Health Service Planning Subarea Boundaries

1. A set of health service planning subareas (HSPSs) includes primary, secondary, and tertiary areas.
2. Primary HSPSs should have a minimum population of 30,000. Secondary HSPSs should have a minimum population of 100,000. Tertiary HSP areas should have a minimum population of greater than 1,000,000. Exceptions may be made in cases where a geographical (e.g., mountains, poor roads, water) or an economic barrier exists or where extreme isolation pertains.
3. HSPSs should conform to a hierarchical distribution in space, i.e., primary HSPSs should be nested within secondary HSPSs, which in turn should be nested within tertiary HSP areas.
4. At any one level (i.e., primary, secondary, or tertiary) HSPSs should be mutually exclusive. Mutually exclusive HSPSs should cover the whole population of the state.
5. Service area boundaries should be contiguous with (i.e., they should not violate) city or town boundaries and with human service district boundaries. Regional or area planning groups or advisory councils should also be reflected in HSPS designation.
6. HSPSs should reflect to some extent current utilization patterns as well as existing roads, transportation systems, and geographical barriers.

Source: Rhode Island Dept. of Health, "Proposed Health Services Planning Subareas for Rhode Island," Health Planning and Resources Development, Tech. Report No. 7, March 1977, p. 30.

Hierarchical structuring has the following advantages: It requires minimum data (population, sometimes distance); it is easily understood; it involves a minimum of analysis; and it establishes service areas based on what should be rather than what exists. Its disadvantages include the fact that utilization patterns (that is, behavioral considerations) are ignored and that no attempt is made to deal with accessibility (travel time and distance).

Poland-Lembcke method

This method[11] has been widely used. The following issues are pertinent:

1. Patient-origin data are required.

259

Table 8-2 Survey of Minimum Population Bases for Tertiary Health Services

Service(s)	Population Base	Source of Suggested Population Base
Swedish Regional Hospital	1,000,000	Arthur Engel, *Perspectives in Health Planning* (London: The Athlone Press, 1968), pp. 75–76. Services provided include plastic surgery, thoracic surgery, and neurosurgery.
Level III (High Risk) Perinatal Center	670,000–1,000,000	Committee on Perinatal Health, *Toward Improving the Outcome of Pregnancy: Recommendations for the Regional Development of Maternal and Perinatal Health Services* (White Plains, N.Y.: The National Foundation—March of Dimes, 1976) p. 1. The population base is derived from the recommended regional number of births (8,000–12,000) and an assumption of a birth rate of 11.9 births per 1,000 population (the 1975 birth rate in Rhode Island). New Jersey and Wisconsin have adopted the midpoint (10,000 births) of the Committee's range as a standard.
Comprehensive Hemophilia Care Center	6,000,000	World Federation of Hemophilia standard quoted in: Petit and Klein, *Hemophilia, Hemophiliacs, and the Health Care Delivery System* (National Heart and Lung Institute, 1975) p. 51.
	2,000,000	Health Plan Development Services, New Jersey Department of Health, *Proposed Standards and General Criteria for the Planning and Certificate of Need for Regional Hemophilia Care Centers*, 1976, p. 8.
Renal Transplant Center	1,500,000	Division of Health Policy and Planning, State of Wisconsin, *Planning Standards for End Stage Renal Disease Kidney Transplantation and Dialysis* (Madison, Wisconsin, 1976) p. 3.
	750,000	DHEW minimum cited in: Michigan Department of Public Health, *Minimal Criteria and Guidelines for Kidney Transplantation Services* (Lansing, Michigan, March, 1972) p. 17.
	2,000,000	Roger Platt, "Planning for Dialysis and Transplantation Facilities, *Medical Care* (May–June, 1973) p. 209.
	2,000,000	Kidney Advisory Committee, "Optimal Criteria for End-Stage Renal Disease Care," *Journal of the American Medical Association* 226 (October 1, 1973) p. 47.
End-Stage Renal Disease Network	3,500,000	Social Security Administration, Department of Health, Education and Welfare, "End-stage Renal Disease Conditions for Coverage," *Federal Register*, Volume 40 (July 1, 1975), Washington, D.C., p. 27786. The proposed ESRD network would include at least two transplant centers to allow for patient choice and sufficient peer review.
Burn Center	4,500,000	Division of Health Policy and Planning, State of Wisconsin, *Planning Standards for Burn Centers* (Madison, Wisconsin, 1975) p. 3.
CAT Scanning	465,000	Health Planning and Resources Development, Rhode Island Department of Health, *The Acquisition and Use of Computer Assisted Tomography in Rhode Island* (January, 1976) p. 14. Prepared for RIDH by Health Planning Council, Inc., this report proposes there to be an initial need for two CAT scanning units in Rhode Island.

Source: "Proposed Health Planning Subareas for Rhode Island," Rhode Island Department of Health, Health Planning and Resources Development, Technical Report Number 7, March 1977, pp. 45–46.

2. Postal addresses are allocated to geographic units (townships or minor civil divisions) to match rural census data.
3. Determinations are made manually in rural areas.
4. Metropolitan areas may use ADMATCH or UNIMATCH in association with DIME (see the Chapter on Computer Graphics).

From the patient-origin study, data on admissions to each geographic unit are used to calculate the percentage of admissions to each hospital. The following possibilities may result:

1. If 50 percent or more of the admissions from one geographic unit are to one hospital, the geographic unit is in the service area.
2. If less than 50 percent of the admissions are not to one hospital, the geographic unit is not in the service area.
3. If several hospitals draw large proportions from the geographic unit but none gets 50 percent or more, the service area line passes through the geographic area.
4. No admissions are from the geographic unit.

Given these possibilities, the following rules for setting boundaries apply:

1. There must be one continuous service area boundary line which cannot cross another line.
2. The geographic unit in which the hospital is located is, with rare exceptions, assigned to the hospital.
3. Drawing service area lines through the geographic unit is a matter of judgment. If a population center exists in a township, the line should be drawn through the center to indicate it is not clearly in the hospital's service area.

Relevance and Commitment Indexes

By far the most popular, relevance and commitment indexes are probably the easiest to apply because they are based on patient-origin data, they reflect behavioral and distance components, and they are based on the premise of a gravity model (to be illustrated in a subsequent method).

Relevance and commitment indexes have also been labeled resident dependency (RD), provider dependency (PD), or market penetration (MP).

The following steps are required to create relevance and commitment indexes:

1. Formulate a matrix of admissions to each hospital by geographic unit, as in Table 8-3.
2. Compute relevance index (RI):

$$RI = \frac{\text{Tract A}^{\text{Adm}}}{\text{Total Tract Adm.}}$$

$$RI = \frac{\text{Admissions from Tract A to Hospital 3}}{\text{Total Admissions to All Hospitals from Tract A}}$$

$$RI = \frac{TA^{A}}{TA} = \frac{30}{100} = .30 \times 100 = 30\%$$

The relevance Index for Tract A to Hospital 3 is 30 percent, representing the percentage of total admissions from Tract A which go to the study hospital (Hospital 3). The complete set of RIs show the tendency of each tract to use the study hospitals.

3. Compute the commitment index (CI):

$$CI = \frac{\text{Tract A}^{\text{Adm}}}{\text{Total Hospital}^{\text{Adm}}}$$

$$CI = \frac{\text{Admissions from Tract A to Hospital 3}}{\text{Total Hospital Admissions from All Tracts}}$$

$$CI = \frac{TA^{Adm}}{\text{Total Hosp.}^{\text{Adm}}} = \frac{30}{70} = .43 \times 100 = 43\%$$

Commitment Index for Hospital 3 to Tract A is 43 percent, representing the percentage of total hospital admissions (Hospital 3) from Tract A.

4. Identify tracts by various levels of relevance to identify service areas (Figure 8-4).
5. The service area population for the study hospital can be determined by multiplying the population of each tract by the RI for that tract. For example:

Tract population	= 20,000
RI	= 30%
Service area population for Hospital 3	= 20,000 × .30 = 6,000 people

102

Table 8-3 Matrix of Admissions to Hospitals

Tract \ Hospital	1	2	3	4	5	6	Total per Tract
A	10	20	30	10	20	10	100
B			10				
C			10				
D			0				
E			20				
Total per Hospital			70				Total Admissions

Source: Modified from P. Reeves, Applied Statistic Training Institute, Executive Series IV, Myrtle Beach, South Carolina, December 1978, p. 16.

SERVICE AREA DELINEATION

The Three-Tiered System

An approach used in Ontario, Canada, uses patient-origin information and the concept of hierarchical structuring or, as it is called in Ontario, the three-tiered system.[12] Relevance and commitment indexes are calculated, and these in turn are related to the community, district, and regional centers. Service areas for each hospital center are then delineated on the basis of this three-tiered system. Thus, while a community center is surrounded only by a community service, the regional center is surrounded by all three areas. The three tiers are defined as follows:

1. Community centers, usually containing hospitals with less than 100 beds, handle maternity cases, less complicated medical cases, and surgical cases of a relatively minor nature.
2. District centers, generally containing hospitals with 100 to 500 beds can render more complicated care because of their more specialized staff and facilities. These centers supply their own needs on the community level and also accept some referral cases from nearby community hospitals.

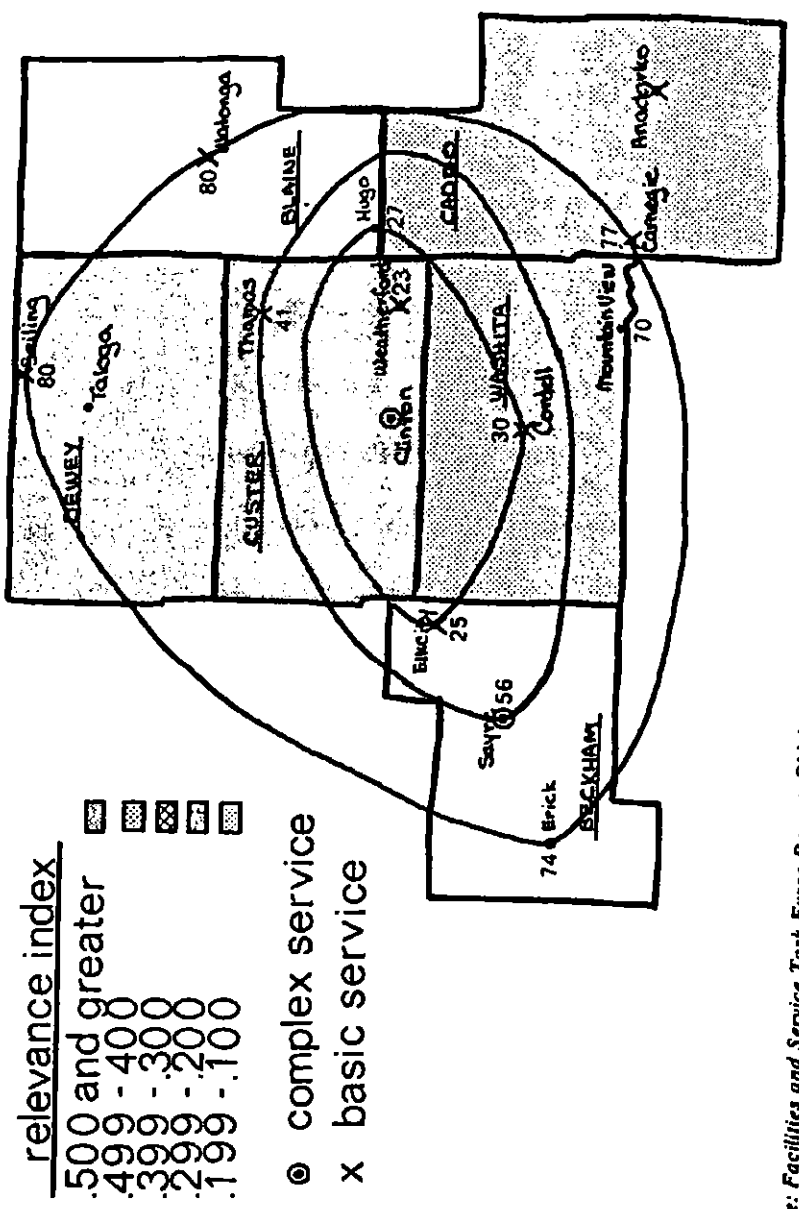


Figure 8-4 Hospital Service Area, Clinton Regional, by Patient Origin and Time/Distance Accessibility

Source: Facilities and Service Task Force Report, Oklahoma Health Systems Agency, April 1979, p. 11.

3. Regional centers are usually university teaching centers. Hospitals in these centers are equipped and staffed to provide very specialized treatment, such as heart surgery, brain surgery, and cancer clinic work and are thereby able to accept cases referred from both community and district hospital centers.

The admission-discharge forms for patients discharged from hospitals are of primary importance in the above delineation. From information on these forms, it is possible to learn the origin and destination of patients being moved, the present referral pattern, and the volume of patients involved. Each township area (that is, the municipal township and the urban communities located within each) is considered individually. Population is allocated on three levels to the hospital centers serving each township area, in proportion to the actual patterns of use; the total township area population is assigned to each of the three levels.

The Ontario experience indicates that, in general, 80 percent of the total active-treatment needs can be provided on the community level, with the remaining 20 percent divided equally between the district and regional levels. The normal planning standard of five beds per 1,000 population designated for active-treatment care (excluding active psychiatric cases) is therefore divided into a community component of four beds per 1,000 and district and regional components of one-half bed per 1,000 for each component. The overall standard of five beds is based on 1,550 days of care per 1,000, and on an assumed occupancy rate of 85 percent. If the population of any area utilizes hospital facilities at a rate higher than this, the demand for beds will be greater than the number indicated as necessary when normal planning standards are used.

Tables 8-4, 8-5, and 8-6 show the delineation of hospital service areas in three sample township areas. In each table, the hospital centers to which township area cases go are listed at left. The number of township residents separated from each center are listed in the first column as a percentage of total township separations. Adjacent columns show, in parentheses, the unadjusted percentages and, without parentheses, the percentages related to the regional, district, and community levels of care. In each instance, the percentage not accounted for on one level is transferred to the next, since regional centers are able to provide care on the district and community levels as well as the regional, and district centers can provide care on both community and district levels. For example, in Township Area 1, Regional Center A provides district as well as regional care, while District Center B provides community as well as district care. Table 8-7 presents the extent of service areas, population service, and beds needed in Regional Center A.

Table 8-4 Delineation of Hospital Service Areas, Township Area 1, Population 15,000

Hospital Centre		% of Total Twp. Separations	Regional Level %		District Level %		Community Level %	
Regional	Centre A	12.0	(10.0)	100	(2.0)	20		
District	Centre B	12.8			(8.0)	80	(4.8)	6
Community	Centre C	24.8					(24.8)	31
Community	Centre D	50.4					(50.4)	63
Township Total		100.0	(10.0)	100	(10.0)	100	(80.0)	100

Source: Modified from a letter from Dr. Frank Houser, District Health Officer, Georgia Department of Human Resources, Dalton, Georgia, 1978.

Table 8-5 Delineation of Hospital Service Areas, Township Area 2, Population 20,000

Hospital Centre		% of Total Twp. Separations	Regional Level %		District Level %		Community Level %	
Regional	Centre A	10.0	(10.0)	100				
District	Centre B	26.0			(10.0)	100	(16.0)	20
Community	Centre C	33.6					(33.6)	42
Community	Centre D	30.4					(30.4)	38
Township Total		100.0	(10.0)	100	(10.0)	100	(80.0)	100

Source: Modified from a letter from Dr. Frank Houser, District Health Officer, Georgia Department of Human Resources, Dalton, Georgia, 1978.

Because it has been found that more care is required by those in the older age categories, modifications are made in the active standards to take into account the age structure of the population of the specific hospital service area under consideration. Thus, the five beds per 1,000 have been subdivided, based on experience in the province as a whole, to show the number of beds per 1,000 population which are required for each age group. Bed standards per 1,000 population increase as the general age level of the population increases (for ages 0–14, 2.6 beds per 1,000 population; for ages 65 and over, 14.7 beds per 1,000 population). These "age refined" standards are then applied to the particular age composition of the service area under study. In an area where the population is relatively

Table 8-6 Delineation of Hospital Service Areas, Township Area 3, Population 25,000

Hospital Centre		% of Total Twp. Separations	Regional Level %		District Level %		Community Level %	
Regional	Centre A	30.5	(5.0)	50	(25.5)	31	(22.4)	28
Regional	Centre E	5.0	(5.0)	50				
District	Centre B	54.9			(54.9)	69	(48.0)	60
Community	Centre C	6.4					(6.4)	8
Community	Centre D	3.2					(3.2)	4
Township Total		100.0	(10.0)	100	(80.4)	100	(80.0)	100

Source: Modified from a letter from Dr. Frank Houser, District Health Officer, Georgia Department of Human Resources, Dalton, Georgia, 1978.

Table 8-7 Extent of Service Areas, Population Served, and Beds Needed in Regional Centre A

Township Area	%	Regional Population	%	District Population	%	Community Population
I	100	15,000	20	3,000	—	—
II	100	20,000	—	—	—	—
III	50	12,500	31	7,750	28	7,000
Total Population		47,500		10,750		7,000
Active Standard		0.5/1000		0.5/1000		4.0/1000
Active Bed Needs		23.75		5.38		28.00

Total active need, excluding psychiatric—57.13 beds.

Source: Modified from a letter from Dr. Frank Houser, District Health Officer, Georgia Department of Human Resources, Dalton, Georgia, 1978.

much older than the provincial average, more beds are required when needs are weighted to age; where the population is younger, fewer beds are needed.

The flexibility of this system must be emphasized. Since service areas are delineated according to actual use, they may vary with changing conditions. The opening of a new hospital, the addition of a new wing, an expansion to provide a greater range of services—all may contribute to produce a shift in service areas. For example, allowance is made for hospital centers situated at provincial boundary points to take into account the needs of out-of-province residents who regularly use hospital facilities in Ontario. If Ontario residents indicate that they prefer to go to hospitals located outside the province, the population of the local hospital service area is reduced.

Flexibility is important in another way as well. If a district center provides more than the usual district level services, allowance is made for this through the allocation of some regional level beds to the district center, in proportion to actual patterns of use.

The Nearest Neighbor (Proximal) System

This is probably the simplest way to define a service area.[13] It is based simply on the nearest place that has the medical service being analyzed. In this system, area boundaries do not reflect patient behavior in seeking health care. The geographic unit utilized will be dependent on population density: A high-density population will use a small unit; a low-density population will use a larger unit, probably a county. With this method, distances to all surrounding places are calculated, and the area is assigned to the nearest place having the service. Other measures which may be computed are the average and the maximum distances between people and the service if data are available for small geographic units (townships, enumeration districts). A computation of this sort would include these steps:

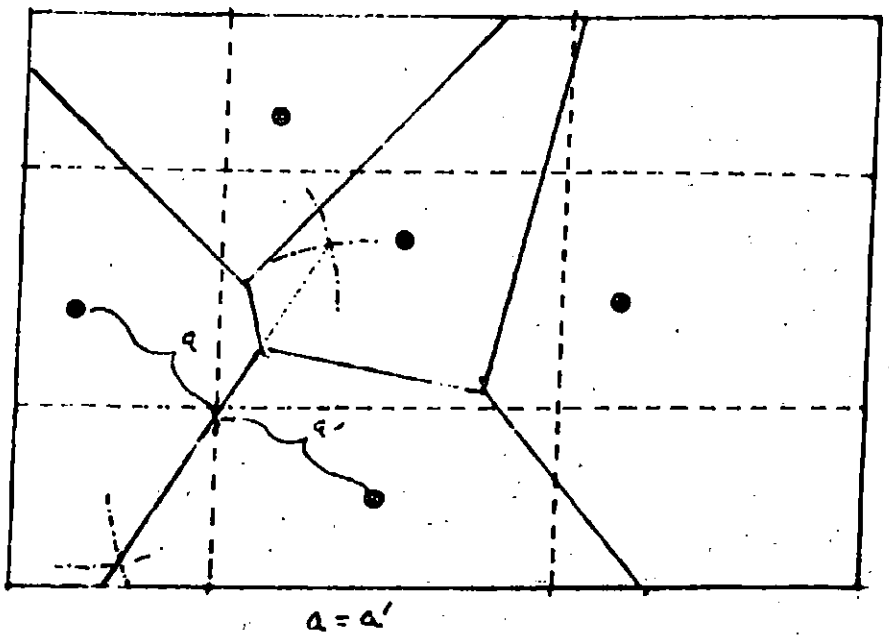
1. Calculate the distance from the center of each area to its nearest service.
2. Multiply by the area's population.
3. Sum the result for each area unit in the service area.
4. Divide the total by the total population of the service area.

If only those communities having a particular health care resource are known, proximal areas may be determined graphically, using either of two procedures:

Procedure 1. Construct perpendicular bisectors between two adjacent points having the resource and assign the region to the closest point. The result is a series of "theisson polygons" (see Figure 8-5).

Procedure 2. Use cartesian (X, Y) coordinates and calculate the distance from consumers to various resource centers. Assign consumers to their closest resource center. The population is assigned in proportion to area served (usually determined by grid squares or planimeters).

Figure 8-5 Graphic Delineation of Service Areas



Source: Meneley et al., *Methodologies for Delimiting Health Service Areas*, Iowa City, Iowa, Health Services Research Center, University of Iowa, January 1977, p. 3.

Gravity Models

The primary use of gravity models[14] to determine service areas has been in rural areas; there has been only limited application to urban areas. The major hypothesis of the gravity model is that the interaction between two places is directly proportional to the mass of the two places and inversely proportional to the distance between the two places. Thus:

$$I_{ij} = P_i P_j / D_{ij}^b$$

where:
- I_{ij} = the interaction between two places (i and j)
- P_i = mass of city i
- P_j = mass of city j
- D_{ij} = distance between two places (i and j)
- b = empirically estimated exponent

With basic modifications, the gravity or interaction model may be changed to:

$$P_{ij} = (A_j^s / D_{ij}^b) / \sum_{j=1}^{m} A_j^s / D_{ij}^b$$

where:
- P_{ij} = the probability that a person in the i^{th} area will use services in the j^{th} place (this is equivalent to the proportion of people in the i^{th} area who will be in the service of the j^{th} place).
- i = 1, 2, ... n (areas for which demographic data exist: enumeration districts, townships, counties).
- j = 1, 2, ... m (area where health services exist, or hospitals, beds, occupancy rates, or some other attraction force).
- A_j = attraction of health resources at the j^{th} place.
- D_{ij} = distance from the i^{th} to the j^{th} place.
- b = function of distance for utility of place.
- s = the rate of increase in attraction of place j with an increase in health resources.

The above equation is used to compute expected probabilities of the use of a given center by each geographic unit in the local area. The size and shape of the service area depends on three factors: (1) the health resources (quantity) at the center and at other competing centers around the tributary area, (2) the attraction influence of additional health resources, and (3) the effect of distance as a deterrent factor.

Patient-Origin Studies

Many of the methods used to delineate hospital service areas are based on patient-origin studies. The following describes some of the advantages and disadvantages of such studies.

1. *Advantages*
   a. The patient-origin study is a reasonable way to gather data about the existing medical market.
   b. Data can be used to make location and allocation decisions (informally using maps and charts and estimating the most logical locations and patterns).
   c. Indexes derived from patient-origin studies can be incorporated into formal decisions (rules).
   d. Data can be used as input to more sophisticated problems of location or allocation.
   e. Areas based on actual flow of medical trade will be identified; that is, the association between location of supply of medical services and the demand for them.

f. Ratios of resource to population will reflect more accurately the existence of shortages or surpluses since they will be based on actual demand for services by the population at risk.
g. Given sufficient data resources, areas can be developed for different types of diagnostic categories (service specific).
h. Planning of facilities can be based on known characteristics of the population being served.
i. A rationale can be developed for determining an adequate mix of services needed in an area, based on patient commuting behavior patterns.

2. *Disadvantages*
   a. Patient-origin studies are not equally useful for all kinds of health services or communities.
   b. In general, tertiary and specialized types of care (not geographically constrained as with primary care) lend themselves to this approach.
   c. Data for dense inner cities are less meaningful than data for communities of "average" population density.
   d. The approach is not suitable for providers who serve populations not geographically contained (HMO members).
   e. Distortion in demand for services may appear because of the absence of resources in a particular area, causing the need to go elsewhere.
   f. Social, psychological, and economic dimensions of seeking health care may divert demand out of local area.
   g. Selected population groups (nonusers of services) may be omitted from data used to determine areas.
   h. The necessary data resources depicting community patterns may be more difficult to obtain on a uniform basis.
   i. Other data resources, for example, sociodemographic characteristics, may not be available in desired geographic units.
   j. Problems may occur because of lack of congruence with other planning areas based solely on geopolitical boundaries.

No single method, of course, can be applied perfectly to all hospital service areas. The planner must make an informed choice and clarify for decisionmakers assumptions about each method. Within this context, the mechanics of a patient-origin study are quite important. A good patient-origin study includes several steps and should involve several members of a team. Also, patient-origin data must be updated as shifts occur in the population and as people change their patterns of seeking care. Physicians' affiliations and locations will also contribute to changing patterns.

Steps to Complete a Patient-Origin Study

Selection of the Data Elements of Interest. The geographic identifiers used to describe patient residence must be delineated at an early stage. Frequently used geocodes include county, zip, municipality, and census tract. It is often desirable to collect multiple geocoded data at different levels of aggregation or, if one wants cross-checks on the geocoding, at the same level of aggregation. The patient-origin study is often used to collect other related information, such as patient's age and sex, length of inpatient stay, and mode of payment. These additional elements should be defined and justified early in the study.

Selection of the Mode of Capturing Data. Preferably, data should be captured in "real time" from all providers in the area. In inpatient studies, data questionnaires should be completed upon discharge for all inpatients discharged during the sample period. Similarly, in the case of ambulatory services, data questionnaires should be completed upon discharge for all outpatients and emergency patients treated during the sample period.

Determination of Sampling Method and Sample Size. The most statistically valid method of patient sampling is with the simple random sample. Other commonly used sampling methods are cluster and systematic sampling. The differing effects of these sampling techniques, that is, how they affect the reliability of the results, should be considered before the survey starts. Equally important is the question of sample size, which determines the probable error range in the resulting data. It may be desirable to consult a statistical expert on these issues so that data gathered during the patient-origin study will be as statistically valid as possible, given cost and other constraints.

Design of Data Questionnaires. Data questionnaires should be designed to make them easy to complete and to manipulate in the data processing task that follows. If the data are to be tabulated manually, keypunched, and then processed through automated means, the requirements for the data questionnaire design will be quite different from those needed if the data are to be read on optical scanning equipment.

Design of Administrative Procedures for Management of the Survey Process. It is important to have well-defined procedures for securing and managing the flow of data from the completion of the survey form through the data processing stage, particularly in a project collecting data on a large scale. Potential problems in data collection and processing should be defined at an early stage, refined through discussion with others who have conducted similar surveys, and further refined in the study pretest. Writ-

ten procedures are necessary, not only to assure the efficient flow of survey forms during the study, but also to assure that the data collection forms are completed in a consistent manner and that questions arising during the data collection process are answered in a consistent manner.

Securing Provider Cooperation and Orienting Participants. Providers must be informed at an early stage about the survey to be conducted. Provider cooperation must be secured to assure complete survey coverage. The active participation of the local hospital council in survey design and study development can also facilitate this task. Once their cooperation has been secured, the providers should be fully oriented as to the nature of the study to be conducted, their role in data collection, and their individual and cooperative benefits from the study. The latter is particularly important in patient-origin studies conducted by planning agencies. The providers must realize the study is being conducted not only to aid the regional planners but also to help the providers reach a better understanding of their relationships to their particular communities and to each other (thus facilitating their individual market analyses).

Pretest Data Collection Method. Even with careful presurvey planning, it is imperative that the data collection scheme be pretested with a limited number of providers. This will confirm the feasibility of the method as designed and act as an aid in identifying weaknesses in the surveying instruments and accompanying administrative procedures. The results of the pretest should be used to redesign survey instruments and alter administrative procedures to increase the probability of a successful survey.

Processing and Storing of Information. The scale of the study and the type of processing system employed will dictate the requirements for processing and storage. If automated data processing is used, computer programs should be written before the survey begins. Arrangements for keypunching or optical scanning must also be made at an early point. Data processing procedures and contracts should be completed at the time data are ready to be keypunched or scanned; the processing should not be delayed by inadequate preparation. The mode of storage—whether on hard copy, punched cards, optical scan forms, or magnetic tape—should also be defined prior to the receipt of completed data forms.

Design and Generation of Output Reports. At the time that the data elements are identified, the basic output reports desired from the study should begin to take shape. If fundamental output needs are described at an early point, the computer programming necessary in an automated data processing system for the production of output reports can be completed before data capture. The computer programs can then be pretested, revised, and readied for report production when the data processing is complete. Information from at least the basic descriptive output reports should be made available for use by providers and planners as soon as possible.[15]

Summary

The patient-origin study is a useful tool to gather data about patient-to-provider flow patterns, thus helping the health planner to conceptualize the health care delivery system for a geographic area. The study can also be used to gather data tangential to the basic purpose of the patient-origin study, thus providing a cost-effective way to generate data valuable for other health planning or research purposes.

ACCESSIBILITY ANALYSIS

Standards and Guidelines

Increased accessibility to health services is a major consideration for most health advocates. Public Law 93-641 suggests guidelines and standards in planning for health services which may be adjusted in terms of accessibility. The standards are both mandatory and discretionary. Discretionary adjustments, in turn, may be generic or standard-specific. The major generic adjustment which can be made in applying national standards (but only if relevant situations cannot be avoided through standard-specific or mandatory adjustments) would result in the denial of access by area residents to necessary health services. Generic adjustments may also significantly increase costs of care for a substantial number of patients in the area and deny care to persons with special needs stemming from moral and ethical values. Such adjustments must be supported by detailed analysis and justification.

Many variables can justify standard-specific adjustments: time and distance to service (accessibility), seasonal changes in use of service, high-risk areas, geographic remoteness (accessibility), and age of the population (see Table 8-8). Standard-specific adjustments must be based on analysis. The level of analysis will depend upon the adjustment proposed. At the very least, however, the agency must: (1) determine the data required for analysis, (2) retrieve such data, (3) analyze and interpret the data, and (4) produce a justifiable base of the results. As indicated by the federal guidelines, it is apparent that, on the one hand, accessibility is a very important factor in the determination of an adequate number of

Table 8-8 Standard-Specific Adjustments

Category	Standard	Basis for Adjustments	Criterion, If Any
General hospital beds	4 beds/1000 population	Age	Persons over 65 constitute more than 12% of population
		Seasonal population fluctuations	"Large" seasonal variations
		Rural areas	"Reasonable" access time
		Urban areas	Variations within an SMSA
		Referral hospitals	Beds utilized by referral patients residing outside area
	80% occupancy	Seasonal population fluctuations	"Large" seasonal variations
		Rural areas	"Normal fluctuations" in admissions
Obstetrical services	Regional planning	None	None
	1500 births per year in level II and level III units	Varying range of service	None
		Moral and ethical preference	None
	75% occupancy in units with greater than 1500 births/ year	None	None
Neonatal units	Regional planning	None	None
	4 beds/1000 live births	High rate of high risk pregnancy	None
	15 beds/level II or III units	Travel time	None
Pediatric services	20 beds/unit in urbanized area	Travel time	More than 30 minutes for 10% or more of the population
	Occupancy rates by unit size	None	None
Open heart surgery	200 procedures per year in adult units	None	None
	100 procedures per year in pediatric units	None	None
	Minimum operating levels prior to opening new units	None	None
Cardiac catheterization	300 procedures per year in adult units	None	None
	150 procedures per year in pediatric units	None	None
	No new units in facilities not performing open heart surgery	None	None
	No new adult units unless existing units have 500 procedures per year	None	None
Radiation therapy	300 cases per year in an area with 150,000 population	Geographic remoteness	None
	No new units unless existing units performing 6000 treatments per year	Geographic remoteness	None
CT Scanners	2500 medically-necessary procedures per year operating goal	Multi-scanner installations	Average of 2500 scans/machine
		Clinical trials	None
		Specialized applications	Seriously sick or pediatric patients
	No new scanners unless existing scanners are operating at 2500 per year levels	Same as above	Same as above

Source: Application of the *National Guidelines for Health Planning* and *Application and Adjustment Procedures Manual*, Midwest Center for Health Planning, September 1978, p. 4.

facilities, but that, on the other hand, attempts must also be made to contain health care costs. The methodology presented here can aid health planners in dealing with the factors of accessibility, availability, and cost of the health care system.

Because accessibility is one of the major factors in health care, it makes sense to analyze the relationship between time and distance to the service, the type of facility and service, and the population served. The primary goals are to minimize the travel time or distance to the health service and to maximize the proportion of the population to be served by the facility. During project review, health planners are faced with the basic question, Is there a need for a new facility or service, such as radiation therapy units, CAT scanners, or obstetrical service?

In a completely interactive system, accessibility analysis is critical to the analysis of bed needs. The analysis outlined below can show that a new service or facility may not be required because travel time and distance are not changed and the population to be served is not increased. The results have cost implications; the advantages of developing a new service or facility do not change either accessibility or availability. In this case, by rejecting the new proposal, costs are kept under control.

Steps in Accessibility Analysis[16]

First, a base map of the area to be analyzed must be digitized. Digitizing is done for every link in the transportation network, that is, distance is determined between each population center and is weighted according to the type of road. For example, a major four-lane highway is assumed to have a value of one, with 55 miles equaling one hour of travel time. Another possible approach is to use the Department of Transportation's time zones which are generated from a small sampling of the distance links traveled to determine the approximate travel time. From the sample, all distances are assigned time values. The correlation between the two methods is so high that either will give quite accurate results. The result is a time/distance matrix which can be used to calculate average accessibility to the service.

Second, population data are required for every town or city that is coded or given a time/distance value. Towns of less than 2,500 population are assigned to a central point in the county.

Third, the facility data must be assigned to one of the many towns in the network. Thus, if 26 CAT scanners are needed, they would be assigned to their respective 26 locations. In this step, urban areas should be analyzed separately from the state analysis.

Fourth, because of the potential for travel by individuals outside the state for service, buffer zones outside or surrounding the state or area being analyzed must be determined, based on the above three steps.

Finally, an algorithm is used to calculate the shortest distance from each town (or node) to the town (or node) that has the service being analyzed. These distance values may be summed for the entire county. The result is a value representing average accessibility in the county to the service being analyzed. The procedure can be done for each census tract or other geographic unit where population data are available.

Graphing Analytical Output

The output from the above analysis (Figures 8-6 and 8-7) shows in graphic and tabular form the following: (1) average accessibility, by county, to each service or trade area defined by accessibility limits (which can be specified by the user); (2) unserved nodes (that is, those towns beyond the specified maximum distance); (3) node assignment and distance (for example, Node 1, population 2,134, is assigned to Node 173, population 37,331, which is the nearest node in terms of time/distance); (4) spatial arrangement of each node to the closest service center (allowing one to define trade areas for the service being analyzed); (5) cumulative proportions reflecting the proportion of population being served with each increment of time/distance to the service for the first, second, third, and fourth nearest locations. (See Figures 8-8 and 8-9.)

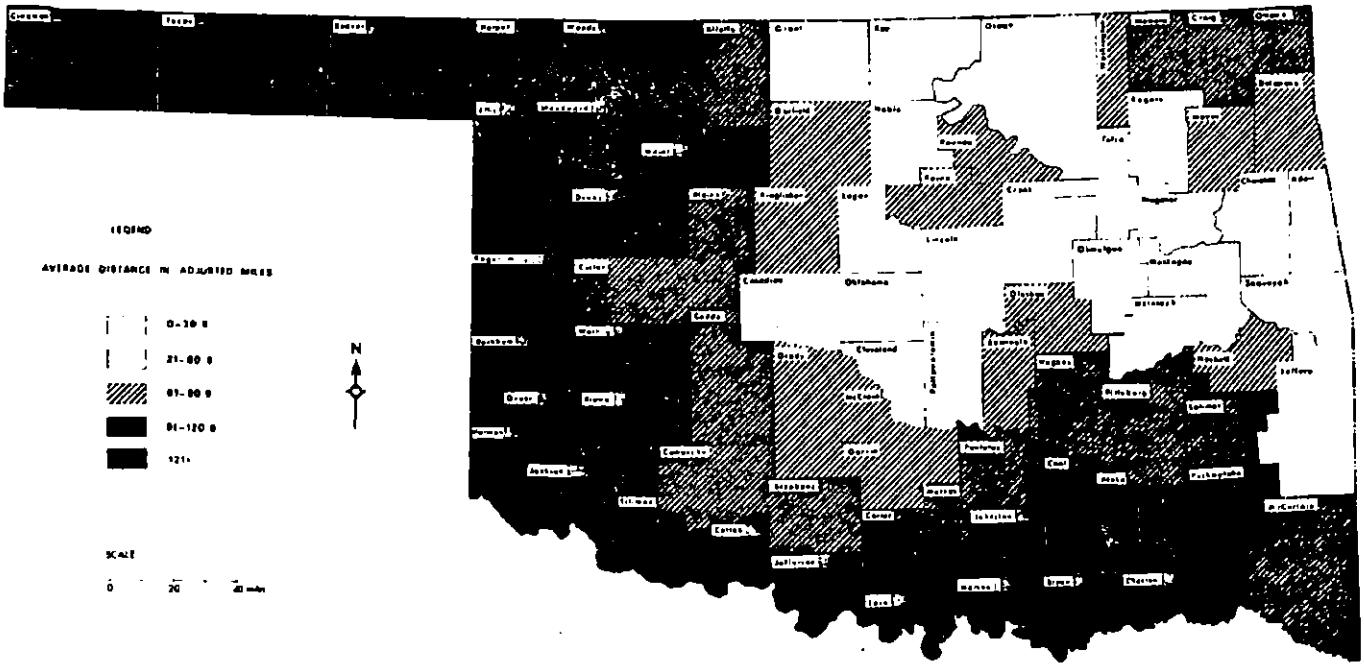
These results may be updated using new population estimates, and the process may be repeated when an agency is asked to review an application for a new service at a new location. For instance, a new map or graph can give the project review officer information showing that the requested service increase does nothing to increase accessibility positively or to serve a greater proportion of the population. If these two objectives have not been attained, the new service location would increase medical care costs without improving accessibility or availability. Thus, the trade-off between a new facility or service on the one hand and accessibility and availability on the other can be made quite clear.

Trade-Off Analysis

Eaton and others suggest that in a trade-off analysis of a set of objectives one can describe how much of one objective must be sacrificed to achieve given values of other objectives.[17] Such an analysis produces trade-off curves that can be very helpful to a project review office. The trade-off curve is a statement of the change in the rate at which one

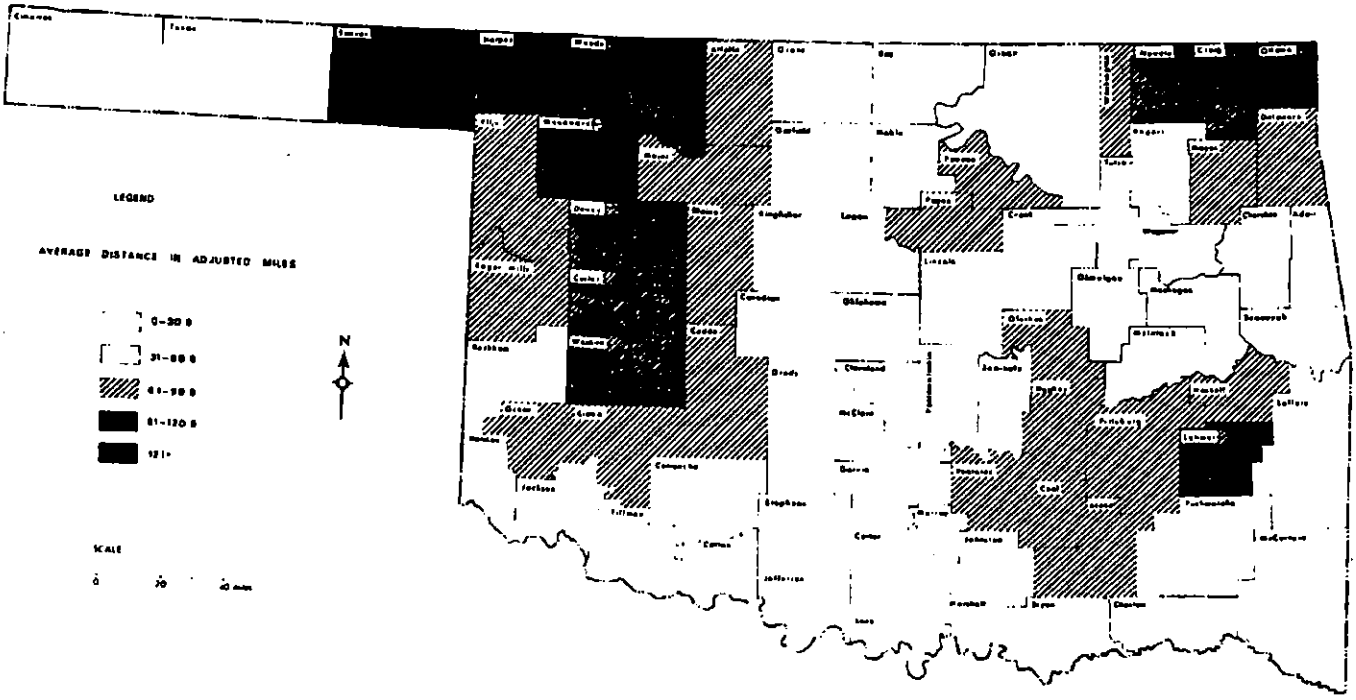
203

Figure 8-6 Accessibility to Oklahoma Hospitals Providing CAT Scanner Services



Source: Oklahoma Health Systems Agency, Health Services Analysis, Inc., Atlanta, Georgia, Health Services Research Center, Iowa, 1978.

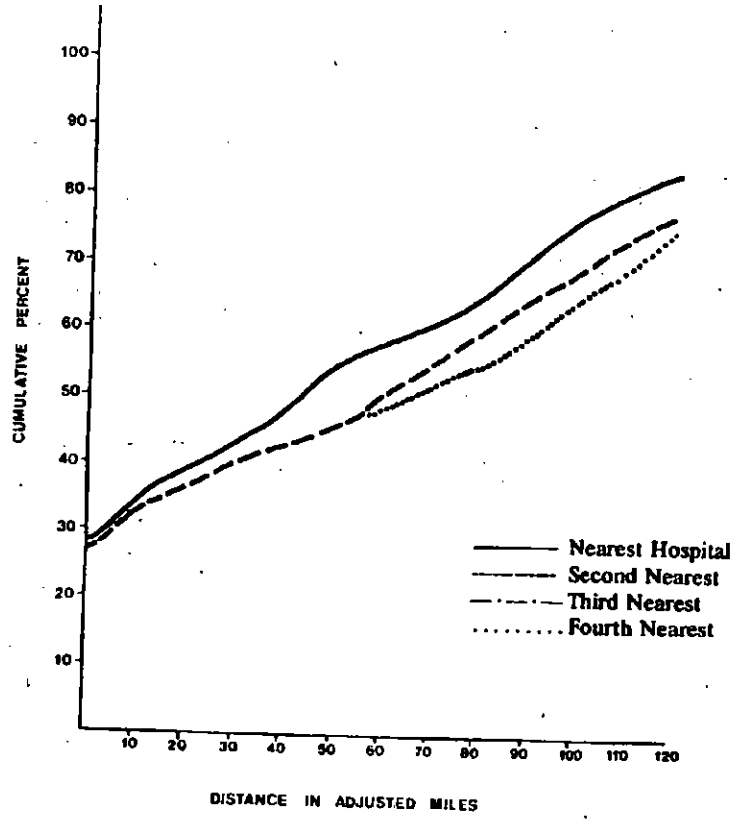
Figure 8-7 Accessibility to Oklahoma Hospitals Providing Radiation Therapy



Source: Oklahoma Health Systems Agency, Health Services Analysis, Inc., Atlanta, Georgia, Health Services Research Center, Iowa, 1978.

OIE

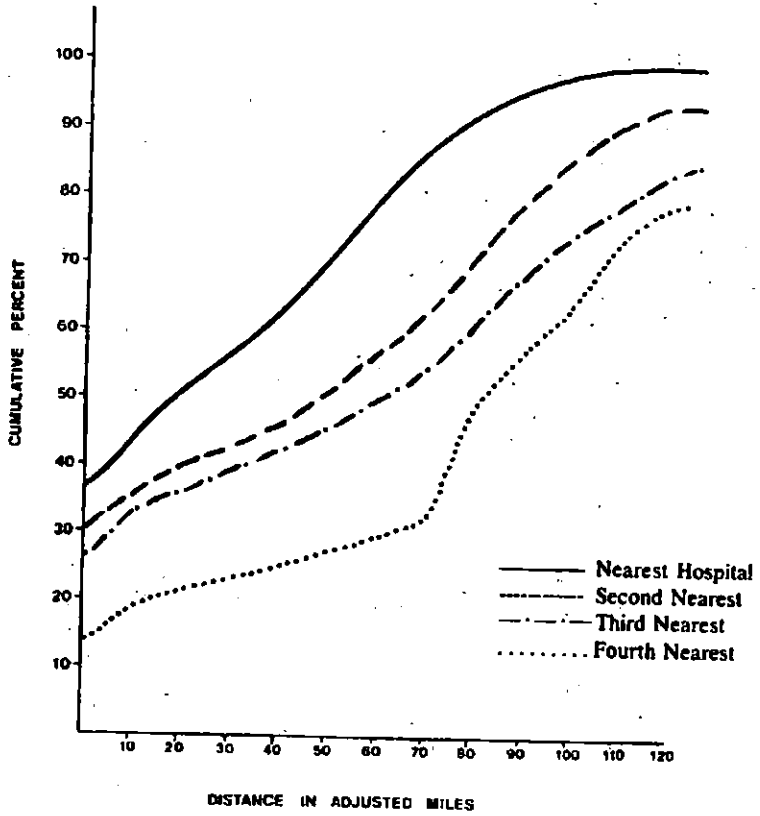
Figure 8-8 Accessibility to Oklahoma Hospitals Providing CT Scanner Services



Source: Oklahoma Health Systems Agency, Health Services Analysis, Inc., Atlanta, Georgia, Health Services Research Center, Iowa, 1978.

objective is improved at the expense of other objectives. If there are two objectives, the trade-off function appears as a curve. In Figure 8-10, the trade-off is between accessibility (time/distance) and availability (number of facilities). Accessibility may be defined as the proportion of the population served within a time/distance constraint (the graph gives a value of two hours, that is, S = 2). Availability may be defined as the maximum number of facilities/services needed to achieve the access. For instance, one objective might be to have enough facilities to provide 90 percent of the population access within 120 minutes. From the curve in Figure 8-10, it can be seen that 12 facilities can cover 100 percent of the population in

Figure 8-9 Accessibility to Oklahoma Hospitals Providing Radiation Therapy

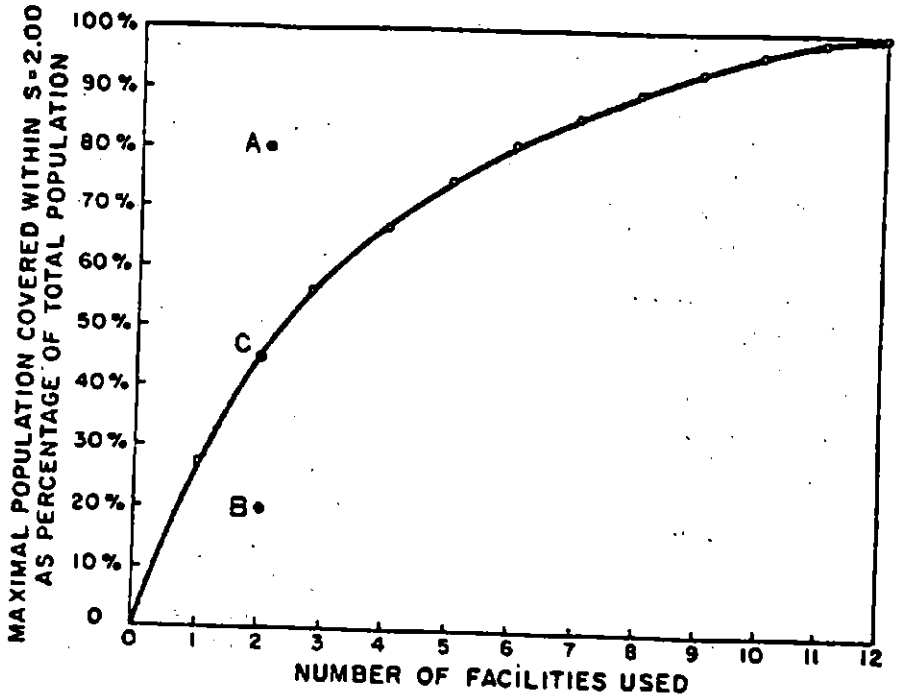


Source: Oklahoma Health Systems Agency, Health Services Analysis, Inc., Atlanta, Georgia, Health Services Research Center, Iowa, 1978.

120 minutes, while one facility can serve only about 30 percent of the population. Thus, a trade-off occurs between accessibility and availability.

The trade-off curve shows that point A is not feasible and that point B is inferior to some other point on the curve, such as C. Consequently, only points on the curve represent plausible policy alternatives. The curve makes explicit the implication of any course and, therefore, can help one to arrive at a decision. For instance, by adding one facility to the existing two facilities, accessibility increases from 46 to 57 percent or 11 percent. The first increases in population covered are less costly than later in-

Figure 8-10 A Trade-Off Curve Between Accessibility and Number of Facilities



Source: D. Eaton, R. Church, and C. Revelle, "Location Analysis: A New Tool for Health Planners," Agency for International Development (A.I.D.), Washington, D.C. (mimeo), No. 53, May 1977, p. 5.

creases, that is, an increase from 11 to 12 facilities increases the gain only 2 percent (98 percent to 100 percent). The planner and policy analyst thus can gain insight about the cost-effectiveness of establishing a new facility or service.

Figure 8-8 (accessibility to CAT scanners) represents 10 facilities with CAT scanners. From the figure we can determine that 10 facilities within 120 minutes travel time serve about 80 percent of the population. A new location for a CAT scanner would increase the number of facilities to 11. If the accessibility were reanalyzed with the additional CAT scanner, we would be able to determine what percentage changes would occur in the population served and the time/distance to the service. Several runs of the analysis at various locations could aid in determining what locations would have the most cost benefits to the population.

Rushton et al. have produced a series of trade-off curves showing the effect of adding new centers for primary medical care (Figure 8-11).[18] The results suggest that after 20 minutes travel time, the options for centers range from 64 to 101 sites. Over 90 percent of the population is served by each service plan; after this point the increase is minimal (the curves flatten out). Clearly, to satisfy the two objectives, it would be best to choose the 64 sites to control the rising health care costs.

Although only the rudiments of trade-off analysis are presented here, the methodology can become quite complex so as to include alternative plans, sites, facilities, and services based on patient behavior, health resource index (required if a new service can only be located where services presently exist), and an accessibility index. The interplay of these factors results in compromise plans by trading off the accessibility index with the resource index. The results of the accessibility analysis and the trade-off function analysis may determine reasonable standard-specific adjustments based on accessibility, efficiency, and level of care.

HOSPITAL PLANNING

Demand-Based Planning

The prediction of future bed needs is based on current levels and trends in service utilization. The disaggregation of data can be very detailed but still not increase the precision of the estimate. Factors such as age, sex, and type of service may have to be separated out for a more specific analysis. An illustration of the use of demand-based methodology is given by McClure:

> As a simple example, suppose a community of 50,000 population currently uses 1400 patient days per 1000 population. Current "demand" is then (50,000 persons × 1400 days per 1000 persons) or 70,000 patient days. Now further suppose that the population is growing 1% per year and the use rate is increasing 0.2% per year (these are approximately national average growth rates). In five years, the population will increase 5% and the utilization rate by 1%. Then future demand in five years can be projected as (50,000 × 105%) × (1400 days per 1000 persons × 101%) or 74,235 patient days. If we apply an average occupancy factor, utilization can be converted to bed needs. Suppose current occupancy is 75% and it is desired to increas-

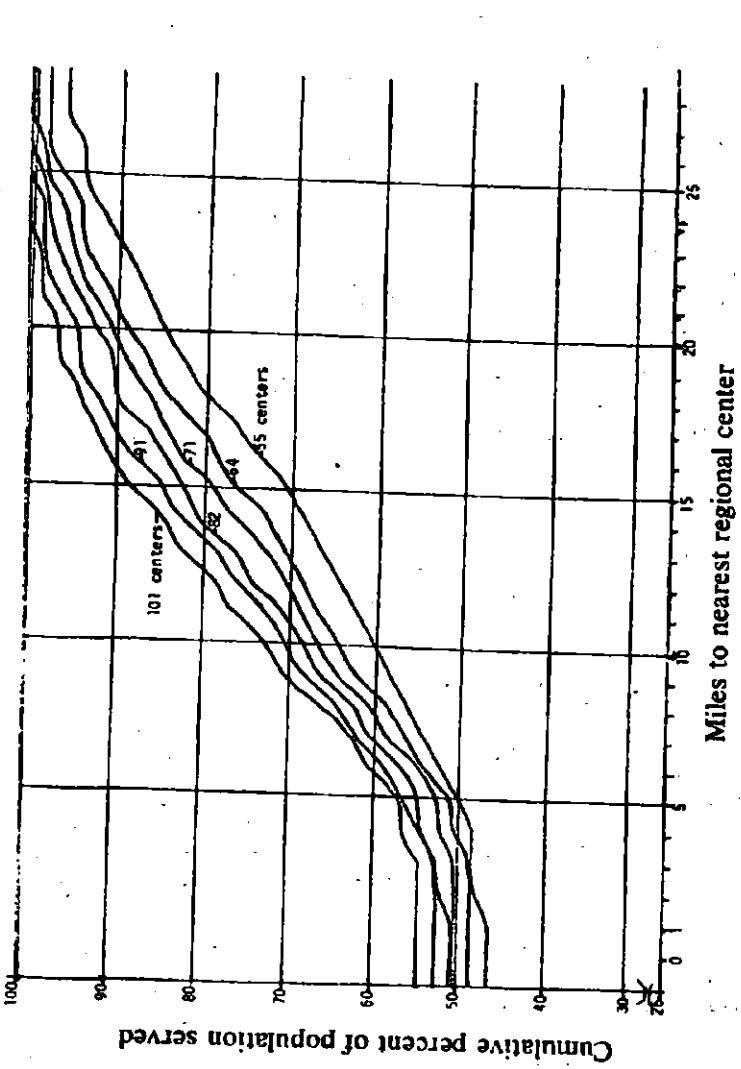


Figure 8-11 Distances to Nearest Regional Centers by Percentage of Population Served

Source: G. Rushton et al., *A Statewide Plan for Regional Primary Medical and Dental Care Centers in Iowa*. Institute for Urban and Regional Research and Health Services Research Center, University of Iowa, Iowa City, Iowa, August 1976, p. 53.

occupancy to 77% in five years. The current bed supply is given by (70,000 patient days ÷ by 365 days per year ÷ 75% occupancy) or 256 beds. This is a bed rate of (256 ÷ 50,000 persons) or 5.1 beds per 1000 persons. Similarly, the future bed need is (74,235 patient days ÷ 365 days per year ÷ 77% occupancy) or 264 beds. This would be a bed rate of (264 beds ÷ 52,500 persons) or 5.03 beds per 1000 persons. According to this simple demand-based method then, this community with a current bed rate of 5.1 beds per 1000 persons would need eight new beds in five years.[19]

McClure also notes that a review of the literature on demand-based planning reflects four approaches or models: formula models, regression models, stochastic approaches, and simulation. His position is that demand-based planning is self-fulfilling and inflationary and that planning should shift toward a population-based approach, requiring population-based data systems as well as training and improved methods for planners.

Population-Based Planning

Population-based planning is a major tenet of PL 93-641. Basically, one sets objectives or goals reflecting a population-based rate, that is, the rate is expressed using a population figure as the denominator. Population-based planning proceeds on the assumption that need is determined by health status and demography and that these population characteristics may be translated into general standards for utilization and capacity rates (that is, four beds per 1,000 population).[20] The purpose is to determine the best allocation of services (utilization and capacity) to population groups so that maximum equity and efficiency in availability, accessibility, and cost can be achieved. McClure illustrates this approach using the same information as in the demand-based model. Assume:

> . . . a typical community of 50,000 persons with a utilization rate of 1400 patient days, and a bed rate of 5.1 beds per 1000 population. (For simplicity, we assume this community has typical age, sex, and other risk characteristics, so that such adjustments can be ignored for the moment.) The research evidence shows that for such a typical population, the range of acceptable medical practice varies from an uncertain minimum of 600–800 patient days to an uncertain maximum 1600–1800 patient days per 1000 population. Since the research evidence shows that comparable

health levels are achieved anywhere within this range, then from a health standpoint the policymaker may arbitrarily choose any utilization rate standard within this range as the planning target for the community. If he is interested only in patient and provider convenience, he might choose a higher rate, say 1500 patient days per 1000 persons. But if he is interested in cost containment, he will choose a lower rate, say 1300 patient days per 1000 persons or less. (He could equally well choose 1200 or 1000 days or less; all are equally safe from a health standpoint. He can choose arbitrarily on the basis of whatever is politically and pragmatically feasible to obtain.) At the specified occupancy of 77%, a goal of 1300 patient days translates to 4.6 beds per 1000 persons. Thus, this population-based assumption would produce a reduction in existing capacity down to 241 beds or less. This is in contrast to the increase to 264 beds projected by the demand-based assumptions. . . .[21]

McClure makes it quite clear that the population-based planning approach should set overall use and capacity rates and then monitor these variables to see that goals are achieved.

We know, for example, that hospital use rates exceeding 600–800 patient days per 1000 population are more than sufficient to produce the optimum health levels achievable by well organized medical care. But as we work down to specific services there are fewer rigorous studies to guide us, and planning judgments become more subjective. For example, we are less certain how many of these hospital days should be intensive care unit days. The important finding that different types of beds are substitutable for each other . . . protects us from reasonable errors in planning judgment and also obviates the need for community health planning at too detailed a level. *Except for extremely high cost services, such detailed planning is better left to individual hospitals*. . . .[22]

This allows the planner to analyze the problem in terms of the total number of beds, not the various types of beds. Many times the health planner stalls the analysis or does an incomplete analysis because data by bed type are not available. Or, worse, hospital administrators and physicians criticize the methodology because bed types were not accounted for. On this point, McClure is correct in stating that the determination of overall bed needs should be based on population-based planning, letting the hospital administration determine bed mix, except of course in the case of very high cost services.

The results of population-based planning by health planners are poor but improving. Data systems at the local level are ineffective and inadequate to provide knowledge about utilization, capacity, and expenditures. With publication of federal guidelines, standards, and reasons for adjustment to the standards, however, the health planner can proceed rapidly with population-based planning.

Demand/Population-Based Planning

Obviously, data constraints preclude the development of a true population-based planning model. For this reason, it is necessary to combine the demand-based model with the population-based approach. For example, occupancy rates (percentages) and average length of stay per 1,000 population may be used in a demand-based model. Other criteria might be patient days and beds per 1,000 population. McClure illustrates this approach with the following example:

. . . . Suppose a state averages 1300 patient days and 4.75 beds per 1000 persons, based on 160 admissions per 1000 persons, 8.1 days length of stay, and 75% occupancy. Suppose the five-year demand-based projection states utilization will rise to 1360 days due to a rise in admissions to 170 offset by a slight fall in length of stay to 8.0 days. At 75% occupancy, this would require an increase in the bed rate to 4.97 beds per 1,000 persons. But suppose the state now makes the easily understood, simple argument that 75% occupancy represents costly, excess underutilized capacity and proposes an occupancy goal of 85%. At the projected use rate of 1360 patient days, only 4.4 beds per 1000 persons will be required. Another simple argument, with lots of research to support it, is that length-of-stay is excessive and could be lowered from 8.0 to 7.8 days (the national average). Then the five-year demand-based projection would be modified to 1325 patient days, which at 85% occupancy would require only 4.25 beds per 1000 persons. Thus, with two simple, defensible arguments—efficient occupancy and length-of-stay standards—the state could justify a reduction in its demand-based projected bed rate from 4.75 to 4.25 beds per 1000 persons, a decline of 10%. The reduction in current bed rate will produce about a 4% decline in use. Then, over five years, this reduction in capacity, if achieved, should in fact produce a fall in patient

314

days from 1300 to 1250 days per 1000 persons—considerably less than original modified demand projection of 1325 days. At this use and bed rate, occupancy would rise to only 80%. Since 85% remains the occupancy target, 1250 patient days will require only 4.0 beds per 1000 population, permitting a further reduction in projected bed needs.[23]

The purpose of this approach is to aid communities in determining where reduction in capacity should be given high priority. In Oklahoma, the reduction of hospital capacity is based on both accessibility analysis and the above type of analysis, particularly, in the latter case, on the demand/population-based planning approach. The Oklahoma program is quite unique, with a tremendous potential for driving down, or at least controlling, the cost of hospital care.

Hospital Planning—The Oklahoma Example[24]

Hospital planning in Oklahoma is based on a four-part analysis. The analysis produces (1) utilization estimates, (2) estimates of hospital performance and bed needs, (3) an identification of hospital location based on accessibility criteria, and (4) estimates of the effects of reallocation of capacity and beds due to possible closure of a hospital. The results are relevance indexes based on demand estimates and population-based estimates, such as beds per 1,000 population and patient-origin data.

Utilization Estimates

In the Oklahoma program, two primary objectives were specified, each focusing on the determination of utilization rates. The first projected the number of discharges by county, service area, and state for 1977 and 1985, utilizing patient-origin data from 1976. The second projected the number of patient days by county, service area, and state for 1977 and 1985, utilizing the above discharge estimates and different estimates of average length of stay (ALOS). The following data were required for this process:

1. Patient-origin data for 1976
2. County populations for 1976, 1977, and 1985
3. Service area boundaries and nodes (towns) within the service areas
4. Current ALOS
5. An estimated ALOS for 1985

Once this part of the model is operational, the health planner can vary the ALOS. That is, if from current literature an appropriate standard can be gleaned, the agency can use that standard as a targeted objective such as ALOS per 1,000 population. Of course, if data are available on different discharge diagnoses, they can also be used.

Utilizing patient-origin data and population estimates, the health planner aggregates total discharges by county service area and state and divides by 1976 population to produce 1976 discharge rates by county, service area, and state. Thus:

1. 1976 County Discharge Rate = Total County Discharges from All Hospitals/1976 County Population × 1,000
2. 1976 Service Area Discharge Rates = Total Service Area Discharges from All Hospitals/1976 Service Area Population × 1,000
3. 1976 State Discharge Rate = All Discharges/1976 State Population × 1,000

The resulting discharge rates (1976) are applied to the 1977 and the 1985 population projections to estimate new discharges. These discharges are multiplied by current ALOS to provide patient days for 1977 by county, service area, and state. In addition, two estimates of patient days for 1985 are calculated, using two estimates of ALOS multiplied by the estimated number of discharges. The products from this step in the analysis, formatted as a table, would be generated three times: once for 1977, using current ALOS, and twice for 1985, using two estimates of ALOS. The results from this step serve as input to the next part of the four-part methodology.

Hospital Performance and Bed Need Estimates

Whereas the focus in the first part of the methodology was on the geographic unit (county, service area, and state), the focus in this part is on the institution (hospital). The primary objective here is to estimate present and future bed needs for the state. There are several subobjectives in this phase of the methodology:

The first subobjective is to estimate discharges by hospital for 1977 and 1985. To accomplish this subobjective, the patient-origin data are used to establish a relevance index matrix which provides relevance indexes for each hospital in 1976. Hospital relevance indexes from 1976 applied to estimated county discharges in 1977 will furnish estimated discharges by hospital for 1977. Similarly, 1985 hospital relevance indexes are applied to northeast service area hospitals and 1976 hospital relevance indexes to the remaining hospitals to estimate hospital discharges in 1985. This step is done three times, in order to account for the impact on the surrounding

hospitals of a new hospital, City of Faith, that Oral Roberts in Tulsa, Oklahoma, will be building. Three cases are identified. In Case A, Oklahoma residents represent 10 percent of City of Faith discharges, in Case B they represent 30 percent, and in Case C they represent 70 percent. The results of the analysis of Cases A, B, and C are shown in four tables developed for 1977 and 1985.

The second subobjective is to estimate hospital performance, that is, expected occupancy rates, given projected utilization and total number of beds for 1977. The utilization estimates from the first part of the methodology by county, service area, and state will now be used. In addition, hospital utilization rates (patient days) will be calculated by multiplying the projected number of hospital discharges (Subobjective 1) by the three estimates of ALOS. Based on this information, expected occupancy rates by hospital can be calculated. The "givens" are the number of beds and projected number of patients for all hospitals, service areas, and the state. Obviously, occupancy rates on a county level are meaningless in this situation, given the methodology for determining county patient days. To obtain the expected occupancy rates, the projected patient days must be divided by 365, resulting in an average daily census (ADC). The ADC is divided by the number of beds, giving the expected occupancy rate by hospital. To get expected occupancy rates by service area and state, the individual hospitals are aggregated. The results will be portrayed seven times: once for 1977 and six times for 1985 (three estimates of City of Faith Hospital times two ALOS).

The third subobjective is to calculate the number of beds needed and the number of beds underutilized, by hospital service area, county, and state. The occupancy rate used will be the standard of 80 percent plus median occupancy rates, by size of hospital and service area. To calculate the number of beds needed, the projected patient days are divided by 365 to give an ADC. Then the ADC is divided by the standard occupancy rate. The result is the number of beds needed. By comparing this value to the number of beds currently available, overutilization or underutilization will be shown. By varying the occupancy rates and the projected patient days (based on ADCs), the methodology becomes interactive. Specifically, if a hospital provides a value reflecting a certain ALOS and occupancy rate, the number of beds to add or delete can be determined.

The fourth subobjective is to estimate bed/population ratios after excess beds are subtracted. Bed/population ratios are calculated by hospital, county, service area, and state. It is possible that the first reduction in beds is too large (that is, more beds have been subtracted than is necessary), and the bed/population ratio will be well below the federal guidelines of 4.0 per 1,000 population. At this point, successive runs can

be made utilizing different bed/population ratios in association with various standard occupancy rates. This provides maximum flexibility for the project review officer to determine new beds or services for the hospital. The results of this analysis are critical to development of the third part of the methodology.

Identification of Hospital Locations Based on Accessibility Criteria

The major objective in this part of the analysis is to determine the best location for cutting the number of underutilized beds. Some constraints, however, must be considered before beds are eliminated. A major constraint is that beds (or hospitals) cannot be cut if accessibility to the hospital meets the criteria specified in the accessibility analysis. Also preference will be given to locations with higher levels of service and performance based on relevance indexes and occupancy rates. Thus, if a hospital or bed were dropped, it would not change the accessibility criteria, and the relevance index and occupancy rates would be low. Here, a decision rule must be established, for example, 30 minutes travel time, a relevance index of .25, and an occupancy rate of .45. Any hospital or hospital beds meeting these criteria could potentially be deleted. The results of this analysis will provide values or locations and potentials to reduce the number of beds.

Effects of Bed Closure and Reallocation

The basic objective in the final part of the analysis is to estimate the effects of bed or hospital closure and of the subsequent reallocation of utilization to their remaining locations. This analysis may be approached in two ways. The first method uses information regarding underutilized beds and the results of an accessibility analysis (see the previous section). The procedure is as follows:

1. Add together all underutilized beds which are not exempted due to accessibility criteria by service area and subtract them from the total number of available beds in that service area.
2. Recalculate a projected service area occupancy rate for the remaining beds, utilizing total patient days for the service area as estimated previously for the years specified.
3. Compare this total occupancy rate with the standard rate applied in part two of the methodology. Then compare the resulting total beds per population ratio with federal guidelines to assess compliance.

The second method is essentially the same as the first except that it is more detailed and institution-specific. The procedure is as follows:

1. Subtract excess beds by node and institution based upon specified criteria.
2. Develop revised relevance matrices, utilizing institutions which remain following the cut in Step One. Reallocate total discharges by county over the remaining hospitals and calculate new relevance indexes.
3. Reestimate the total number of patient days per hospital, utilizing the new estimates of hospital discharges calculated in Step Two and multiplied by the specified ALOS. Next, recalculate projected hospital performance utilizing the new bed complement and new utilization estimates to produce projected occupancy rates by institution. At this point, compare the resulting occupancy rates with the standard imposed in part two to see if the elimination of underutilized beds has brought the state into compliance with the federal guidelines criteria. If expected occupancy rates are not in compliance with the standard imposed in part two, the process can be repeated until occupancy rates and bed/population ratios satisfy these guidelines.

Hospital Planning under Rural Conditions

The national guidelines' standards provide for adjustment of occupancy and bed supply targets under rural conditions.[25] Criteria for adjustment are "rural areas where a majority of the residents would be more than 30 minutes travel time from a hospital" and "rural areas with a number of small (fewer than 4,000 admissions per year) hospitals." Both provisions imply an adjustment of occupancy (and thus the bed supply) targets.

By ensuring "accessibility," an area's utilization is distributed over several small facilities rather than concentrated in one single large facility. If a single large hospital can operate more efficiently than several small ones, the expected occupancy of one large facility will be greater than the aggregated occupancy of several small facilities. For example, if one were to apply the Poisson statistical estimate (at the 99 percent confidence level) to an area with an ADC of 100, the beds needed in one single facility would be $100 + 2.33\sqrt{100} = 124$ beds. If the ADC were equally distributed in four institutions, the beds needed would be $4(25 + 2.33\sqrt{25}) = 4 \times 37 = 148$ beds. The expected occupancies would be 80.0 percent and 67.5 percent, respectively.

Acceptance of occupancy targets due to the size of the facilities is shown in the allowance of a lower occupancy for a large number of hospitals with less than 4,000 admissions per year. The Poisson statistical estimate at the 99 percent confidence level (an assurance that the ADC will not exceed the beds available 99 out of 100 days) does not permit an 80 percent occupancy level until the ADC is 88. An average daily census of less than 88 would mean a lower occupancy target using the Poisson statistical estimate (at the 99 percent confidence level). An ADC of more than 88 would permit an occupancy of more than 80 percent.

Thus, an adjustment for rural conditions is, first, an adjustment to the occupancy target and, second, an adjustment to the bed-supply target. The methodology outlined below uses the Poisson statistical estimate for determining an optimal occupancy in an area.

Step One

Project the ADC for institutions or service areas using whatever methodology is employed by the state C/N standards to forecast bed need.

Admissions/1,000 population × ALOS = patient days/1,000 population

Patient days/1,000 population × population (thousands) = patient days

Patient days ÷ by 365 = ADC

Note: It is assumed that all states use a form of Hill-Burton to project ADC. This form takes current use for a current population and forecasts future use depending upon increases or decreases in the population. Some states also make assumptions about changes in the use rate regardless of population changes, but this does not really affect the basic concept.

Step Two

Apply the Poisson statistical estimate to the ADC to determine the number of beds needed.

$$ADC + Z\sqrt{ADC} = \text{beds needed}$$

Z refers to a confidence level (or interval) that statistically estimates, on the basis of the Poisson distribution, the number of days that the ADC will not exceed the beds needed. Health planning agencies have some flexibility in setting this figure.

Z	Confidence Level	# Times ADC ≤ Beds Needed
1.65	95%	95 out of 100
2.33	99%	99 out of 100

Step Three

Determine individual hospital or service area occupancy targets.

ADC/beds needed = Occupancy

Step Four

Determine HSA occupancy target by aggregating individual hospital or service area ADCs and the number of beds needed.

HSA ADC/HSA beds needed = HSA occupancy

Barring other discretionary adjustments, this becomes the national-guideline, "adjusted" occupancy target.

Step Five

Adjust the bed supply target (4.0 beds/1,000 population) to reflect the adjusted occupancy target.

Bed supply target × Occupancy target/Adjusted occupancy target = Adjusted bed supply target

Barring further discretionary adjustments, the adjusted bed supply target will be the target used to compare with number of beds needed. For example, there are six hospitals in HSA 1, one in each of the six planning subareas. All of them are considered to be needed to ensure reasonable accessibility. The bed need forecasting methodology is as follows:

Current pt. days/Current population × Projected population/365 = Projected ADC

The area population is expected to increase 10 percent, from 75,000 to 82,500 in five years. The analysis and target adjustment would appear as shown on the chart on the next page.

The beds needed per 1,000 population in the following example are not equal to the adjusted bed supply target (4.25) because the anticipated use rate of the population is 1,332 days of care per 1,000 persons. Adjustment to the occupancy and/or bed supply target for other conditions (for example, age, seasonal population fluctuations) might be possible and would bring the bed supply target closer to the beds-needed figure.

	Step 1				Step 2	Step 3
Hospital	1977 ADC	1977 Pop. (000s)	1982 Pop. (000s)	1982 ADC	1982 Beds Needed	1982 Occ.
A	20.7	10	11	23	35	66%
B	42.3	10	11	47	63	75%
C	91.8	20	22	102	126	81%
D	50.4	15	16.5	56	74	76%
E	31.5	10	11	35	49	71%
F	34.2	10	11	38	53	72%
Total	290.9	75	82.5	301	400	Step 4

Step 4. 301/400 = 75.3% = Adjusted Occupancy Target

Step 5. 4.0 beds/1,000 pop. × .80/.753 = 4.25 beds/1,000 pop. = adjusted bed supply target.

The Poisson statistical estimate is an adjustment for size, a condition implied by the "rural," standard-specific conditions. It may only be used where state C/N bed-need methodologies include it as part of the formula for determining bed need or where a bed-need projection formula does not exist. Formulas for projecting bed need which include or state an occupancy target directly would apparently not allow the use of the Poisson statistical estimate or any other adjustment for rural conditions unless specifically provided for in the formula.[26]

METHODS FOR DETERMINING RESOURCE REQUIREMENTS

Below are four methods (provided by the Region IV HPDC) for determining health resource requirements.[27] The four methods, which do not exhaust the list of available approaches, are summarized in Table 8-9 in terms of data requirements and general applicability to the resource standards of the initial national guidelines. Appendix 8A presents additional methods for determining resources for several of the standards noted in the national guidelines.

8/2

Table 8-9 Data Requirements of the Four Basic Methods and Applications

Method	Basic Data Requirements: Current Population	Projected Population	Resource Ratios	Standard of Service Use by Time	Production Capacity	Health Status Indices	Disease Counts	Volume of Treatment	Duration of Aggregate Services	Annual Service Production Capacity	Annual Use Rates	Applications: To Initial National Specific Standard	Applications: Guidelines	Resources Concerns: General
1. Resource/ Population Ratio Method	X	X	X									• Acute care pediatric hospital beds	1. General hospital beds 2. Neonatal intensive/ intermediate care 3. Pediatric beds	
2. Service Target Method	X	(X)		X	X							• CT scanner equipment	1. Radiation therapy units 2. Renal dialysis units 3. CT scanner units 4. Cardiac catheterization/ open heart surgery	
3. Health Needs Approach						X	X	X	X	X		• Radiation therapy equipment	1. Radiation therapy units 2. Renal dialysis units 3. CT scanner units 4. Cardiac catheterization/ open heart surgery	
4. Effective Demand Approach	X	X									X	• General hospital bed capacity and occupancy rates	1. General hospital bed capacity 2. Neonatal intensive/ intermediate care 3. Pediatric beds	

Source: William E. DeBardlaben, "Basic Methods for Determining Community Resource Requirements," *Methodological Approaches to Address the National Guidelines for Health Planning in Local and State Health Plan Development, A Workbook for Planners*, Health Planning/Development Center, Inc., Atlanta, January 1979, p. 62.

51E

Resource/Population Ratio Method

The resource/population ratio method determines resource requirements by multiplying population by a selected ratio:

Resource requirement = Population × Selected ratio

Current requirements are calculated by multiplying present population by the selected ratio; future requirements, by multiplying projected population by the same selected ratio. For this calculation, the planner would need the following:

1. Current population in pediatric age group (<15)	227,063
2. Projected 1983 pediatric age population	233,100
3. Current selected ratio of beds/population	1.3/1,000
4. Desired selected ratio of beds/population	1.0/1,000

With the above data set, the computation is:

Current status = 1.3/1,000 × 227,063 = 295
Present requirements = 1.0/1,000 × 227,063 = 227
Future requirements = 1.0/1,000 × 233,100 = 233

Thus, based on the desired selected ratio of 1.0 beds per 1,000 population, the HSA in 1983 would require 233 acute care, pediatric hospital beds. The following comment on this method is noted:

> The simplicity of the method is its greatest advantage. Relative to other methods, data requirements are minimal, and the statistics are easily obtained; the estimates can be prepared in short order at low cost; the methodology requires only modest staff expertise. It is useful as a descriptive device, as an input to more sophisticated methodologies, as a validation of estimates derived by other means, and as an initial datum to be used in producing more thoughtful judgments within the community.
>
> The weaknesses of the method are serious and may be overriding. Inevitable changes in the future involving socioeconomic conditions, technological and biomedical advances in health care and the configuration of the delivery system are ignored, although they affect the amount of services the population demands and the amount of services health personnel will provide. To ignore these changes and focus on population growth may be feasible in the very short run, but is perilous for long-term projections.[28]

Service-Target Method

The service-target method focuses on the service produced and the volume of the resources as the central determinant of resource requirements. Occasionally referred to as the "normative approach," this method has as key parameters the norms or standards of the services required by the community and of the services produced by health resources. Thus, the methodology quantifies the public's health demands and the health care system's resource outputs:[29]

Resource requirement = Population × Services per person/
Resource productivity (average output per unit of time)

In a mathematical format, the formula is:

$$R_t = \frac{V \times P}{q \times a} \text{ or } \frac{N}{q \times a}$$

where: R_t = resource requirement for year t
V = standard or norm for quantity of service required per year, per person
P = population, current or projected
N = $V \times P$ = total number of services per year for a given population
a = optimal standards of resource performance or production capacity
q = maximum production capacity of a specific resource

A typical problem facing an HSA might be the determination of projected requirements for CAT scanners for a health service area by 1983. By obtaining the following inputs, the planner can apply the model to estimate the number of CAT scanners needed:

1. Projected area population in 1983 (total) = 888,000
2. Service target = .03116 scans per person per year
3. An assumed optimal scanner utilization rate of 85 percent (with allowance for down-time, institutional peculiarities, and so forth)
4. An assumed maximum theoretical capacity per scanner of 4,000 procedures per year, 16 procedures per day, and 250 days per year.

Given the above information, the planner applies the formula as follows:

$$R_t = \frac{V \times P}{q \times a}$$

where: V = .03116 scans per person per year (based on national estimates of 6.7 million scans per year)
P = 888,000 projected population, 1983
a = 85% optimal utilization (task force consensus)
q = 4,000 production capacity (theoretical)

then: R_t = .03116 × 888,000/4,000 × .85 = 8.1 scanners
R = 8 scanners by 1983 (assuming optimal utilization of existing capacity)

The results suggest that eight scanners will be sufficient by 1983. It should be noted that the method has nothing to do with the location of the scanners based on availability standards (see the section on accessibility analysis which demonstrates the process of planning facilities in optimum locations).

The service-target method has been described in the following terms:

> The strength of the service target approach to estimating requirements is its focus on the central issue of providing services and, therefore, on the importance of the efficient and effective functional organization of the delivery of care, not spatial. Attention is directed to utilization and its impact on productivity. The implications of work flow and referral patterns can be studied, allowing the planner to theoretically test the effect of alternative mixes on resource requirements.
>
> The search for understanding of the underlying factors and relationships that are required in the service target approach is one of its greatest strengths. The planner is faced with the Achilles' heel of this method when he attempts to quantify the variables dealing with service targets and productivity. He will almost certainly find important data gaps. Should he decide to collect primary data, he must be prepared for a long-term effort, considerable expense, and the need for expert technical assistance. The planner should be wary of the temptation to elaborate the study to such an extent that the detailed findings are only of academic interest and of no practical value in the light of his community's actual policy options.
>
> By far the greatest danger of this method is the use of improper criteria for setting the service standards, in terms of the demand for services and the productivity of the resource. If the standards used are not valid, the estimates may be grossly unrealistic. (Eight scanners would not reflect the present emphasis on cost containment) . . .[30]

Health-Needs Method

A more sophisticated approach to resource requirements employs the biologic needs of the community as the fundamental determinant of resource requirements. The starting point is to identify and quantify the community's health care needs, using normative judgment of what constitutes good health care. Reflecting its basic orientation, the health-needs approach has been labeled the "biologic care" and "professional standards" method. Thus,

> [w]hen asked, "how much and what types of health resources do you need to provide optimal care for your community," the planner turns to the health needs approach. His concern is with what "ought to be" and what "might be," not with "what is" or "what is likely to be." As a planner, he is thinking in terms of the ultimate goal and highest target of health care planning to assure that level of preventive, diagnostic, or therapeutic care which will obtain the optimal health status for the community. . . .
>
> The health needs approach may be viewed as a particular application of the service target method, in which the standard for consumer demand for services is set by professional judgment of the care needed. These standards can be set at any level: optimal, minimal, or acceptable care somewhere in between. . . .[31]

The health-needs method is based on the following formula:

$$\text{Resource requirements} = R_t = P \times C \times V \times D/W$$

The computations are as follows:

R_t = resource requirements in year t
P = population that needs a given type of care for a specific health problem in year t
C = the average number of specific conditions per year per person
V = the average number of a given kind of service per person/per condition/per year, based on need
D = average time duration required per unit of service
W = average annual operating time of a given resource

A problem the agency may be required to solve is, What are the required resources projected for radiation therapy for a health service area by 1983? Utilizing the above formula,

where: P = 1,332 cancer patients in need of treatment (obtained by assuming a total projected population base is 888,000 for 1983, a crude annual average incidence rate of .003 cases per person, and a 50% rate of new cases needing radiation therapy)
C = one condition
V = 30 treatments for each new case per year
D = 11 minutes per treatment
W = 115,200 minutes per year (based on 8 hours per day, 5 days per week, 48 weeks per year)

then:

$$R_t = P \times C \times V \times D/W$$
$$= 1{,}332 \times 1 \times 30 \times 11/115{,}200$$
$$= 3.8$$

Thus: R = 4 radiation therapy units in 1983.

The following further comment on this method is noted:

> The logical coherences of the health needs approach—that needs are determined by health care required—are very satisfying for health resources planners. One can easily conceptualize the analytical framework: types of health conditions, such as acute, chronic, preventive; required types of care, such as ambulatory, hospital based, long term; and health professionals providing the care, such as physicians, nurses, therapists. When the community's needs are the proper criteria, and data on health status and appropriate treatment are obtainable, this approach is excellent.
>
> The planner should be aware that the technical difficulties in defining and quantifying health needs, "acceptable" modes of care, and resource output are formidable but not insurmountable. Professionals do not agree; the health status of a population changes over time; medical practices advance; and last, but not least, published statistics are out of date, inapplicable, or incomplete.
>
> Most serious of all, criticism is directed at the method's failure to take into account the patients' willingness to seek care in the first place and the community's ability to pay for health services. The assumption is that there are no financial, psychological, or social constraints to seeking care. The concept of "demand as need" has the inherent danger of overestimating requirements.[32]

Effective-Demand Method (Population-Based)

This dynamic approach assumes that the composition as well as the size of the population are major determinants of the types and quantity of services used. In turn, it is the population's demand for particular services that creates resource requirements.[33] The effective demand method is expressed in the following formula:

$$R_t = \sum R_{tp} = \sum V_{tp} \times N_{tp}$$

where: R_t = resource quantity needed in target year t
$\sum_{p=1}^{n}$ = sum over the n population cohorts (≤15, 15–44, 45–64, ≥65)
R_{tp} = resource quantity needed in target year by the p^{th} cohort
V_{tp} = utilization rate of the service in the target year by the p^{th} cohort
N_{tp} = number of individuals (1,000s) in the target year in the p^{th} cohort

Since we assume that utilization rates, by cohort, remain constant over time:

$$V_{tp} = V_{bp} = R_{bp}/N_{bp}$$

where: V_{bp} = utilization rate of the service in the base year by the p^{th} cohort
R_{bp} = resource quantity utilized in base year by the p^{th} cohort
N_{bp} = number of individuals (1,000s) in the base year by the p^{th} cohort

Substituting this in the first formula:

$$R_t = \sum_{j=1}^{n} [(R_{bp}/N_{bp}) \times N_{tp}]$$

The determination of projected requirements for hospital beds by age and sex for the health service area by 1983 is a major problem for a health

agency. The above formula, applied in the following manner, can provide an estimate of bed needs in 1983 (Table 8-10).

$$R_t = R_{bp}/N_{bp} \times N_{tp}$$

$$= 40{,}787 + 105{,}632 + 124{,}971 + \ldots + 193{,}435 = 1{,}013{,}545$$

Hospital beds = 1,013,545 (patient days)/365 (days) × .80 (optimal occupancy) = 3,471 beds

Thus, 3.5 beds per 1,000 population will be required by 1983.

These comments on the effective-demand method are relevant:

> The strong suit of this method is its reliance on measures of effective demand—the utilization rate—and its detailed study of the behavior of groups of people seeking health care. The degree of disaggregation inherent in this method permits the planner to base resource estimates on the health care demands of each segment of the population, reflecting the cultural and physical characteristics of each group of people. This estimate represents the change likely to occur as a result of one important factor, namely, the change in population over time.
>
> However, the assumption that other important economic factors, such as income, price, and third-party financing, remain unchanged is very limiting and unlikely to represent reality. The assumption that population and the demand for services are so related that a change in population brings a proportionate response in the demand for services is very restrictive and needs to be validated. Furthermore, the assumption that present utilization patterns are the proper standard for the future is questionable, since we realize that there are many unmet health needs in our country as well as inappropriate uses of limited health resources.[34]

Table 8-10 Population and Patient Days for the Health Service Area, 1977 and 1983

Age-Sex Cohorts	Base-Year Population (1977) (N_{bp})	Base-Year (1977) Demand (R_{bp})	Average Annual Utilization Rate (V_{bp}) ($V_{bp}=R_{bp}-N_{bp}$)	Projected Year (1983) Population (N_{tp})	Project Year (1983) Demand ($V_{bp}XN_{tp}$) (Total Days)
MALES					
<15	115,564	39,731	343.8	118,637	40,787
15–44	180,266	102,896	570.8	185,059	105,632
45–64	81,570	121,735	1,492.4	83,738	124,971
65>	36,676	149,455	4,075.0	37,651	153,428
FEMALES					
<15	111,499	30,439	273.0	114,463	31,248
15–44	195,836	213,109	1,088.2	201,043	218,775
45–64	91,430	141,506	1,547.7	93,861	145,269
65>	52,159	188,424	3,612.5	53,546	193,435
TOTAL	865,000	987,295	1,141.38	888,000	1,013,545

Source: William E. DeBardlaben, "Basic Methods for Determining Community Resource Requirements," *Methodological Approaches To Address The National Guidelines for Health Planning in Local and State Health Plan Development, A Workbook for Planners*, Health Planning/Development Center, Inc., Atlanta, January 1979, p. 58.

Summary

Clearly, the selection of a method to determine resource needs should be approached cautiously. There is no standard method, but there are standard guidelines. Thus, the assumptions underlying the methods should be critically examined and thoroughly explained to the individuals who make the decisions. None of the four methods approaches the problem of optimum utilization through optimum geographical distribution of services. For this reason, the methods require one further step, illustrated in the question, If the need is for eight CAT scanners, where should they be located? The accessibility analysis and the hospital performance analysis reviewed earlier are needed to complete this final step.

LOCATION AND ALLOCATION ISSUES IN HEALTH PLANNING

Location and allocation issues may be subdivided into four problem categories: (1) locational problems, (2) allocation problems, (3) location-then-allocation problems, and (4) allocation-then-location problems.[35] The main concern in locational problems is to locate one or more health services or facilities in the most desirable place. This procedure can involve running or relocating a facility in the existing system. The major concern in allocational problems (relatively rare in health planning) is to have boundary lines which tie consumers to specific service sites or providers. In this case, the federal guidelines concerning obstetrics and neonatal services suggest that regionalization of the services must be considered. Location-then-allocation problems are concerned with which

sites are selected and with the way consumers or the population are "districted" for the new sites. This type of planning has been prevalent in school districting. The concern in allocation-then-location problems, the reverse of the location-then-allocation problem, is to first draw the districts and then to define the best location for the services or providers within the district.

The first requirement is a set of points, consumers or population, distributed in a geographic area. A weighted value may be attached to each point that would make the location more or less desirable. A set of facilities without predetermined locations must then be located within the geographic area. The major objective in this process is to find locations for the central facilities and to allocate each of the points (consumers or population) to one of the facilities. The destination points may represent counties, communities, census tracts, or neighborhoods. The weights used may be estimates of demand, for example, the population may be used as the demand value. The central facilities would be either hospitals, primary care centers, emergency facilities, obstetrical services, or neonatal services.

Location/allocation problems are solved by optimization or heuristic techniques. If the issues represent an optimization problem, optimization techniques can solve the problem. The major purpose of this approach is to identify an objective function which will be maximized or minimized. If heuristic techniques are applied to an optimization problem, the results will be suboptimal. Optimization techniques include linear programming, integer programming, and the tree-searching method (branch and bound). Each method will produce the best solution based on the stated objective function and constraints. Solving by optimization techniques, however, is not always feasible. There are three reasons for this: First, many problems cannot be expressed as a series of linear equations because of too many complex relationships among the variables. Second, the techniques may require too much data or inaccessible data. Third, some optimization problems are so complex that excessive and expensive computer time is needed to determine an optimal solution.

Heuristic methods provide potential alternatives to optimization methods. The solution of optimization problems heuristically requires a set of decision rules (algorithms) to find good solutions; optimality is not guaranteed. The practical side of this approach is that it reflects more of "real world" constraints. The results from heuristic methods are quite satisfactory and hold the greatest promise for health planners.

Many issues complicate location, allocation, and accessibility problems. First, primary, secondary, and tertiary service areas must be determined for hierarchical services. Second, multiple services with different travel distance requirements can be obtained from the same facility. Third, different distance and travel measures may be used, for example, straight-line, actual, time, cost, and subjective. Fourth, location and allocation problems may be formulated as static problems (these assume that location/allocation assignments can be implemented immediately) or as dynamic problems (these provide for phasing in the implementation). Finally, in both static and dynamic models, there is a conflict between short-run and long-run solutions. A brief listing of significant contributions to the solution of location and allocation problems is provided in Appendix 8B.

SUMMARY

Health care resources analysis involves methods for determining appropriate service areas and trade-off functions, for doing accessibility and location/allocation analysis, and for hospital planning (demand-based, population-based, and a combination of the two). By applying these resource analytical methods to problems facing our country, the health planner can be greatly rewarded. The need to reduce health care costs and to find more equitable and efficient solutions poses a continuous challenge to health care planners.

324

APPENDIX 8A*

DETERMINATION OF RESOURCE NEEDS BY VARIOUS METHODS

Details of eight methods to determine health resource needs are presented below.[36]

Table 8A-1 Calculations for Short-Stay Hospital Bed Requirements for 1983, Resident Population

A) *Indigenous Urban Population*

P = 829,056
PD = 829,056 × 1,141.38/1,000 = 946,268
W = 1 percent of patient days
Q = .80

$$N = \frac{D - W}{365 \times Q} = \frac{946{,}268 - 9{,}463}{365 \times .80} = 3208$$

B) *Indigenous Rural Population*

P = 55,944
PD = 55,944 × 1,141.38/1,000 = 63,853 patient days
W = 1 percent of patient days
Q = .70

$$N = \frac{D - W}{365 \times Q} = \frac{63{,}853 - 639}{365 \times .70} = 247 \text{ beds}$$

*Source: This entire appendix is from William E. DeBardlaben, *Methodological Approaches to Address the National Guidelines for Health Planning in Local and State Health Plan Development. A Workbook for Planners*, Health Planning/Development Center, Inc., Atlanta, January 1979.

Table 8A-1 continued

C) *HMO Population*

P = 3000
PD = 3000 × 856.0
W = 0
Q = .80

$$N = \frac{D - W}{365 \times Q} = \frac{2{,}568 - 0}{365 \times .80} = 9 \text{ beds}$$

D) *Referral Population*

D = 30,000 patient days
W = 0%
Q = 80%

$$N = \frac{D - W}{365 \times Q} = \frac{30{,}000 - 0}{365 \times .80} = 103 \text{ beds}$$

E) *Transient Population*

P = 25,000 annual average visitors
D = 25,000 × 570.69/1,000 = 14,267
W = 0
Q = .80

$$N = \frac{D - W}{365 \times Q} = \frac{14{,}267 - 0}{365 \times .80} = 49 \text{ beds}$$

F) *Total Beds Needed* = 3,208 + 247 + 9 + 103 + 49 = 3,616

P = Population
PD = Patient Days
W = Wellness Factor
Q = Occupancy Standard
N = Number of Beds

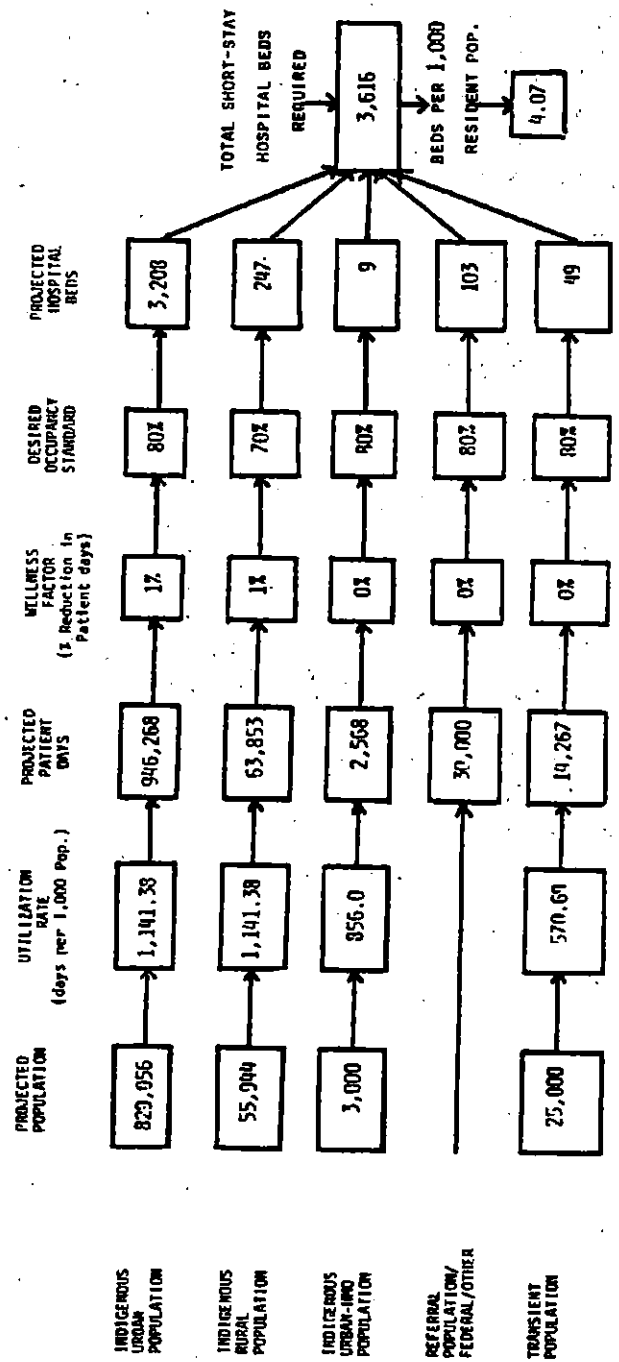


Figure 8A-1 Short-Stay Hospital Bed Requirements for 1983

Table 8A-2 Calculations for Obstetrical Resource Requirements, 1983

RESIDENT POPULATION

A) *Indigenous Urban Population*

Projected Child-bearing Population = 186,970

Projected Birth Rate = 66.0 per 1,000 pop.

$$\text{Projected Births} = \frac{186{,}970 \times 66.0}{1{,}000} = 12{,}340$$

$$\text{Projected Admissions} = 12{,}340 + (.02 \times 12{,}340) = 12{,}587$$

$$\text{Projected Beds} = \frac{4 \times 12{,}587}{365 \times .75} = 184$$

B) *Indigenous Rural Population*

Projected Child-bearing Population = 14,073

Projected Birth Rate = 60.0 per 1,000 pop.

$$\text{Projected Births} = \frac{14{,}073 \times 60.0}{1{,}000} = 844$$

Projected Admissions = 844

$$\text{Projected Beds} = \frac{4 \times 844}{365 \times .60} = 15$$

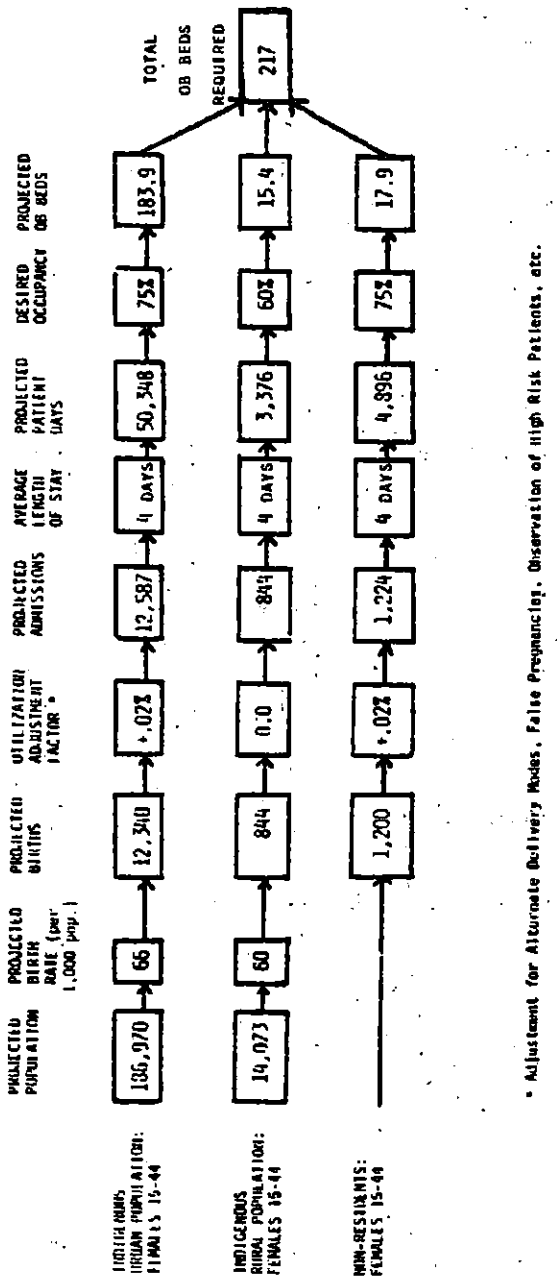
NON-RESIDENT POPULATION

Projected Births = 1200

Projected Admission = 1200 + (.02 × 1200) = 1224

$$\text{Projected Beds} = \frac{4 \times 1224}{365 \times .75} = 18$$

Total Projected Beds = 217



* Adjustment for Alternate Delivery Modes, False Pregnancies, Observation of High Risk Patients, etc.

Figure 8A-2 Obstetrical Resource Requirements for 1983, Health Service Area 1

Table 8A-3 Calculations for Neonatal Special Care Resource Requirements, 1983

RESIDENT POPULATION

Projected Births	=	12,340
Low Weight Birth Rate (per 100 live births)	=	7.5
Projected Admissions $= \frac{7.5 \times 12{,}340}{100}$	=	926
Projected Neonatal Beds $= \frac{12 \times 926}{365 \times .75}$	=	41

NON-RESIDENT POPULATION

Projected Births	=	844
Low Weight Birth Rate (per 100 live births)	=	7.5
Projected Admissions $= \frac{7.5 \times 844}{100}$	=	63
Projected Neonatal Beds $= \frac{12 \times 63}{365 \times .75}$	=	3

HIGH RISK NEONATES

Projected Admissions	=	85
Projected Neonatal Beds $= \frac{30 \times 85}{365 \times .90}$	=	8
Total Neonatal Beds	=	51

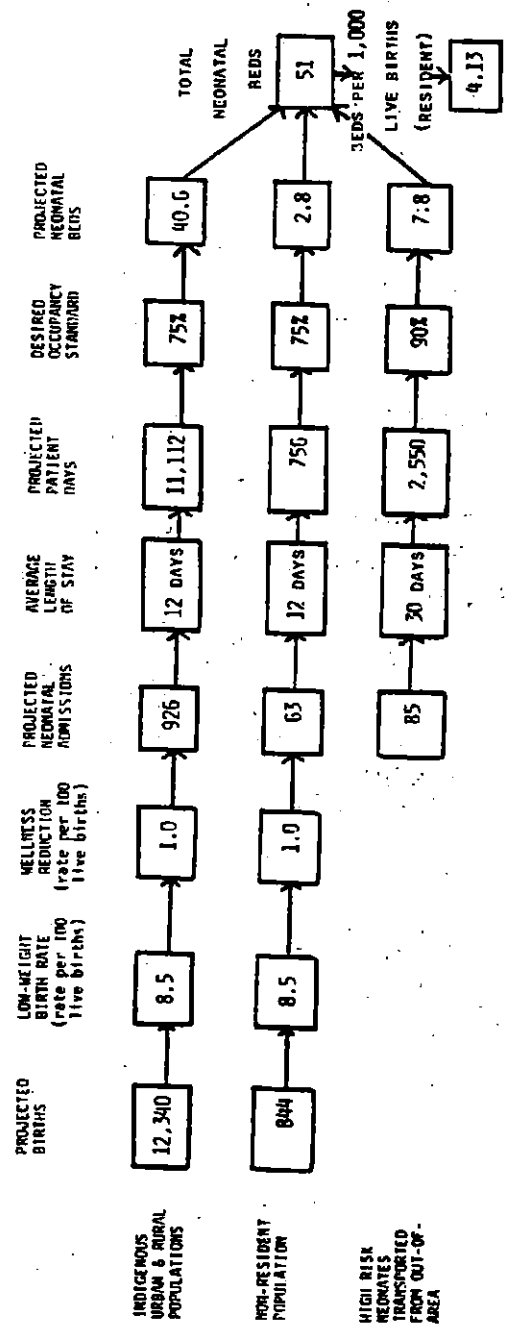


Figure 8A-3 Neonatal Special Care Resource Requirements for 1983

Table 8A-4 Calculations for Pediatric Bed Requirements, 1983

A. INDIGENOUS URBAN AND RURAL (EXCLUSIVE OF HMO ENROLLEES)

Population		= 232,311
Patient Days =	$\frac{309.03 \times 232,311}{1,000}$	= 71,791
Wellness Factor =	$.01 \times 71,791$	= 718
Occupancy Factor		= 70%
Beds =	$\frac{71,791 - 718}{365 \times .70}$	= 278.17

B. HMO POPULATION

Population =	$(26.3\%) \times 3,000$	= 789
Patient Days =	$\frac{789 \times 309.03}{1,000}$	= 244
Wellness Factor =	$.25 \times 244$	= 183
Occupancy Factor		= 70%
Beds =	$\frac{244 - 61}{365 \times .70}$	= .71

C. REFERRAL POPULATION

Patient Days		= 10,450
Occupancy Factor		= 100%
Beds =	$\frac{10,450}{365 \times 1}$	= 28.63

D. TOTAL BEDS

$278.17 + .71 + 28.63$ = 307

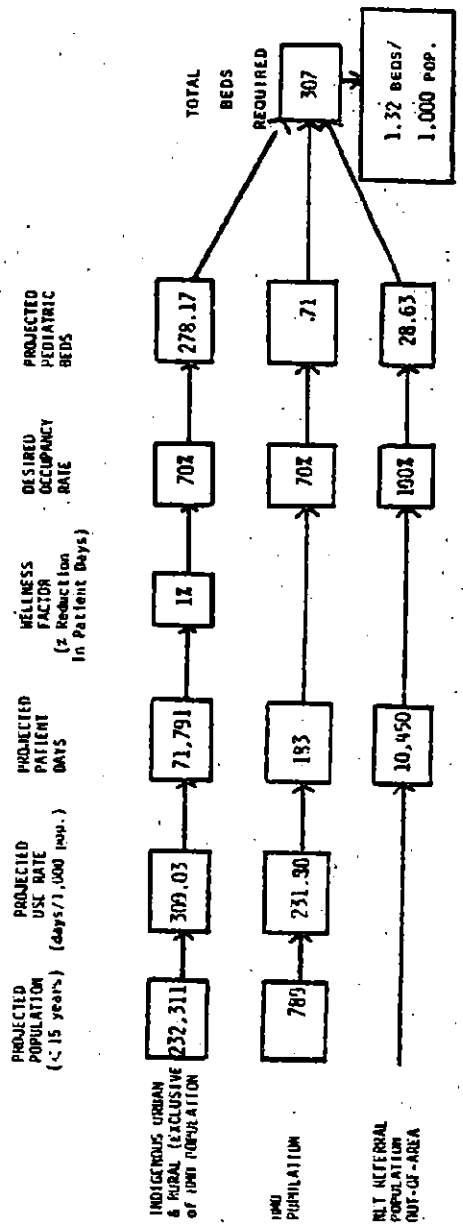


Figure 8A-4 Pediatric Bed Requirements for 1983, Health Service Area

Table 8A-5 Calculations for Cardiac Catheterization Requirements, 1983

A. RESIDENT POPULATION (≥15 years)

Population 15 years and older = 654,900
Resource Level = 1.25 Labs/300,000 pop.

$$\text{Resource Requirements} = \frac{1.25 \times 654{,}900}{300{,}000} = 2.73 \text{ units}$$

B. RESIDENT POPULATION (≤15 years)

Population under 15 years = 233,100
Resource Level = 1 Lab/300,000 pop.

$$\text{Resource Requirements}_{\text{s}} = \frac{1 \times 233{,}100}{300{,}000} = .78 \text{ units}$$

C. REFERRAL POPULATION (out of area, all ages)

Population (from studies) = 800
Resource Level = 1 Lab/600 patient studies

$$\text{Resource Requirements} = \frac{1 \times 800}{600} = 1.33$$

D. TOTAL RESOURCE REQUIREMENTS

$2.73 \times .78 + 1.33 = 4.84$ units

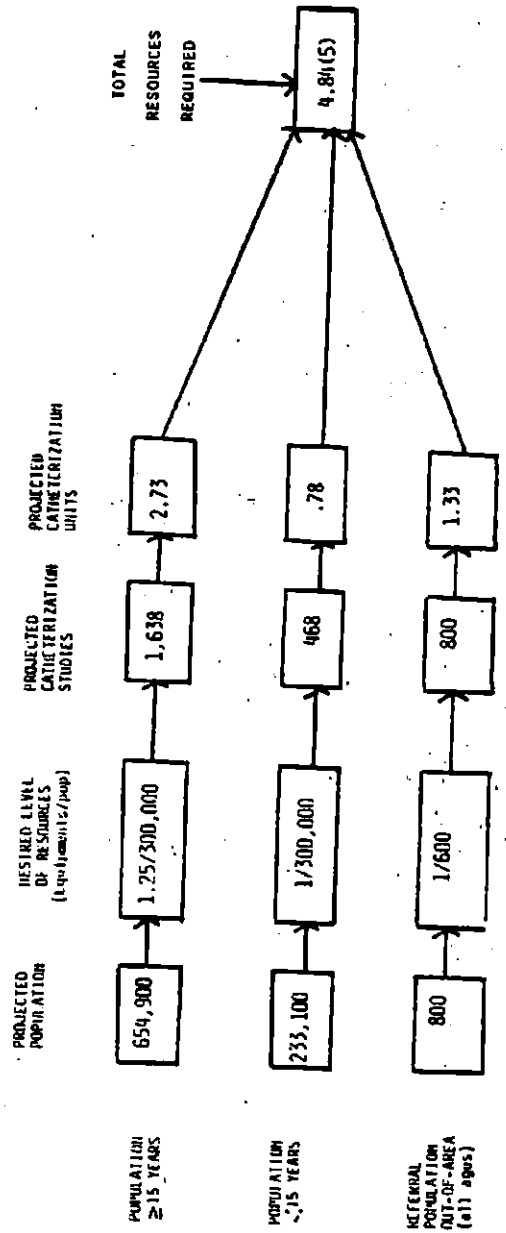


Figure 8A-5 Cardiac Catheterization Requirements for 1983, Health Service Area 1

Table 8A-6 Calculations for Open Heart Surgery Requirements, 1983

A. RESIDENT POPULATION (≥15 years)

Projected Cath Procedures		= 1,638
Open Heart Procedures Generated per Cath Study		= 1 per 4
Number of Open Heart Procedures generated	= .25 × 1,638	= 410
Maximum Number of Open Heart Units	$= \frac{410}{350}$	= 1.17

B. RESIDENT POPULATION (<15 years)

Projected Cath Procedures		= 468
Open Heart Procedures Generated per Cath Study		= 1 per 4
Number of Open Heart Procedures Generated	= .25 × 468	= 117
Maximum Number of Open Heart Units	$= \frac{117}{130}$	= .90

C. REFERRAL POPULATION

Projected Cath Procedures		= 800
Open Heart Procedures Generated per Cath Study		= 1 per 4
Number of Open Heart Procedures Generated	= .25 × 800	= 200
Maximum Number of Open Heart Units	$= \frac{200}{350}$	= .57

D. TOTAL OPEN HEART UNITS

1.17 + .90 + .57 = 3 units

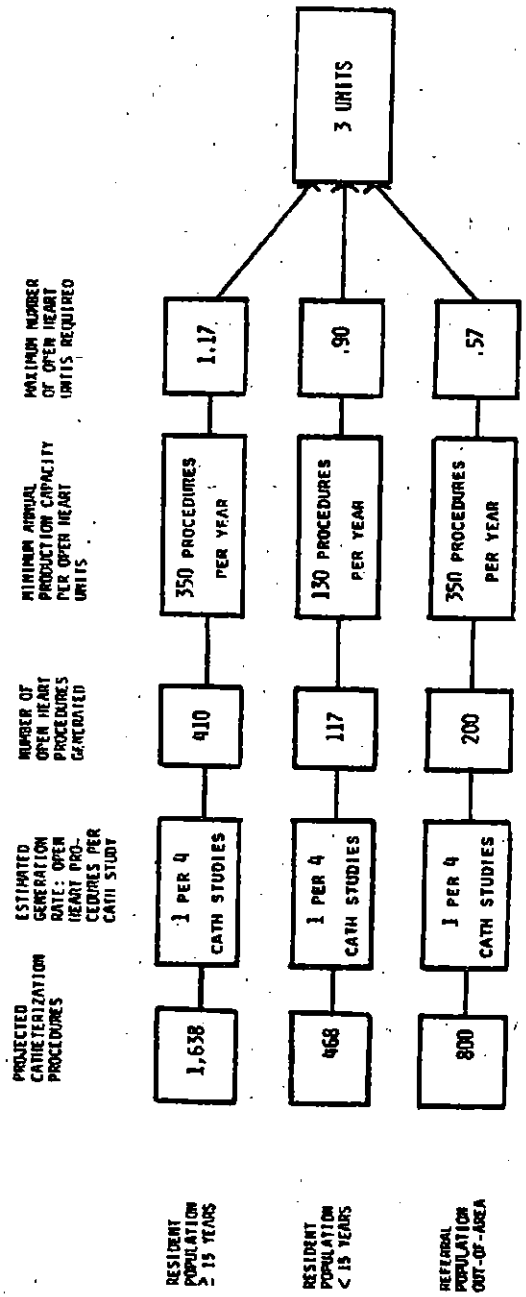


Figure 8A-6 Open Heart Surgery Resources Requirements for 1983, Health Service Area 1

Table 8A-7 Calculations for Radiation Therapy Requirements, 1983

A. RESIDENT POPULATION

Population		= 888,000
New Cancer Patients	$= \frac{3 \times 888{,}000}{1{,}000}$	= 2,664
Patients Requiring Treatment	= .50 × 2,664	= 1,332
Projected Number of Treatments	= 30 × 1,332	= 39,960
Maximum Number of Megavoltage Units	$= \frac{39{,}960}{10{,}000}$	= 3.99 or 4

B. REFERRAL POPULATION

Patients Requiring Treatment		= 333
Projected Number of Treatments	= 30 × 333	= 9,990
Maximum Number of Megavoltage Units	$= \frac{9{,}990}{10{,}000}$	= .99 or 1

C. TOTAL UNITS REQUIRED

4 + 1 = 5 units

331

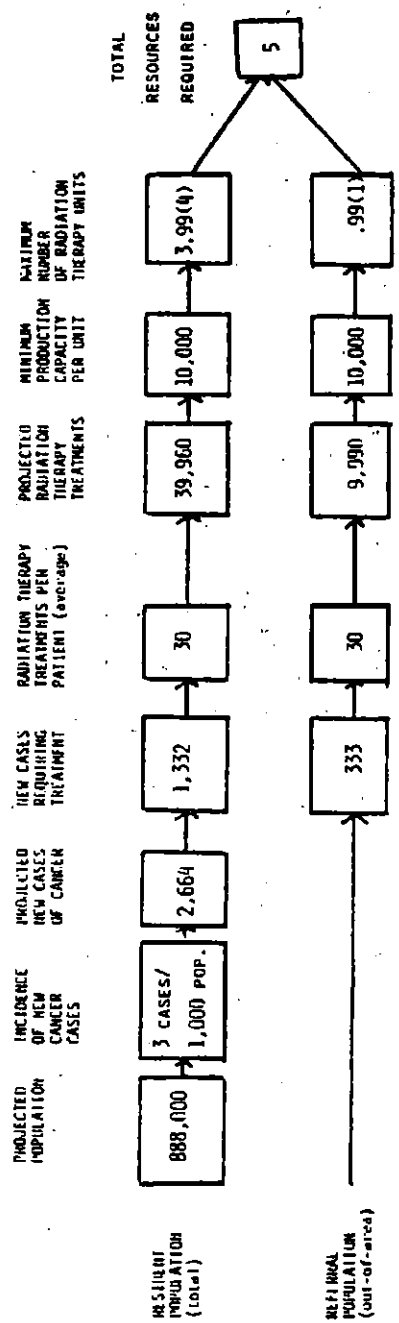


Figure 8A-7 Radiation Therapy Requirements for 1983, Health Service Area 1

Table 8A-8 Calculations for CT Scanning Resource Requirements, 1983

A. RESIDENT POPULATION

Projected Population = 888,000
CT Scanner/Population Ratio = 1/200,000

$$\text{CT Scanners Required} = \frac{888{,}000 \times 1}{200{,}000} = 4.44$$

B. OUT-OF-AREA POPULATION

Projected Population = 4,300
CT Scanner/Population Ratio = 1/3,000

$$\text{CT Scanners Required} = \frac{4{,}300 \times 1}{300{,}000} = 1.43$$

C. TOTAL CT SCANNERS REQUIRED

4.44 + 1.43 = 5.87 or 6 CT Scanners

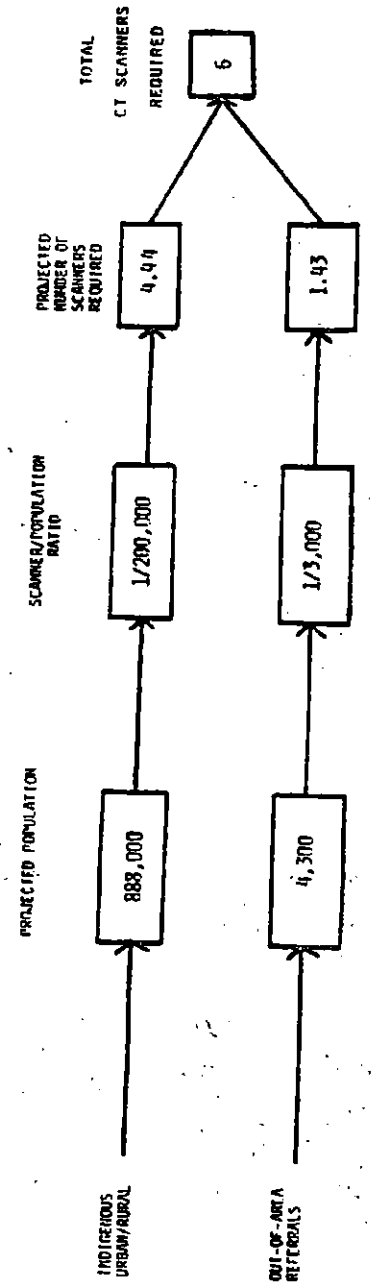


Figure 8A-8 CT Scanning Resource Requirements for 1983, Health Service Area 1

NOTES

1. The bulk of the material for this section is abstracted from Government Studies and Systems, *State-of-the-Art Report, Methods to Determine Geographic/Population Boundaries for Specific Health Services*, DHEW Contract No. HRA-231-77-0055 (Philadelphia: April 1978), pp. 5–6. The author was a consultant in the production of this volume.

2. G. J. Meneley et al., *Methodologies for Delimiting Health Service Areas* (Iowa City, Iowa: Health Services Research Center, the University of Iowa, Jan. 1977), p. 53.

3. Ibid.

4. Ibid.

5. Ibid.

6. D. L. Drosness and J. W. Lubin, "Planning Can Be Based on Patient Travel," *Modern Hospital* 106 (April 1966), pp. 92–94.

7. G. W. Shannon and G. E. A. Dever, *Health Care Delivery—Spatial Perspectives* (New York: McGraw-Hill Book Co., 1974), p. 139.

8. Ibid.

9. Rhode Island Dept. of Health, "Proposed Health Services Planning Subareas for Rhode Island, Health Planning and Resources Development, Tech. Report No. 7," Mimeographed (March 1977), p. 69.

10. Ibid., pp. 45–46.

11. Eleanor Poland and P. A. Lembcke, *Delineation of Hospital Service Districts*, U.S. Public Health Service, No. W-14 (1967), p. 895.

12. This section is modified from a letter from Dr. Frank Houser, District Health Officer, Georgia Dept. of Human Resources, Dalton, Georgia, 1978.

13. Meneley et al., *Methodologies*, p. 53.

14. J. N. Meade, "A Mathematical Model for Deriving Hospital Service Areas," *International Journal of Health Services* 4, no. 2 (Spring 1974): 353–364; and Meade, "Validity of Reporting Patient Origin Studies for Rural Hospitals," *Public Health Reports* 91, no. 1 (Jan./Feb. 1976): 62–66.

15. Government Studies and Systems, *State-of-the-Art Report*, pp. 23–25.

16. I would like to express appreciation to Gerry Rushton and Joe Meneley for their contributions in completing an accessibility analysis for Oklahoma that provided the nucleus for this section.

17. D. Eaton, R. Church, and C. ReVelle, "Location Analysis: A New Tool for Health Planners," Agency for International Development (A.I.D.), Washington, D.C., no. 53 (May 1977), p. 42.

18. G. Rushton et al., *A Statewide Plan for Regional Primary Medical and Dental Care Centers in Iowa*, Institute for Urban and Regional Research and Health Services Research Center (University of Iowa, Aug. 6, 1976), p. 82.

19. W. McClure, *Reducing Excess Hospital Capacity*, Bureau of Health Planning and Resource Development, DHEW (Oct. 15, 1976), p. 71.

20. See *National Guidelines for Health Planning*, DHEW (HRA) 78-643, March 28, 1978, for criteria reflecting population-based standards. Also, see Table 8-8 for standard-specific adjustments which may be applied to population-based planning.

21. McClure, *Reducing Excess Hospital Capacity*, p. 75.

22. Ibid., p. 76.

23. Ibid., p. 79.

24. This section on Oklahoma Hospital Planning could not have been accomplished without the services of Ann Ashmore and Karen VanWagner of the Oklahoma Health System Agency and Gerry Rushton and Joseph Meneley of the University of Iowa, Iowa City, Iowa.

25. Midwest Center for Health Planning, 702 North Blackhawk Avenue, Madison, Wisconsin 53705, "Application of the National Guidelines for Health Planning" and "Application and Adjustment Procedures Manual," Sept. 1978, 59, pp. 24–28. The Midwest Center is to be commended for producing such a fine document.

26. Ibid., pp. 26–28.

27. William E. DeBardlaben, "Basic Methods for Determining Community Resource Requirements," *Selected Methodological Approaches to Address the National Guidelines for Health Planning in Local and State Health Plan Development, A Workbook for Planners*, Health Planning and Development Center, Inc., Region IV, Atlanta, Georgia, January 1979, pp. 41–63, 186. I would like to thank the HPDC for permission to reprint selected parts of this document.

28. Ibid., pp. 45, 46.

29. Ibid., p. 46.

30. Ibid., p. 50.

31. Ibid., p. 51.

32. Ibid., p. 54.

33. Ibid.

34. Ibid., p. 60.

35. Government Studies and Systems, *State-of-the-Art Report*, pp. 29–61.

36. This appendix is abridged from William E. DeBardlaben, "Essential Analytic Steps by Specialty Areas," *Methodological Approaches to Address the National Guidelines for Health Planning in Local and State Health Plan Development, A Workbook for Planners*, Health Planning and Development Center, Inc., published by HPDC, Atlanta, Georgia, January 1979, pp. 91–134. I would like to thank the HPDC for permission to reprint selected parts of the document.

APPENDIX 8B
REFERENCES LOCATION/ALLOCATION PROBLEMS

1. Abernathy, William J., and Hershey, John C. "A Spatial-Allocation Model for Regional Health-Services Planning." *Operations Research* 20 (1972): 629–642.

2. Chapelle, R. A. "Location of Central Facilities, Heuristic Algorithms for Large Systems," Ph.D. Dissertation, Norman, Okla.: University of Oklahoma, 1970.

3. Church, Richard, and ReVelle, Charles. "The Maximal Covering Location Problem." *Papers of the Regional Science Association* 32 (Fall 1974): 101–118.

4. Dee, Norbert, and Liebman, Jon C. "Optimal Location of Public Facilities." *Naval Research Logistics Quarterly* 19 (1972): 753–759.

5. Ellwein, Leon B., and Kalberen, John T., Jr. "Optimal Location of Cancer Centers on the Basis of Population Access." *Federation Proceedings* 34 (1975): 1411–1416.

6. El-Shaieb, A. M. "A New Algorithm for Locating Sources Among Destinations." *Management Science* 20 (Oct. 1973): 221–231.

7. Francis, R. L., and Goldstein, J. N. "Location Theory: A Selective Bibliography." *Operations Research* 22 (1974): 400–410.

8. Garfinkel, R. S.; Neebe, A. W.; and Rao, M. R. "An Algorithm for the M-Medium Plant Location Problem." *Transportation* Science 8 (1974): 217–236.

9. Health Planning Research Services. *Introduction and Sample Methodology* and *A Report on the State of the Art. A Manual for Evaluating the Location of a Health Service*, vols. 1 and 2. Washington, Pa., 1976. Mimeographed.

10. Lea, A. C. "Location-Allocation Systems: An Annotated Bibliography." Discussion Paper No. 13. Dept. of Geography. Toronto: University of Toronto, May 1973.

11. ReVelle, Charles; Marks, Davis; and Liebman, Jon C. "An Analysis of Private and Public Sector Location Models." *Management Science* 16 (1970): 492–707.

12. Revelle, Charles; Toregas, Constantine; and Falkson, Louis. "Applications of the Location Set-Covering Problem." *Geographical Analysis* 8 (1976): 65–76.

13. Scott, Allen J. "Dynamic Location-Allocation Systems: Some Basic Planning Strategies." *Environment and Planning* 3 (1971): 73–82.

14. Scott, Allen J. "Location-Allocation Systems: A Review." *Geographical Analysis* 2 (April 1970): 95–119.

15. Shannon, Gary W., and Dever, G. E. Alan. *Health Care Delivery: Spatial Perspectives*. New York: McGraw-Hill Book Co., 1974.

16. Shannon, Gary W.; Spurlock, Carl W.; Gladin, Steven T.; and Skinner, James L. "A Method for Evaluating the Geographic Accessibility of Health Service." *The Professional Geographer* 27 (February 1975): 30–36.

17. Shuman, Larry J.; Hardwick, C. Patrick; and Huber, George A. "Location of Ambulatory Care Centers in a Metropolitan Area." *Health Services Research* 8 (Summer 1973): 121–138.

18. Tapiero, Charles S. "Transportation-Location-Allocation Problems Over Time." *Journal of Regional Science* 11 (December 1971): 377–384.

19. Toregas, Constantine, and Revelle, Charles. "Optimal Location under Time or Distance Constraints." *Papers of the Regional Science Assn.* 28 (Fall 1972): 133–143.

20. Toregas, Constantine; Swain, Ralph; Revelle, Charles; and Bergman, Lawrence. "The Location of Emergency Service Facilities." *Operations Research* 19 (1971): 1363–1373.

21. Wesolowsky, George O. "Dynamic Facility Location." *Management Science* 19 (1973): 1241–1248.

22. Wesolowsky, George O., and Truscott, William G. "The Multiperiod Location-Allocation Problem with Relocation of Facilities." *Management Science* 22 (September 1975): 57–65.

23. Zimmerman, James P. "Service Areas and Their Needs Must be Reassessed." *Hospitals* 49 (Sept. 1975): 46–48.

Chapter 9

Computer Graphics in Health Planning

BACKGROUND

Many trends are impacting on the changing scene of health care in this country: the shift from medical to social epidemiology, the shift from infectious to chronic disease patterns, the emergence of community health with emphasis on population groups in communities rather than individuals in clinics, and the use of computer graphics as a tool for community administrative decision making. The present chapter is concerned with the latter trend.

Medical care in this country has traditionally focused on the patient or the individual. This tradition has evolved from an infectious disease pattern in which cause-and-effect relationships were single and simple. Disease patterns are no longer infectious, however. Now they are chronic and noninfectious, and the causes and effects are not single and simple, but multiple and complex. Consequently, the dimension of the community had to be added to medical care.

In the 1980s, the focus on the community—the new health recipient—will be crucial to the holistic and wellness policies outlined previously. This new direction, however, suggests the need for new approaches to present community health information. One such approach, computer graphs, is not new, per se, but its application to community health usually has been limited and not well represented.

The practical use of computer graphics has been pioneered predominantly through avant-garde engineering technology by persons associated with the aerospace industry.[1] In the past, the medium of graphics has created the representation of geographic phenomena primarily for visualization. More recently, however, some administrative and planning applications of computer graphics output have surfaced in the health and health-related literature.[2 3 4] Researchers, planners, and administrators in

335

Chapter 6

Promoting Implementation

Tiré de: SPIEGEL, A.D. et HYMEN, H.H.
Basic Health Planning Methods.
Aspen, 1978.

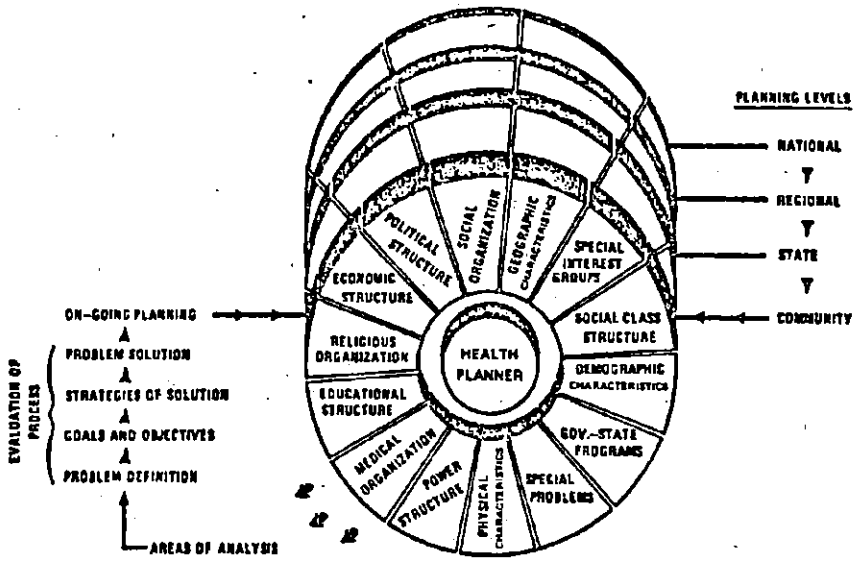
After all the discussion and the detailed energy involved in identifying the problem, finding the resources, considering the alternatives and determining priorities, the time comes to implement a program. Simply stated, implementation means that you are now going to utilize your materials, methods, and resources to do what you said you were going to do about a particular problem.

FIRST CONSIDERATIONS

In almost all implementation strategies, consideration will have to be given to the desired changes that the planner hopes to achieve. This consideration requires the ability to step back and see the whole forest, as well as the individual trees. An ecologic approach to comprehensive health planning as represented in Exhibit 6-1 suggests the broad spectrum of individuals and organizations that could be involved in the implementation process.[1] The author uses the term "ecologic" to indicate the teamwork that will be required between man and his environment to achieve comprehensive planning—note the elements in the wheel in Exhibit 6-1. Attention is directed in the diagram to the planning levels (national, regional, state, and community) as well as the evaluation of the process and the areas for analysis in the evaluation process. Of course, the overwhelming number of agencies and organizations could lead to the conclusion that a "philosopher-king," or a "Solomon" is needed as a planner. But social change manages to take place, though there may be a chance of reducing the "cottage industry" image through sustained planning efforts.

In planning for change, the health planner must consider the following dilemmas:[2]

Exhibit 6-1 An Ecologic Approach to Comprehensive Health Planning



Source: Reprinted with permission from J. G. Bruhn, "Planning for Social Changes: Dilemmas for Health Planning," *American Journal of Public Health* 63 (July 1973): 604.

- He must be an advocate for the beneficial effects of planning in a society that has been uninterested in social planning.
- He must plan in the context of a changing system while he is trying to effect change, realizing that his insights and knowledge about the system are also changing.
- He cannot deny his professional role and cannot become too entangled in the politics of planning, yet he must be a participant with the community in the planning process.
- The perpetual cycle of planning, which creates new needs and thus more planning, places him in the difficult position of needing to know when to leave the community system.

From the viewpoint of R. H. Murray,[3] "muddling through" with irrational, hit-or-miss planning is not enough. Four *sine qua nons* were cited to create conditions for change:

1. We must recognize that there will be a pervasive reformation of the health care system.

2. We must accept the federal government's major role in this change but work to restrict its power.
3. We must develop an open organization for the creation of a coherent national health policy and a strategy for implementation.
4. National health policy should be implemented through local agencies by managing the muddling-through process.

Although these *sine qua nons* could refer to a national health insurance plan, a PSRO, or an HSA, they could just as easily be related to any implementation effort. Certainly the system will undergo reformation to some degree; there will be an interplay among organizations involved in the change; a need will exist for an overall implementation strategy; and, based on past history, the local agencies will want to control their own destiny.

Taking a behavioral approach, another concept of implementing change lists the characteristics of change that need consideration in the day-to-day working situation:[4]

- Above all, the model of implementing change should be practical.
- Parts of the model should be manipulatable.
- Economy of use should be a primary consideration.
- Ease of communication is important.
- The model should be comprehensive.
- Synergism—the force of factors working together—is important to consider. A general weakness of most models is that they seem to view, or at least report, changes as linear, rational processes.
- The model should lend itself to intervening in phases.
- Differential investment in working with the components of the model should be possible.
- The model should call attention to how the change process influences the rest of the system.
- The model should be flexible and versatile enough to apply to different organizational systems.
- The model should provide a basis for a subsequent evaluation of the effectiveness of change.
- The model should recognize the human qualities of the participants involved.

Incorporating these characteristics of change into a working model, a checklist for change was developed and is reproduced in Exhibit 6-2. The acronym "A Victory" is taken from the first letters of the eight factors considered to be required for organizational change.

Exhibit 6-2 A Checklist for Change Through Research Utilization

This checklist is intended to serve as a guide rather than as an outline for a systematic plan to bring about change. All factors interact, so that a given manipulation to increase the probability of desired results could influence more than one factor.

Ability

_____Are staff skills and knowledge appropriate to accommodate the desired change?

_____Are fiscal and physical resources adequate for the change?

Values

_____Is the change consonant with the social, religious, political, ethnic values of the beneficiaries?

_____Is the change consonant with the philosophies and policies of the program supporters?

_____Is the change consonant with the personal and professional values of staff?

_____Is the top man in the organization in support of the desired change?

_____Are the characteristics of the organization such as to render change likely?

Information

_____Is information on the desired change clear?

_____Does information about the idea bear close relevance to the improvement needed?

_____Is the idea behind the desired change one that is "tryable," observable, of demonstrated advantage, etc.?

Circumstances

_____Are conditions at this setting similar to those where the idea was demonstrated to be effective?

_____Does the present situation seem to be conducive to successful adoption of this particular plan?

Timing

_____Is this a propitious time to implement this plan?

_____Are other events going on or about to occur which could bear on the response to this change?

Obligation

_____Has the need for this change been ascertained through sound evaluation?

_____Has the need for this change been compared with other needs in this program?

Resistances

_____Have all reasons for *not* adopting this change been considered?

_____Has consideration been given to what may have to be abandoned if this plan is launched?

Exhibit 6-2 Continued

———Has consideration been given to all who would lose in this change?

Yield

———Has the soundness of evidence about the benefits of this proposal been carefully assessed?

———Have possible indirect rewards for this change been examined?

Source: Reprinted with permission from S. E. Salasin and H. R. Davis, "A Practical Model for Planned Change." Paper presented at AHPA meeting, November 13, 1972.

An illustration of the "A Victory" approach can be seen in the following example for a mental health project:[5]

> A "weekend hospital" was proposed for patients who needed intensive help but who couldn't come regularly for a variety of reasons. Services would be provided from Friday evening through Sunday afternoon and continued for a period of 16 weeks. Techniques would include saturational group therapy, a marathon approach.
>
> What factors caused the hospital to adopt and implement the project or to not adopt and implement the project?
>
> Now, the checklist for change can be applied to this situation and the results of the examination noted.

Ability	Two kinds of resources needed; staff competence in conducting marathon therapy and up to $9,000 for staff time and incidental costs. These modest resource needs constituted the greatest barrier to implementation.
Values	Therapists had to commit themselves to be away from their families for 16 straight weeks, and this clashed with the personal values of most potential participants. Therapy methodology also provided some conflict in values as well as some agencies who were rigidly orthodox in their values and their consideration of implementing innovative change.
Information	Initially only 5% of the potential users were familiar with the weekend hospital idea de-

	spite good traditional information dissemination. Adding an attractive brochure with concrete examples raised the information level to 55%. Information was felt to be limited in terms of relevance, relative advantage, and credibility.
Circumstances	The weekend hospital concept proved to be more relevant to families in suburban neighborhoods and areas where transportation was difficult.
Timing	New budget cycles, reorganizations, changes in decision-makers and crises offer opportunities for timely implementation of change. At the time the weekend hospital was promoted, the timing did not appear particularly appropriate.
Obligation	Hospitals did not indicate they were aware of any pressure from patients for an alternative to the traditional hours of service. A few staff members felt there was a need and others reflected a need to start something new. There was virtually no felt need to implement a program such as the weekend hospital.
Resistances	Almost everybody resists upsetting the apple cart when they can rationalize the good of their services. A new program will require that resources be allocated from existing programs and administrators resist losing budget money and/or staff. Therefore, resistance is built in to most systems except where the new service is added on. Then different resistances enter.
Yield	Few rewards were implied for implementing a weekend hospital. Job security, creativity, writing papers and the like didn't interest staff members. Some noted that the weekend hospital might be more convenient for certain patients.

This analysis of the responses of community mental health centers to the weekend hospital idea clearly indicates that implementation at

that time did not have a reasonable chance of success. Implementation could not be advanced despite the fact that the weekend hospital had been developed, demonstrated, and thoroughly evaluated in a controlled design over a period of years. Patients served in the weekend hospital improved significantly more—both statistically and practically—than did control patients who were provided traditional outpatient treatment over the same period.

From another vantage point, that of the health systems engineer (also referred to as industrial or management engineers), barriers to implementation were noted in a survey of 240 health systems experts[6] (see Exhibit 6-3). Furthermore, four major problem areas were identified, along with the frequency that these problems act as barriers to the implementation and attainment of the goals of the health system. Note that planning and policy problems emerged as the most prevalent problem with 70 percent of the frequency (see Exhibit 6-4). Although the listing of barriers in Exhibit 6-3 concerns itself with aspects of quality, cost, and accessibility of health care services, this listing directly applies to the implementation of a project since those aspects are usually part of any project. In Exhibit 6-4, the health systems engineer, though mainly concerned with operational problems, replied that he spent more time on areas outside of operational functioning. Of interest to planners should be the fact that the health systems engineers are being called upon to fulfill many of the tasks and functions of health planners, for better or worse. All other problem types cited constitute an overwhelming majority of topics that the health planner should rightly be concerned with: educational problems 50 percent, research and development problems 20 percent, and planning and policy problems 70 percent.

In a specific example relating to the implementation of a university hospital ambulatory care evaluation, the impediments to implementation were documented with an emphasis on judging quality assurance. Six barriers were cited:[7]

1. The condition of clinic records and individual charts.
2. The lack of established criteria for care.
3. Problems of provider intercommunication during the evaluation process.
4. Manpower availability.
5. Choice of evaluation method.
6. The method of implementing resulting plans for corrective action.

The sixth barrier, implementation, pervades all the others and indicates the necessity for reflecting on implementation during all phases

Exhibit 6-3 Composite List of Important Barriers To Improving the Quality, Cost, and Accessibility of Health Care Services

Barriers	Problem Impact		
	Access	Cost	Quality
Consumers lack understanding of preventive measures.	•	•	•
Consumers lack understanding of when services are needed.		•	•
Third-party payment policies do not adequately discriminate between poor and good quality care.			•
Incentives for quality care provided by competitors, governing boards, third-party payers, and consumers are minimal.	•	•	•
Lack of quality care standards.			•
Consumers and providers view service as a reactive rather than a preventive function.	•		•
Organization and coordination of all health services is inadequate.		•	•
Inability to agree on health goals.	•		•
Consumers lack means of providing incentives to improve quality.		•	•
Services are not accessible when needed.			•
Consumers hesitate to take issue with the physician.		•	•
Information and measures causally relating treatment interventions to outcomes are largely absent.	•	•	•
Fee for service system encourages an increase in the volume of services provided but not in the quality of services.	•	•	•
Consumers lack understanding of availability of services.	•		•
Providers lack skills to evaluate proposed innovations.	•	•	
Needed cost benefit information does not exist.		•	
The increasing and changing nature of demand for health services.		•	
Inflation.		•	
Implementation skills for cost reduction projects are lacking.		•	
Cost reduction projects have had the wrong forms.		•	
Providers lack management skills.		•	
Consumer problems in financing their health care.	•		
Incentives do not exist for physicians to locate in rural or ghetto areas.	•		
Services not reimbursed by third-party payers are not sought.	•		
Legal and educational barriers to use of paramedical personnel.	•		
Design skills needed to improve accessibility are not present.	•		
Economic barriers prohibit wide distribution of high cost/low demand services.	•		

Source: D. H. Gustafson, G. L. Rowse, N. J. Howes and R. S. Shukla, "Opportunities for Improvement in Health Systems Engineering," *Inquiry* 14 (March 1977): 97. Reprinted, with permission of the Blue Cross Association, from INQUIRY. Vol. XII, No. 1, p. 44. Copyright © 1975 by the Blue Cross Association. All rights reserved.

of the health planning process. Why spin wheels if a method cannot be prepared to achieve the desired result? It might be better to let things go on as is until something can be done.

Another revealing example deals with an analysis of factors influencing the implementation of family planning services in the United States. A survey of hospitals and health departments in New York state showed that a majority of the hospital administrators (43 percent) and the health officers (63 percent) did not implement family planning services because they believed that the services were provided by another agency.[8] Additional reasons for nonimplementation

Exhibit 6-4 Problem Types and Extent of Prevalence

Problem type	Percent prevalence
Operational problems	25
Those that require the refinement of existing procedures to improve productivity or satisfaction in health systems. Projects of this type occur primarily in delivery settings and focus on techniques for establishing standards, improving work methods and personnel utilization, accomplishing facilities layout, developing simple communication and monitoring procedures, and locating and providing needed information for decision-makers.	
Planning and policy problems	70
These involve the need for identifying appropriate policies and procedures that alter the health system at an organization-wide or interorganizational level. Projects dealing with these problems are currently being undertaken primarily in planning agencies, but should in the future also be undertaken at the upper management levels of delivery systems. Usually these projects are of a complexity that would benefit from considerable data collection, analysis, and the interaction and agreement of many diverse types of professionals. Outcomes of these projects include identification and redistribution of inequitable services across delivery areas, increased emphasis on the provision of preventive services to consumers, better definition of and agreement on health goals, new consumer financing mechanisms, innovative reward and incentive systems for health professionals, improved types of delivery units, provision of new computer and other management service mechanisms, and better coordinative links among various health organizations.	
Educational problems	50
These involve the need to improve the knowledge and skills of consumers or providers involved with the health system.	
Research and development problems	20
Those for which considerable data need to be collected under rigidly controlled conditions in order to develop effective and optimal solutions.	

Source: D. H. Gustafson, G. L. Rowse, N. J. Howes and R. S. Shukla, "Opportunities for Improvement in Health Systems Engineering," *Inquiry* 14 (March 1977): 97. Reprinted, with permission of the Blue Cross Association, from INQUIRY. Vol. XII, No. 1, p. 44. Copyright © 1975 by the Blue Cross Association. All rights reserved.

are listed in Exhibit 6-3. Among hospitals, the low community demand and the feeling that family planning was an inappropriate responsibility to assume indicate that health planners would need to mount a promotional information campaign prior to implementing a service-type of implementation project.

Implementation planning is being affected by the continuing dissatisfaction expressed by a variety of individuals and organizations with the health care system. As innovations or changes are considered, the health planner must be aware of three dominant themes consistently identified in the literature:[9]

1. People, both professional and lay, are more frequently questioning commonly used medical procedures. Tonsillectomy and hysterectomy are among the surgical procedures increasingly criticized as ineffective, overly expensive, and, occasionally, as dangerous.
2. Consumers are demanding more services from hospitals. Services

Exhibit 6-5 Reasons for Nonimplementation of Family Planning Services Named by New York State Hospital Administrators and Health Officers

Hospitals			Health Departments		
Reasons	No.	%	Reasons	No.	%
1. Provided by another agency	21	42.9	1. Provided by another agency	12	63.2
2. Low community demand	11	22.5	2. Did not answer	2	10.5
3. Inappropriate responsibility to assume	6	12.2	3. Lack of appropriations	1	5.3
4. Community needs met	4	8.2	4. Lack of board approval of funds	1	5.3
5. Opposition by board	2	4.1	5. Community needs met	1	5.3
6. In process of instituting	2	4.1	6. Opposition by physicians	1	5.3
7. Not enough time to institute	1	2.0	7. Not enough time to institute	1	5.3
8. Personal opposition by administrator	1	2.0			
9. Did not answer	1	2.0			

Source: Reprinted with permission from J.T. Gentry, et al, "A Comparative Analysis of Factors Influencing the Implementation of Family Planning Services in the United States," *American Journal of Public Health* 64 (April 1974): 376.

such as geriatric maintenance, counseling, and humane basic primary care are merely examples.

3. The health needs of the community are receiving increased attention, as well as the concept of preventive health services. Malnutrition, housing, stress, and a wide variety of additional factors affecting the health of the community will have to be dealt with by health organizations.

Implementation projects have to be considered within the environment created by these three dominant themes. Obviously, the chances for success in implementing a project are much better if the contemplated change is in keeping with current feelings among the individuals and organizations involved in the project.

Three types of organizational change were identified that relate to implementation planning: technical, adjustment, and adaptational.[10] Technical changes are those that are primarily changes in the normal and usual activities of a hospital, but not in the basic goals. These changes may vary in their cost and their impact on a hospital. A decision to replace a four-test blood analyzer with a twelve-test one presents a cost increase with little impact on the overall function of the laboratory. However, a decision to install a cobalt machine represents a substantial cost increase as well as making an impact on the distribution of resources and utilization of the services of the cobalt machine. Nevertheless, both of these changes are essentially technical in nature dealing with the same services and the integration of new technology.

Adjustment changes are those that require the currently existing mechanisms for delivering health care to change to provide a service not offered in the past. Movements of the hospital to act as a primary care provider to the community or to establish a nontherapeutic abortion service mean that the hospital will shift existing resources and personnel to provide the services. The means remain the same; the ends change.

Adaptation changes are those that involve changing not only the means but also the ends. This is the most drastic type of change because it means a complete modification of the existing system. A community hospital that shifts its emphasis from acute and chronic treatment services to preventive health services is an example of adaptive change. The hospital would have to reorient its staff, reallocate its services, and engage in community outreach, case finding, social work, family medicine, and prevention.

Tools that the health planner can use to implement change were categorized by A. Etzioni[11] as utilitarian, normative, and coercive controls. Utilitarian control has as its base the control of resources such as knowledge, skills, and/or economic capabilities. Obviously, any change that has to be implemented requires that the planner have something to say about utilitarian control of resources.

Normative control is based on the fact that the goals, values, and/or ends are shared and agreed to in essense by those doing the planning. There can be no doubt that if a consensus exists among the parties involved in any implementation activity, the activity will proceed smoothly securing whatever resources are needed to initiate the project.

Coercive control rests on the actual control or access to control of force that influences any decision. Every health planner will probably

come up against some use of pure, brute force in implementation situations. It might be of a political, economic, or zealous nature. But the planner will have to abandon the rational decision for the coercive one since alternatives may not be forthcoming.

All three types of control seriously affect health planning and the promotion of implementation methodology. Bickering among bureaucrats is a common occurrence when utilitarian control is exerted. Traditional values that refuse to budge battle with the "young turks" trying to reform the health care system by bringing in the grass roots consumers to run services. Or, a local politician throws all the planning priorities out of kilter by demanding that a building and services be established in his home territory before anything else happens. All these examples have actually occurred and will occur again in the future. These types of control may be used subtly or with raw élan; but, in any event, the implementation strategy has to be readjusted, redirected, or modified to accommodate the planning goals.

Yet, despite the need to concentrate on the variables involved in implementation and the resultant change, some planners suffer setbacks for reasons beyond their control or influence. In a discussion about delivering health care amid change, J. J. McGuire[12] noted the following about changes in the system:

> The greatest problem we face is how to control change in a complex environment so that some of the fundamentals we stand for are not destroyed or whittled away by political ambition playing upon consumer expectation like a fine instrument.

In relating the above to implementation of health programs, H. M. Sapolsky[13] considers the market and the regulatory strategies. Both are being pursued to control costs, to promote quality care, and to redirect activities into ambulatory, preventive, and long term care—away from the traditional emphasis on acute care. Market strategy aims to induce physicians to conserve resources. Regulatory strategy aims to compel physicians to reorder their professional priorities. Sapolsky contends that neither strategy can succeed and suggests an alternative—bureaucratic competition, the pitting of one professional interest against another. Implementation of this strategy in health would allow three or four governmental agencies to compete for the management control of specific federal health programs. Each agency would have a specific value bias, such as preventive care, chronic care, primary care, or inpatient acute hospital care. Thus, each agency would be implementing a program with its own value system, and the

possibility of competitive choices to be decided by the public or its representatives seems more likely. In theory, these preferences could resolve the health crises that were engendered with the professionals making the decisions alone.

In view of the emphasis on change discussed here, do providers and health care professionals have to gird themselves for a topsy-turvy health world? M. F. Doody[14] reviewed the literature in 1975 and considered the question of whether providers change the system. He concluded:

> The 1975 literature on health care delivery is largely "more of the same" and indicates a high degree of acceptance of the status quo by providers. To some extent, this situation is not suprising, considering the general economic climate and the uncertainty about impending legislative and regulatory proposals.

Past history appears to support the status quo concept. Health care providers have been able to maintain considerable influence over what does and does not get implemented. Health planners until now have been able to anticipate and meet most objections. Will the new legislation and regulatory proposals change all prior history? It is too early to tell, but this certainly bears watching as developments take place.

IMPLEMENTATION WORKSHEET

An implementation worksheet can be used somewhat like the traditional approach utilized by newspaper reporters. Rudyard Kipling summed it up:

> I keep six honest serving-men
> (they taught me all I know);
> Their names are WHAT and WHY and WHEN
> And HOW and WHERE and WHO.
> I send them over land and sea,
> I send them east and west;
> But after they have worked for me,
> I give them all a rest.
>
> "The Elephant's Child," Rudyard Kipling

For each implementation activity, the following questions have to be answered:

- Who does what for whom?
- In what order and when?
- With what resources?

Utilizing the implementation worksheet approach, the planner can be quite clear about the specific tasks, the time period required for completion, and for what individual or agency is going to be responsible. Of course, it is vital that the planner know the end goals. Exhibit 6-6 lists the steps in the development of an implementation strategy. Note that the logical determination of the details also includes dates and the identification of responsible persons.[15]

Two examples of worksheets are illustrated in Exhibits 6-7 and 6-8. One deals with the development and distribution of a training manual for patient educators; the other shows plans to implement a countywide detection and treatment center. Both worksheets operate within a deadline and cover more than eight months in the implementation process.

The implementation worksheet method is probably the most traditional, although the use of a worksheet adds a bit more structure. Generally, the planner takes the decision to implement the activity and figures out the details in a logical manner, answering the appropriate questions about who, what, where, when, and how. Why has already been answered by the policy decision to go ahead with the activity.

Another related technique calls for the development of an action plan (see Exhibit 6-9), which sets down the short term objective, the specific activities, identifies who is responsible, defines the purpose of the activities, and lists the methods and the media to be used in the implementation.[16]

An action implementation plan assumes that for every action there will be a reaction—if something is provided then there will be some outcome. Implementation in an action plan, as is true in considering all implementation plans, must weigh the available resources, timing, and effort required for success. Again, the approach has a kinship with the newspaperman's guidelines for writing a headline story:

- why—the effect of the objective to be achieved;
- what—the activities required to achieve the objective;
- who—individuals responsible for each activity;
- when—chronological sequence of activities and timing in relation to other community events;
- how—materials, media, methods, techniques to be used;

Exhibit 6-6 Development of Implementation Strategy

1. Specify clearly (who does what for whom).
2. Be sure someone is responsible for the whole activity and coordinates individuals who may carry out the different tasks.
3. Identify all preparatory steps prior to doing that activity (e.g., prepare materials, write article, acquire equipment, train volunteers, prepare training manual, determine treatment protocol).
4. List steps in the order in which they must occur.
5. Check for missing steps that must be added.
6. Determine when (date) each steps should begin and end.
7. Check your dates to make sure that the correct amount of time has been allowed.
8. Consult with organizations affected by the activity; identify potential problems, opportunities, etc.
9. Specify what resources will be needed and their source.
10. Specify what constraints need to be addressed.
11. Make sure all people involved know what is expected of them and by when.

Source: National Heart, Lung and Blood Institute. *Handbook for Improving High Blood Pressure Control in the Community* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 36.

- cost—an estimate of materials and time; and
- feedback—when and how to tell if the activities are working and if adjustments are needed.

Sometimes a linear timetable is used to place the activities along the line and give visual shape to the overall program. In any case, most planners have the available data on hand to implement a program. The worksheet method or action plan technique opens the door for an examination of other techniques that require structured planning for implementation such as flow charts, network diagrams, decision trees, PERT, and gaming simulations.

Flow Diagrams

At times, the planning for implementation will involve services to clients in a clinic, hospital, health center, or other type of facility. When this occurs, a diagram showing the flow of the clients through the system is helpful in the implementation process. As the flow chart

Exhibit 6-7 Worksheet for Implementation Strategy: 1

Worksheet for Implementation Strategy

Activity: Develop and distribute training manual for patient educators

Constraints: Deadline of 7/30. Must be accepted by Committee.

Implementation Strategy:

	(who does what)	(when)
1.	Committee develops methodology	11/10
2.	Educator prepares draft manual	12/10
3.	Committee reviews draft manual	12/20
4.	Nine practicing educators review draft manual	1/10
5.	Educator prepares final draft	1/31
6.	Artist prepares graphics	2/15
7.	Nine practicing educators test manual	4/15
8.	Educator evaluates test and makes revisions	5/15
9.	Committee approves manual	6/1
10.	Manual is printed	6/30
11.	Manual distributed to nurses and educators	7/30
12.	Educator evaluates training manual	7/30
13.		
14.		
15.		
16.		
17.		
18.		
19.		
20.		
21.		
22.		
23.		
24.		
25.		

Source: National Heart, Lung, and Blood Institute, *Handbook for Improving High Blood Pressure Control in the Community* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 33.

is being put together, the logic of the steps can be mentally tested to see if the steps reach the end result.

Consider Exhibit 6-10, a listing of five alternatives in a decision grid for an executive heart care program. Elements under consideration include the effect on mortality and disability, cost, feasibility, desirability and acceptance, and a priority decision. Alternative 3 received the first priority. Note that an exercise test, a serology exam, a psychological test, and a maintenance regimen along with referral to the client's private physician are all part of the proposed program.[17]

Now examine Exhibit 6-11, which illustrates a flow chart for new outpatients. Does this flow diagram meet the requirements of the new executive heart program? Questions could be raised as to where does

Exhibit 6-8 Worksheet for Implementation Strategy: 2

Worksheet for Implementation Strategy	
Activity: Develop a county detection and treatment center Constraints: Residents of the county have to be stimulated to come in and to undertake follow-up. Deadline 1/1 of next year.	
Implementation Strategy: (who does what)	(when)
1. Director of Ambulatory Services develops proposal for detection/treatment center	1/1
2. County Board funds proposal	3/1
3. County renovates facilities	6/1
4. Director hires staff, including administrator and medical director	6/15
5. Administrator and medical director develop protocols and procedures, including special effort to motivate residents to participate and to follow-up	8/15
6. Administrator acquires equipment and supplies	8/15
7. Data specialist develops patient record forms	9/15
8. Health educator develops education materials	9/15
9. Staff and volunteers are trained by medical director, health educator and head nurse	10/1
10. Administrator tests methods and materials with county residents	10/15
11. Staff begins detection and treatment services to county residents	11/15
12. Administrator evaluates detection and treatment services	11/15

Source: National Heart, Lung, and Blood Institute, *Handbook for Improving High Blood Pressure Control in the Community* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 37.

the exercise testing, laboratory work, and the psychological testing take place. Perhaps in the box labeled clinic or service.

Another example is Exhibit 6-12, which depicts an automated multiphasic health testing sequence. Does this flow chart meet the requirements of the executive heart care proposal? Perhaps a combination of the two answers the need. The executive heart care proposal

Exhibit 6-9 Implementation with an Action Plan

An action plan detailing activities, responsibilities, and methods might be set up in the following fashion.

Short Term Objective	Activity	Responsibility	Purpose	Methods & Media
Gain support of community leaders	Interviews, conferences, written documents and articles	Health leaders, social scientists	Bring facts to attention; enlist support in solving problems	Kit of materials with statistical evidence, case histories, examples
Gain support of service groups	Talks, discussions, programs	Educators, health professionals, youth leaders, agency-based members	Bring facts to attention, involve service groups in the study and solution of problems, enlist support	Talks reinforced with written materials, audio-visuals, newsletters
Help health and welfare agencies to anticipate change & begin planning change	Seminars, workshops, study groups, conferences	Administrators, experts, policy-level board and staff	Discuss data, study present practices and needed change, enlist community for change	Wide variety of media: face-to-face and small group discussions, audiovisuals, printed materials.

Source: S.G. Deeds, *A Guidebook for Family Planning Education* (Columbia, Md.: Westinghouse Population Center, 1973), p. 56.

does have some elements of multiphasic testing and routine clinic operations.

For implementation purposes, the flow diagram logically leads to the consideration about what manpower, materials, and funding are needed at each step in the process.

Look again at Exhibit 6-11, the flow chart for new outpatients. Start with the direct travel arrows and respond to questions such as:

Exhibit 6-10 Decision Grid, Executive Heart Care Program—Target Population 5,000

Alternative	Effect on Mortality	Effect on Disability	Relative Cost	Feasibility	Desirability and Acceptance	Summary	Decision
1. Present system: treatment of acute cardiac episodes vs. prevention; inadequate early detection by private physicians	120 deaths per year, 60% considered preventable	800 disability days per year, 50% considered preventable	Dollar cost of preventable portion of morbidity and mortality as expressed in actual dollars expended for medical service: $640,000 per year		Unsatisfactory because known medical and social interventions not used, upping cost	All interested parties persuaded to modify present practices	Not recommended
2. Early detection through screening 80% of target, serology exercise tests, referral of high risk patient to own physician	10% reduction of preventable deaths	15% reduction of preventable disability	Cost: $75,000 Return: $69,000	Excellent technology and personnel available	Poses minor acceptance problems with patients, but none with attending physicians	Cost high compared with return	Not recommended
3. Above screening plus psychological testing plus recommendation of maintenance regimen to own physician	25% reduction of preventable deaths	30% reduction of preventable disability	Cost: $125,000 Return: $300,000	Requires recruitment and training of clinical psychologists to give screening test	Some resistance on part of patient to be overcome	Greatest return compared with cost. Least coercion of patient or disturbance of doctor-patient relationship	First priority

Exhibit 6-10. Continued

Alternative	Effect on Mortality	Effect on Disability	Relative Cost	Feasibility	Desirability and Acceptance	Summary	Decision
4. Above program plus group education of high risk patient in personal care program, including diet, exercise, stress prevention, plus continuing education of attending physician	30% reduction of preventable deaths	50% reduction of preventable deaths	Cost: $150,000 Return: $350,000	Requires development of education aids, such as films and videotapes. Some delay in implementation	May be difficult to obtain full participation because of lack of interest	Embraces patient responsibility, greatest potential return if screening program discontinued since patient will be motivated to continue	Second priority
5. Above program except quarterly examination, supervised group exercise, education of spouse on diet	35% reduction of preventable deaths	65% reduction of preventable disability	Cost: $300,000 Return: $415,000	Places heavy burden on available manpower	Excessive control and coercion of patient. Voluntary participation not likely to remain at level needed to obtain results	Greatest total return but burden on system and patient excessive	Not recommended

Note: This grid can be used to arrive at a decision on any problem once alternatives are defined and elements selected. At option of the user the elements may be weighted. For example, the relative cost element could be assessed a weighting factor of 2.5 times any other element, which would support alternative 3 as being of first priority.

Source: American Hospital Association, *The Practice of Planning in Health Care Institutions* (Chicago: AHA, 1973), pp. 56–57. Reprinted with permission. © 1973 American Hospital Association.

Is there enough parking space for the additional clients?
Do we need public conveyances to travel to the clinic?
If the clinic is to operate at night, do we have lights?
Is a parking attendant needed?
Do we need signs to direct the clients?
Is there an information person in the vestibule?
Does the control officer check for insurance coverage?
How does the control officer communicate with the clinics?
How many interview cubicles are needed?
Who does intake screening?
Can we use physician assistants in the clinic service?
Is medical social service available to all clients?
Is there a discharge officer at check out?

These are only a few of the questions triggered by the flow diagram that relate directly to the implementation of the program. One method of plotting the implementation needs is to list the steps and then fill in the chart under each heading:

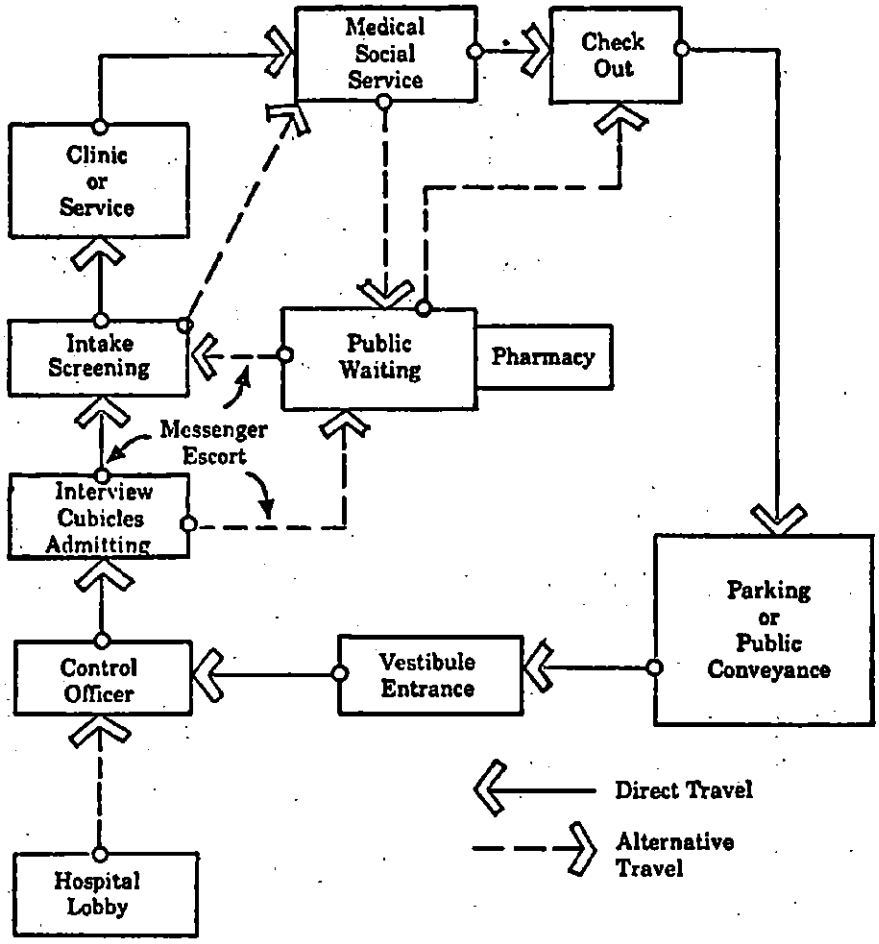
Step	Manpower	Materials	Funding
Parking	Attendant	Minibus Lights	Hospital budget
Vestibule	Information clerk	Signs Mimeo directions Phone	Hospital budget Volunteers
Control	Control officer Messenger	Phone Files Forms	Hospital budget

As each step is charted, the planner can add or delete activities and requirements as ideas occur. When the planner finishes following the arrows in the flow chart, the implementation process should emerge fully documented as to what is needed for the total effort.

Mentally, the staff can go through the process using the flow chart and the accompanying data about manpower, materials, and money and give the implementation plans a dry run. This should help find any rough spots and point out needed adjustments. Of course, once the program starts, the planner should be on hand to evaluate the validity of the flow chart and the other aspects of the implementation plan.

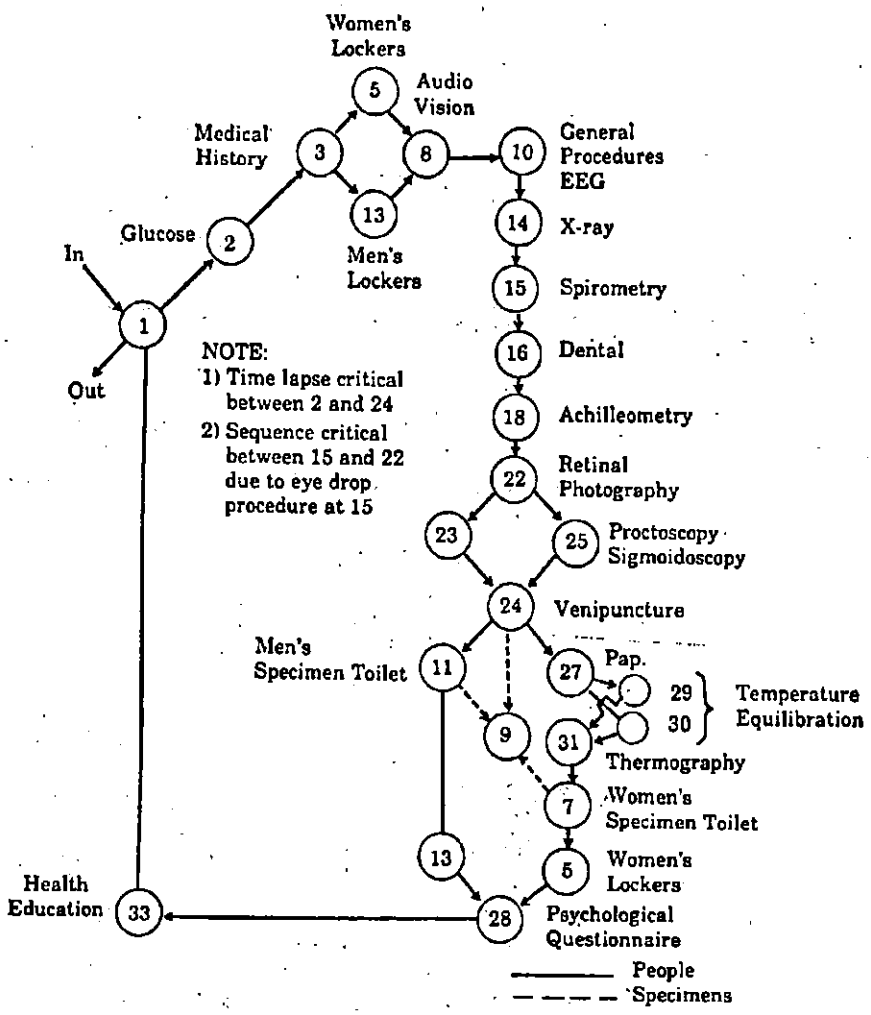
A variation of the flow chart is shown in Exhibit 6-13. This chart shows the client flow through a rehabilitation system with the four

Exhibit 6-11 Flow for New Outpatients



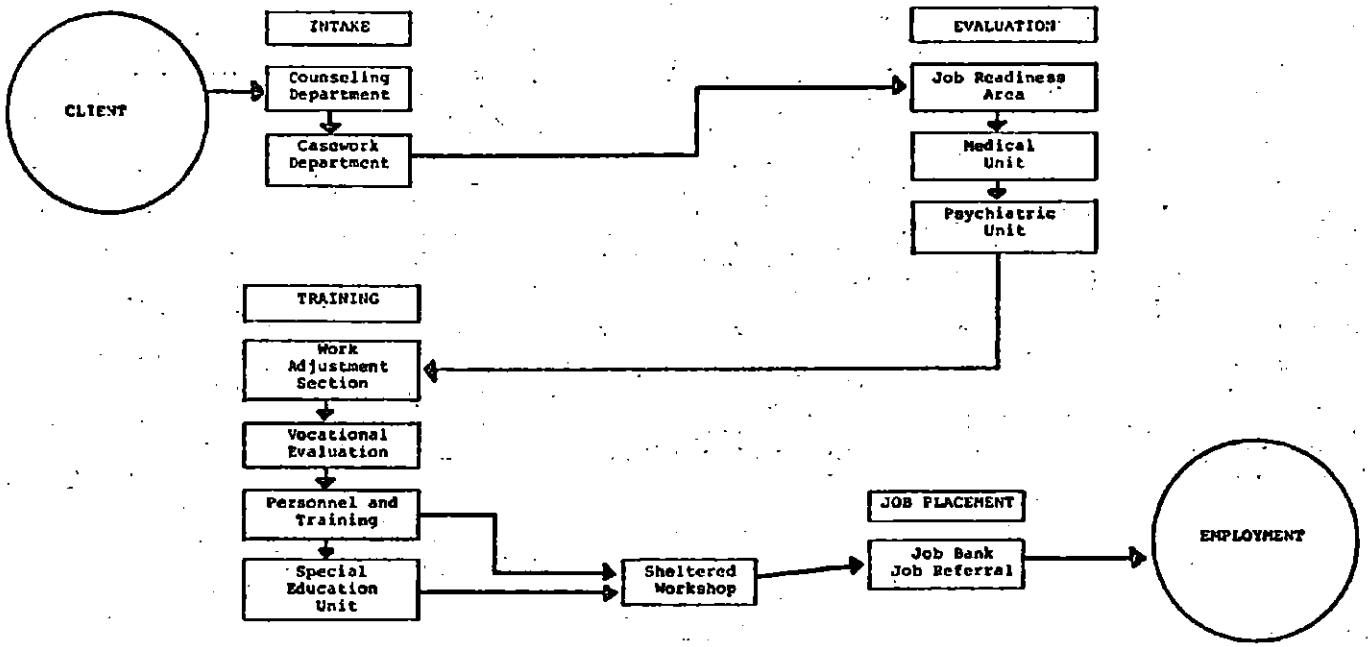
Source: Health Resources Administration, *Guidelines for Functional Programming, Equipping, and Designing Hospital Outpatient and Emergency Activities* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 135.

Exhibit 6-12 Automated Multiphasic Health Testing Sequence



Source: Health Resources Administration, *Guidelines for Functional Programming, Equipping, and Designing Hospital Outpatient and Emergency Activities* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 152.

Exhibit 6-13 Client Flow Through Rehabilitation System



Source: G.T. Horton, V.M.E. Carr, and G.J. Corcoran, *Illustrating Services Integration from Categorical Bases,* Human Services Monograph Series of Project Share, No. 3, November 1976, p. 73.

components of intake, evaluation, training, and job placement. Each of the boxes under each of the four components could, and probably does, need a separate flow chart of its own. Each box has its own type of unit and is related to the others, yet each unit preforms a task that is different enough to require a separate method of thinking about implementation. Furthermore, each unit can have interactions with organizations outside the immediate facility, clinic, or hospital. Therefore, using a flow chart to depict an entire health care system including the public sector, the private sector, the volunteers, the social services, and all the actors in the system becomes much more complicated. In addition, many of the variables are no longer under the control or influence of the planner at the hospital or at any single agency. Nevertheless, it might be appropriate for a task force group of concerned planners to meet and work on the implementation process together. Then the same procedures utilized with the simpler flow charts could be used to everybody's advantage.

Another flow chart (Exhibit 6-14) shows the medical care appraisal model used in the state of Wisconsin and indicates the links to the PSRO and to the various committees.

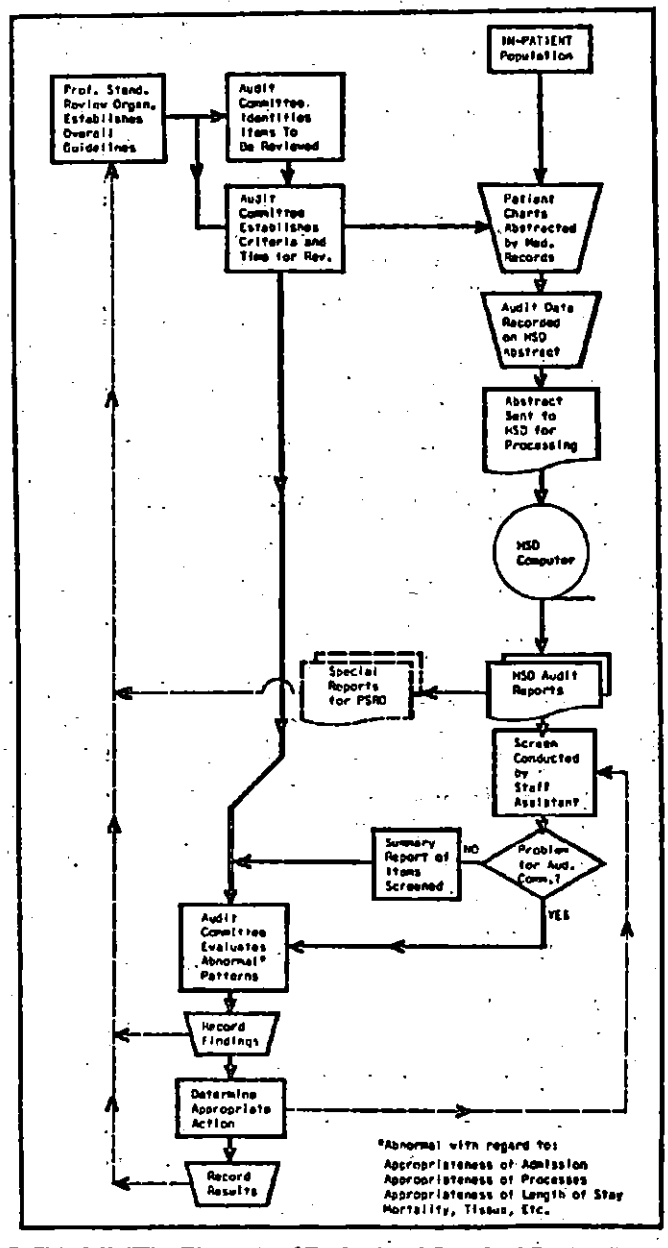
A flow chart for optimum office efficiency has been developed by Hilton P. Terrell of Anderson, South Carolina.[18] Terrell, with tongue in cheek, diagrams the path of patients through a family practice center with attention to the efficiency pathway option, a few decision trees, and a time factor. Exhibit 6-15 shows this effort.

Going from a concrete example with a humorous twist to the generic abstract, Exhibit 6-16 illustrates a problem-solving technique flow chart. Five steps are covered in the flow chart: problem specification, fact finding, goals and criteria for solutions, evaluation of solutions, and implementation. Note the use of the yes/no decision and the resultant flow, as well as the use of the dead-end flow that occurs when parts of the process are rejected.

REFERRAL NETWORK DIAGRAM

A referral network diagram identifies all the organizations involved with a specific problem and diagrams the referral relations between them. In addition, the services provided by each of the agencies in what has now become a health care system are identified. For implementation purposes, this type of referral network diagram is important because all the actors and their services have to be integrated for a program to be sucessful. Furthermore, the points at which the client can enter this health care system are identified.

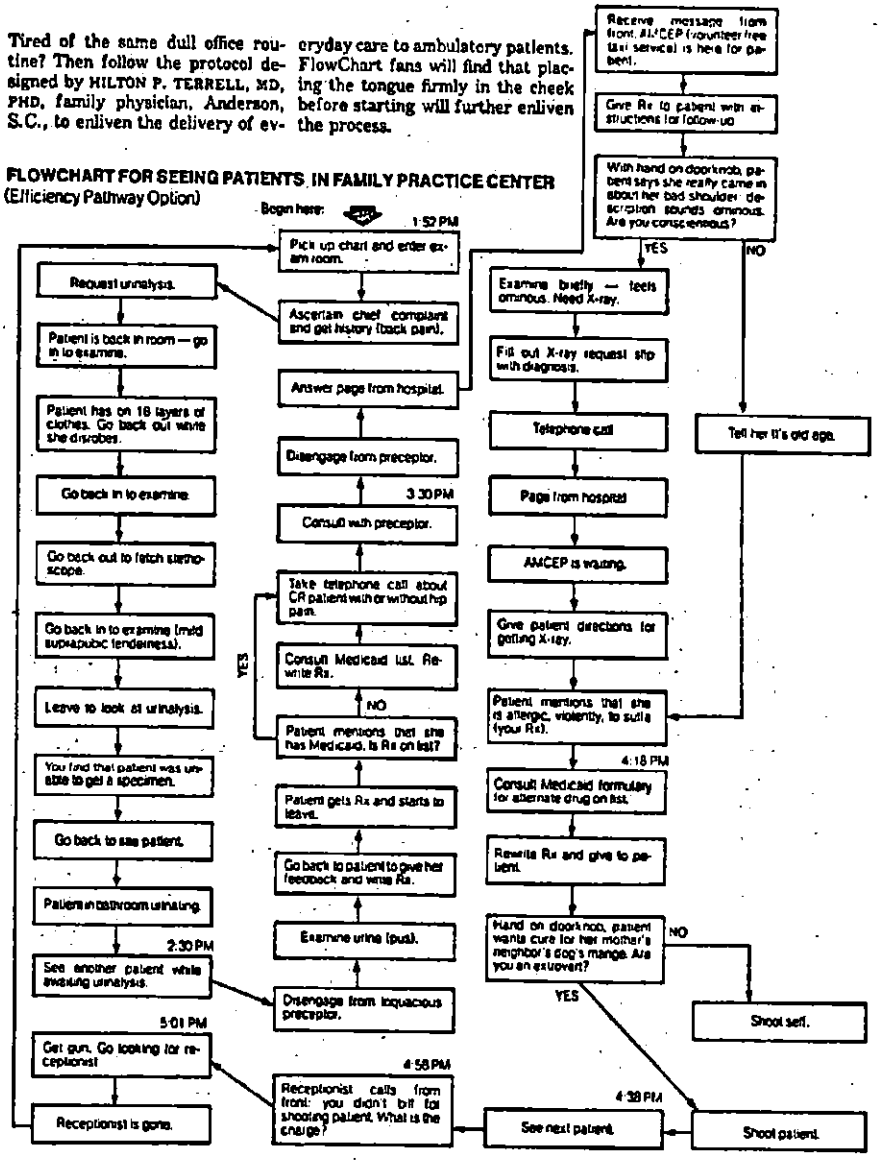
Exhibit 6-14 Health Service Data of Wisconsin Medical Care Appraisal Model



Source: S. Shindell, "The Elements of Professional Standard Review," *Journal of Legal Medicine* 1 (September-October 1973): 37.

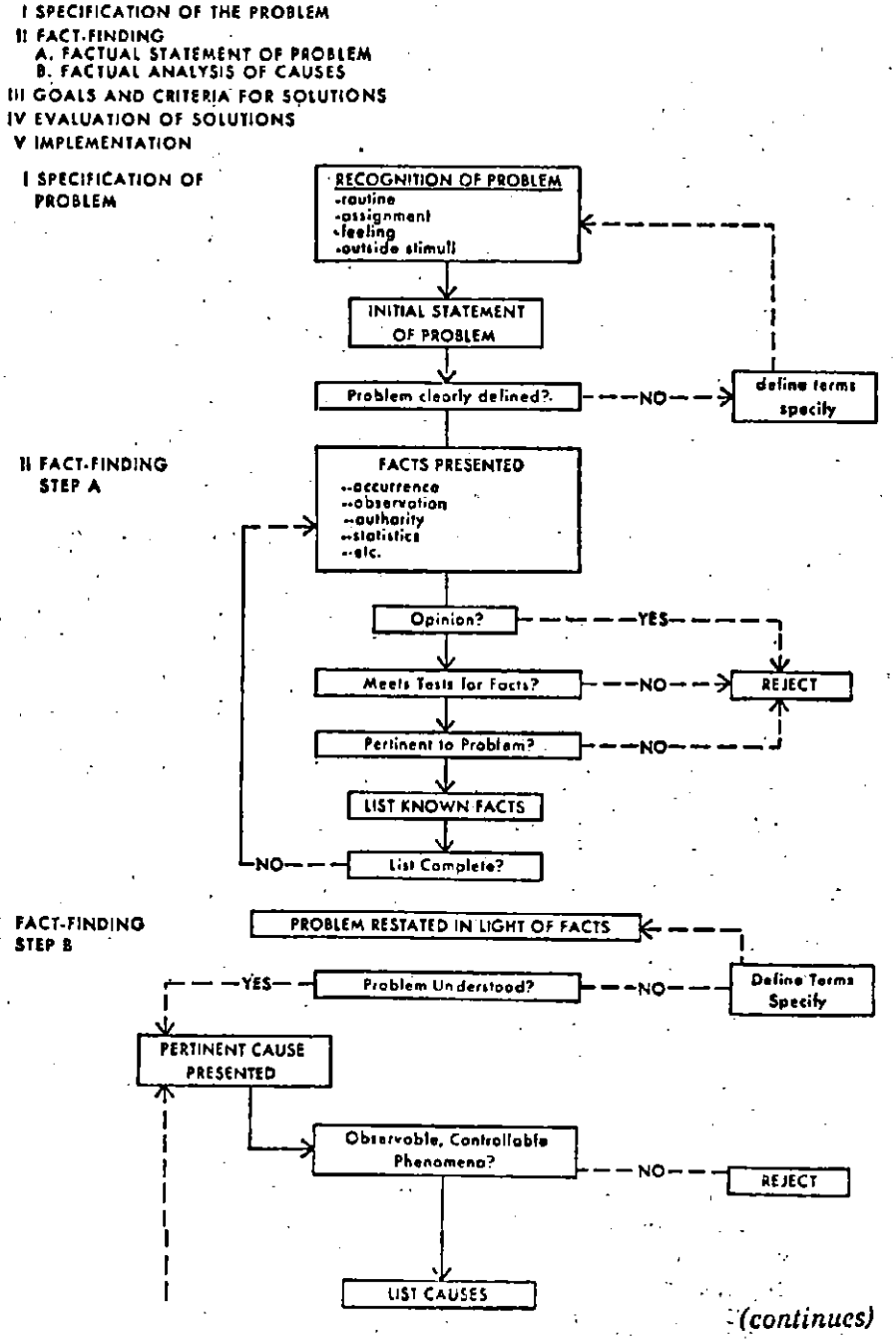
Exhibit 6-15 Flow Chart for Optimum Office Efficiency

Tired of the same dull office routine? Then follow the protocol designed by HILTON P. TERRELL, MD, PHD, family physician, Anderson, S.C., to enliven the delivery of everyday care to ambulatory patients. FlowChart fans will find that placing the tongue firmly in the cheek before starting will further enliven the process.



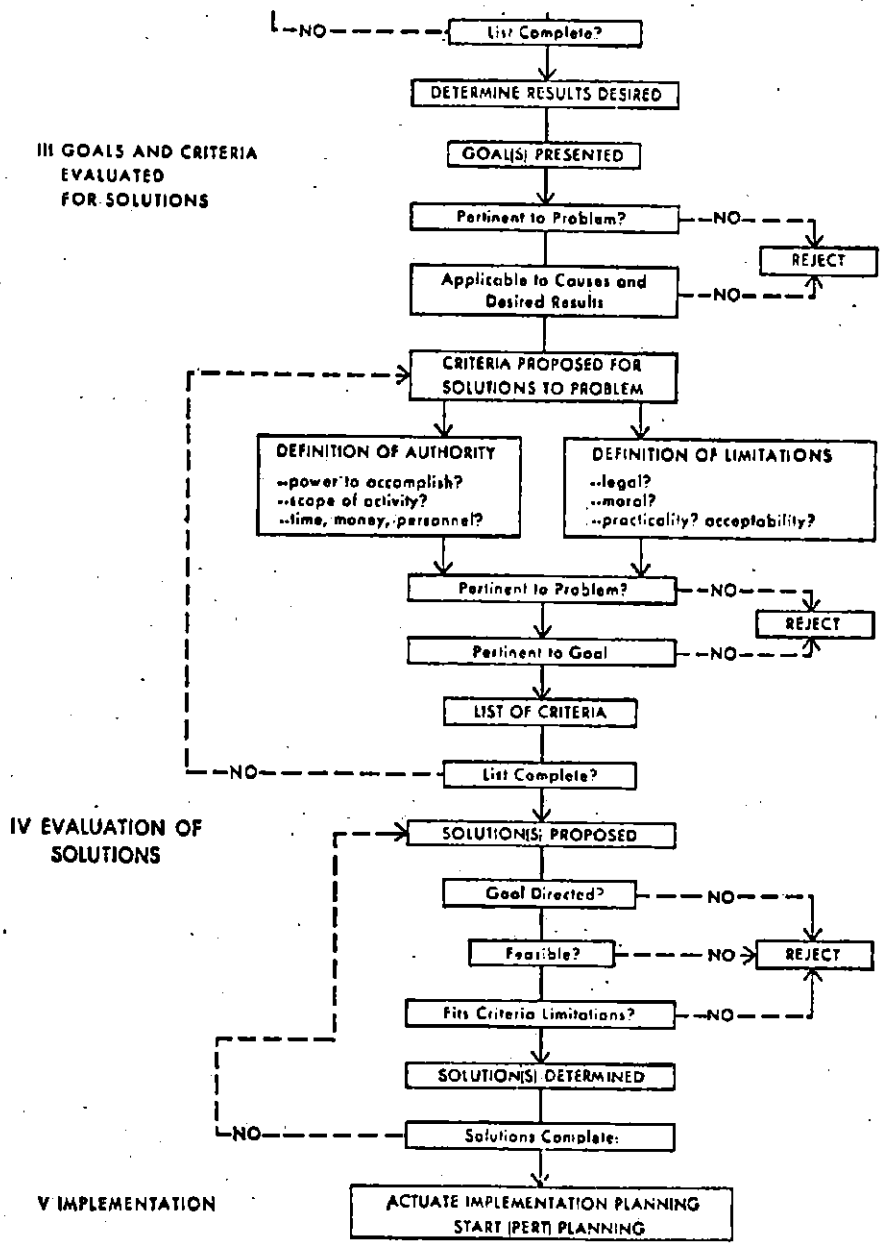
Source: H.P. Terrell, "Flow Chart for Optimum Office Efficiency." From PATIENT CARE, July 15, 1977. © 1977, Miller and Fink Corp., Darien, Conn. All Rights Reserved.

Exhibit 6-16 Problem-Solving Technique Flow Chart



(continues)

Exhibit 6-16 Continued



Source: M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966), pp. 28–29. Reprinted with permission. © 1966, American Public Health Association.

Consider Exhibit 6-17, a referral network dealing with illegitimate births. Note the two-way referrals in this system as well as the points of entry such as schools, welfare or health departments, and private physicians. Also note the criminal element in this system. (This does not mean that the health planner has to deal with the criminal element, but that the planner should be aware of its existence and possible effect on the implementation program being considered.) Services provided by each of the organizations are listed.

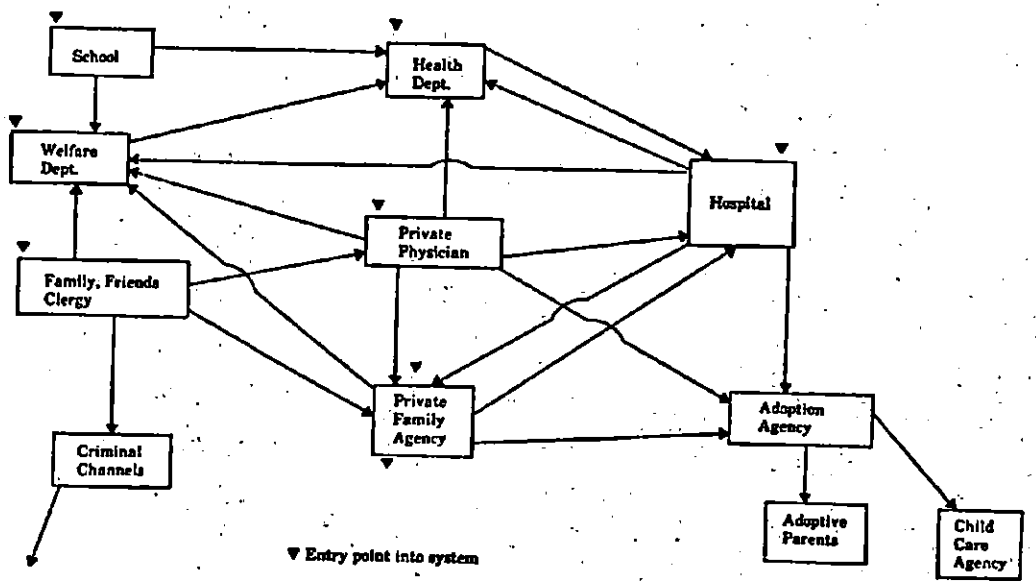
With the referral diagram, the points of entry into the system, and the services of each agency identified, the planner can consider the problem of implementing a program to improve health care for the newborn illegitimate child. Private physicians might be the focal point, except that the woman with an illegitimate child might not seek care early enough. Health and welfare departments also might be late entrants into the care picture. Schools could implement an educational campaign, but that would require postponed action since the girls might be too young or not ready to take action until they became pregnant. In this case, the focal point might be the hospital. True, the hospital would not see the women until they were ready to give birth, but this could still be the point at which all the other referral units cross and a logical program could be implemented. The implementation activity would have to deal with the newborn after birth and hopefully with the future health of mother and child. However, the hospital could interact with the other organizations and take the lead in initiating a program that would touch the entire health care system for this particular problem.

In utilizing a referral network approach, the planner secures more assistance toward determining the focus of an implementation effort rather than the actual details of implementation itself. Yet, this is valuable since a program that is launched from the wrong base might never succeed.

DECISION NETWORK

The decision network approach considers that implementation must involve the examination of all the decisions related to a specific project or program. A decision chart, logical flow diagram, branched logic chart, protocol, decision tree, or algorithm is diagramed in flow chart style with a beginning and an ending point within which is encompassed a complete action or sets of action where decisions are made. At each decision point, a "binary logic" form is used to answer questions with a

Exhibit 6-17 Referral Network — Illegitimate Births



Services

Health Dept.	– prenatal care, well baby clinic, education, disease control, birth & death certificates
Welfare	– income, adoption, guardianship, medical care, counselling
Hospitals	– prenatal care, medical care, well baby care, counselling
Physician	– prenatal care, medical care, counselling, adoption
Private Agency	– income, unwed mothers home, counselling, adoption, medical care
Schools	– referral source, education
Criminal	– income, medical care, baby placement, homes

Source: *Health Planning and Public Accountability Workbook* (Germantown, Md.: Aspen Systems Corporation, 1976), p. 270. Reprinted with permission.

yes or no, present-absent, more than-less than or another alternative to guide the patient's path through the clinic, service, or office.

Problem-action guidelines[19] for a physician's assistant illustrate a decision flow chart network. These decision networks help to accomplish the following:

- classify the problems presented by patients to the health care facility;
- logically display important questions and examination steps to be followed by the health worker;
- isolate the most appropriate actions and treatments for the problem, using a limited set of available drug supplies;
- encourage referral of a difficult problem to the appropriate workers or physicians; and
- provide a basis for supportive supervision and inservice training.

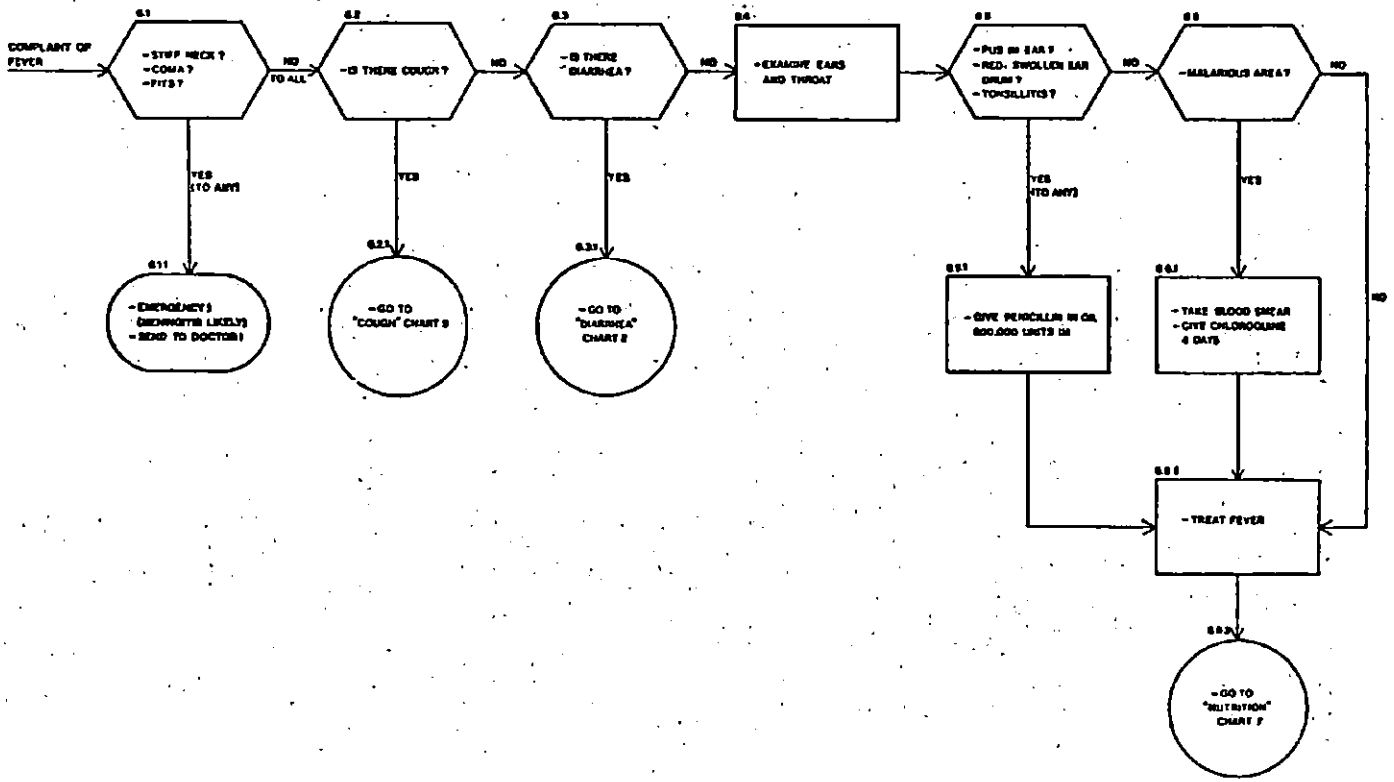
A problem action guide for a patient coming with a complaint of fever is shown in Exhibit 6-18. Note the yes/no decision at each step as well as the referral to other guides when required. At the first decision, a yes response leads to an emergency referral to the physician (shaded elipse). Actions for the physician's assistant to take are indicated in the rectangular boxes, and the arrows indicate the movement of the patient (with the time after the action noted). Six-sided boxes contain questions to ask the patient. Circles contain final instructions and directions to proceed elsewhere.

What does a decision network mean for implementation purposes? Referring to Exhibit 6-18, the fever decision network, implementation of this type of service means including attention to the following:

- having physicians available for referral of emergencies;
- having the other decision referral charts available for the health workers;
- providing examination rooms for ear and throat examinations along with the required equipment;
- having penicillin available with equipment for injections;
- having chloroquine available and packing materials;
- providing materials to do blood smears; and
- providing materials to treat fevers.

These aspects of the implementation relate only to the specific actions noted on the decision network. However, further consideration should raise other implementation activities, such as:

Exhibit 6-18 Fever Flow Chart



Source: Reprinted with permission of Management Sciences for Health, a nonprofit foundation, from N. Hirschhorn, J. Lamstein and R. O'Connor, *Problem-Action Guidelines for Basic Health Care,* 1974.

- charts required for medical records,
- additional personnel needed for laboratory analysis,
- how to store drugs and equipment,
- cleaning up the facilities,
- what to do about those allergic to penicillin, and
- how to assure inpatient admission for emergencies.

These implementation problems are merely illustrative; they indicate how a decision network chart allows for the consideration of a strategy for initiating the activity. The referral charts should be consulted to see if they raise additional elements to be integrated into any implementation program. The cough referral chart is shown in Exhibit 6-19. Immediately it can be noted that there are additional laboratory tests, medications, another emergency situation, and referrals to additional decision chart networks.

Utilizing a decision network means that the logical flow must be internally consistent, completely unambiguous, and the major decisions should be diagramed with all the yes/no decisions brought to a logical conclusion. All the alternatives must be plotted as well to create a comprehensive treatment diagram. Once this is accomplished, the planner can proceed to prepare the resources, money, manpower, and materials required to implement the diagramed project.

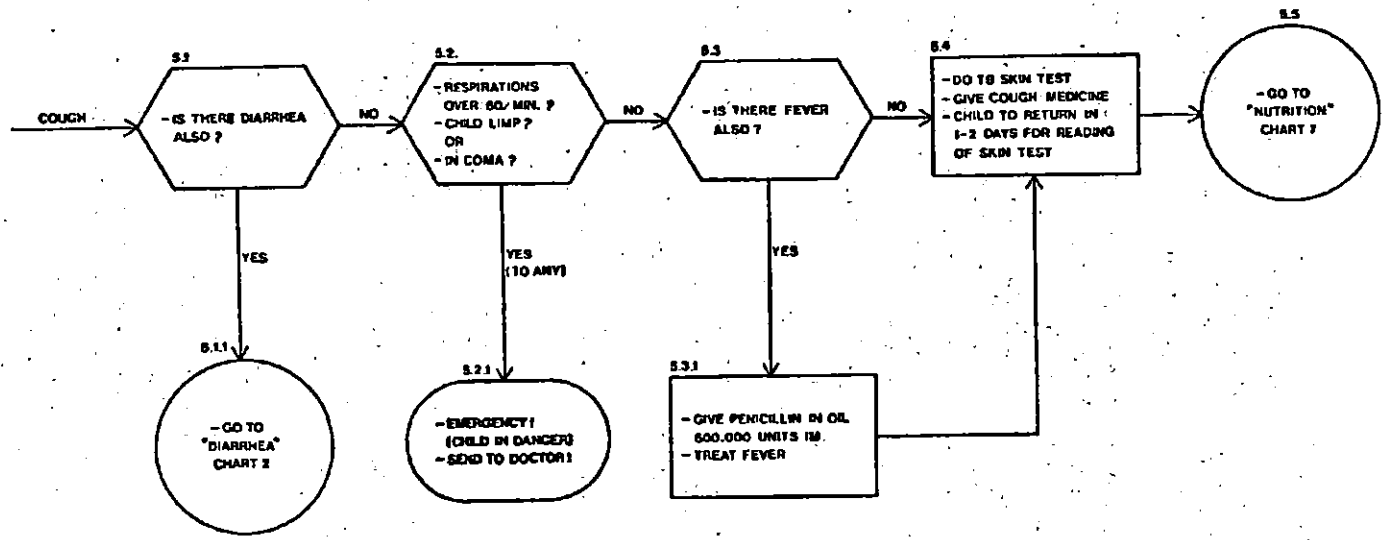
A PLANNING AND DEVELOPMENT FRAMEWORK (PADFRAME)

Although PADFRAME was designed for classifying organized medicine, the methodology can be applied to any existing group. PADFRAME[20] is designed to help organize, to clarify purpose, to set future directions, and to delineate specific activities, programs, and services. A framework and a process for implementation are provided by three main determinants:

- identification of segments of the health system,
- listing of program characteristics, and
- specification of implementation approaches.

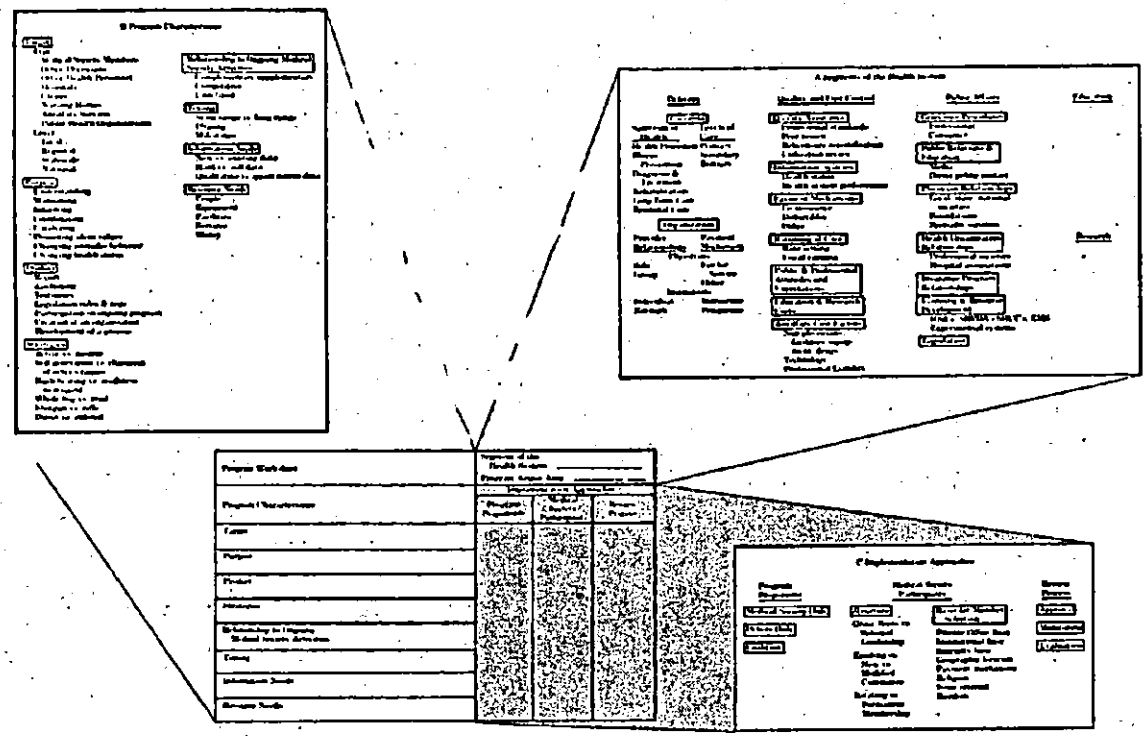
Segments of the health care system have been cast in numerous classifications for a variety of reasons. In this case, the planning organization can choose one that suits it best, as illustrated in Exhibit 6-20. What this identification of the segments of the health system does is merely indicate the possible areas of interest for future im-

Exhibit 6-19 Cough Referral Flow Chart



Source: Reprinted with permission of Management Sciences for Health, a nonprofit foundation, from N. Hirschhorn, J. Lamstein and R. O'Connor, *Problem-Action Guidelines for Basic Health Care*, 1974.

Exhibit 6-20 Program Development Worksheet and PADFRAME Dimensions. A, segments of health system; B, program characteristics; C, implementation approaches.



Source: G. Stuehler, Jr., "Organizing Organized Medicine Through a Planning and Development Framework (PADFRAME)," *Journal of the American Medical Association* 237 (1977):1591. Reprinted with permission. © 1977, American Medical Association.

plementation activities. Note that there are major categories with subsets and further divisions.

Program characteristics take a more direct approach to the implementation of activities. Exhibit 6-20 illustrates the type of information on a listing of program characteristics. Note that a target audience is listed along with the geographic level at which the program takes place. Under the heading of purpose, a generalized concept of the goal or objective can be noted. In addition, this listing calls attention to timing, resource needs, information needs, relationships, the product, and the strategies that might be used in program implementation.

Six strategies are noted, and each deserves some explanation. Each strategy is cited as the end point of a continuum, but there are, of course, shades in between that can take place.

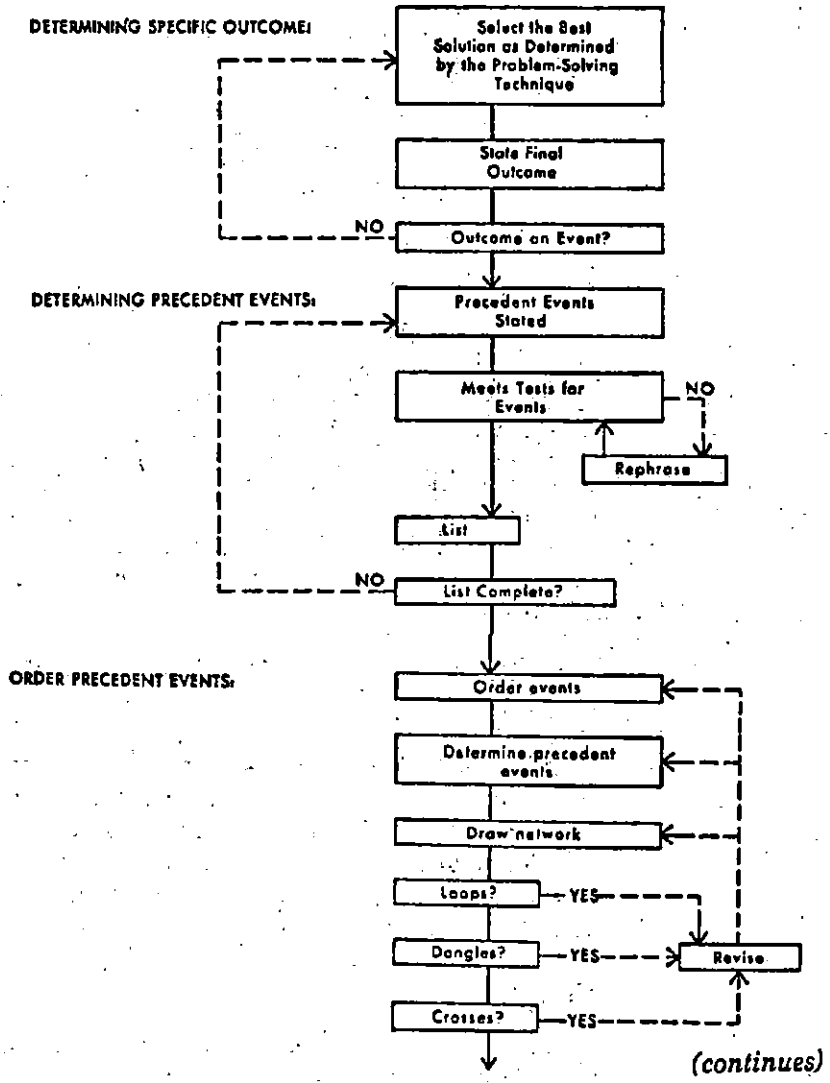
- An *active* program means that the organization initiates and continues the program; a *passive* activity would generally respond to outside controls and influences.
- A *self-generated* activity would be conceptualized by the organization, led by it, and publicly identified with the agency. A *champion of other's causes* means that the organization supports the efforts of others, either publicly or privately.
- A *bush-beating* program tries to drum up interest and concern for the issue the agency wishes to address; a *readiness to respond* strategy requires that data and positions be developed in anticipation of questions or invitations to respond to the issue.
- Taking on the entire program at the begining is a *whole hog* strategy; a *trial* basis allows for taking on a part of a program on a small scale.
- A program directed at a large indiscriminate group or a number of different aspects indicates a *shotgun* approach. A *rifle* approach aims at a highly selective target or at a specific aspect of a program.
- Attacking an activity head-on seeking to affect it *directly* or attacking the problem *indirectly* through a secondary goal that eventually reaches the desired effect are two other strategies.

As noted, any of these strategies could be part of implementation planning; all should be weighed for possible use in a program.

Specification of implementation approaches takes into account the proponents of the program and the participants from both a structural basis and the basis for selection of participants (see Exhibit 6-20). Note the types of structures indicated. In addition, the implementation approaches include the review process, activities of monitoring, approval, and evaluation. In any implementation program, there must be a

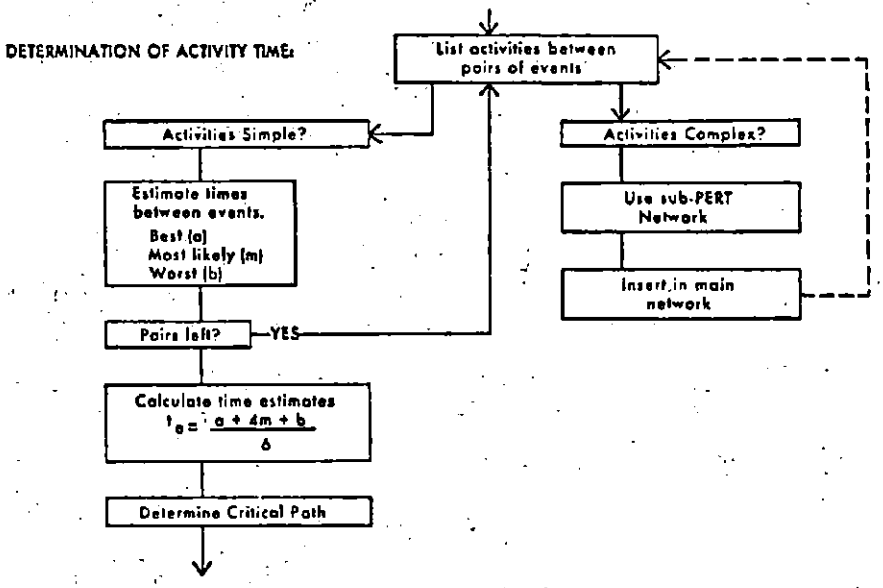
Exhibit 6-21 Program Implementation Through PERT

1. DETERMINING SPECIFIC OUTCOME
2. DETERMINING PRECEDENT EVENTS
3. ORDER PRECEDENT EVENTS
   Include a Network
4. DETERMINATION OF ACTIVITY TIME
   Sub-PERT if necessary
5. ALLOCATION OF PRIORITY EVENTS
   Include Major Computations
6. COMPLETION OF PROGRAM



(continues)

Exhibit 6-21 Continued



Source: M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966), pp. 44-45. Reprinted with permission. © 1966, American Public Health Association.

built-in ability to monitor the process and to make any required adjustments.

In totality, the PADFRAME technique acts as a stimulus to planning and as a creative effort toward the organization of the implementation activity.

PROGRAM EVALUATION AND REVIEW TECHNIQUE (PERT)

PERT is a tool to be used in planning. It is not a panacea or mechanized answer to how to do health planning. Generally, PERT is most useful in planning the implementation phase of a project. In essence, PERT allows the planner to develop a map plotting future activities that result in a network-type diagram so the flow of events can be seen in totality.

According to E.T. Alsaker,[21] PERT's graphic network analysis consists of:

1. the development of a model or activity network of a proposed program or part of a program;
2. the evaluation of the network and adjustment of it to provide a

degree of assurance that there will be a minimum of risk in reaching the objective on time and within the limits of acceptable cost if the implementation PERT plan is followed; and

3. the use of the network to monitor and control the activity represented.

Yet, PERT is a tool that emphasizes the means to an end already agreed upon by the policy makers. Management tools, no matter what they are called, cannot resolve the intricate questions that relate to political-value problems and decisions. There will always be executives who review the results produced by sophisticated hardware and then make their own decision based upon their judgment and experience.

However, once the decision is made, tools such as PERT provide the structured framework to produce a logical, ordered plan to carry out the decision. A flow chart showing program implementation through PERT is shown is Exhibit 6-21. This chart will be understood more clearly after checking the PERT terminology list (Exhibit 6-21A); explanation of the loops, dangles, and crosses (Exhibit 6-21B); and the step-by-step development of a PERT network (Exhibit 6-21C). Note that during the step-by-step development, there are additional references to other exhibits that illustrate how the steps are carried out.

In going over the steps in developing a PERT network, it is suggested that readers familiarize themselves with the exhibits and diagrams. This will certainly allow for an easier understanding of the methodology. In fact, as the steps are presented, it might be enlightening to refer back to the exhibit or diagram rather than to continue on through to the end of the procedure.

A PERT network for the development of an open heart surgery unit is illustrated in Exhibit 6-25. Starting with the decision to develop the capability, the PERT diagram follows through to the final event of having the equipment and team tested and ready. Note the critical path and the time estimates.

GAMING SIMULATIONS FOR IMPLEMENTATION

Games and simulations are common everyday situations. Monopoly is a familiar game and a part of the cultural scene, with jokes about owning Boardwalk or Park Place. Similarly, simulations became everyday occurrences as the space program progressed and the television stations broadcast simulations of rocket flights and moon landings.

Exhibit 6-21A PERT Terminology List

Final event or specific outcome—The event that signifies the completion of the program or the attainment of the goal.

Event—A factual statement of completion of an activity or series of activities. Events take no time or action and are not descriptions of an activity in progress. "Hospital clinic opened" is an event; "ordering supplies for clinic" is an activity in progress. Events are shown by numbers in network diagrams.

Activity—Actions between events that determine the time, effort, and personnel necessary to move from one event to the next. Represented on network diagram by the arrows between events.

Precedent event—An event that comes immediately before the event in question. All events but the first in a program have one or more precedent events.

Successor event—The event that comes immediately after the event in question. All events but the last in the program (final event) have one or more successor events.

Sequence number—The position of each event in the entire program and its relation to the other events.

Network—A system of diagramming the flow of events from the begining event to the final events. A network diagram consists of (1) sequence numbers indicating the sequence of events; (2) arrows designating activities required before proceeding to the next event; and (3) parallel paths representing a series of events taking place during a specific time sequence and an independent series of activities necessary in relation to the attached events.

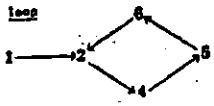
Critical path—Series of activities and events that require the most time.

Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association), pp. 36-37. © 1966, American Public Health Association.

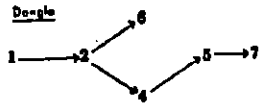
In combining games with simulations, the planner attempts to duplicate the real world and yet maintain the attraction of a game. Through the use of a gaming simulation technique, the health planner can test out a variety of implementation ideas and learn from a miniature dry run of the proposed activity. A number of alternative implementation strategies can be tested as well as a number of suggested implementation activities. With the knowledge acquired from the gaming simulation, the planner can then proceed to set the implementation process into the real situation. Gaming simulations can reveal the bugs and quirks of the proposed activity and allow the planner to

Exhibit 6-21B Loops, Dangles, and Crosses

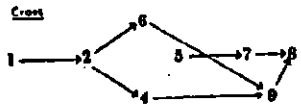
Three problems may occur in plotting the network diagram: loops, dangles, and crosses.



Events have not been clearly ordered. Successor events must be done before a precedent event can be done. The solution for a loop is to reorder the events in a more logical sequence with better determination of preceding events.



There is no successor event for event number 6. Therefore, event number 6 does not attach itself to the rest of the network diagram. As 6 stands now, it is not required for the completion of the program. Solution is to reevaluate the preceding events to locate a successor event. If 6 is not needed, it should be dropped.



This may be more mechanical than a problem with the program. The network diagram should be redrawn to eliminate the cross.

Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966) pp. 36-37. © 1966, American Public Health Association.

Exhibit 6-21C Steps in Developing a PERT Network

1. Specification of final event	A final event indicates the conclusion of the activity. No more action should take place after the final event. Furthermore, the final event should be tangible and physical, such as "diabetes clinic opened" or "immunizations given to 75 percent of the children."
2. Determination of precedent events	Working backward from the final event, simply list all the other events that have to take place prior to the final event. Do not worry about the order of events now but just list every activity that must take place to lead up to the final event. In the diabetes clinic example, events might include hiring staff, renovating the facility, securing a budget, training staff, publicity and information activities, and printing forms. Activities may be added by experts or consultants as the list is circulated for comments and suggestions.
3. Ordering precedent events and diagraming the network	Take the list of events prepared in step 2 and start to arrange the events in a logical sequence starting with the final event (determined in step 1) and work backward to the beginning event, which will be step 1. As you rewrite the events, to the left of each event write the number indicating the order in which these events must occur. Now you have a sequence listing of the events to the left of the events. To the right of each event write the number of the event or events that must come immediately before the event in question if that event is to take place. At this point, consider only the event immediately preceding the specific event under question (see the example in Exhibit 6-22). Note that there is some interdependence relative to several events such as # 10, 13, 24, and 25. Taking

Exhibit 6-21C Continued

	the sequenced events, a network of flow diagram can be prepared (see Exhibit 6-23).
4. Determination of activity time	This step is concerned with how much time it will take to move from one event to another (Exhibit 6-24). Here it is vital to have expert knowledge about all the activities required to reach the next event. Since there will be differences between the planners as to the amount of time required to perform the activities, a mathematical formula is used as follows: $t_e = \frac{a + 4m + b}{6}$, where t_e is time estimate; m is normal time, a is minimum time, and b is maximum time.
5. Determine critical path	By selecting the priority events, the critical path of the PERT network can be determined. This is usually marked by a heavier line tracing the events from initiation to the final event. Without the events taking place in the critical path as noted, the final event cannot take place.
6. Complete program	Initiate the activities, follow the network diagram, monitor the events, make adjustments as required, and reach the final event.

Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966) pp. 36–37. © 1966, American Public Health Association.

prepare resolutions for the difficulties. However, gaming simulations used by health planners differ radically from those played solely with a computer. Most often, the health planners' gaming simulation will use human participants; therefore, the technique is open to all man's frailties. Despite this constraint, the gaming simulation can be a valuable technique in planning for implementation.

Gaming simulations come in a variety of shapes and sizes with a broad spectrum of purposes. However, a number of common structural

Exhibit 6-22 Ordering Precedent Events

Program: Augment DPT-Polio immunization services to lower socioeconomic groups in the county.

Final Event: 50 percent of the preschool children in lower socioeconomic groups in the county immunized against DPT-Polio.

Sequence No.	Event	Preceding Event
27	Final event (outcome)	25, 26
26	Vaccine made available to private physicians	17
25	Outside clinic sites opened	18, 19, 20, 24
24	Health dept. hours expanded	22, 23
23	Medical supplies obtained	14
22	Home visits organized	21
21	Mobile teams established (Dr., nurse, health educator, clerk)	16
20	Volunteers recruited and oriented	13
19	Support of community organizations obtained	13
18	Mass media support obtained	13
17	Support of private physicians secured	13
16	Educational materials produced	15
15	Priorities established for previously identified target groups	13
14	General departments oriented	13
13	Program personnel trained	11, 12
12	Office supplies and equipment obtained	10
11	Personnel recruited	10
10	Board of supervisors approval secured	9
9	County budget submitted	8
8	County budget prepared	7
7	County budget discussed with county administration	6
6	Project application approved by state health department	5

Exhibit 6-22 Continued

5	Application for funds submitted	4
4	Application for funds prepared	3
3	Plan developed	2
2	Decision made to apply for federal grant	1
1	Meeting held to consider program	0

Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966) pp. 36–37. © 1966, American Public Health Association.

components can be found in most that are a central part of the technique. Key components are emphasized in the following requirements:

- participants cast in *roles*;
- *interactions* between those participants in their roles;
- *rules* governing the interaction between the participants while playing their roles;
- *goal(s)* with respect to which the interactions occur; and
- a *criterion* for determining the attainment of the goal, which also indicates the end of the game or simulation.

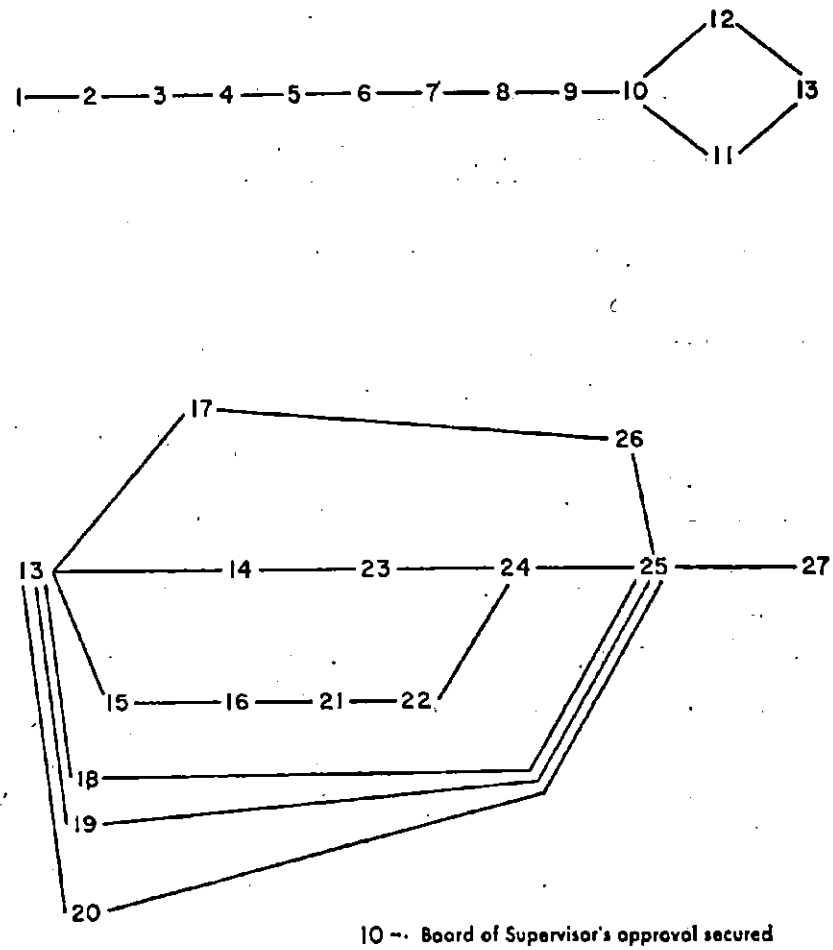
A game has also been defined as any contest (play) among adversaries (players) operating under constraints (rules) for an objective (winning). A simulation is a representation of reality in which participants assume roles that are assigned to them, and through a series of scheduled meetings they attempt to define the nature and extent of the problem and the goals of the community for dealing with that problem. Both the requirements and the definitions embrace the same principles, which require vital elements in the system to be identified and created in a manner that represents a simplified slice of reality.

Roles refer to the identities the participants are asked to assume for the duration of the gaming simulation. Planners must provide the participants with enough of a description of the role that the player can really assume the part. Players must be cautioned to maintain their role at all times and not revert to their real-life situation. Exhibit 6-26A depicts the role of a blood banker in the hemophilia Blood Money game[22] (Exhibit 6-26). Appendix B details a community drug education game; the role of a psychiatrist is outlined for a participant.

Interactions are the vital part of the gaming simulation because the planner is looking to learn how the real world participants will react in

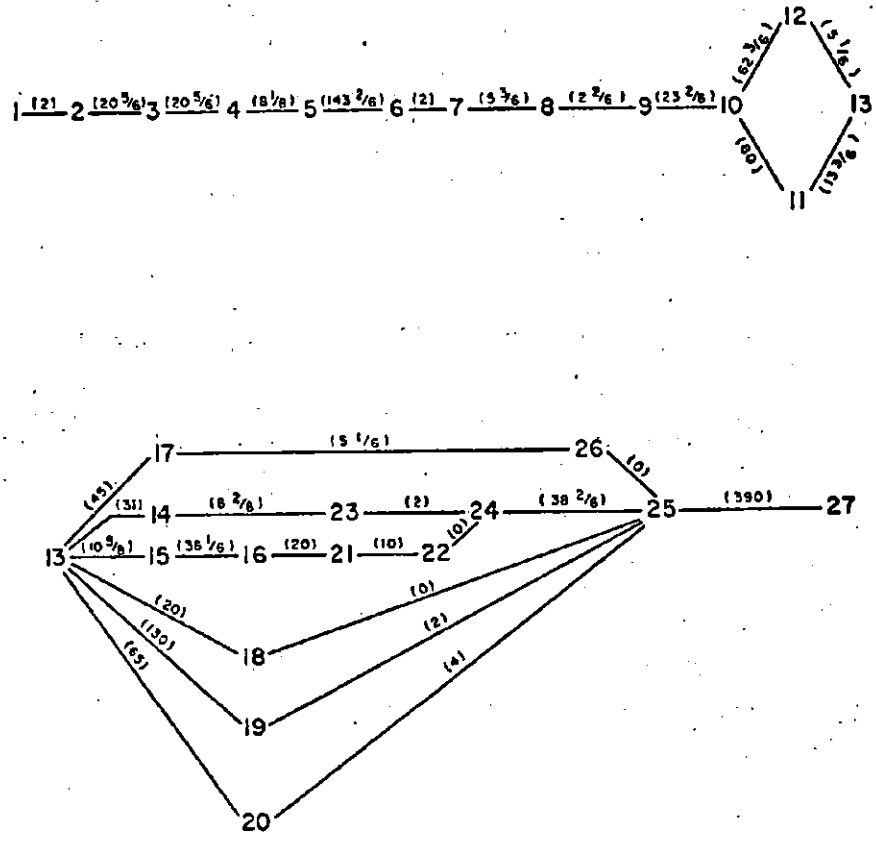
(Text continues on page 298.)

Exhibit 6-23 Diagramming the PERT Network



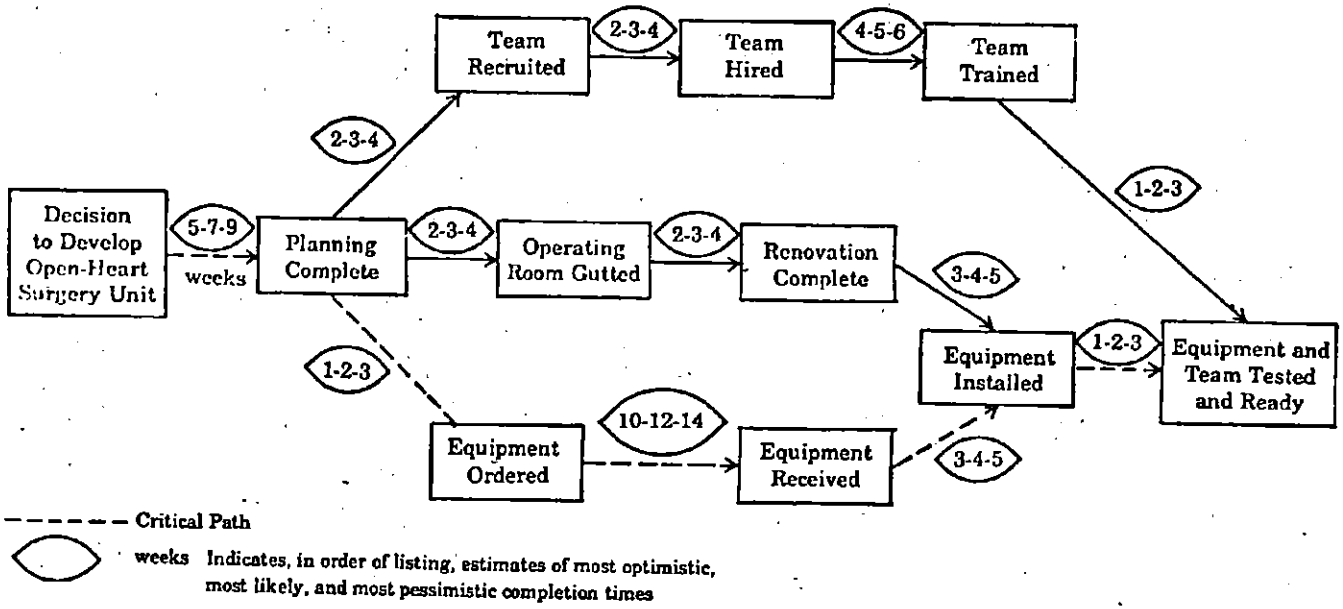
Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT (San Francisco: American Public Health Association, 1966), p. 38.*
© 1966, American Public Health Association.

Exhibit 6-24 PERT Network for Immunization Program with Activity Time Noted



Source: Reprinted with permission from M.F. Arnold, ed., *Health Program Implementation Through PERT* (San Francisco: American Public Health Association, 1966), p. 43. © 1966, American Public Health Association.

Exhibit 6-25 PERT Network for Development of Open Heart Surgery Unit



Source: B.B. Longest, Jr., *Management Practices for the Health Professional* (Reston, Va.: Reston Publishing Company, 1976), p. 101. Reprinted with permission. © 1976, Reston Publishing Co.

384

Exhibit 6-26 Blood Money Gaming Simulation

BLOOD MONEY is a gaming-simulation, but as is often the case, we shall refer to it in the following pages in simplified terms as a "game." It should be clear, however, that it is a game of a special sort, promising not only the fun of typical game play, but the learning possible from active involvement with a model of a segment of reality. It combines, therefore, the playfulness of a game and the seriousness of daily life. We attempt to use the former to illuminate the latter.

B. What's in This Game: Subject and Purpose

BLOOD MONEY simulates on an abstract level the general character of the experience of the hemophiliac and those who provide him with care and blood. It is designed to convey a general understanding of issues and problems in hemophilia care so players can more fully comprehend the worlds of the patient and of those working in the health care sector when these are described in subsequent talks or readings.

In some ways, play of this game, like play of others, is fun. In other ways BLOOD MONEY is not as enjoyable, since it models or simulates experiences that are frustrating and difficult for those who live them in real life. In any case, it has proven an interesting and valuable learning experience for those who have participated, and we are sure you will find it so as well.

C. Who Can Play?

While the specific disease around which the game is built is hemophilia, those who have participated in early runs have suggested that there are many general elements of the patient and care systems in the game that pertain to other diseases. Thus it provides a learning vehicle for those whose concern is not with hemophilia per se, but with the nature, operation, and dilemmas of the health care delivery system.

Appropriate audiences of players include:

1. Those who deal with hemophiliacs and desire a greater understanding of their problems and perspectives—e.g.:
   a. medical and paramedical staff;
   b. secretarial and other staff members;
   c. clinic, hospital, and blood bank administrators and staff;

Exhibit 6-26 Continued

   d. parents and other relatives of hemophiliacs; and
   e. employers, teachers, school administrators, etc.
2. Those who will be dealing with patients—i.e., those in training for or anticipating any of the roles above, such as medical and nursing students.
3. Others interested in understanding hemophilia—e.g.:
   a. blood donors or potential donors, and
   b. other citizens.
4. Students of health—i.e., those interested in problems of the delivery of health care services or elements of the "sick role:"
   a. upper elementary school, junior high school, and senior high school students studying these questions;
   b. college and graduate school students in sociology, psychology, health, etc., courses; and
   c. psychologists and psychiatrists interested in the relationships between physical and mental illness.
5. Policy makers:
   a. local officials,
   b. state and national health care administrators, and
   c. state and national representatives and administrators charged with the development or passage of health care legislation.

The game remains exactly the same for play with these disparate groups. Differences in the experience and knowledge of the participants is brought to bear in the differential character of the postplay discussion-critique and in the follow-up materials presented. It should be noted that the game is *not* designed to be played by persons with hemophilia, although we have, on occasion, invited them to participate by playing the roles of those in charge of delivery of care, while "real life" doctors and others in the delivery system played patients.

D. Number of Players

The game is played optimally with a group of 24–28. As will be described later, however, groups of 20–23 or of 29–35 can be accommodated.

E. Running the Game—Personnel Needed

Two operators are required to run the game. Both should be familiar with it from a thorough reading of this manual. It is strongly recommended that an assistant be recruited to take over some of the respon-

Exhibit 6-26 Continued

sibilities of Operator No. 2. This assistant need not understand the game in its entirety nor be aware of the underlying assumptions of the model. He or she can be "trained" in five minutes, and thus one of the players can be recruited to play this role if numbers permit.

F. Time Requirement

The context of use will vary depending upon the audience (players). For students, it will comprise part of a larger educational program, and be run within normal or extended class periods. For others, it will most often be used as a self-contained program of approximately two and a half hours duration.

The following indicates approximate times needed for play:

1. Introduction: 15 minutes;
2. Play of two rounds: 1 hour 30 minutes;
3. Discussion: minimum 30 minutes;
4. Follow up: as much time as desired, or via printed material that can be taken home and read.

Although it would be best to include all steps in the same time period, it is possible to schedule steps 1 and 2 for one day (one class) and hold the discussion and follow-up soon thereafter (the next class).

G. Space Requirement

One large room is required for play of the game. The furniture must be movable. A number of chairs and approximately seven small tables are needed. A chart for room setup is provided later in the manual.

H. Materials Requirement

The game has been constructed to consist of standard materials that can be easily located and cheaply purchased. All paper materials needed (e.g., role descriptions and charts) are provided in this manual. The number of copies of each page you will need to photocopy is indicated at the top of the page. All purchased materials and most of the paper materials are reusable. Players may, however, occasionally write on the paper materials and you may have to reproduce more for subsequent runs of the game.

Source: C.S. Greenblat and J.H. Gagnon, *Blood Money: A Gaming-Simulation of the Problems of Hemophilia and Health Care Delivery Systems* (Washington, D.C.: U.S. Government Printing Office, 1976).

Exhibit 6-26A Role Sheet

BLOOD BANKERS

Goals: You are in charge of the very limited supply of available blood and are responsible for disseminating it. The blood is represented by *red* chips. You will get new supplies from Operator #2 at the beginning of each round, but the amount of chips will vary depending on the number of donors to the system as a whole (beyond your control) and on your efficiency in dissemination and collection of payment. Inefficiency will lead to cuts in supplies.

You operate one blood bank together and should make joint decisions and divide the work. At least one of you should be present at the bank at all times. Your joint goal is to operate the bank as well as possible to serve the community.

Charges: Yours is a nonprofit operation. Although the blood has been donated, there have been considerable administrative costs incurred in its collection, analysis, storage, treatment, etc., resulting in the necessity to charge for each unit. The basic (and minimum) charge is 2 white chips for each unit of blood (each red chip). The check on your efficiency will be to ascertain whether you have collected twice as many white chips as the number of red chips you have distributed, or have arrangements to collect them soon. You will have to set up your own accounting system.

Welfare agents may ask for proof that their clients have received blood.

Dissemination: Patients should be given blood *only* upon presentation of a prescription from a doctor. These prescriptions should be on yellow papers and must be written with green felt markers and signed by the doctor.

If, at any time, you feel you do not have enough blood on hand to fill the prescription that a citizen presents to you, you may give the patient *less* than is prescribed or send the citizen back to the doctor for a reduced prescription. You should understand, however, that the amount of blood prescribed is *critical* to the further health of these citizens, and that they are in need of getting it in the *right amount* and *very soon* after an attack if they are to avoid extensive further disability.

Source: C.S. Greenblat and J.H. Gagnon, *Blood Money: A Gaming-Simulation of the Problems of Hemophilia and Health Care Delivery Systems* (Washington, D.C.: U.S. Government Printing Office, 1976).

Exhibit 6-26B Scenario Material

HEW NEWS

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

PRESS RELEASE
Monday, March 10, 1975

National Institutes of Health
William E. Sanders (301) 496-5343

WASHINGTON, D.C.—Congress today overwhelmingly approved an all-inclusive Catastrophic Disease National Health Insurance Act. In a carefully orchestrated move by Senate and House leaders, the compromise bill, which had gone to a conference committee to resolve differences, was introduced and passed by the two Houses in the same day. Leaders of both Houses stated that the large congressional majority assures the President will sign the new compromise bill into law, or face a certain veto override.

Never again, said the bill's supporters, will Americans be made to suffer financially from debilitating disease. All U.S. citizens and residents will be covered under the Catastrophic Disease Health Insurance program.

NEW PAYROLL TAX

The new program will be financed by a 3.5% tax on employers' payrolls, a 1% tax on employees, and 2.5% on the self-employed up to $20,000, with the remainder coming from general revenues. All monies collected are to be paid into the National Catastrophic Health Insurance Fund. The Fund will be administered by the newly created Catastrophic Disease Health Insurance Board established under the Department of Health, Education, and Welfare.

The Board will be advised by the new Catastrophic Disease Health Insurance Council composed of a representative group of consumers, providers, health organization personnel, and other interested parties.

MEDICARE AND MEDICAID STAY

Both the Medicare and Medicaid programs that provide preventive and catastrophic benefits will remain intact, except that the same benefits established under the Catastrophic Disease National Health Insurance Act will be extended to include these programs. It is likely that Federal and/or State representatives involved in administering these programs will sit on the Catastrophic Health Insurance Council.

Exhibit 6-26B Continued

BENEFITS

The Act provides that the first $500 spent annually on medical care must be borne by the family or individual suffering from the disease. Thereafter, the government will pay for necessary disease-related treatment—all physician care, hospital days, outpatient care, nursing home stays, supportive services, rehabilitation needs, psychiatric care, inpatient and outpatient drugs including blood and blood products, and prosthetic devices. The program does not cover health needs unrelated to the disease.

REIMBURSEMENT

Reimbursement will be similar to the current Medicare program. The physician bills the patient who files a claim for the reasonable cost allowed by Medicare. Under the Catastrophic Disease program, reasonable cost can be determined by the Catastrophic Disease Health Insurance Board in advisement with the Catastrophic Disease Health Insurance Council.

$25 MILLION FOR PLANNING

The actual start-up date for the Catastrophic Disease National Health Insurance program will be January 1, 1977. Until then, the Act calls for the expenditure of $25 million over a two-year period to plan in detail how the program will work. It is hoped that with adequate planning the program will run smoothly once it is implemented.

Instructions on Questions To Be Addressed and Statements To Be Prepared by Players

Keeping in mind the role you are playing, there are three questions you need to consider with respect to the news you have just heard. At the end of 45 minutes, we need from each group a detailed statement about these three questions to be presented to the group as a whole for its discussion and debate. Following that there will be a short period for reformulations in light of what others have said; and, finally, a brief discussion of these reformulations will take place before we move to the next event.

The questions all have reference to the *first year following the event* and are as follows:

Exhibit 6-26B Continued

1. How will this event affect you?
2. What will you do or try to do because of it? (i.e., How will you alter your behavior, operations, etc.?)
3. What should others in the system do about it? (Product, Care, Payment, and Research and Education interests).

Please try to be as specific as you can in responding to these questions. We hope you will work to define the common elements of your responses through discussion in each group, but minority reports on any dimensions are perfectly appropriate where there is dissent.

Let me also note that we are placing an observer at each table to help collect more detailed reports than you will be able to make in the open discussion. These people will not be interfering in the discussions.

SESSION II
April 24, 1975

I. Instructions of Players

Today's events are a continuation of the events that occurred in the game you participated in last month. For those of you who were not here then, you represent yourself and your type of interest or constituency.

To bring you up to date and to remind everyone of what has occurred before, I will read a summary of events that have happened to medical benefits for hemophilia. As you know, on March 10, 1975, Congress passed the all inclusive Catastrophic Disease National Health Insurance Act. The Act provided for full coverage of hemophilia and all hemophilia-related health needs once the patient or his family paid for the first $500 of care. The program financed by a tax on employer's payrolls and a smaller tax on the self-employed and employees, with the remainder coming from general revenues paid into a Trust Fund. It is administrated through the Department of Health, Education, and Welfare by the Catastrophic Disease Health Insurance Board, which is advised by the Catastrophic Disease Health Insurance Council composed of a representative group of consumers, providers, health organization personnel, and other interested parties. The Act was put into effect in January 1977, following a two-year planning period.

Once the program began, many more hemophiliacs sought care than had been expected. In fact, 13,000 hemophilic patients previously unknown to the system came for treatment. Half of these patients were

Exhibit 6-26B Continued

moderate hemophiliacs who wanted more adequate and frequent treatment. The other half were severe hemophiliacs suffering the chronic effects of hemophilia. These severe hemophiliacs as a group were older, poorer, less skilled, and mainly rural people who had severe disabilities and were seeking adequate treatment for the first time.

Both the moderate and the severe hemophiliac stressed the hemophilia care system very directly. The older, severe group required a great deal of orthopedic surgery, dental care, and psychological and vocational rehabilitation. The moderate patients, like the severe patients, required full diagnostic work-ups and, in the beginning, considerable monitoring of their disease.

The influx of these 13,000 additional hemophiliacs into the system was felt by Congress during the third year of the Catastrophic Disease Health Insurance Program. Congress became alarmed over the amount of money that was going to pay for hemophilia treatment in proportion to the rest of its health care dollar. Rather than changing the payment mechanism or controlling the quality of treatment, Congress decided that the government would continue to pay either to the patient or the patient's doctor a fixed upper limit of $5,000 per hemophilic patient in any given year.

II. Presentation of Scenario 2

Now, with this in mind, I will read the first "News Bulletin."

15 HEMOPHILIA CENTERS TO BE ESTABLISHED

Responding to pressure to improve care within the $5,000 limit set for each hemophiliac, Congress has authorized the Department of Health, Education, and Welfare to establish 15 comprehensive care centers across the United States. Each center will offer comprehensive care, have a complete diagnosis unit able to do both diagnostic tests and genetic testing, and will engage in research, education, and training.

Forty-five million dollars has been appropriated by Congress to establish the 15 centers. Each center will cost approximately $3 million to build. Operating expenses will be funded in part by government money—the first $5,000 of care each hemophiliac receives under Catastrophic National Health Insurance. Ten percent of the income obtained by each center must be spent on research. Each center is free

to decide how much of its budget it wishes to allocate for education and training.

Attendance at these centers will be voluntary; neither the patient nor the doctor will have to use the services provided at these centers. The doctor-patient relationship, if it is satisfactory to both parties, would not be disturbed.

Each center will have a patient population of approximately 2,500 hemophiliacs. The 2,500 patient population figure is based on the known number of hemophiliacs seeking care: 38,500. Thus, if 2,500 patients all require the government allotted benefit of $5,000 per hemophiliac, the center's total operating budget would be $12.5 million. For all 15 centers, the budget would be approximately $187.5 million.

The number and location of hemophilia centers in a region were determined on the basis of the known geographical distribution of hemophiliacs. Within certain limits, individual hemophiliacs will be able to attend the facility nearest to them. The 15 cities chosen are Seattle, Washington; Los Angeles, California; Denver, Colorado; Minneapolis, Minnesota; Kansas City, Kansas; Houston, Texas; Chicago, Illinois; Ann Arbor, Michigan; Nashville, Tennessee; Rochester, New York; Pittsburgh, Pennsylvania; Chapel Hill, North Carolina; Atlanta, Georgia; New York City, New York; and Boston, Massachusetts.

---

1. Group Statements by Participants to the Establishment of 15 Regional Centers

a. Care Sector Report

We have problems because, supposedly, we have $15.5 million for each of these hemophilia centers, but we do not really have this money because the patient has the ultimate decision on where to spend his money. This is a great challenge: How can we build these new facilities when these facilities may never be utilized? The patient may choose not to use our services, then where are we? What is required is an ongoing operating budget to allow the center to operate before it has a full complement of patients.

In order to create an operational budget, we recommend initiating a requirement for a yearly comprehensive examination for all hemophiliacs to be given only at the center. This yearly examination would become the qualification for support by the center for that year or this examination report would be sent to a private physician. The

private physician, once he received this report, could treat this hemophilic patient for a year.

Considering the staff and the laboratory facilities that are needed for a thorough evaluation, $500 is not an unreasonable fee. This fee is not excessive once it is realized that the evaluation is being conducted by the best people available. After all, the physician will be writing a clear-cut evaluation of the patient's status and making recommendations for care during the following year. This initial $500 examination is not unprecedented; it is similar in principle to having support coming only through qualified programs.

Now, coming to grips with the actual money to build. The proposed National Hemophilia Commission should be charged with finding sites for these 15 comprehensive centers. These sites should be selected on the basis of patient flow, accessibility, and the availability of expert staff. Selection, moreover, should be on the basis of competition. There may be great competition among those cities that have been chosen, for instance, in New York.

Source: C.S. Greenblat and J.F. Gagnon, *Blood Money: A Gaming-Simulation of the Problems of Hemophilia and Health Care Delivery Systems* (Washington, D.C.: U.S. Government Printing Office, 1975).

the staged situation. By creating the role for each participant, the health planner provides human diversity and, it is hoped, represents all the participants and their viewpoints. By providing a scenario for the players, the health planner describes the overall environment in which the players act out their roles. In the Blood Money gaming simulation the participants use a dart board, barter exchange, and a scenario (Exhibit 6-26B) that employs a news release from the federal government. In the drug education game, a front page story and an editorial from the local newspaper are used to set the stage (Appendix B).

Rules and/or guidelines are usually presented to the players as instructions for the gaming simulation. In some situations, the rules detail what the players are allowed to do, and in other games the rules state what the players cannot do and leave the field open for all other actions. Health planners may wish to use both techniques. To stimulate interactions, the planner may wish to set rules that restrict the players to a specific situation and concentration upon a narrow goal; the broader guidelines could be used for a more open gaming simulation where the health planner may be seeking to identify all the ramifications.

Goals can be stated as objectives in the rules or can be incorporated in the role profiles given to each player. An objective might be to arrange for equitable distribution of supplies of blood to hemophiliacs, while the player's role profile might instruct him to gain control over all the hospital beds in the area. Sometimes the goal is stated in terms of what must be achieved to "win" the game. Exhibit 6-26 explains the subject and purpose of the Blood Money gaming simulation. Appendix B does the same for the drug education game.

Criteria for ending the gaming simulation can be when somebody "wins," or an artificial time limit can be imposed—the game ends with whatever resolution has taken place. Exhibit 6-26 and Appendix B both indicate that time limits are being used.

Gaming simulations are particularly relevant in health planning because of the legal mandates to include consumers and providers in the planning process, with the consumers in the majority. To achieve the desired objectives, health planners must be prepared for the human interactions that will take place regarding the implementation of activities. Therefore, it would be prudent to try to anticipate any possible deviations from the proposed implementation process and to prepare for the barriers that may arise within the professional community as well as among the general public. A gaming simulation can provide the health planner with an imitation of the real world working on the

activity and suggesting solutions and approaches to achieve cooperation and effective participation toward the desired goals.

Illustrative abstracts from the two gaming simulations already mentioned, the drug education simulation game and the Blood Money hemophilia game, will indicate concretely how the use of a gaming simulation fits into the promotion of implementation strategies. To direct attention to pertinent techniques, each abstract will be listed with highlights (Exhibit 6-26C and Appendix B).

IMPLEMENTATION AND SOCIOPOLITICAL INTERACTION

Up to this point consideration of implementation has been confined to the rational, scientific approach to putting decisions into the action phase. However, the real world does not always abide by rational logic, and the health planner must be aware of social and political interactions that influence any implementation program. As previously noted, the gaming simulation can be used as a technique to prepare the planner for the human factors that arise in setting a project into motion.

L. Fontana and A. Beckerman[23] contend that the major interest groups in the health sector are the "professionals," represented most significantly by physicians in group or private practice; the "corporate sector," represented by hospitals and nursing homes, medical schools, insurance companies, drug companies, and public health agencies; and the "consumers." These divisions within the health care area will obviously lead to disagreement when it comes to implementation activities. Yet, the rules of the game are structured to reach a decision that conflicting interest groups can accept without suffering any serious damage.

However, the point is made that the inherent inequality of bargaining power within the interest groups is often not considered in efforts at representation.[24,25] Politicians, for example, are more inclined to respond to the demands of special interest groups rather than the actual needs of the entire population. Carrying this concept further, a study of prior health planning efforts and the structure of the new Health Systems Agencies led to the forecast that the HSAs will have little impact on the health care crisis because the HSAs fail to take into account the dynamics of the health sector. In addition, the researchers believed that "health planning as presently instituted has been in the basic interest of the health providers . . . it is mistaken to believe that health planning is inevitably in the public interest." It was pointed out

Exhibit 6-26C Blood Money and Implementation Activities

What's in this game—subject and purpose	Aims to experience how the hemophiliac gets his care and blood.
Who can play?	Identifies the participants and indicates representativeness.
Number of players	
Running the game—personnel needed	Gives manpower requirements.
Time requirement	Sets sequence of meetings by time periods.
Space requirement	
Materials requirement	
Blood bankers	This is the role profile for the player. Note the goals, the cost charges, and the dissemination of the blood supplies.
Press release from HEW	This is the technique used to set up the scenario in which the game takes place.
Instructions on questions to be addressed and statements to be prepared	This moves the players into a consideration of the implementation of various activities.
Session II—instruction of players	Adds to scenario background for the next session.
Scenario 2	This directs the players into another set of circumstances. The technique of adding additional information to scenarios can be used to redirect the players to focus upon another implementation problem.

Source: C.S. Greenblat and J.H. Gagnon, *Blood Money: A Gaming-Simulation of the Problems of Hemophilia and Health Care Delivery Systems* (Washington, D.C.: U.S. Government Printing Office, 1976).

that the HSA should not be considered a neutral device despite the representation of different segments of the health care sector.[26] A study of health planning in eight model cities projects strongly suggested that providers dominated the decisions in every city.[27] Another analysis[28] concluded that it was unlikely that the HSAs would generate radical changes in the health care system. Rather, the HSA is more likely "to provide an institutional forum for legitimizing existing patterns of power distribution and to accede as slowly as they can to those irresistible exogenous forces that would have produced major change in any event."

To counter this idea, a hospital administrator noted in *Hospitals*, the journal of the American Hospital Association, that planning should address society's needs, and operations should be consonant with the greater social consciousness.[29]

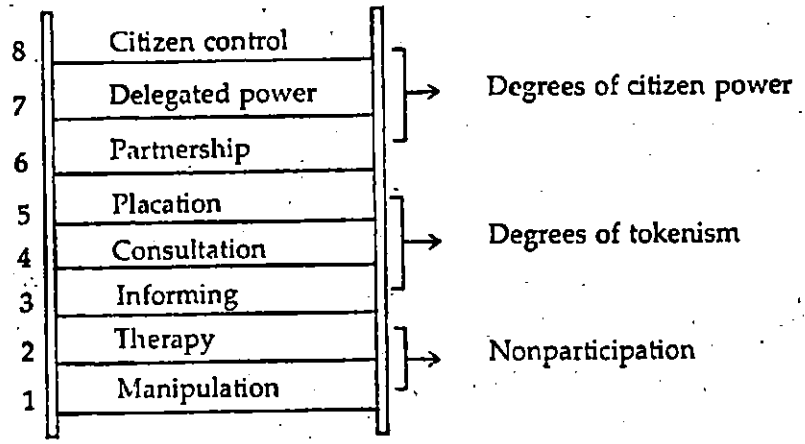
In view of the legislative mandate to have a majority of consumers on the boards of HSAs, the sociopolitical interactions that could affect implementation activities can be increased as the professionals and the corporate sectors work with the nonprofessional consumer. In a report on a training program for consumers in policy-making roles in health care projects,[30] the consumers generally criticized professionals for using complicated, technical language to express simple points. One board member said, "Doctors can rattle on for hours; if you try to ask a simple question, they just take it and go on and on." This type of attitude will seriously affect implementation of a program. However, a basic question raised by professionals concerns whether the public really knows what it wants. Two conflicting points of view have been stated in this respect. One holds that the public, be they consumers, patients, or clients, is unknowing and cannot diagnose their own needs or discriminate among the range of possibilities for meeting them. The other view contends that the consumer is the best judge of his own welfare.[31] In a balance sheet style, the conflict can be presented as shown:

Unknowing	*Best Judge*
Man is ignorant	Man is wise
Protectionism prevails	A free market exists
Man is potentially altruistic	Man is primarily self-interested
Physicians seek out patients	Patients go to the physicians

Fuel for conflicts about implementation activities are evident in this balance sheet of viewpoints.

A ladder of citizen participation developed by S.R. Arnstein[32] illustrates some of the techniques utilized in planning situations, wherein professionals worked with the consumers. Note the choices of style and alternatives offered.

Ladder of Participation

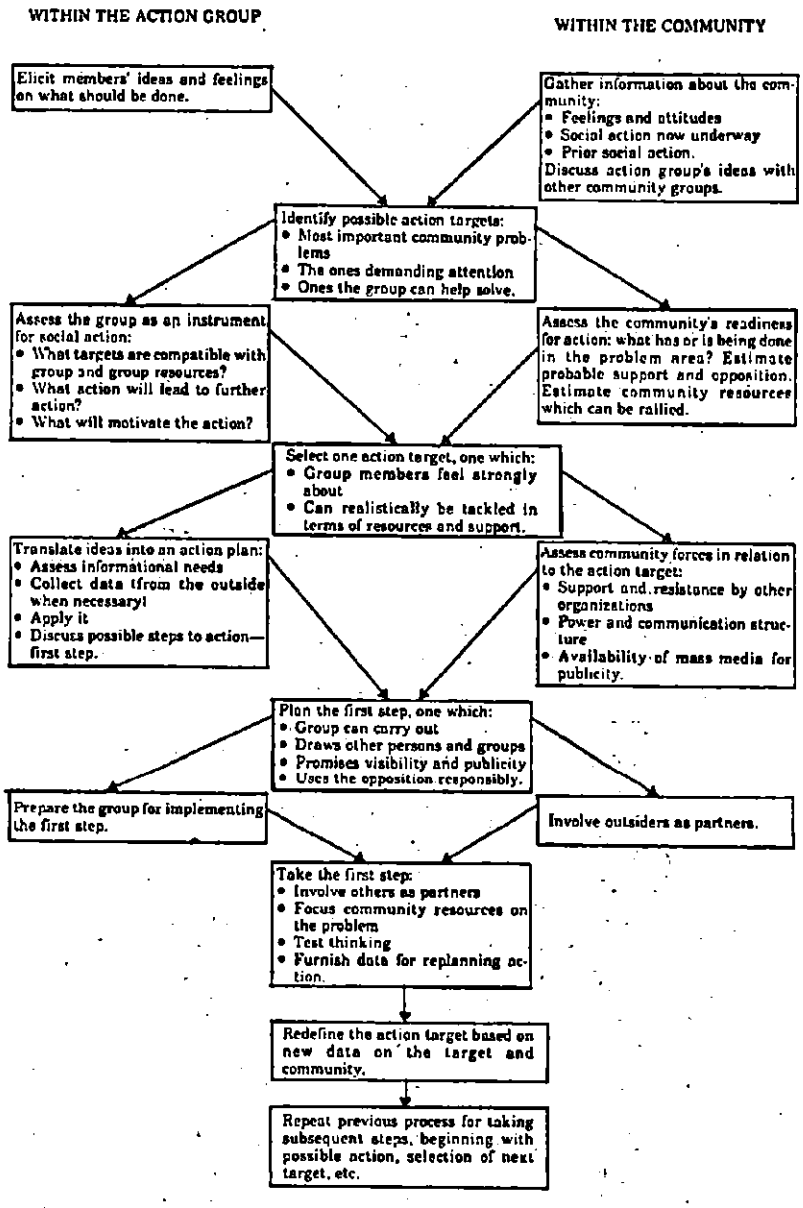


Source: S.R. Arnstein, "A Ladder of Citizen Participation." Reprinted by permission of the *Journal of the American Institute of Planners*, vol. 35, July, 1969.

In planning for the implementation of an activity where the cooperation of the consumer is desired, the level of participation can be crucial in success or failure. As C.B. Galiher and her associates noted, "What is clear is that the health professional needs the consumer to achieve a relevant and responsive health care system."[33]

Given the discussion of the roles of the providers and the consumers relative to health planning implementation, specific attention can be devoted to what the planner must be alert to in the interactional stages. Exhibit 6-27 illustrates the group and community interaction in the planning and implementation processes. Questions for consideration in relating the community structure to health action and the implications of these structures demonstrate the need for the health planner to discover the community power structure. Sources of influence should be kept in mind as the planner mounts the implementation activity and has need of assistance. Then, consideration must be given to the constrictions on health planning implementation of activities. Public versus private interests must be recognized and faced in a manner that does not wreck all program possibility. Still, it is evident that activities do get implemented and programs do come into being. What moves an idea from being significant and important to being urgent

Exhibit 6-27 Group and Community Interaction in the Planning and Implementation Processes



Source: Sigrid G. Deeds, *A Guidebook for Family Planning Education* (Columbia, Md.: Westinghouse Population Center, 1973), p. 85.

and immediate? Do health planners have a part to play in making a suggested activity urgent?

The following are examples and illustrations of community structure problems, influence sources, constrictions, and determinations of significance and urgency. Some techniques and methods are suggested, and others are implied, for meeting these factors in implementing activities.

Community Structure and Health Action

Actions for health must be related to the structure of the community. Situations cited below are examples of how the community structure could interact with the implementation of a health activity.

- Is health a major problem of concern to the community?

  Public opinion surveys demonstrate that health is not the first concern of the public. Crime control, public safety, and the economy have all ranked higher than health problems. Community leaders tend to place health concerns on a lower level and do not press for many activities. Of course, a possible epidemic or other health threat does mobilize the community, but this is a passing interest and does not remain too long after the threat disappears.

- Why upset the status quo?

  Community leaders tend to avoid conflicts that can affect their own power base. A status quo mentality keeps activities that may be controversial from being implemented. Things have worked reasonably well without this new activity, and the community shows no desire to push something new or different.

- What's the role of business and banking leaders?

  Any activity involving the expenditure of sums of money appears to be dominated by community people associated with industry and with banks. Voluntary health organizations look to these people for advice and for professional help. Generally, these community leaders tend to be conservatives when it comes to implementing a new health activity.

- Is there a payoff for political leaders?

  Payoff does not refer to graft but to whether it is worthwhile for political leaders to become involved. As one politician noted, "The do-gooders will vote for me regardless. Those anti people will vote against me. Why should I antagonize them?" Obviously, the politi-

cal leaders could be strong influences in the legislative and regulatory aspects of any implementation activity.

- What's the influence of the professionals?

  Professionals base their influence upon their narrow technical expertise. Community leaders may request advice on that aspect but not on political or social implications. Professionals are often treated more as expert consultants than participants.

Implications for Health Planners

What does this analysis of the community structure and health action mean to the health planner? What are the implications for implementation? Clearly, the health planner has to identify the community leaders and work with them. Activities will be much easier to implement with the active cooperation of the community structure.

Community leaders can be identified by reputation, by official position and by decision-making ability.

A reputation technique uses one leader to identify others in the community. Then, those named by the first are also asked to identify other leaders, and so on. Eventually the names that appear most often emerge as the agreed upon community leaders by consensus. To use this technique, the health planner could start with a local politician, the president of a large industry in the community, or a bank president.

Decision-making community leaders may be more difficult to identify because these individuals often prefer to remain out of the limelight. Community leaders having the real influence could be political party chairmen or bosses, extremely wealthy individuals, so-called power brokers, socially prominent people, or aides to some of these decision makers. The cooperation of these political forces would make any implementation project infinitely smoother.

Assuming that the community leaders can be identified, how does the health planner motivate the obvious community leaders to participate in the implementation activity? Rational approaches may not really stimulate the community leaders to participate. Tradeoffs may be the situation arising most frequently. What can the health planner do for the community leader in return for active participation? E. C. Banfield[34] noted in his book on political influence that a variety of ways exists to influence people:

- Obligation—the individual feels obligated to participate because it is morally right or personally a duty.

- Friendship—individuals participate because a friend asked them to do so.
- Selling—the individual is responsive to a selling pitch.
- Coercion—the individual is politely blackmailed into activity.
- Rational persuasion—the power of the data and the weight of the presentation persuades the individual.
- Inducement—either positive or negative inducement is used to bring the individual over to your side.

All these incentives must be tailored to fit the specific community leader in the particular situation. The same message does not move all people with equal force.

Sources of Influence

Sources of influence that can be vital to an implementation activity are listed below.

- Money/credit—it is easy to buy what you need if you have the resources.
- Control of jobs—people that feel threatened with the loss of their jobs will be inclined to vote with you.
- Control of information—without all the facts, people may move in the direction desired based on the available information only.
- Social standing—some people are influenced by those whom they consider to be in the social elite.
- Ethnic solidarity—some individuals may feel compelled to participate because they want to help one of their own.
- Legality/constitutionality—the power of the law to urge people to cooperate, usually backed up by sanctions of some type.
- Personal (human) energy—Sheer work and ability to stick with the problem may influence others over time.
- Knowledge/expertise—people can be influenced by people whose intellectual ability they admire and respect.
- Popularity/esteem/charisma—That something extra that cannot be defined, but it charms people into cooperating.

Constrictions on Health Planning Implementation

Most health planners will have to deal with the conflicting self-interests of various individuals and organizations as the planning of implementation activities gets underway. At this point, there may be some thought about the public interest versus the private interest.

- Public interests are served if the decision furthers the ends of the whole public rather than the ends of some specific sector.
- Private and special interests are served by decisions that further the ends of some part of the public at the expense of the ends of the larger public.

Health planners might assume, perhaps naively, that individuals and organizations have the public interest in mind in decisions that revolve around implementation activities. However, a consideration of the types of ends that people could be favoring point out the continuum of real life situations that may constrict the health planner's implementation project. Participants may hold views that are[35]

1. community regarding rather than self-regarding,
2. stable rather than transitory,
3. general rather than particular in reference,
4. pertain to the role of citizen rather than to some private role,
5. common or statistically frequent rather than idiosyncratic or infrequent, or
6. logically or morally justified rather than emotionally or not justified.

Sources of conflict can easily arise in situations where these values are represented at random among those who are needed to cooperate in the implementation program. Some might wish to emphasize coordination over competition. Others might favor their own particular constituency over the general public. Still others might concentrate on cost containment over the environmental impact. In any case, the conflict can be bitter and impose severe constraints on the final implementation effort. Some of the more common restrictions imposed upon the health planner are illustrated below:

- Powerful interest groups strongly represent their adherents to make sure that their area of interest is not encroached in any implementation activity. If a voluntary health organization feels that education of the public is their domain, they will resist any activity that will have others engaging in education of the public about their specific disease. In many instances, the voluntary health agency will have its educational campaign directly linked to its fund-raising drive. Therefore, an implementation activity that intrudes in the educational area threatens their lifeline of donations. Since many of these powerful interest groups have boards of directors with highly placed individuals serving on them, considerable pressure can be brought to bear on influencing the decisions.

- Rapid action may be deferred because of the need to seek a consensus and to avoid conflicts. It takes time and effort to win people over to accept a least common denominator where everybody gains something and nobody loses too much. In the example about education, one agreement that did take place allowed the voluntary health agency to cosponsor all the educational activities with the name of the agency standing out boldly in any educational endeavor. Actually, this compromise probably favored the voluntary health organization since they were so solidly connected with the educational effort by the public anyhow. Many people assumed that the voluntary health agency was the prime power behind the implementation effort.
- Lack of consensus as to purpose tends to result in oiling the wheel that squeaks the most first. Health planners become so tied up in efforts to maintain the status quo that there is a failure to deal with root causes in implementation activities. With divergent viewpoints, the end result might be the pouring of old wine into new bottles to collect grant funds or other funding.
- Insufficient authority to implement activities can quickly force the health planner to a dead stop. This lack of authority could be related to legal powers, to political influence, to a lack of money or manpower, or to a lack of status in the health care community. Can a public health planner have enough authority to move private physicians into participating? One has only to read of the history of the comprehensive health planning agencies (CHPs) to garner a rather extensive listing of difficulties of this nature.
- Confusion concerning health planning models can also constrain implementation efforts of health planners. All the reasonable, logical data in the world might not persuade the individual who knows from "the seat of his pants" how to run an implementation program. Although planners can use many of the techniques presented in this book or more sophisticated technology, there is no universal agreement on what is the best method. Planning is not an exact science. Nevertheless, planning methods and techniques are constantly being developed, reviewed, and reevaluated; and the health planner must take and use the good and discard the methods that do not work in his particular situation.

WHAT MAKES AN IMPLEMENTATION ACTIVITY SIGNIFICANT OR URGENT?

Data and analysis may suggest a host of implementation activities for the health planner to undertake in the community. In the chapter

on priorities, the rationale for making decisions was emphasized. Yet, the health planner also has to realize that sociopolitical interactions will come to bear on the decisions as well as any weighting procedure or other priority determination mechanism.

Many activities will fall into the category of being significant health problems that ought to be attacked. Fewer will be classified in the category needing urgent attention. This does not mean that the significant problems are any less important, but that the urgent problems can be less important statistically or logically but higher on the scale sociopolitically.

An implementation activity can be significant for the following reasons:

- People affected—a large silent majority may make the activity numerically important as their lives are affected. There may only be a small number of people affected, but they may be the vocal few that make the activity significant.
- Facilities—if the facility being affected by the activity is on Main Street, it may be a significant program. However, if the facility is "out of sight" then it is "out of mind" and not as important.
- Costs—the higher the cost, the more significant the implementation activity becomes. Questions of tax monies and accountability color the significance.
- Risks/losses—if the activity falls into an all or nothing situation regarding implementation, the activity can engender more importance than if the risk were small.
- Social disruption—the more social disruption that the implementation activity causes, the more significant the project. If the activity escapes the public interest with a yawn rather than a riot, it will remain important but not urgent.
- Effects/benefits—if the implementation activity calls for a high powered rifle to kill a mouse, the activity will not be significant enough to stir action while still retaining its importance.
- Participants—the identity of the participants involved in implementation planning can bring significance to the activity with the weight of their names, their social position, or their connections.

To move an implementation activity from the significant category to the urgent category requiring immediate attention raises the consideration of the following factors:

- Mass communication—all types of mass media provide extensive

coverage of the proposed activity and keep the project in the limelight until action is taken to meet the urgent need.

- Crisis labeling—simply labeling the activity as an answer to a crisis situation prompts action. A self-fulfilling prophecy situation occurs. Government officials urging fluoridation of the water supply to prevent dental decay arouses a strong public outcry from those who feel that their freedom of choice is being impinged. Calling something urgent makes it so.
- Pressure (flak)—individuals and organizations maintain constant pressure through letters, phone calls, personal talks, and even picketing, sleep-ins, and boycotts until the urgent situation is rectified. Letters to the editor and strong complaints also create flak.
- Moral values—an appeal to deeply held moral values can create an urgent situation. Welfare recipients stealing the city blind and driving Cadillacs arouses a strong public cry to do something about those people who do not work and sponge off others who do. Moral values are offended and demand action.
- Power—power brokers in smoke-filled back rooms decide that something is urgent, and the activity is implemented quickly.
- Influence—individuals who have the ability to get others to think, feel, or act as they wish influence the decisions that create the urgency.

Many of the human interaction situations that have been discussed have been incorporated into a health planning game that simulates policy-making scenarios. This health planning game was developed by the Columbia University School of Public Health.[36]

TECHNIQUES FOR MEETINGS, GROUPINGS, AND PRESENTATIONS

It is possible that some implementation activities, particularly those with heavy sociopolitical overtone, might involve the use of task forces, study groups, ad hoc committees, or formal presentations to professionals and lay groups. Health planners should be aware of the basic styles and the relation between the intended purpose and the technique. Planners might consider this sequence of questions:

1. What specific objective has to be accomplished?
2. What type of meeting, grouping, or presentation is called for?
3. What subgroupings are needed to achieve the objective?

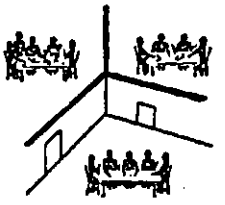
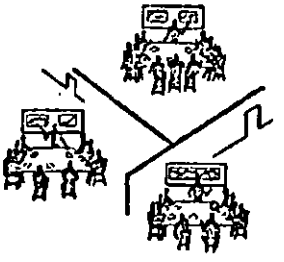
4. What presentation methods should be used?
5. How can the audience or participants become more involved?

A majority of the meeting styles are listed in Exhibit 6-28 by the purpose, the type, and relevant features of that technique. Although there might not be universal agreement on the exact terminology, the general concept of each method is fairly well accepted.

With the advent of the Health Systems Agencies, it is quite probable that health planners will have numerous situations where an implementation activity will be the calling together of experts in committees to develop areawide plans. These methods and techniques can be adapted to meet the mandates of participation, sociopolitical interaction, representation, and the production of the final product—the recommendations. Most planners probably know the methods and techniques, but a listing highlighting the purposes and the features can serve as a quick and easy reference.

Exhibit 6-28 Meetings, Groupings, and Presentation Techniques

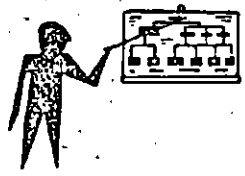
TYPE OF MEETING

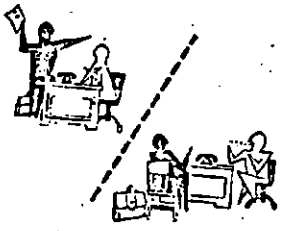
If This Is Your Purpose ▼	This Type ▼	Has These Features ▼
To plan, get facts, solve organization and member problems	WORK CONFERENCE 	▼ general sessions and face-to-face groups (15 or less) ▼ usually high participation ▼ provides more flexible means for doing organization's work
To train each other to gain new knowledge, skills or insights into problems	WORKSHOP	▼ general sessions and face-to-face groups ▼ participants are also "trainers" ▼ trainers can be brought in, too
To share experience among "experts"	SEMINAR	▼ usually one face-to-face group ▼ discussion leader also provides expert information
To train in one or several subjects	INSTITUTE 	▼ general sessions and face-to-face groups ▼ staff provides most of training resources

KINDS OF GROUPINGS

If This Is Your Purpose ▼	This Kind ▼	Has These Features ▼
To give orientation or information to total group, transact official business	**GENERAL SESSION**	▼ includes total meeting group ▼ useful for demonstrations, speeches, lectures, films ▼ can be subdivided for limited face-to-face group activities
To consider special interests of various members by means of exchange of opinions, experience, and ideas EXAMPLE: "Let's plan our golf tournament." "How can we get more members interested?" "Can we get an expert on this subject in addition to?"	**SPECIAL INTEREST GROUPS**	▼ composed of those interested in subject ▼ usually no action required, but findings may be reported ▼ size varies widely; best if kept small
To help new members get acquainted, explain conference mechanics and plan of operation EXAMPLES: "What about getting mail?" "Where can we go after hours?" "What's the schedule?"	**ORIENTATION GROUPS**	▼ mixed membership from total group ▼ member of planning staff as resource on all questions and to help them get acquainted ▼ exists for brief period at start of conference only

PRESENTATION METHODS

To present complex technical info, such as organization structure, processes, etc.	**SPEAKER WITH VISUALS** 	▼ more thorough and more certain communication but slower and more costly ▼ usually creates greater interest than speaker only
To present information from several points of view	**SYMPOSIUM**	▼ two or more speakers—each usually makes short talk ▼ speakers can help audience get full understanding of specific subject ▼ chairman summarizes and directs questions ▼ audience usually does not participate verbally

If This Is Your Purpose ▼	This Method ▼	Has These Features ▼
To present information, often controversial, from several points of view	**PANEL**	▼ panel participates—each states views and holds discussion with one another ▼ panel members usually rehearse briefly ▼ discussion is guided by moderator ▼ questions and commentary with audience
To develop several opposing sides of an issue	**FORUM**	▼ two or more speakers take opposing sides on an issue; address audience rather than each other ▼ moderator summarizes points of view and leads discussion ▼ audience usually limited to asking questions
To help audience analyze individual or group action in "natural" setting	**SITUATION PRESENTATION**	▼ members of group present role play, vignette or case-study (example: showing problems of coordination in staff meeting) ▼ commentator may call attention to specific points as play progresses ▼ audience gains a common experience for discussion afterwards
To dramatize the outer or inner forces that clash in a human situation	**CONFLICT PRESENTATION**	▼ members of group present role play or staged skits ▼ "ghost voice" or "alter ego" talks out loud to show inner thoughts of each character (example: outside pressures on two people in conflict) ▼ audience gains insights into problems through emotional appeal
To demonstrate techniques or skills and show relative effectiveness	**SKILL PRESENTATION** 	▼ members of group (or live actors) demonstrate several ways of handling a situation (example: ways to sell difficult customer) ▼ audience observes, discusses advantages and disadvantages of various approaches

Source: Reprinted with permission from Hugh Gyllenhaal, "For the Modern Meeting Planner: Guide to Styles, Groups, Methods," *Sales Meetings* 7(January 4, 1957): 19.

REFERENCES

1. J.G. Bruhn, "Planning for Social Change: Dilemmas in Health Planning," *American Journal of Public Health* 63 (July 1973): 602.

2. *Ibid.*

3. R.H. Murray, "Muddling Through Isn't Enough," *Journal of the American Medical Association* 234 (October 20, 1975): 287.

4. S.E. Salasin and H.R. Davis, "A Practical Model for Planned Change," paper presented at the annual meeting of the American Public Health Association, Atlantic City, N.J., November 13, 1972. Also cited in this paper: E.M. Glaser and H.L. Ross, "Increasing the Utilization of Applied Research Results," final report to the National Institute of Mental Health, Grant No. 5 R18 MH09250-02 (Los Angeles: Human Interaction Research Institute, 1971).

5. *Ibid.*

6. D.H. Gustafson, G.L. Rowse, N.J. Howes and R.S. Shukla, "Opportunities for Improvement in Health Systems Engineering," *Inquiry* 14 (March 1977): 87.

7. G.A. Goldberg, "Implementing University Hospital Ambulatory Care Evaluation," *Journal of Medical Education* 50 (May 1975): 435.

8. J.T. Gentry, A.D. Kaluzny and J.E. Veney, "A Comparative Analysis of Factors Influencing the Implementation of Family Planning Services in the United States," *American Journal of Public Health* 64 (April 1974): 376.

9. A.D. Kaluzny and J.E. Veney, "Types of Change and Hospital Planning Strategies," *American Journal of Health Planning* 1, no. 3 (January 1977): 13.

10. *Ibid.*

11. A. Etzioni, *A Comparative Analysis of Complex Organizations* (Glencoe, New York: The Free Press, 1961), pp. 130-142.

12. J.J. McGuire, "Delivering People Care Amid Change," *Journal of the Medical Society of New Jersey* 72 (June 1975): 473.

13. H.H. Sapolsky, "A Solution to the 'Health Crisis'," *Policy Analysis* 3, no. 1 (Winter 1977): 115.

14. M.F. Doody, "Will Providers Change the System?" *Hospitals* 50 (April 1, 1976): 51.

15. National Heart, Lung and Blood Institute. *Handbook for Improving High Blood Pressure Control in the Community* (Washington, D.C.: U.S. Government Printing Office, 1977), p. 36.

16. S.G. Deeds, *A Guidebook for Family Planning Education* (Columbia, Md.: Westinghouse Population Center, 1973), p. 40.

17. American Hospital Association. *The Practice of Planning in Health Care Institutions* (Chicago: AHA, 1973), p. 56-57.

18. H.P. Terrell, "Flow Chart for Optimum Office Efficiency," *Patient Care* 11, no. 13 (July 15, 1977): 165.

19. N. Hirschhorn, J. Lamstein and R. O'Connor, *Problem-Action Guidelines for Basic Health Care: A Tool for Extending Effective Services Through Auxiliary Health Workers* (Cambridge, Mass.: Management Sciences for Health, 1974).

20. G. Stuehler, Jr., "Organizing Organized Medicine Through a Planning and Development Framework (PADFRAME)," *Journal of the American Medical Association* 237 (April 11, 1977): 1589.

21. E.T. Alsaker, "The Basic Technique: Network Analysis," *PERT: A New Management Planning and Control Technique*, by G.N. Stilian *et al.* (New York: American Management Association, 1962), p. 37.

22. C.S. Greenblat and J.H. Gagnon, *Blood Money: A Gaming-Simulation of the Problems of Hemophilia and Health Care Delivery Systems* (Washington, D.C.: U.S. Government Printing Office, 1976).

23. L. Fontana and A. Beckerman, "Interest Groups and Community Health Planning: Cooperation or Cooptation?", paper presented at annual meeting of the American Sociological Association, New York, August 1976.

24. D.F. Mazziotti, "Underlying Assumptions of Advocacy Planning: Pluralism and Reform," *Journal of the American Institute of Planners* 40 (January 1974): 38.

25. P. Clavel, "Planners and Citizen Boards: Some Applications of Social Theory to the Problem of Plan Implementation," *Journal of the American Institute of Planners* 34, no. 3 (May 1968): 130.

26. Fontana and Beckerman, *op. cit.*

27. C. W. Douglass, "Consumer Influence in Health Planning in the Urban Ghetto," *Inquiry* 12 (June 1975): 157.

28. B.C. Vladeck, "Interest-Group Representation and the HSAs: Health Planning and Political Theory," *American Journal of Public Health* 67 (January 1977): 23.

29. E.L. McClendon, "Societal, Organizational Concerns Should be addressed in Planning," *Hospitals* 49 (December 1, 1975): 35.

30. L. Meisner, A. Parker, L. Austin, C. Orr and M.L. Ortega, "A Training Program for Consumers in Policy-Making Roles in Health Care Projects," (Berkeley: School of Public Health, University of California at Berkeley, May 1970).

31. W.C. Richardson and D. Newhauser, "First Question in Health Planning: Does the Public Know What It Wants, or Not?", *Modern Hospital*, 110 (May 1968): 115.

32. S.R. Arnstein, "A Ladder of Citizen Participation," *Journal of the American Institute of Planners* 35 (July 1969): 216.

33. C.B. Goliher, J. Needleman and A.J. Rolfe, "Consumer Participation," *HSMHA Health Reports* 86 (February 1971): 99.

34. E.C. Banfield, *Political Influence. A New Theory of Urban Politics* (Glencoe, New York: The Free Press, 1961), pp. 3-6.

35. M. Meyerson and E.C. Banfield, *Politics, Planning and the Public Interest* (Glencoe, New York: The Free Press, 1955), pp. 322-329.

36. N.M. Clark and M. Pinckett-Heller, "The Health Planning Game: A Simulation in Policy Making," paper presented at the annual meeting of the American Public Health Association, Washington, D.C., November 1977. (Prototype Health Planning Game available from N.M. Clark at Columbia University School of Public Health, 21 Audubon Avenue, New York, N.Y. 10032. $15 plus $1.50 postage and handling).

LE SYSTEME DE SANTE DU QUEBEC: OBJECTIFS DE SYSTEME OU OBJECTIFS DE SANTE?

Raynald Pineault, M.D., Ph.D., F.R.C.P.(c)
Directeur
Département de médecine sociale et préventive

et médecin conseil
Département de santé communautaire
Cité de la Santé, Laval

Richard Lessard, M.D., M.P.H., M.A., F.R.C.P.(c)
Chef du département de santé communautaire
Cité de la Santé, Laval

et professeur adjoint
Département de médecine sociale et préventive
Université de Montréal

SOUMIS POUR PUBLICATION à l'Union Médicale du CANADA

MAI 1983

REMERCIEMENTS

Les auteurs tiennent à remercier leurs collègues, André-Pierre Contandriopoulos et Georges Desrosiers, pour leurs commentaires critiques et judicieux sur les premières versions de cet article.

RESUME

En dépit de la réforme majeure qui a été amorcée dans les services de santé et les services sociaux au début des années 1970, le Québec n'a pas vraiment réussi à se doter d'une véritable politique de santé, si l'on excepte certains thèmes tels que la périnatalité, la nutrition et la santé dentaire. Les objectifs d'accessibilité, de continuité, de globalité et de qualité qui ont été mis de l'avant dans le cadre de cette réforme peuvent s'avérer des instruments efficaces dans la poursuite d'objectifs de santé, mais ils n'en sont pas nécessairement une garantie. Les conséquences de l'absence de politique de santé ne sont pas sans causer d'inconvénients sérieux au niveau de la gestion du système de soins puisque les critères de choix dans le contexte de la décroissance, plutôt que de se reposer sur la capacité du système à améliorer la santé de la population, doivent bien souvent faire appel à des considérations d'ordre administratif ou économique.

SUMMARY

In spite of the vast reform in health and social services that was launched in the early 70's, Quebec has not been fully successful in establishing an overall health policy except in limited sectors such as nutrition, perinatal and dental health. Objectives such as accessibility, continuity, comprehensiveness and quality that have been put forward are potentially but not necessarily instrumental in achieving health objectives. The absence of health policy can create problems in the management of the health care system since cost containment decisions tend to be based on administrative or economic considerations rather than on the capacity of the system to improve the health of the population.

Introduction

Le début des années 70 marquait un point de départ important pour le Québec dans le secteur de la santé et des services sociaux. Le contexte socio-politique allait faire en sorte qu'une réforme importante de ce secteur allait s'amorcer. Le projet mis de l'avant allait s'accoler l'étiquette "santé communautaire" pour bien signifier qu'un virage vers la santé allait s'effectuer. Dix ans après le début de cette réforme, il convient de s'interroger sur les points suivants: (1) Y a-t-il eu vraiment virage vers la santé au Québec? (2) Ce virage s'est-il concrétisé dans l'élaboration de politiques dont le point de départ est davantage des objectifs de santé plutôt que de soins? (3) Dans quelle mesure l'importance accordée aux politiques du système de soins plutôt que de santé peut-elle avoir des conséquences sur l'organisation et la gestion du système de santé et de soins? (4) Enfin, quelles mesures peut-on proposer pour qu'une véritable politique de santé soit mise de l'avant et qu'elle rallie les divers intervenants du milieu des affaires sociales? Voilà les questions auxquelles cet article veut s'adresser. Bien sûr, il n'a pas la prétention de pouvoir apporter des réponses définitives à toutes ces questions, mais il espère pouvoir amorcer des réflexions utiles dans ce sens.

Cadre de référence pour la planification en santé

Afin de mieux situer nos propos, il est utile de présenter brièvement un cadre conceptuel de la planification en santé. D'abord précisons tout de suite ce que nous entendons par système de santé et système de soins. Nous reprendrons ici les définitions adoptées par le Rapport du Comité sur la rémunération des professionnels de la santé du Québec. "Le système de santé est ainsi le fonctionnement de l'ensemble des éléments qui déterminent l'état de santé de la population; il est composé du système de soins et aussi d'aspects ou d'éléments du système économique et social, du système culturel, du système politique, etc..." "Le système de soins est formé par un ensemble de ressources (les professionnels, les établissements, etc.) structurées en vue de fournir des services et des biens à la population dans le but d'améliorer son état de santé" (1).

A la figure 1; on retrouve les trois éléments qui constituent le système de soins: les ressources, les services et l'état de santé. Les ressources produisent des services qui sont utilisés par la population pour maintenir, améliorer ou restaurer son état de santé (2). La tâche du planificateur de la santé comprend plusieurs étapes. D'une part, tel que représenté sur la partie gauche de la figure 1, il doit procéder à une analyse critique du système de soins et voir si les ressources déployées sont utilisées de façon productive et appropriée, et si elles sont accessibles. De plus, il doit se poser la question

fondamentale à savoir dans quelle mesure les services produits et utilisés contribuent à l'amélioration de l'état de santé. D'autre part, tel que représenté à la figure 1, la tâche essentielle de la planification revient à formuler des objectifs et à élaborer des politiques. Ces objectifs peuvent porter sur l'état de santé (politiques de santé), sur les services utilisés ou produits (politiques de soins) sur les ressources (politiques de ressources) ou encore sur les autres déterminants de la santé (politiques de promotion de la santé). La notion importante à retenir pour nos propos, c'est qu'une politique de ressources ou de soins devrait découler logiquement d'une politique plus globale de santé. Cette relation est illustrée à la figure 1 par les flèches reliant les différents niveaux. En d'autres termes, il est peu logique de formuler des objectifs de soins ou de ressources sans se demander si les ressources ou les services additionnels requis vont dans le sens des objectifs de santé.

Au trois niveaux, les besoins sont définis par l'écart entre les objectifs et la situation actuelle. Encore ici, l'on peut voir que les besoins définis en termes de ressources ou de services n'ont de véritable valeur que s'ils sont reliés à des besoins de santé. De façon plus concrète, par exemple, une politique de soins axée uniquement vers des objectifs d'accessibilité et de continuité des soins, ou encore pire, une politique de ressources uniquement centrée sur la répartition "équitable" des ressources entre les régions ne peuvent conduire à une efficacité sociale du système de soins qu'a la condition qu'elles s'articulent avec une logique de santé, c'est-à-dire qu'elles s'inscrivent dans une politique de santé cohérente qui relie les objectifs de ressources et de soins à des objectifs de santé.

Enfin, si l'on réfère au système de santé plutôt qu'au système de soins, l'on voit que le lien entre l'état de santé et les autres déterminants de la santé est important et que le système de soins, bien qu'il occupe la majeure partie de notre schéma, ne représente qu'un facteur parmi les autres nombreux déterminants de la santé. A ce titre, le système de soins ne constitue qu'un sous-système de l'ensemble du système de santé et une véritable politique de santé doit inclure les actions sur les autres déterminants de la santé au même titre que les actions qui portent sur le système de soins. Pour mieux illustrer de façon concrète la distinction que nous établissons entre des objectifs de santé, nous présentons au tableau 1 des exemples pour chacune de ces deux catégories.

Politique de soins ou de santé

Maintenant que nous avons précisé la signification que nous donnons aux différents termes que nous allons employer et surtout que nous avons insisté sur l'importance d'inscrire les politiques de soins ou de ressources à l'intérieur d'une politique plus large de santé, nous allons nous poser la question suivante: Existe-t-il au Québec une véritable politique de santé?

Pour répondre à cette question, il nous faut d'abord reculer dans le temps et bien préciser le contexte dans lequel les politiques en matière de santé ont été formulées au Québec et au Canada et de voir comment elles ont évolué dans le temps. Un facteur important à considérer dans l'analyse des systèmes de santé et des soins provinciaux au Canada est l'influence déterminante des programmes fédéraux sur les régimes des provinces. En effet, on observe que le gouvernement fédéral a été le promoteur de mesures sociales qui mettaient l'accent d'abord et avant tout sur les

soins. En d'autres termes, depuis le début des années 50, les préoccupations qui ont été traduites dans les programmes fédéraux de soins, notamment l'Assurance-hospitalisation (en 1958) et l'Assurance-maladie (en 1968), visaient en premier lieu à rendre accessibles des soins globaux de santé à tous les Canadiens. Sensibles aux avantages économiques reliés à ces programmes fédéraux, les provinces ont certes été encouragées à adhérer à ces régimes. Ce n'est qu'en 1974 qu'on observe un changement brusque dans l'orientation du gouvernement fédéral qui, par le rapport Lalonde, annonce clairement et explicitement son intention de mettre dorénavant l'accent sur la promotion de la santé et sur les modifications de comportement (3). Sans entrer dans le détail de l'analyse de ce document et surtout dans toutes les considérations strictement économiques derrière cette façade de la promotion de la santé, il faut admettre que le mérite principal de cette politique a été de rendre explicite la position du Ministère de la santé à l'égard de la promotion de la santé et de l'importance des problèmes de santé reliés au comportement (4).

Revenons maintenant au Québec. Comme les autres provinces, mais avec un peu d'hésitation et de retard, le Québec adhère aux deux programmes de l'Assurance-hospitalisation (en 1961) et de l'Assurance-maladie (en 1970). Cette dépendance des politiques provinciales sur les programmes fédéraux se retrouve même au niveau de la Commission Castonguay-Nepveu dont le travail commencé en 1966 se terminait en 1970 par la production de son rapport final. Bien qu'il soit mentionné dans le rapport de cette commission que toute la réforme sur la santé

et les services sociaux au Québec vise ultimement une amélioration de la santé de la population, ce sont néanmoins les objectifs de soins qui y prédominent. En effet, on constate que le premier objectif de la réforme en est un d'accessibilité économique, que le deuxième en est un de contrôle des coûts et que les autres concernent la continuité des soins, leur globalité et leur qualité (disponibilité). La promotion de la santé et la prévention sont considérées dans ce rapport comme parties intégrantes du système de soins, et aucune attention particulière n'y est réservée.

Réalisant que des propositions concrètes n'avaient pas été présentées par la Commission à l'égard des programmes de prévention, le Ministère des Affaires sociales confiait en 1971 à un groupe de travail le mandat précis de faire des recommandations sur l'organisation du secteur de la santé publique. Encore ici, on ne peut dire qu'il s'agissait d'une véritable préoccupation de santé de la part du Ministère mais plutôt d'une volonté de réorganiser le système des unités sanitaires et des services de santé municipaux. La recommandation principale de ce groupe de travail de créer les départements de santé communautaire se concrétisa dans les règlements à la Loi des services de santé et des services sociaux.

Des exemples de politiques de santé

Jusqu'ici nous avons montré qu'une véritable politique de santé n'a jamais été réellement adoptée par le Ministère des Affaires sociales. Cette constatation ne signifie pas une complète absence de préoccupation dans ce sens. En effet, dans certains secteurs notamment la périnatalité, la nutrition, la santé dentaire et quelques autres, le Québec a établi de véritables politiques de santé. En matière de périnatalité, la politique mise de l'avant en 1973 visait trois objectifs:

(1) réduire de moitié le taux de mortalité maternelle en 1973, afin de le porter en 1980 à 1.1 par 10,000 naissances;

(2) abaisser le taux de mortalité périnatale à 12 par 1,000 naissances vivantes en 1980, comparativement à 18.3 en 1972;

(3) enfin, ramener à 4.5% le taux de prématurité comparativement à 8.24% observé en 1970 (5).

Sans entrer dans l'interprétation complexe des résultats obtenus, le point important à noter pour notre discussion est que les actions qui furent alors proposées s'inscrivaient dans la poursuite d'objectifs de santé, plutôt que d'être seulement des objectifs de soins, coupés de liens explicites avec l'amélioration de la santé.

De la même façon, suite aux résultats de l'enquête Santé-Canada en 1971, le Ministère des Affaires sociales lançait en 1977 sa politique de nutrition dont les trois principaux objectifs étaient les suivants:

(1) réduire de 25% l'apport en matière grasse;

(2) réduire de 50% l'apport de sucre;

(3) augmenter l'apport en fibres alimentaires (6).

Encore ici, dans les actions proposées par cette politique, l'on sent la préoccupation de les relier à la poursuite de ces objectifs plutôt que de les situer simplement par rapport à une logique d'activités et de services.

Enfin, la politique de santé dentaire est un autre exemple qui illustre la possibilité d'établir des objectifs de santé et de mettre en place des actions qui vont dans le sens de ces objectifs (7). Incidemment, c'est dans le cadre de cette politique que le Ministère a pu prendre certaines décisions à l'égard du transfert de certaines ressources du secteur privé au secteur public. Sans porter de jugement sur le bien fondé d'une telle décision, l'on est forcé d'admettre qu'elle est fondée au moins sur une réflexion à l'égard d'objectifs de santé et qu'elle s'inscrit dans une politique de santé dentaire. Elle est, cependant incomplète en ce qu'elle néglige de considérer des facteurs extérieurs au système de soins et, par conséquent, elle omet dans ses actions la mesure reconnue comme la plus efficace et la moins coûteuse en matière de prévention en santé dentaire, soit la fluoruration des eaux.

Les trois exemples que nous venons de présenter illustrent bien que le Québec n'est pas complètement démuni de politiques de santé.

Il serait injuste de croire que ce sont là les seuls domaines où une politique formelle de santé a été mise de l'avant. Ainsi, par exemple, les personnes âgées ont fait également l'objet de beaucoup d'attention de la part des planificateurs et les documents publiés par le Ministère constituent de véritables politiques des personnes âgées (8,9). Cependant si l'on regarde de plus près, l'on s'aperçoit que l'objectif général de maintien à domicile découle d'abord d'une préoccupation de services et que le lien entre les services de maintien à domicile et celui du maintien de l'autonomie des personnes âgées n'est pas toujours clair. Ce point est d'ailleurs très bien souligné dans la conclusion d'une étude récente réalisée par Béland pour le compte du Ministère: "Une politique qui a pour but de préserver l'autonomie des personnes âgées n'est donc pas nécessairement une politique qui cherche à les garder à domicile... La politique de maintien à domicile doit être réorientée vers une politique de maintien de l'autonomie". (10).

Il y aurait certainement lieu de mentionner d'autres problèmes de santé qui ont fait l'objet d'une démarche pouvant conduire à l'élaboration d'éléments d'une politique de santé, notamment le tabagisme, l'alcoolisme, l'activité physique. Le point que nous voulons faire ressortir maintenant est que ces exemples demeurent peu nombreux, reflétant bien l'absence d'une politique globale de santé et d'objectifs spécifiques à atteindre pour des problèmes de santé particuliers.

Et pourtant, les exemples d'élaboration de véritables objectifs de santé à l'intérieur desquels doivent s'inscrire les soins de santé et les ressources ne manquent pas. La section américaine de l'O.M.S. a récemment publié un document qui trace les grandes lignes de ce que pourraient être les objectifs d'une politique de santé et le cadre de référence pour les différentes actions dans le domaine de la santé (11). Encore plus près de nous, les Etats-Unis, par l'intermédiaire de leur département de la santé publique, se sont préoccupés de plus en plus pendant les dernières années d'établir des objectifs de santé à l'intérieur desquels différentes actions préventives et curatives pourraient s'inscrire. Plusieurs documents sont nés de cette réflexion et l'un d'eux, en particulier, reflète bien ce que peut être une politique de santé par opposition à une politique de soins (12). Qu'on ne se méprenne pas. Nous ne voulons pas prétendre ici que les Etats-Unis ont une véritable politique de santé. En effet, le document dont nous parlons n'a certes pas été endossé par le gouvernement américain actuel et son opérationalisation est loin d'être une réalité.(13) Ce que nous voulons dire, c'est qu'il représente un bon exemple à suivre pour ceux qui veulent doter leur pays d'une véritable politique de santé. Nous irions même jusqu'à avancer que, compte tenu du contexte du Québec et notamment du réseau de la santé communautaire déjà existant, une telle politique pourrait retrouver au Québec un accueil beaucoup plus favorable qu'aux Etats-Unis.

Conséquence de l'absence d'une politique de santé

Nous avons jusqu'ici tenté de montrer qu'au Québec, la préoccupation majeure a été le système de soins plutôt que le système de santé et qu'en conséquence, les objectifs formulés par le Ministère des Affaires sociales, à quelques exceptions près, reflétaient cette préoccupation.

La question que nous voulons nous poser maintenant est la suivante: quelles sont les conséquences de l'absence d'une politique d'ensemble de santé?

De façon générale, si l'on se réfère à notre cadre conceptuel pour la planification de la santé, l'on peut affirmer que l'absence d'une politique de santé, c'est-à-dire d'objectifs de santé, fait en sorte que les interventions et les soins de santé de même que les ressources qui y sont affectées tendent à être planifiées selon une logique qui leur est propre et qui ne se situe pas nécessairement par rapport à une finalité de santé. De façon plus spécifique, nous pouvons identifier les conséquences suivantes:

(1) l'absence de cadre de référence pour les actions de prévention et de promotion de la santé;
(2) l'importance exagérée accordée aux structures et professionnalisme;
(3) la confusion et le manque de direction dans la décentralisation;
(4) l'incohérence des actions de santé et la mauvaise utilisation de l'expertise;

(5) la difficulté, voire l'impossibilité d'évaluer les résultats des interventions en santé;

(6) et enfin, la difficulté de gérer de façon rationnelle le développement ou la décroissance.

Revenons sur chacune de ces conséquences. Une première conséquence est le manque de cadre de référence pour les actions prévention et de promotion de la santé. En effet, comme l'illustre la figure 1, pour réaliser des objectifs de santé, les actions doivent souvent se situer à l'extérieur du système de soins, au niveau des "autres déterminants" de la santé. C'est un fait bien connu que les facteurs socio-économiques sont de plus en plus importants pour expliquer les problèmes reliés aux maladies dégénératives et chroniques. En conséquence afin de diriger nos actions sur les facteurs de risque, il nous faut bien souvent aller au-delà du système de soins et considérer les habitudes de vie, les facteurs de l'environnement physique, social ou de travail. L'absence d'objectifs de santé peut faire en sorte que ces actions ne sont pas considérées ou sont minimisées au point de vue de leur pertinence. Cette situation contribue à la grande fragilité des programmes de prévention et des institutions qui en sont responsables, d'autant plus que les effets de la prévention se manifestent bien souvent à long terme. Ainsi, pour les gouvernements dont la demi-vie est la plupart du temps limitée, la rentabilité politique d'un investissement en prévention n'est pas toujours évidente,

Cette première conséquence est amplifiée par la seconde, l'importance exagérée accordée aux structures et le professionnalisme. Dans un système où il n'y a pas d'objectifs de santé, les structures organisationnelles et les professions règnent de façon suprême parce qu'elles tendent naturellement à répondre à leurs intérêts propres plutôt qu'à ceux d'un système. En effet, les structures et les organisations visent en tout premier lieu leur survie et les professions doivent voir à leurs intérêts. Dans un tel contexte, toute pratique novatrice, comme par exemple, une préoccupation de tous les instants pour la prévention ou une pratique plus axée sur la prévention, qu'elle s'accompagne ou non d'un changement de rémunération, a de la difficulté à voir le jour parce qu'elle vient modifier des pratiques professionnelles existantes et devient concurrente sur le plan des ressources. Par ailleurs, ces pratiques novatrices ont besoin du temps pour afficher leur utilité. Il est intéressant de noter, à cet égard, que le Ministère est bien souvent plus exigeant pour ces pratiques "nouvelles" qui ont à faire la preuve de leur efficacité que pour des pratiques bien établies qui n'ont jamais été soumises à une évaluation formelle et qui, pourtant, représentent une grande partie des investissements publics en santé. Si des objectifs précis existaient, comme par exemple en rapport avec le problème de l'hypertension ou tout autre problème de santé pour lequel les connaissances actuelles et les ressources dont nous disposons rendent possible une intervention efficace, la profession médicale pourrait, dans ce cas, articuler sa pratique avec les politiques de santé et contribuer à la réalisation de ces objectifs de santé.

Une autre conséquence de l'absence de politique de santé se ressent par un certain degré de confusion et par un manque de direction au niveau de la décentralisation. En effet, l'existence de politiques de santé pour l'ensemble du Québec permettrait aux régions d'inscrire leurs objectifs à l'intérieur d'un cadre général plutôt que de reprendre à leur compte l'élaboration de politiques de santé. En d'autres termes, décentraliser sans donner d'orientation peut conduire au chaos. En l'absence de politique de santé, le Ministère se voit bien souvent forcé de donner des orientations sur les services et les ressources. Les circulaires émanant du Ministère témoignent de cet état de fait puisque les mandats conférés aux CRSSS concernent généralement la gestion des ressources et des services et ne portent que très rarement sur des problématiques de santé. Il nous semble bien qu'une façon plus logique de procéder serait pour le Ministère de formuler des orientations de santé et de laisser les régions choisir elles-mêmes les moyens (services et ressources) pour se conformer à ces orientations.

L'exemple des DSC en témoigne également puisque, en l'absence de politiques de santé et de prévention adéquates, ces structures ont consacré beaucoup d'énergie à définir des problématiques de santé et à établir des programmes de santé. Sans verser dans l'excès contraire qui voudrait que tous les programmes de santé soient élaborés au niveau provincial, nous croyons que les DSC pourraient travailler davantage collectivement à définir les programmes touchant les problèmes de santé communs à toutes les régions. Ce travail commun devrait également

permettre une plus grande cohésion non seulement au niveau de la programmation, mais également au niveau de la réalisation de ces programmes. A cet égard, l'exemple de la programmation présenté dans un des documents américains mentionnés plus haut pourrait s'avérer utile (12).

Comme nous venons de le souligner, la décentralisation en l'absence de politiques de santé peut conduire à une certaine incohérence dans les actions et également à une dilution et à une mauvaise utilisation de l'expertise. Cette situation peut certes être aggravée par le manque de communications entre les différents niveaux du réseau des affaires sociales. Ce phénomène de manque de communication se retrouve notamment dans les structures gouvernementales où les technocrates sont généralement menacés par un tel mouvement de déplacement vers les structures régionales ou sous-régionales.

L'absence de politique de santé rend difficile sinon impossible l'évaluation des résultats des interventions en santé et c'est là une autre conséquence malheureuse. L'évaluation de résultats doit se

faire nécessairement par rapport à des objectifs. Lorsque des objectifs de santé ne sont pas clairement énoncés, l'évaluation peut difficilement porter sur les résultats obtenus, et elle doit se limiter à apprécier les efforts déployés, c'est-à-dire la quantité et le type de ressources, et la justesse des services produits et utilisés (figure 2). Evidemment, ces deux derniers types d'évaluation ne peuvent apporter de réponse à la question suivante: dans quelle mesure les ressources mobilisées et les services produits ont-ils contribué à améliorer la santé de la population visée par ces actions?

Pour mieux illustrer ce point, prenons deux exemples. Le premier est celui de la périnatalité. Sans entrer dans les questions méthodologiques reliées à la recherche évaluative, il est juste de reconnaître que le fait d'avoir mis en place une politique de périnatalité nous permet de tirer certaines conclusions sur l'efficacité de cette politique suite au progrès réalisé dans ce secteur tel que tendent à le montrer les indicateurs de mortalité et de morbidité (14).

D'autre part, dans le domaine des maladies cardiovasculaires, on a noté au Québec, au Canada et dans plusieurs autres pays occidentaux, une amélioration marquée durant la dernière décennie sur le plan de la mortalité (15,16). L'une des difficultés inhérentes à l'interprétation de ces données est précisement que, contrairement à la périnatalité, ces changements sont survenus sans que l'on ait préalablement réussi à mettre ensemble dans le cadre d'un programme cohérent les différentes actions

mises en place pour améliorer les résultats dans ce domaine. En somme, l'utilité d'une politique de santé est de fournir un cadre logique d'analyse permettant d'associercertains résultats de santé à des actions bien identifiées.

Reliée à cette difficulté d'évaluer les résultats de ces actions, une dernière conséquence, c'est la difficulté, voire l'impossibilité de gérer de façon rationnelle soit le développement, soit la décroissance. Dans la période de prospérité relative que nous avons connue dans les années 60, nous ne nous sommes pas posés beaucoup de questions sur la cohérence du développement des services de santé. Comme nous l'avons mentionné plus haut, le développement des services de santé a été dicté principalement par une dépendance intéressée et presque complice des politiques provinciales à l'égard des programmes fédéraux de soins. Depuis le début des années 70, la situation a bien changé. La prospérité a fait place à la récession. Dans ce contexte, comment gérer la décroissance? Si l'on prend l'exemple de la loi 27, et sans mettre en doute le bien-fondé de cette législation, il semble bien que l'absence d'une véritable politique de santé rend difficile l'application de mesures de restriction et de décroissance. D'autres législations et décrets vont dans le même sens. Ce qui manque dans l'ensemble, c'est une logique de santé. Nous avons l'impression que la logique des mesures visant la décroissance se limite à une perspective de réduction des coûts en intervenant sur les services et les ressources alors qu'elle devrait poser la véritable question: quels services ou quelles ressour-

ces pouvons-nous diminuer sans crainte de porter atteinte à la santé de la population? Dans un contexte où il y a absence d'un tel cadre de référence, les décisions prises par nos élus politiques ou leurs technocrates tendent à se fonder principalement sur des critères administratifs, économiques et politiques.

CONCLUSION

L'élément central qui ressort de notre analyse est l'absence au Québec d'une politique globale et cohérente de santé. Cette constatation doit être cependant interprétée et comprise à la lumière de plusieurs considérations. Premièrement, le Québec n'est pas complètement dépourvu d'objectifs de santé comme nous l'avons démontré plus haut. De plus, le Québec n'est pas un cas unique puisque cette constatation pourrait bien s'appliquer aux autres provinces canadiennes et à la plupart des pays industrialisés. En ce qui concerne les pays en voie de développement, compte tenu des ressources limitées dont ils disposent et des problèmes de santé énormes auxquels ils doivent faire face, c'est presque par nécessité que la plupart d'entre eux ont dû se doter d'objectifs de santé.

D'ailleurs, il faut bien reconnaître que le système de soins du Canada et du Québec, en particulier, est considéré par plusieurs analystes comme l'un des meilleurs au monde (17,18). Le contrôle des dépenses de santé réalisé par le Canada, et le Québec tout particulièrement, au cours des dernières années est exemplaire en comparaison avec d'autres pays industrialisés (19).

Enfin, il est évident qu'en arrière des objectifs de soins se cachent, bien sûr, des objectifs de santé. L'accessibilité aux soins est une condition peut-être nécessaire, mais certainement non suffisante pour améliorer la santé de la population. En d'autres termes, la

relation entre l'objectif de soins et l'objectif de santé semble implicite, et c'est à ce défaut de logique que nos critiques s'adressent. La question que nous nous posons au terme de cet article est la suivante: n'y aurait-il pas lieu de formuler de façon explicite des objectifs de santé qui soient plus spécifiques et plus opérationnels que celui trop général d'amélioration de l'état de santé et qui puissent être utilisés pour orienter nos choix dans l'allocation des ressources de soins? N'y a-t-il pas lieu, par exemple, lorsque l'on pense à ajouter des établissements de santé dans le réseau, de considérer au-delà des questions d'accessibilité géographique, l'impact qu'aura cet établissement sur la prévalence et l'incidence de problèmes de santé précis?

Les naissances de bébés de petit poids, la prématurité et les malformations congénitales, une mauvaise alimentation et un taux élevé de caries dentaires et de maladies parodentaires ne sont pas les seuls problèmes de santé sur lesquels on doit s'attendre à une action coordonnée de l'ensemble du système. Il y aurait lieu d'établir également des objectifs de santé pour d'autres problèmes de santé importants et pour lesquels la technologie actuelle nous permet d'intervenir efficacement.

A cet égard, les départements de santé communautaire auront un rôle tout à fait déterminant à jouer de par le mandat que leur a confié le législateur et de par les ressources qu'ils possèdent.
Le rôle du DSC, rappelons le, consiste:

1) à déterminer conjointement des objectifs de santé spécifiques que le réseau et ses intervenants doivent atteindre; 2) à suggérer les moyens de les atteindre (programmation); 3) à participer à la coordination des efforts nécessaires à l'atteinte des objectifs (coordination), et; 4) à évaluer le degré d'atteinte des objectifs (évaluation). S'ils devaient accomplir cette mission, et ce conjointement avec leurs partenaires du réseau, les D.S.C. pourraient devenir l'élément qui pourrait être moteur dans la transformation du système de soins en un instrument efficace d'amélioration de la santé des Québécois. Une utilisation de ces objectifs de santé serait de définir les priorités d'investissement pour les CLSC en matière de prévention et de permettre à la profession médicale de joindre ses efforts à la réalisation de ces objectifs.

FIGURE 1

CADRE DE REFERENCE POUR LA PLANIFICATION DE LA SANTE

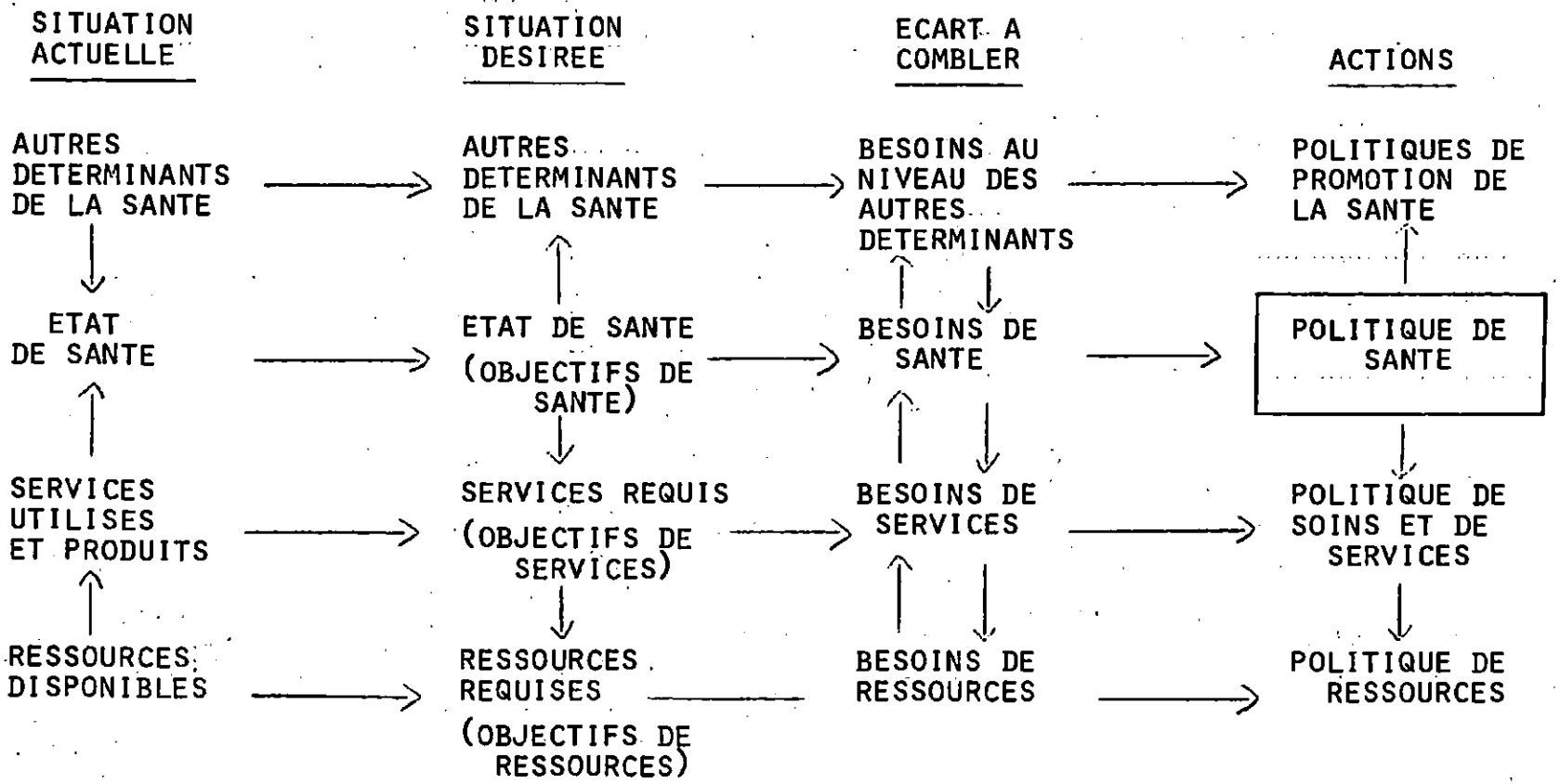
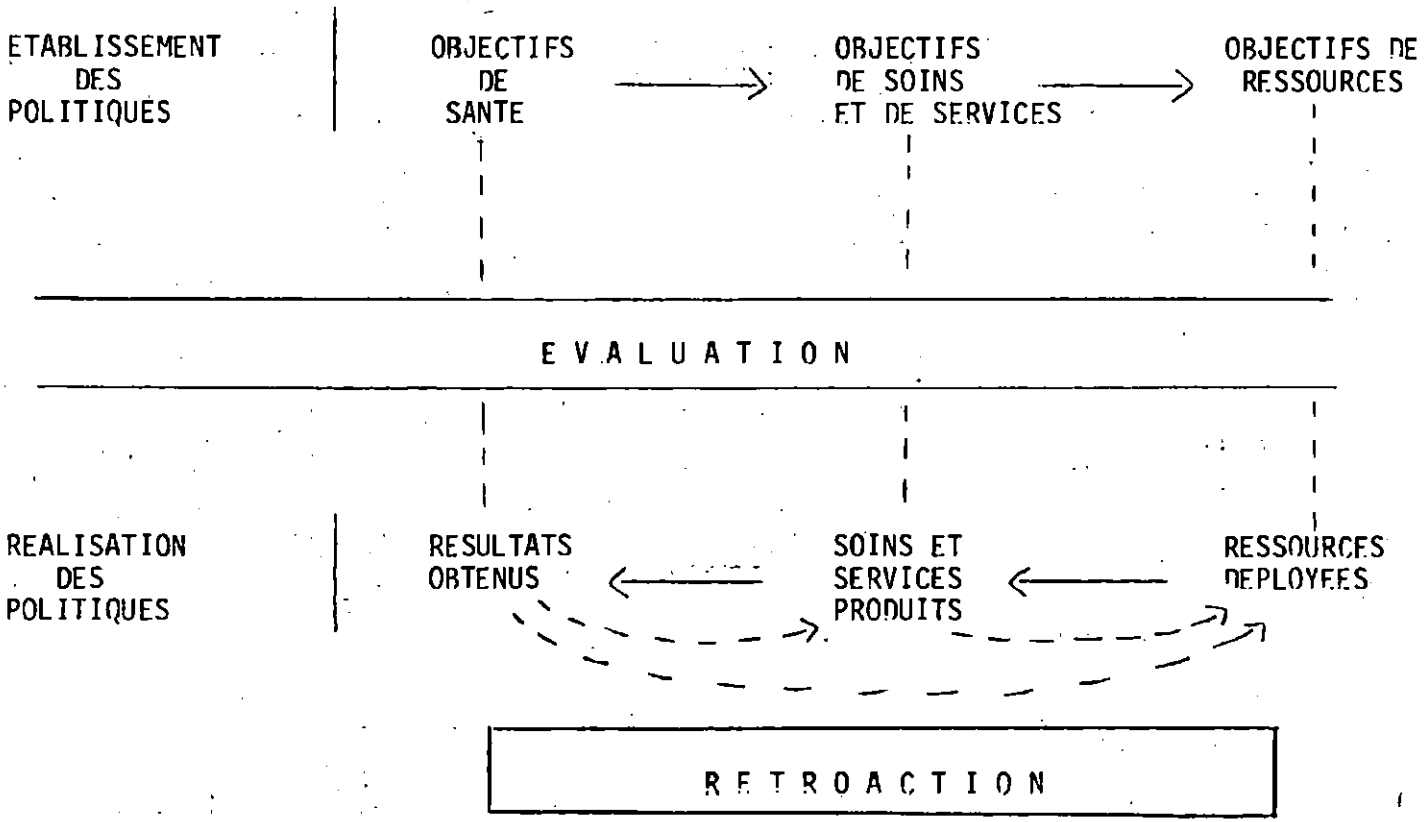


FIGURE 2

Cadre de référence pour l'évaluation des politiques de santé



BIBLIOGRAPHIE

1. Le Système des Honoraires Modulés. Rapport du Comité sur la Rémunération des Professionnels de la Santé du Québec, Québec, 1980.

2. Pineault R.: La planification des services de santé: une perspective épidémiologique. Adm. Hosp. Soc., 1979; 25 : 6-13.

3. Lalonde M.: Nouvelle Perspective de la Santé des Canadiens. Gouvernement du Canada, Ottawa, 1974.

4. Evans R.: A retrospective on the "New Perspective". J. of Hlth Policies, Policy and Law, 1982, 7: 325-344.

5. La Périnatalité: Une Politique du Ministère des Affaires sociales. Ministère des Affaires sociales, Québec, 1973.

6. Une Politique Québécoise en Matière de Nutrition. Ministère des Affaires sociales, Québec, 1977.

7. Politique de Santé Dentaire au Québec. Ministère des Affaires sociales, Québec, 1981.

8. Le Ministère des Affaires Sociales et les Personnes Agées. Ministère des Affaires Sociales, Québec, 1977.

9. Pour mieux Vieillir au Québec. Ministère des Affaires sociales, Québec, 1980.

10. Béland F.: Les Principaux Résultats de l'Analyse des Désirs d'Hébergements de trois Echantillons de Personnes Agées au Québec. Ministère des Affaires sociales, Québec, 1982.

11. Health for All by the Year 2000: Plan of Action for the Implementation of Regional Strategies. Pan American Health Organization, Official document no 179, Washington, 1982.

12. Promoting Health, Preventing Disease: Objectives for the Nation. U.S. Dept. of Health and Human Services, Washington, 1980.

13. Robbins A.: Can Reagan be indicted for betraying public health? AM. J. Pub. Hlth., 1983; 73: 12-13.

14. Blanchet M. et Levasseur M.: Périnatalité: bilan et prospective. Carrefour des Aff. Soc., 1980; 2: 10-28.

15. Mortalité, Morbidité et Facteurs de Risque reliés aux Maladies du Système Circulatoire au Québec. Ministère des Affaires sociales, Québec, 1978.

16. Tendances de la mortalité et de la morbidité liées aux maladies du coeur. Mal. Chron. au Can., 1981; 2: 9-12.

17. Monnier J. et coll. Santé Publique: Santé de la Communauté. SIMEP, Villeurbane, France, 1981; 434-438.

18. Rodwin V.G.: Health Planning in International Perspective. Dans: Planning for Health, Henrik L. Blum, Human Sciences Press, New York, 1981; 396-444.

19. Bennett J.E. et Krasny J.: Health Care in Canada. Dans: Health and Canadian Society, D. Coburn et coll. (eds), Fitzhenry et Whileside, 1981; 40-66.

TABLEAU 1

EXEMPLES D'OBJECTIFS DE SOINS ET D'OBJECTIFS DE SANTE

OBJECTIFS DE SOINS	OBJECTIFS DE SANTE
- ASSURER UNE DISPONIBILITE ADEQUATE DES RESSOURCES	- DIMINUER LA MORTALITE DUE AUX MALADIES CARDIO-VASCULAIRES
- RENDRE LES SERVICES DE SANTE ACCESSIBLES A UNE POPULATION	- DIMINUER LA MORTALITE ET LA MORBIDITE RELIEES AUX ACCIDENTS
- FAIRE EN SORTE QUE LA QUALITE DES SOINS SOIT CONFORME AUX NORMES PROFESSIONNELLES	- DIMINUER L'INCIDENCE DES MALADIES INFECTIEUSES DE L'ENFANCE
- ATTEINDRE UNE GLOBALITE DES SOINS, C.A.D. OFFRIR UN EVENTAIL COMPLET DE SERVICES SOCIO SANITAIRES	- DIMINUER LE POURCENTAGE DES ENFANTS QUI PRESENTENT DES CARIES SUR LEURS DENTS PERMANENTES
- MAINTENIR LA CONTINUITE DANS LES SOINS	- AUGMENTER LA PROPORTION DES PERSONNES AGEES QUI MAINTIENNENT UN DEGRE ADEQUAT D'AUTONOMIE
- AUGMENTER LE DEGRE DE PRODUCTIVITE DANS LA FOURNITURE DES SERVICES	- DIMINUER L'INCIDENCE DES MALADIES TRANSMISES PAR VOIE SEXUELLE

LECTURE POUR MARDI LE 14 JUIN 1983

La modification de comportements reliés à la santé

I — Revue de la littérature théorique

Michel O'Neill[1]

Résumé

Cet article passe en revue les principales théories expliquant les mécanismes par lesquels des habitudes de vie reliées à la santé peuvent être influencées. Une analyse critique de l'utilité pratique de ces théories est présentée, ainsi que quelques suggestions concrètes pour la construction de programmes de promotion de la santé. Finalement, quelques réflexions sur la nature particulière des programmes d'éducation sanitaire sont proposées.

Au Québec, comme ailleurs au Canada[1] et dans la plupart des pays industrialisés, on accorde une importance grandissante aux activités d'éducation sanitaire et de promotion de la santé. Dans le sillage du document *Nouvelle Perspective de la santé des Canadiens*[2], plusieurs professionnels de la santé, oeuvrant en milieu préventif aussi bien que curatif, sont maintenant convaincus qu'il faille modifier un certain nombre d'habitudes de vie qui ont un effet direct sur la santé de la population et sur l'utilisation de soins médicaux de plus en plus coûteux. Force nous est de constater cependant que bien peu de ces professionnels sont familiers avec les théories du changement de comportement, ou encore avec les stratégies visant à influencer des habitudes telles une mauvaise alimentation, le tabagisme, la surconsommation d'alcool, le manque d'exercice physique ou la conduite dangereuse d'un véhicule automobile.

C'est pour combler en partie cette lacune que cet article a été rédigé. Il vise d'abord à fournir aux praticiens une vision générale des principales théories existantes pour expliquer les mécanismes par lesquels un comportement humain relié à la santé peut être modifié de façon durable. Il propose ensuite certaines considérations plus pratiques et dégage des suggestions concrètes de ces théories pour la construction de programmes de promotion de la santé. Finalement, en cours de texte, quelques réflexions sont présentées sur la nature particulière des programmes visant à changer des habitudes de vie.

Au cours des pages qui suivent, on emploiera indifféremment les vocables "éducation sanitaire" ou "promotion de la santé" pour désigner le champ de connaissance qui vise à influencer des comportements humains reliés à la santé. Pour reprendre une définition classique:

"l'éducation sanitaire est un processus qui vise à informer, motiver et aider la population à adopter et maintenir de saines habitudes de vie, à promouvoir les modifications à l'environnement facilitant ainsi l'acquisition de ces habitudes et à stimuler la formation professionnelle et la recherche dans cette optique."[3]

Soulignons que divers moyens incitatifs et coercitifs peuvent être employés pour influencer les habitudes de vie d'un individu ou d'une collectivité[4], et rappelons, pour conclure ces remarques préliminaires, que les "éducateurs sanitaires" s'attaquent à des habitudes et comportements tant au niveau de la prévention primaire qu'à ceux des soins curatifs et de la réadaptation. En effet, même si, en santé communautaire, l'on tend à mettre davantage d'emphase sur des populations non malades chez qui l'on tente de prévenir l'apparition de problèmes de santé, il ne faut pas oublier les importants travaux effectués auprès de personnes malades, comme par exemple ceux de Sackett et Haynes[5,6] sur la fidélité à divers régimes thérapeutiques.

Revue des principales théories expliquant la modification du comportement relié à la santé

L'éducation sanitaire se situant aux confins des sciences médicales et sociales, il n'est donc pas étonnant que, dans un récent article, Simonds[7] ait dénombré plus de vingt-cinq modèles théoriques – provenant de champs aussi divers que la psychologie sociale, la sociologie, l'anthropologie, la psychologie, l'économie politique, la médecine behaviorale, la géographie et les sciences de la communication – cherchant à expliquer la modification de com-

1) Chargé de cours, Département de médecine sociale et préventive, Université de Montréal, étudiant au Doctorat, Département de Sociologie, Boston University.

portements reliés à la santé. Il existe évidemment plusieurs autres théories visant à comprendre et prédire l'évolution des comportements humains en général; la revue de littérature qui suit ne tient cependant compte que des modèles ayant pour objectif spécifique la compréhension d'habitudes ou de comportement reliés à la santé, ces modèles étant déjà suffisamment abondants et diversifiés.

1. Une théorie psycho-sociologique: le "Health Belief Model" (HBM)

La psychologie sociale est probablement la discipline des sciences sociales qui a le plus marqué l'éducation sanitaire depuis une trentaine d'années. Historiquement, cela s'explique en partie par le fait que les premiers scientifiques à se pencher de près sur la modification de comportements reliés à la santé étaient des psychologues sociaux. Il faut de plus souligner que le modèle théorique que plusieurs d'entre eux ont commencé à élaborer au début des années 1950, le "Health Belief Model", a été depuis l'objet de raffinements continus et a inspiré plusieurs dizaines de travaux empiriques. Le HBM n'est évidemment pas le seul modèle psychosociologique se préoccupant de santé, mais il possède plusieurs des caractéristiques communes à la plupart des autres modèles[8] et revêt à ce titre une importance particulière; c'est donc cette théorie qui sera considérée ici.

Becker et ses collègues, qui ont depuis quelques années pris la relève des Hochbaum et autres pionniers du HBM, résumaient les caractéristiques du modèle en ces termes:

"... la décision d'adopter un comportement adéquat est influencée par: 1) la motivation générale d'un individu pour les questions de santé; 2) la perception qu'a un individu de sa vulnérabilité à une maladie; 3) la perception qu'il a de la sévérité d'une maladie et des séquelles qu'elle laisse; 4) les croyances envers l'efficacité relative de diverses actions entreprises face à une maladie; 5) l'évaluation faite par l'individu des différents coûts économiques, physiques et psychologiques associés à ces actions; 6) les réactions de l'individu à divers incidents ou événements (v.g. mort d'un ami atteint d'un problème de santé similaire) qui peuvent servir de "déclencheur" à son action. Il a de plus été observé que certaines caractéristiques personnelles (par exemple: l'âge, le sexe, le statut socio-économique, l'étendue des connaissances, la capacité de faire face aux difficultés) pouvaient influencer la motivation, la menace que représente la maladie ainsi que la perception des bénéfices de l'action préventive."[9]

Cette approche théorique met donc l'emphase sur la perception qu'une personne se fait de la réalité. Aussi subjective que cette perception puisse paraître par rapport à certains critères dits "objectifs", elle est néanmoins considérée comme le point critique pour influencer le comportement. Comme nous le verrons plus loin, le HBM a été abondamment critiqué pour cette approche de ne miser que sur des éléments internes à l'individu, afin de tenter d'influencer sa conduite en privilégiant des stratégies individuelles faisant peu ou pas de cas de l'environnement. Malgré son importance historique et son état relativement structuré, le HBM ne rallie donc plus autant l'unanimité qu'il l'a fait durant plusieurs années.

2. Modèles théoriques provenant de la recherche en "communication persuasive"

Au cours des dernières décennies, des sommes astronomiques ont été investies dans des projets de recherche visant à fournir aux compagnies, aux partis politiques et à d'autres groupes ou individus les moyens de "persuader" la population d'adopter un produit, un candidat, une idée. Tel que le propose Jacques de Guise[10] dans sa revue de ce champ de recherche, trois types de théories ont servi à encadrer les efforts des chercheurs: les théories de la balance, les théories du behaviorisme et les théories psychanalytiques.

Dans une de ses conclusions les plus frappantes, de Guise soutient qu'indépendamment du modèle théorique utilisé par les propagandistes ou les publicistes, il n'est pas possible dans nos sociétés pluralistes de faire adopter par la population n'importe quel produit simplement en y consacrant davantage de publicité, télévisée ou autre. Consciemment ou non, le consommateur développe toute une série de mécanismes de défense dont les plus connus, dans le jargon psychologique, se nomment "l'exposition sélective", la "perception sélective" et la "rétention sélective". On a donc généralement tendance à surestimer de beaucoup la capacité réelle de la publicité à "créer des besoins" et à faire mécaniquement adopter des comportements nouveaux. C'est d'ailleurs cette constatation que les spécialistes du "marketing social" rappellent régulièrement[11,12] à ceux qui pensent qu'il suffit d'inonder les ondes de messages favorables à la santé pour qu'automatiquement la population adopte de bonnes habitudes de vie.

Un autre des apports importants de la recherche en "communication persuasive" a été d'ébranler fortement l'idée qui prévalait à l'effet qu'un changement d'attitude chez une personne soit une condition nécessaire et suffisante à son changement de comportement. En effet, tel que le souligne Green à l'issue d'une analyse fouillée de la question,

"... les attitudes et les croyances sont fréquemment en contradiction avec les comportements; un changement d'attitude ou de croyance ne conduit pas inexorablement à un changement de comportement, de plus, même des décisions fermes, des engagements publics et des désirs fortement manifestés ne sont pas suivis de réelles modifications du comportement."[13]

En accord sur ce point avec de nombreux autres chercheurs, Green souligne qu'il faut non seulement se préoccuper des éléments internes, tels la motivation ou la volonté de changer, mais aussi des facteurs externes qui influencent l'individu. Selon lui, on se doit d'intervenir particulièrement auprès des différents groupes primaires (famille, collègues de travail, amis) auxquels une personne appartient. Butler et Paisley[14], après une exhaustive revue de la littérature, en arrivent à une conclusion relativement similaire: les éducateurs sanitaires, s'ils veulent agir efficacement pour modifier les comportements de leurs clientèles, doivent intervenir non seulement sur des facteurs personnels (perception, connaissance, affectivité, motivation) mais également sur les processus d'intégration d'un individu à divers groupes sociaux (famille, école, etc.) et sur son environnement social plus global, en particulier via les media qui véhiculent les normes et les valeurs de la société ambiante.

Les recherches en communication persuasive ont donné lieu à deux autres développements théoriques d'importance. D'une part, Kar[15] fait le bilan des recherches portant sur les mécanismes par lesquels une innovation technologique se répand – que ce soit l'adoption d'un nouveau médicament par les médecins d'une municipalité donnée, ou encore l'adoption du tracteur par les fermiers d'un pays en voie de développement. Parmi les découvertes importantes

de ces recherches, notons le fait qu'il existe dans chaque communauté des personnes ayant un accès privilégié à l'information et qui la décodent et l'interprètent pour le reste de la communauté – mécanisme désigné sous le nom de "communication à deux paliers" ou "two step flow". Ceci fait ressortir une fois de plus l'importance de l'entourage d'un individu que l'on veut influencer et fait de ces communicateurs privilégiés des personnes de choix pour une intervention éducative, l'effet multiplicateur étant ici potentiellement très important.

Finalement, l'utilisation de messages éveillant la peur chez un individu est souvent suggérée par des professionnels de la santé comme un moyen potentiellement efficace pour susciter des changements de comportement. De nombreuses recherches sur le sujet ont été faites depuis les premières expériences de Janis et Feshback[16] au début des années 1950. Richards concluait récemment, après avoir considéré les quelque vingt ans de recherches effectuées sur le sujet, que "... faire appel à la peur a parfois changé quelque chose et parfois non."[17] Il a été relativement bien démontré cependant qu'il existe un seuil au-delà duquel les personnes sont si effrayées par le message qu'on leur présente qu'elles en oublient même la solution qu'on leur proposait pour lutter contre un comportement néfaste.

3. Modèle théorique s'appuyant sur l'économie politique

Un des modèles les plus récents et les plus importants pour ceux qui s'intéressent à la modification de comportements reliés à la santé se retrouve dans les travaux d'auteurs tels Navarro, Waterman et Waitzkin, Krauze et quelques autres.[18] Le mérite principal de ces auteurs est d'avoir mis l'accent sur la compréhension de l'échec de la plupart des programmes de modification d'habitudes de vie[19]. Dans leurs écrits, on se rend clairement compte que la santé, comme toutes les autres ressources, est inégalement distribuée entre une minorité riche et en bonne santé et une majorité beaucoup moins riche, et en moins bonne santé.

McKinlay[20], dans la même veine que les auteurs précédents, replace la promotion de la santé dans le contexte de l'économie politique des sociétés capitalistes avancées et soutient que les actions éducatives ont beaucoup trop visé à "blâmer la victime", reprenant en cela l'expression de Ryan[21] et Crawford[22]. En effet, les pressions des valeurs et des normes que véhicule la société ambiante, elles-mêmes modelées puissamment par l'influence de corporations dont l'intérêt premier est de s'accaparer d'un maximum de profit, ne font "... qu'encourager et renforcer les comportements mêmes que nous visons à changer"[23]. Il devient donc crucial pour les éducateurs sanitaires, toujours selon McKinlay, de ne pas seulement chercher à transformer l'individu mais aussi, et surtout peut-être, de voir "... de quelle manière il serait possible de contrôler la puissance démesurée du lobbying de certains groupes d'intérêts dans le monde de la santé"[24].

La compréhension des grandes forces économiques et politiques qui modèlent l'environnement social doit donc venir compléter la compréhension des mécanismes psychologiques de l'individu que l'on vise à changer. L'action éducative, selon ces auteurs, se doit par conséquent d'être exercée autant au niveau de l'environnement social plus général qu'à celui de l'environnement plus immédiat de la famille ou du groupe d'amis, ainsi que le soulignaient plus haut les théories développées dans le champ de la communication persuasive. Comme exemple de changement de cet environnement social général, Roter et Wang[25] soulignent comment des changements de pratiques dans le domaine sanitaire ont été stimulés avec succès par les campagnes massives d'éducation politique visant à façonner les valeurs d'un ordre nouveau en Chine Populaire.

Il est bien évident qu'une telle façon d'envisager la modification d'habitudes de vie est loin de recueillir l'accord général. On reproche souvent à ce genre de modèle d'être de peu d'intérêt pratique, en plus des résistances émotives et idéologiques qu'il rencontre dans plusieurs pays où l'on rejette sans autre forme de procès tout ce qui a une quelconque saveur socialisante. Il n'en demeure pas moins que les débats les plus stimulants des dernières années ont été suscités par ce genre de théories, et qu'elles ouvrent des pistes d'interventions fort intéressantes à explorer.

4. Colorations théoriques particulières à la santé publique et à la santé scolaire

Bien que ce ne soit pas le cas au Québec, les programmes de promotion de la santé sont divisés dans plusieurs pays entre deux groupes bien distincts de professionnels. Le premier de ces groupes est composé d'enseignants et se préoccupe d'éducation sanitaire en milieu scolaire, concentrant ses interventions sur l'information concernant l'hygiène et la santé ainsi que sur l'éducation physique. L'autre groupe, formé principalement de professionnels de la santé, s'occupe de la prévention à l'extérieur du milieu scolaire. Aux Etats-Unis, par exemple, ces deux aspects de la promotion de la santé sont clairement séparés, chaque groupe de professionnels ayant ses propres associations, ses propres revues spécialisées ainsi que ses propres programmes universitaires. Bien que ce ne soit pas absolument tranché, on peut relever chez chacun de ces deux groupes d'éducateurs sanitaires certaines préférences théoriques et idéologiques.

Du côté de la santé publique, l'épidémiologie, la statistique et les principes d'organisation des soins de santé sont les principales disciplines de base utilisées. Il n'est donc pas étonnant de constater qu'une théorie épidémiologique de la modification des comportements ait été suggérée; à la suite d'une analyse des recherches faites au début des années soixante afin d'appliquer le modèle épidémiologique à des problèmes particuliers (prévention des accidents, santé mentale, nutrition) Lauzon[26] a développé une théorie plus globale applicable à n'importe quelle habitude de vie. Reprenant une terminologie familière au monde de la santé publique, il souligne que

► p. 739

l'on doit rendre "l'hôte" (la personne à éduquer) plus résistant, que l'on doit modifier la composition de "l'agent" (cigarette, automobile, alcool, nourriture, etc.) de manière à en diminuer l'effet nocif, et que, finalement, on doit créer une barrière environnementale (au sens physique et social du terme) afin d'empêcher l'agent de venir en contact avec l'hôte.

Une autre transposition directe des approches développées en santé publique est suggérée par Bass[27]; dressant une liste des dix conditions requises pour réussir un programme de contrôle des maladies infectieuses, il suggère d'appliquer les mêmes critères à des problèmes non infectieux en prenant le contrôle du tabagisme comme exemple.

Finalement, bien que les méthodes behaviorales basées sur la théorie de l'apprentissage aient connu de nombreux développements dans le monde de l'éducation en général, les éducateurs sanitaires des milieux scolaires n'y ont pas consacré autant d'intérêt que ceux de la santé publique. Dans un article classique, Pomerleau et coll.[28] présentent de nombreux programmes qui ont été développés pour contrôler divers problèmes de santé grâce aux méthodes dérivées des travaux de Pavlov, Watson ou Skinner; ils dégagent de plus certaines avenues de recherche dans lesquelles les spécialistes du conditionnement devraient s'engager afin d'aider ceux qui le désirent à modifier des habitudes de vie jugées néfastes.

Sans être trop catégoriques là-dessus, les éducateurs sanitaires oeuvrant dans les milieux scolaires semblent pour leur part se rattacher davantage à un type d'intervention plus empathique, où l'implication consciente et totale de l'individu est valorisée. Un exemple typique du genre de théories que l'on retrouve dans les revues spécialisées en "School Health Education" propose à la personne de mettre en lumière les valeurs qui guident son action afin de pouvoir les changer[29]; selon cette stratégie nommée "value clarification", les individus d'un groupe sont encouragés à franchir sept étapes où ils choisissent, discutent et mettent en application une nouvelle habitude de vie.

Un autre bon exemple des théories empathiques qu'on retrouve dans cette littérature est tiré de la théorie écologique de Hoydman[30]: après avoir passé en revue cinq conceptions de la nature humaine (les modèles chrétien, psychanalytique, behavioriste, humaniste et existentialiste), il suggère une approche où l'être humain est perçu comme évoluant, s'adaptant, agissant, coexistant, se réalisant et se transcendant, en relation avec son environnement.

► p. 742

Autres modèles théoriques

Jusqu'ici nous avons vu un certain nombre de théories inspirées par la psychologie sociale, la communication persuasive et l'économie politique; nous avons de plus constaté que les éducateurs sanitaires oeuvrant à l'école et ceux oeuvrant à l'extérieur de l'école avaient des orientations théoriques quelque peu différentes. Ces courants théoriques, bien qu'indiquant les principales tendances de l'heure, n'épuisent nullement l'ensemble des modèles disponibles au praticien qui cherche à se dépêtrer, à travers les dédales de la littérature scientifique, pour trouver des indications utiles à l'organisation de programmes d'intervention.

Par exemple, certains auteurs proposent une approche "holistique"[31], où l'on exalte la capacité des groupes et des individus à prendre eux-mêmes leur santé en main, en ayant recours le moins possible à des professionnels (activités de "self-help", qui ont déjà pris chez nos voisins du sud l'ampleur d'un mouvement social). Des techniques de bio-feedback, de méditation transcendentale, ainsi que d'autres moyens relativement inusités sont aussi proposés par les tenants de cette approche, afin d'aider la population à modifier certaines habitudes de vie.

Les super-marchés de la théorie restent cependant les congrès d'éducateurs sanitaires, où l'on est littéralement enseveli sous une avalanche de nouveaux modèles: l'approche par les besoins, l'analyse transactionnelle appliquée à la promotion de la santé, le modèle des risques calculés, une théorie basée sur la philosophie utilitaire, une panoplie de modèles dit "humanistes", une théorie systémique, une approche de de simulation basée sur la théorie des jeux, etc.[32].

Comment s'y retrouver à travers tout ce fatras? Y a-t-il des tendances générales? Certaines théories sont-elles meilleures que d'autres? C'est ce que la seconde partie de cet article explorera.

NDLR: La seconde partie de cet article paraîtra dans un prochain numéro de L'Union Médicale du Canada.

La modification de comportements reliés à la santé

II. Suggestions pour l'organisation de programmes de santé communautaire

Michel O'Neill[1]

Il est évident que les modèles théoriques évoqués dans la première partie de l'article n'ont pas tous le même degré de raffinement ou de validité empirique. Certains modèles sont conçus avec une vision davantage scientifique – i.e. que les auteurs tentent de vérifier à l'aide de procédures scientifiquement contrôlées – alors que d'autres modèles sont davantage au niveau de la prise de position idéologique ou philosophique, – i.e. que les auteurs ne cherchent pas nécessairement à valider empiriquement et scientifiquement. Au cours des paragraphes qui suivent, nous tenterons donc de voir dans quelle mesure les diverses théories évoquées jusqu'ici peuvent fournir au praticien des indications concrètes pour la construction de programmes de promotion de la santé.

1. Degré d'applicabilité des différentes théories

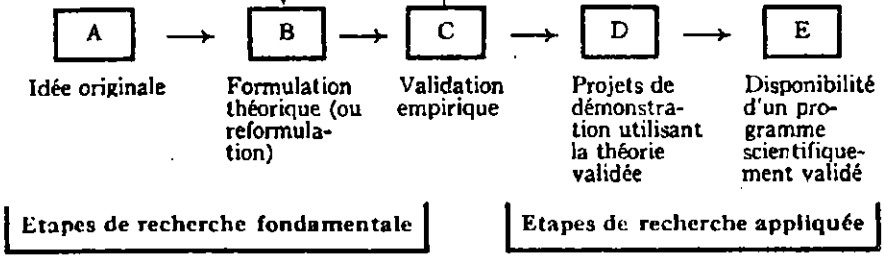
Afin de pouvoir évaluer la contribution possible de divers modèles théoriques à la pratique, il est utile d'établir un certain nombre d'étapes que toute théorie doit franchir pour encadrer de façon systématique des stratégies d'intervention. Un cheminement en cinq étapes est présenté dans le schéma ci-contre.

Prenons le "Health Belief Model", afin d'illustrer l'utilité d'une telle approche.[2] Les tenants du HBM ont d'abord eu l'idée qu'appliquer les concepts du psychologue Kurt Lewin au champ de la modification de comportements préventifs était potentiellement fort valable (étape A). Ils **formulèrent** donc une théorie où un certain nombre de variables indépendantes – susceptibilité perçue à la maladie; vulnérabilité perçue, etc. – devait selon eux permettre d'expliquer et même de prédire la variable dépendante: l'adoption d'un comportement préventif adéquat (étape B). Une première série de sept **études empiriques** fut ensuite mise sur pied afin de voir si réellement la théorie tenait bon lorsque confrontée à la réalité (étape C). Les résultats de ces recherches tendaient grosso-modo à confirmer la théorie, mais aucune recherche à elle seule ne fournissait une confirmation complète du modèle. En conséquence, cela ramena les chercheurs une étape en arrière et ils **reformulèrent** leur théorie, raffinant leurs variables, en ajoutant d'autres (étape B à nouveau) et, mettant en route d'autres projets de validation (étape C à nouveau). Le produit final de ces ajustements successifs de la théorie est le modèle "révisé" tel qu'il existe à l'heure actuelle, après trente années, et dont nous avons exposé plus haut les éléments principaux.

Il apparaît donc que, même si des applications concrètes peuvent être déduites de ces recherches réalisées la plupart du temps avec un souci de recherche fondamentale, bien peu d'efforts ont été investis afin de transformer ces acquis théoriques en **programmes bien concrets** (étapes D et E). Pour formuler autrement ce problème, on sait par

Schéma 1

Etapes à franchir afin de transformer une théorie en programmes scientifiquement validés



1) Chargé de cours, Département de médecine sociale et préventive, Université de Montréal. Etudiant au doctorat en sociologie, Boston University.

NDLR — Nous reproduisons ici la suite d'un article dont la première partie a été publiée dans le numéro de mai 1980 en page 733 et suivantes. Nous prions le lecteur de prendre note des particularités suivantes: 1 — Le schéma 1 reproduit ci-dessus a été publié par erreur dans le corps de la première partie, en page 736. En fait, il n'aurait dû paraître qu'une seule fois, tel que reproduit ici. 2 — La bibliographie se rapportant à l'ensemble des deux parties de l'article se retrouve en totalité dans ce numéro, en pages 927 et 928.

exemple avec certitude, à partir de la recherche plus fondamentale effectuée à l'aide du HBM, que la perception individuelle de la vulnérabilité à la maladie influence certainement la probabilité de modification du comportement; néanmoins, on ne sait pas vraiment **comment**, de façon systématique, un éducateur sanitaire peut augmenter au sein d'une population la vulnérabilité perçue à une maladie, afin de susciter le changement d'habitude de vie souhaité. Et l'on ne connaît pas non plus le poids relatif des différentes variables indépendantes du modèle: doit-on jouer sur la motivation générale envers la santé, davantage que sur la vulnérabilité perçue ou la sévérité perçue? Doit-on faire jouer toutes ces variables à la fois? Se renforcent-elles ou se nuisent-elles l'une l'autre? Voilà autant de questions concrètes que le praticien, aux prises avec un programme à mettre sur pied, voit sans réponse.

Si l'on porte maintenant un regard général sur l'ensemble des théories disponibles en éducation sanitaire, en utilisant le cheminement en cinq étapes proposé plus haut, force nous est de constater qu'un très grand nombre de modèles en sont encore au stade de l'idée potentiellement fructueuse (étape A) et que quelques-uns seulement ont franchi le cap d'une formulation rigoureuse (étape B). Seules quelques rares exceptions en sont rendues à l'étape de validation empirique (étape C), ces projets ayant souvent des carences méthodologiques importantes. Quant aux étapes D et E, on peut presque dire que rien ne se fait de façon systématique à ce niveau. Malgré cet état relativement peu avancé des connaissances, certaines grandes tendances et certains éléments de consensus se dégagent de la littérature et offrent aux praticiens, sinon des recettes infaillibles, du moins des points de repère valables. Quelques remarques préalables sur la nature particulière des programmes visant à changer des comportements et des habitudes de vie s'imposent cependant.

2. Nature particulière des interventions en éducation sanitaire

Dans les domaines de la médecine et de la santé publique, les professionnels ont été habitués à l'efficacité immédiate de types d'interventions à l'aide desquelles un problème de santé spécifique est solutionné par une technique, préventive ou curative, bien déterminée. Un bon exemple de ceci, qui demeure l'archétype de l'intervention médicale, est le vaccin: le microbe étant isolé et connu, le vaccin développé, il ne suffit que d'un petit nombre d'injections et le problème est réglé. Les problèmes de santé auxquels nous avons à faire face aujourd'hui sont cependant très différents de ceux du passé et la façon de les solutionner doit évoluer en conséquence.

En effet, la cible principale des interventions professionnelles n'est plus, dans nos sociétés industrialisées du moins, cette kyrielle de maladies infectieuses contre lesquelles les interventions médicales ont eu un certain succès. Une certaine impatience se manifeste dans le monde de la santé lorsque l'on réalise, comme nous l'avons fait dans les pages précédentes, qu'il n'y a pas de solution simple et mécanique pour intervenir avec succès et efficacité sur les habitudes de vie qui sont reliées aux grandes épidémies modernes (cancers, accidents, maladies cardio-vasculaires). Il n'y aura jamais de "vaccin" contre l'alcoolisme, une mauvaise alimentation ou l'absence d'exercice physique.

Qui plus est, les problèmes éthiques étaient à peu près absents des interventions classiques en santé publique; il y avait bien quelques originaux qui refusaient de se prêter de bon gré aux interventions, comme dans le cas des campagnes de vaccination, par exemple, mais on ne rencontrait pas les vagues massives d'opposition, que des programmes tels que la fluoruration des eaux de consommation ou le port obligatoire de la ceinture de sécurité soulèvent. De nos jours, ces problèmes éthiques deviennent de plus en plus délicats et aigus, à mesure que l'on se mêle de la vie privée des gens. Il devient donc important aujourd'hui de réfléchir sur les limites de l'intervention des professionnels de la santé publique, surtout quand ils s'arrogent le droit, au nom de vérités "scientifiques", d'inciter ou même de forcer la population à adopter des habitudes de vie que souvent ils ne pratiquent pas eux-mêmes.

Il est donc nécessaire de réaliser que le type de problèmes de santé auxquels sont confrontées les sociétés modernes est de *nature* différente et nécessite des interventions de *nature* différente qui n'auront plus jamais la même certitude et la même efficacité mécanique que les vaccins d'antan et, c'est dans cet-indications concrètes présentées ici. te optique qu'il faut interpréter les

3. Cinq suggestions pratiques pour la construction de programmes de promotion de la santé

1) Tous les théoriciens s'entendent pour dire que modifier de façon durable un comportement humain est une tâche **extrêmement complexe**, spécialement lorsque ces comportements sont profondément ancrés dans la culture et quotidiennement valorisés ou renforcés par l'environnement socio-économique. C'est une vérité que l'on oublie trop souvent en organisant des programmes de promotion de la santé dont l'on s'étonne ensuite du peu de succès.

2) L'approche postulant qu'il faut donner de l'information afin de changer les connaissances, ce qui entraînerait un changement d'attitudes menant plus ou moins inévitablement à l'adoption de nouvelles habitudes de vie, est connue sous le nom de "K.A.P." (Knowledge, Attitudes, Practices) dans la littérature. En plus des objections de Green mentionnées plus haut, Simonds, un des chercheurs les plus respectés en éducation sanitaire, concluait récemment qu'il importait "... de voir que l'approche des changements linéaires à partir des connaissances jusqu'au comportement, en passant par les attitudes, est trop simpliste pour planifier la plupart des interventions".[34]

Pourtant, comme on l'a déjà démontré ailleurs[35], c'est le cadre conceptuel auquel, intuitivement, un grand nombre de professionnels de la santé se rattachent encore bien que son inexactitude ait été démontrée **ad nauseam.**

3) Il est important de noter que travailler simplement à changer l'individu n'est pas suffisant, étant donné les résistances extrêmement importantes que l'environnement socio-économiques oppose au changement d'habitudes de vie. Même Hochbaum qui a longtemps été le principal bâtisseur du HBM, modèle qui, rappelons-le, mise quasi essentiellement sur les perceptions propres à l'individu sans faire référence à son environnement, admettait récemment:

"... nous en sommes venus à admettre et à tenir compte de l'importance très grande de facteurs qui ne sont pas aisément modifiables par les mesures traditionnelles d'éducation sanitaire, et qui influencent directement et de façon décisive les comportements; ces facteurs sont les conditions sociales, politiques, économiques et environnementales, incluant même la façon dont les services sont dispensés – en résumé un ensemble

Suite p. 927 ►

de facteurs qui façonnent la vie quotidienne des gens – et ils poussent les individus à accomplir des actions qu'ils veulent parfois éviter tout en rendant des actions qu'ils désirent entreprendre difficiles ou impossibles."[36]

On semble donc s'entendre pour adopter une approche que l'on nommera "holistique", "écologique", ou du nom que l'on voudra, mais où les facteurs environnementaux, autant que les facteurs individuels, entrent en ligne de compte et, font l'objet de programmes. Les auteurs ne tombent pas tous d'accord sur le nombre de facteurs à prendre en considération, ni sur la manière de les regrouper[37,38,39,40], mais l'on s'éloigne de plus en plus des positions dogmatiques où des chercheurs prônaient de ne s'occuper que de l'individu, ou, que de l'environnement.

4) Une approche multifactorielle à la promotion de la santé requiert des interventions simultanées à divers niveaux. La principale conclusion de l'Organisation Mondiale de la Santé à propos des programmes d'éducation sanitaire, mis sur pied depuis une vingtaine d'années dans des pays en voie de développement, souligne l'inutilité de travailler à modifier des comportements reliés à la santé, sans s'attaquer en même temps **aux autres dimensions du développement économique et social**[41]. Ces conclusions sont aussi valables pour les pays développés, où l'on a tendance à oublier l'importance fondamentale des facteurs socio-économiques dans la détermination de l'état de santé d'une population. Si l'on parle parfois de ces facteurs, on s'en désintéresse généralement quand vient le temps d'intervenir, se contentant de petits programmes de changements individuels, plus simples à conduire certes, mais sans effet durable.

Dans le même ordre d'idée, il faut déployer simultanément une multiplicité de stratégies éducatives. La plupart des programmes qui ont eu quelque succès ont été de longue durée, les messages éducatifs y étaient véhiculés par une multitude de canaux qui se renforçaient l'un l'autre et la participation des éduqués, bien qu'elle demeure toujours problématique à obtenir (cf. à ce sujet Bégin et al[42] pour le cas du Québec), était une réalité vécue et n'était pas juste un voeu pieux ou, une façon déguisée de vendre à la population un programme tout prêt d'avance[42].

5) Pour le praticien voulant construire un programme en regard d'une habitude de vie précise, il existe une quantité considérable de recherches empiriques et d'analyses de programmes dans la littérature. Ces recherches sont de qualité inégale et reposent peu souvent sur des théories bien articulées mais elles constituent néanmoins une somme intéressante d'essais et d'erreurs qui ont été faits par le passé. On ne peut s'attendre à y trouver la recette miracle qui garantira le succès de son programme, mais on y retrouvera à coup sûr certaines choses à **ne pas faire.** Ainsi, des erreurs grossières qui ont été faites à répétition un peu partout pourraient être évitées si l'on prenait quelques heures pour réviser cette littérature avant de se lancer dans un nouveau programme. Sur des sujets tels la cessation de fumer, la planification et le contrôle des naissances ou l'alcoolisme, pour ne prendre que quelques exemples, des masses d'informations sont disponibles et ont déjà été synthétisées et classifiées[44].

Conclusion

Nous avons vu que les programmes d'éducation sanitaire et de promotion de la santé sont devenus fort populaires récemment et risquent de le demeurer pour un certain temps encore. Leur capacité à diminuer les coûts des soins de santé commence à être documentée avec plus de précision[45]. Néanmoins, tel que cet article l'a souligné, ce champ attend encore que les diverses disciplines des sciences sociales et médicales produisent les connaissances permettant de construire des programmes efficaces à coup sûr. Pourtant, les quelques suggestions pratiques évoquées plus haut, en dépit de leur apparente simplicité, sont loin d'être connues et encore moins transposées dans la pratique quotidienne des professionnels de santé publique au Québec et ailleurs au Canada. Tel que le soulignait Badgely[46], nous avons fait bien peu de chemin depuis que l'O.M.S., en 1954, publiait son premier rapport d'experts sur l'éducation sanitaire[47].

Il faut donc que la recherche développe des modèles plus solides et davantage orientés vers une application possible par les praticiens de santé communautaire. Il semble cependant encore plus essentiel qu'un profond changement d'attitude survienne chez ces mêmes praticiens; ils doivent cesser de rêver au jour où l'on pourra faire changer le comportement humain à volonté et partager avec les chercheurs les frustrations, les complexités et les dilemmes éthiques qui se posent lorsque l'on travaille à vouloir modifier des habitudes de vie solidement ancrées. Ce changement d'attitude permettrait aux praticiens de faire un meilleur usage de la littérature disponible qui, malgré ses limites, n'en renferme pas moins de précieuses indications. De plus, cette littérature témoigne, tel que le souligne Philip[48], que c'est à la fois sur des prises de positions philosophiques et sur des connaissances scientifiques que se fondent les interventions éducatives. Des changements importants doivent donc être apportés à la pratique de l'éducation sanitaire au Québec sinon, et pour longtemps encore, les programmes de modification de comportements reliés à la santé sont condamnés à échouer lamentablement, comme ce fut si souvent le cas pas le passé.

Remerciements

La production de cet article a été rendue possible grâce à une bourse du "Programme National de Recherche et Développement en matière de Santé" du Ministère fédéral de la Santé et du Bien-être social, qui permet à l'auteur de poursuivre des études doctorales en sociologie de la santé (Boston University). L'auteur remercie Pierre Joubert, Raynald Pineault, Sol Levine et Fernand Turcotte pour leur critique d'une version préliminaire de cet article, ainsi que Marie O'Neill pour son efficace travail de secrétariat.

Summary

This paper reviews the major theoretical models available to understand the mechanisms through which health related behaviors and lifestyles can be influenced. A critical analysis of the relative value of these theories for practice is presented, as well as concrete hints to organize health promotion programs. Finally, a few reflections upon the special nature of health education interventions are offered.

Bibliographie

1. Badgely, Robin F.: Health Promotion and Social Change in the Health of Canadians, in Proceedings of the Seminar on Evaluation of Health Education and Behaviour Modification programs, Canadian Health Education Society and Canadian Council on Health Education, Ottawa, pp. 20-43, décembre 1978.
2. Lalonde, Marc: Nouvelle perspective de la santé des Canadiens, Ministère fédéral de la Santé et du Bien-être Social, Ottawa 1974.
3. Somers, Ann (ed): Health Promotion and Consumer Health Education; New York: Prodist; p. 3, 1976

fun volume du task force "Preventive Medecine U.S.A.") (traduction libre).
4. O'Neill, Michel: La modification des comportements de santé: perspectives d'intervention au Québec. Cahiers santé communautaire, 1: 69-94, 1978.
5. Sackett, B.L. et Haynes, R.B.: Compliance with Therapeutic Regimens: Baltimore; Johns Hopkins University Press, 1976.
6. Haynes, R.B.: The Influence of Health Education on Compliance with Therapeutic Regimens; in Proceedings of the Seminar on Evaluation of Health Education (op. cit.), pp. 43-71.
7. Simonds, Scott; Emerging Challenges in Health Education; International Journal of Health Education, 19 (4): 1-18 Special supplement, 1976.
8. Becker, Marshall et coll.: Selected Psychosocial models and correlates of individual health related behavior. Medical Care, 15 (5): 27-49. Pour une discussion en profondeur du HBM, voir les divers articles in Becker, Marshall (ed), The Health Belief Model and Personal Health Behavior; Charles B. Slack; Thorofare, N.J., 1974.
9. Becker et coll.: op. cit.; p. 28-29.
10. De Guise, Jacques: Communication de masse et persuasion in Cossette, Claude (ed); Communication de masse, consommation de masse; Boréal Express, Sillery pp. 75-111, 1975.
11. Lussier, Denis: Marketing et santé communautaire; Communication présentée au congrès de l'Association pour la santé publique du Québec, Montréal, octobre 1978.
12. Grover, P. et Miller, Jean: Guidelines to Make Health Education Work, Public Health Rep., 91 (3): 249-253, 1976.
13. Green, Lawrence: Should Health Education Abandon Attitude Change Strategies? Perspective from Recent Research, Health Education Monographs, 1 (30): 25-46, 1970 (traduction libre).
14. Butler, Matilda et Paisley, William; Communicating Cancer Control to Public, Health Education Monographs, 5 (1):5-25, 1977.
15. Kar, Snehendu: Implications of Diffusion Research for Planned Change. Internat. J. Health Educ., 17 (3): 192-220, 1974.
16. Janis, Irwing L. et Seymour, Feshbach; Effects of Fear Arousing Communication; J. Abnormal Social Psychol. 48 (1): 78-92, 1953.
17. Richards, David N.: Methods and Effectiveness of Health Education: the Past, the Present and the Future of Social Scientific Involvement. Social Sci. Med. 9 (3): 141-156, 1975.
18. Voir par exemple: Navarro, Vincente; Medicine under Capitalism, Prodist, New York, 1977. Waitzkin, Howard et Waterman, Barbara; The Exploitation of Illness under Capitalism. Boobs-Merrill, New York and Indianapolis, 1976. Kauze, Elliot, A., Power and Illness; Elsevier, New York, 1977.
19. Feudenberg, Nicholas: Shaping the Future of Health Education: from Behavioral Change to Social Change; Health Educ. Monographs, 6 (4): 372-378, 1979.
20. McKinlay, John B.: A Case for Refocusing Upstream: the Political Economy of Illness, in Enelow, A.S. et Henderson B.J. (eds); Applying Behavioral Science to Cardiovascular Risk; New York, American Heart Association, 1975.
21. Ryan, William: Blaming the Victim, New York, Vintage, 1976 (édition révisée).
22. Crawford, Robert: You are Dangerous to your Health: the Ideology and Politics of Victim Blaming. Internat. J. Health Serv., 7 (4): 663-80, 1977.
23. McKinlay, John: Op. cit.; p. 27 (traduction libre).
24. Ibid, p. 28 (traduction libre).
25. Roter, Debra et Li Wang, Virginia: Health Education Contrasts in the U.S. and the People's Republic of China. Internat. J. Health Education, 20 (2): 90-97, 1977.
26. Lauzon, Richard: An Epidemiological Approach to Health Promotion. Can. J. Public Health, 68: 311-316, 1977.
27. Bass, Frederic: A Public Health Approach to Illness Producing Behavior: the Example of Tuberculosis Control Applied to Cigarette Smoking. Consumer Health Education and Health Promotion, op. cit., pp. 148-161.
28. Pomerleau, B. et coll.: The Role of Behavior Modification in Preventive Medecine. New Engl. J. Med., 292: 1277-1282, 1975.
29. Osman, Jack D.: The Use of Selected Value Clarifying Strategies in Health Education. J. School Health, 44: 21-25, 1974.
30. Hoydman, Howard S.: Models of Human Nature and their Impact on Health Education. J. School Health, 44: 374-381, 1974.
31. Brown, John F.: Autres approches à la promotion de la santé, in Education Sanitaire, Santé et Bien-être social Canada, 15 (2-3): 1-24, 1976.
32. Un échantillon de ce que l'on peut trouver en feuilletant les: Abstracts of the IXth International Conference on Health Education, Ottawa, Canada 1976.
33. L'article de Becker et coll.: op. cit., a été utilisé pour dégager les grandes lignes du HBM.
34. Simonds, Scott: op. cit., p. 4.
35. O'Neill, Michel: Vers une problématique de l'éducation sanitaire au Québec; Thèse de maîtrise en Sociologie, Université Laval, 1976.
36. Hochbaum. Godfrey: At the Treshold of a New Era, paper presented at the 25th Meeting of [illegible]; cité par Simonds, loc. cit., p. 16 (traduction libre).
37. Richards, David N.: op. cit., p. 153.
38. Young, Majorie: Review of Research and Studies Related to Health Education Health Education Monographs, 1 (24): 1967. Numéro complet consacré à ces questions.
39. Kar, Senhendu: A Model for Intervention Strategy in Health Programs. International Union for Health Education, Abstracts of the IXth International Conference on Health Education op. cit., p. 110-111.
40. Point de vue de McKinlay, cité par Richards, David N:; op. cit., p. 153.
41. World Health Organization; Health Education: a Program Review, Geneva, 1974, Offset Publication no. 7.
42. Bégin, Clermont et coll.: L'expérience de la participation dans les établissements de santé de services sociaux au Québec: un bilan. Can. J. Public Health, 69 (6): 445-452, 1978.
43. Simmons, Jeannette (ed): Making Public Health Education work, special supplement of the Amer. J. Public Health, 5: 1-48, 1975.
44. Les revues analytiques de la littérature éditées par Majorie Young (Health Education Monographs, vol. 1 nos 23-28 et 34-35) ainsi que plusieurs publications techniques de l'OSM fournissent d'excellents points de départ pour bien des sujets.
45. Rocella, Edward: Potential Reducing of Health Care Costs by Public and Patient Education. Public Health Reports, 91 (3): 223-225, 1976.
46. Badgely, Robin F.: Health Promotion Evaluation in Canada, University of Toronto, 1975 (Research Proposal).
47. Organisation mondiale de la santé; Rapport du comité d'experts de l'éducation sanitaire de la population, Genève, 1954. Rapport technique no. 89.
48. Philip, Lois: Principles of Health Education: a critical review Internat. J. Health Education, 17 (2): 106-113, 1974.

Tiré de : O'Neill, Michel ; Vers une problématique de l'Education sanitaire au Québec, Montréal; Librairie de l'Université de Montréal ; 1979; pp. 18-40.

CHAPITRE II: L'histoire de l'éducation sanitaire.

CHAPITRE II: L'histoire de l'éducation sanitaire.

L'éducation sanitaire, du moins sous l'acception selon laquelle elle est considérée le plus fréquemment à l'heure actuelle, a une histoire relativement courte bien que, comme pour toute autre activité humaine, on puisse toujours remonter à des origines lointaines, où certains tabous religieux des sociétés dites primitives étaient liés indéniablement à des pratiques hygiéniques décourageant certains comportements mettant en péril la survie de l'espèce. Le but du présent chapitre est de brosser à très larges traits les principales étapes de l'organisation plus systématique de groupes et de programmes s'intéressant à l'éducation sanitaire un peu partout à travers le monde, dans un premier temps, et au Québec dans un second. Il faut envisager cette entreprise comme une mise en contexte permettant d'éviter de croire que les années 1970 au Québec sont le moment et le lieu où tout a commencé; ce n'est donc pas une analyse historique minutieuse et fouillée que trouvera le lecteur dans la première partie du chapitre, mais plutôt une esquisse des grandes tendances décelables mondialement.

A) L'éducation sanitaire à travers le monde

Existant depuis toujours à travers les préceptes religieux et les tabous au niveau des collectivités, préoccupation fréquente du médecin qui expliquait à son patient des règles de base d'hygiène ou qui tentait de le motiver à suivre avec précaution le traitement prescrit, l'éducation sanitaire n'a commencé à se définir et à se percevoir comme une fonction spécifique qu'au début du XXème siècle. Suivant les vagues de faveur et de défaveur de la médecine préventive, dont elle est partie constituante, c'est dans la mesure où les gouvernements jugeaient prioritaire d'intervenir pour préserver la santé publique, ou au contraire

dans la mesure où ils laissaient entre les mains des professionnels de la santé la responsabilité de ce champ que l'éducation des masses à une meilleure santé est apparue.

Gilbert (1) souligne que c'est en 1918 qu'un groupe de membres de l'Académie de Médecine de New-York lançait la première campagne d'éducation sanitaire du public en Amérique. C'est surtout dans les écoles que travaillèrent ces premiers éducateurs, ce qui se traduit encore aujourd'hui par le fait qu'une forte proportion des publications en éducation sanitaire concerne des activités en milieu scolaire, si bien qu'un ouvrage américain relatant l'histoire de l'éducation sanitaire chez nos voisins du sud (2) ne traite de fait que de l'introduction lente et laborieuse de cours d'hygiène dans les programmes scolaires des enfants. Les Américains ont donc pris une bonne longueur d'avance dans la mise sur pied de programmes systématiques d'éducation sanitaire des masses. Encore aujourd'hui, par le nombre de milieux universitaires où il est possible d'obtenir une formation en "Health Education" (plus d'une centaine d'universités offrent des diplômes de 1er, 2e ou 3e cycles en éducation sanitaire), par le renom de l'association qui regroupe les éducateurs (SOPHE: Society for Public Health Education) et qui publie depuis 1951 les "Health Education Monographs" où sont consignés des résultats de recherches et des travaux théoriques de grande envergure, par le nombre impressionnant de travaux de recherches qui y sont conduits, par le rayonnement international de leurs interventions auprès des pays du Tiers Monde, on peut dire que nos voisins du Sud jouent un rôle dominant dans ce secteur de l'activité humaine.

(1) Gilbert, Jules: Education Sanitaire, théorie et pratique, Masson et Cie, Paris, 1959, p. 12.

(2) Means, Richard K.: A history of health education in the United States, Lea and Tebiger, 1962.

"L'éducation sanitaire a pris, en France, sa formation organique et pleinement officielle après la libération, (bien que) cette organisation ait été précédée par une longue période d'initiatives et d'efforts privés et publics remontant au début de l'ère pastorienne".(1) Sous l'impulsion d'un certain nombre de personnes qui, tant au plan national qu'au plan international, en ont fait la promotion, l'éducation sanitaire a peu à peu gagné sa place au soleil en France et la structure de son organisation faisait, en 1967, l'objet d'un voyage d'observation d'une délégation québécoise.(2)

Dans les pays de l'Europe de l'Est, la période qui a suivi la seconde grande guerre semble avoir été fertile en réorganisations; contrairement aux pays occidentaux où l'importance et la crédibilité qu'on accorde à la prévention par l'éducation et par diverses autres méthodes fluctuent, parfois même d'une décade à l'autre, elle semble s'intégrer tout normalement à la pratique des professionnels et orienter de façon très efficace l'ensemble des politiques de santé dans les pays communistes. C'est du moins l'impression qui se dégage des écrits qu'on peut lire à leur sujet: "Dans les services de santé en URSS, tous les travailleurs sont tenus de consacrer au moins quatre heures par mois à la propagande éducative (...). Quelle que soit la forme de cette activité, c'est un trait caractéristique du système que l'éducation sanitaire fasse obligatoirement partie des fonctions de tous les travailleurs de la

(1) Viborel, L. (éd); L'éducation sanitaire; cité dans GILBERT, Jules; Education sanitaire, théorie et pratique, Paris, Masson et Cie, p. 18.

(2) De Léan, Pierre ; L'éducation sanitaire et sociale en France , Ministère de la Santé du Québec, 1966.

santé aux divers échelons".(1) C'est aussi l'impression que laissent les communications des participants de ces pays à des congrès internationaux.(2) Cette capacité d'orienter les politiques de santé vers la prévention et de suciter la participation des professionnels de la santé à leur application, bien qu'il faille toujours se méfier de la distance existant entre le discours de personnes ayant avantage à promouvoir leurs réalisations et ce qui se passe dans la réalité concrète, repose probablement sur la capacité qu'a l'Etat, dans ces sociétés fortement centralisées, à établir un plan directeur de l'action collective et à le faire respecter avec des méthodes qui apparaissent parfois comme excessivement contraignantes pour nos sociétés démocratiques.

L'observation des politiques de santé et du rôle important que joue l'éducation sanitaire dans les pays scandinaves tendrait à renforcer cette hypothèse, à savoir que dans les pays ou l'Etat exerce un contrôle important sur la planfication et la mise en application des politiques de santé, on constate l'importance accordée à des politiques où la prévention et l'éducation sont des composantes normales bien intégrées au système qui reçoivent un financement régulier et relativement important. Des mesures législatives soutiennent souvent des campagnes d'information et la lutte anti-tabac menée en Suède est un exemple intéressant de ce genre d'intervention étatique bien orchestrée.

(1) O.M.S.; L'éducation sanitaire en Russie , rapport d'un voyage d'études en URSS par un groupe d'experts internationaux; Genève, 1963, p. 14.

---

(2) Par exemple: Lozan, Bruno; Notre expérience sur l'éducation sanitaire antivénérienne dans la pratique de la médecin générale en Roumanie , ou encore: Grigopol, Fédor; Sur l'éducation sanitaire des asthmatiques en Roumanie; Communications présentées lors de la 9ième Conférence Internationale d'éducation pour la santé, Ottawa, Août-Septembre 1976.

Il est d'autre part intéressant de constater dans un certain nombre de pays du Tiers Monde le même genre de phénomène, qui serait cette fois explicable par le fait que d'une médecine traditionnelle, on passe parfois brutalement à une organisation calquée avec plus ou moins de bonheur sur les modèles de sociétés capitalistes avancées (américaine surtout, en raison de l'abondance des consultants américains sur le plan international et du fait que beaucoup de ressortissants des pays du Tiers Monde reçoivent leur formation supérieure aux Etats-Unis), modèles qui proposent à certains moments l'orientation préventive si difficile à réaliser dans nos pays occidentaux.

Passant en revue quelques pays, nous avons pu constater que c'est surtout aux environs des années 1940 qu'ont vraiment commencé à se structurer des organismes nationaux d'éducation sanitaire, bien que durant les décades précédentes, des initiatives éparses, privées ou publiques, aient témoigné de l'importance de plus en plus grande que l'on voyait à éduquer les "masses populaires" sur diverses questions de santé. Nous avons aussi fait ressortir, constatation générale qui mériterait sans doute une démonstration plus détaillée et plus nuancée, que l'organisation concrète de ces organismes nationaux recevait une faveur qui variait en même temps que l'intérêt accordé à la prévention en général, sauf dans des pays où l'organisation sociopolitique communiste ou socialiste était en mesure d'imposer aux professionnels de la santé la prévention et l'éducation comme une partie normale de leur travail. Nous allons, avant de conclure ce rapide survol de la situation de l'éducation sanitaire à travers le monde en considérant les principaux organismes internationaux pertinents, nous attarder au cas du Canada.

Palko [1] souligne qu'avant le début du XXème siècle, l'éducation sanitaire au Canada était faite par les médecins sur une base individuelle et visait surtout, appuyée par quelques rares initiatives d'éducation de masse, à enrayer les grandes épidémies de maladies infectieuses (variole, tuberculose, etc.) par des mesures d'hygiène appropriées. Face à un besoin croissant, et à cause des pouvoirs limités que la Confédération lui accordait dans le domaine de la santé (auprès des autochtones, des prisonniers, des marins et des voyageurs arrivant de l'étranger), le gouvernement fédéral créait en 1937 une "Division de la publicité et de l'éducation sanitaire", à l'intérieur de son ministère de la santé, dont le rôle premier était de produire du matériel servant à l'éducation de la population. A travers le temps et divers changements de nom, cette division élargit son rôle en stimulant la recherche et en établissant un centre de documentation d'envergure nationale. A cause de l'impossibilité où il était de prendre directement des initiatives auprès de la population pour des raisons constitutionnelles (la santé est de juridiction provinciale), le gouvernement fédéral canadien a donc joué un rôle de promoteur et de coordonnateur des efforts d'éducation auprès des provinces, encourageant différentes activités par des subventions de recherche ou de fonctionnement, assumant la production de matériel de toute nature (affiches, brochures, films, etc.), publiant divers documents spécialisés dont la revue Education Sanitaire qui en est, en 1977, à sa quinzième année de parution, provoquant des rencontres fédérales provinciales sur l'éducation sanitaire, etc., etc..

(1) Palko, Michael E. ; "Significant developments in Canadian Health Education", in Medical Services Journal, Canada, September, 1967, pp.1060-1077.

Ces activités existent depuis plusieurs années, mais on a ressenti la nécessité impérieuse de présenter, en 1975, un projet visant à faire le point sur la situation de l'éducation sanitaire au Canada; on y mentionne que les constatations faites il y a vingt ans déjà par un comité d'experts de l'O.M.S. sur l'éducation sanitaire à travers le monde sont encore tout à fait d'actualité au Canada, et qu'il est temps de faire une évaluation systématique de l'ensemble des efforts consentis ici ainsi que de leur rendement, étant donné que ces efforts sont appelés à prendre une expansion considérable suite au document intitulé Nouvelle Perspective de la santé des Canadiens.(1)

Mentionnons encore que les provinces des Prairies furent les premières à se doter d'un service d'éducation sanitaire (dès 1923 en Alberta) et que les services provinciaux se développèrent à des rythmes et selon des priorités très différents.(2) Soulignons de plus qu'il existe depuis 1962 une "Société canadienne des spécialistes en éducation sanitaire", qui est essentiellement active en milieu anglophone, et que c'est dans la revue de l'association canadienne de santé publique, le Canadian Journal of Public Health, qu'on peut retrouver la quasi totalité des articles scientifiques publiés au Canada sur l'éducation sanitaire.

Sur la scène internationale, c'est en 1950 que l'O.M.S. créait sa "Division de l'Education Sanitaire du Public", et depuis cette date, dix-sept publications techniques ont été réalisées dont onze par des comités d'experts internationaux réunis pour discuter de sujets particuliers

(1) Badgly, Robin; Health Promotion Evaluation in Canada ; Toronto, 1975.
(2) Palko, M.; Op. Cit., p. 1062.

reliés à l'éducation sanitaire (préparation des enseignants, recherche et évaluation, éducation anti-tabac, éducation sur l'hygiène dentaire, etc., etc.); l'O.M.S. nomme aussi un responsable de l'éducation sanitaire dans chacun de ses bureaux (bureau pan-américain, bureau européen, etc.) et peut offrir aux gouvernements qui en font la demande des expertises particulières.

En 1952, une organisation internationale non gouvernementale (à la différence de l'O.M.S. qui, reliée à l'O.N.U., regroupe un certain nombre de pays membres) vit le jour; l'Union Internationale d'Education pour la Santé (U.I.E.S.) :

> "1) sert de lien entre les organisations et particuliers travaillant dans le domaine de l'éducation pour la santé dans les divers pays en les aidant à mettre en commun leurs expériences et leurs connaissances.
>
> 2) facilite les échanges d'information et les confrontations d'expériences à travers le monde pour tout ce qui touche à l'éducation pour la santé et notamment aux programmes, à la formation professionnelle, à la recherche, aux méthodes et aux techniques, aux moyens de communication, etc..
>
> 3) favorise la recherche scientifique et l'amélioration de la formation professionnelle en matière d'éducation pour la santé.
>
> 4) favorise le développement d'une opinion publique avertie de tout ce qui se rapporte à une existence saine; tout celà, dans le but ultime de contribuer à l'amélioration de la santé par l'éducation". (1)

En plus de conférences internationales tenues à tous les trois ou quatre ans depuis 1953 et dont la plus récente eût lieu à Ottawa en 1976, l'U.I.E.S.

(1) Prospectus de L'U.I.E.S., Genève, avril 1976.

offre à ses membres de nombreux services techniques dont en particulier la Revue Internationale d'Education pour la Santé (RIES) qui est, depuis 1958, avec les Health Education Monographs dont nous avons parlé plus haut, le canal international par excellence pour la publication d'articles scientifiques concernant l'éducation sanitaire.

Cette brève histoire de l'éducation sanitaire dans le monde, beaucoup plus descriptive qu'analytique, permet donc de constater que le développement de ce champ de connaissances, bien que relativement récent, s'est déroulé avec rapidité. Des organismes et des canaux de communication internationaux bien organisés, sous-exploités au Québec et qui pourraient être d'un grand profit, se sont structurés de plus en plus systématiquement depuis 1940 et ne demandent qu'à être utilisés.

B) L'éducation sanitaire au Québec

Une rétrospective historique complète pour le Québec devrait tenir compte d'abord des aspects "coercitifs" de l'éducation sanitaire en relevant dans les diverses législations fédérales, provinciales ou même municipales les mesures qui ont visé à modifier certains comportements de la population en regard de sa santé ou à promouvoir des mesures d'hygiène ou de salubrité; elle devrait de plus comporter une analyse de la manière dont ces législations ont été appliquées ou non. On devrait aussi faire l'inventaire des divers types de milieux ou d'organisations (milieux publics ou parapublics de santé ou d'éducation, corporations professionnelles, organismes privés de bienfaisance, etc.) qui se sont préoccupés, utilisant des stratégies plus "persuasives", d'éduquer les individus ou les collectivités à une meilleure hygiène de vie. Nous aurions alors un portrait relativement adéquat des contributions du passé et serions davantage en mesure d'éviter ce que l'on constate trop souvent dans l'agir humain: la répétition, à quelques décades d'intervalle, des mêmes erreurs et des mêmes difficultés de fonctionnement.

Les limites du présent travail, qui vise d'abord à poser le problème général de l'éducation sanitaire au Québec en 1977, ne permettent cependant pas une rétrospective aussi détaillée et ce sont surtout les interventions de l'Etat québécois, par le biais de la "Division de l'Enseignement de l'Hygiène" et du réseau des "Unités Sanitaires" qui retiendront notre attention. On peut considérer que c'est par l'intermédiaire de ces organismes (ainsi que des bureaux de santé municipaux) que le gros des efforts consentis au Québec durant les récentes décades ont été réalisés et leur observation nous permettra de dégager les lignes de force de notre passé en éducation sanitaire. Pour guider ce retour aux sources, un découpage temporel en cinq

périodes est utilisé: les origines de l'éducation sanitaire (avant 1940); l'organisation et la croissance (1940-1955); l'apogée (1955-1960); le ralentissement (1960-1972); la relance (1972-...).

Les différents éléments qui ont servi à la rédaction de cette partie de chapitre proviennent principalement du dépouillement des rapports annuels du Ministère de la Santé (de 1936 à 1976), de six entrevues d'une durée moyenne de deux heures avec les principaux responsables passés et présents de l'éducation sanitaire au Québec, de conversations avec Fernand Turcotte (professeur en Médecine Sociale et Préventive à l'Université Laval et auteur d'un cours sur "l'évolution du système de santé au Canada") ainsi que des analyses précitées de Gilles Dussault concernant l'évolution de la profession médicale au Québec entre 1941 et 1970.

1) Les origines de l'éducation sanitaire: avant 1940.

Jusqu'au début des années 1960, l'Etat québécois ne se préoccupait directement que des aspects "publics" de la santé (hygiène et salubrité du milieu, cueillette des données d'état civil, contrôle des épidémies, etc.), et dispensait lui-même les services requis pour ces aspects. D'autre part, le traitement de la maladie chez l'individu reposait entre les mains des professionnels de la santé (les médecins) à qui l'Etat, en sanctionnant légalement dès 1847 l'existence du Collège des Médecins et Chirurgiens, reconnaissait l'exclusivité de cette tâche.

L'Etat québécois se préoccupa donc principalement au début du XXème siècle, de lutter contre les fléaux collectifs qu'étaient alors les épidémies de maladies infectieuses; pour ce faire, à la suite de démonstrations éloquentes de la compagnie d'assurance La Métropolitaine et de projets pilo-

tes dans quelques coins de la Province, financés par la Fondation Rockerfeller, il décida de se doter dès le début des années 1930 d'un réseau d'Unités Sanitaires qui devaient, par diverses mesures préventives appliquées à des populations, s'attaquer à ces problèmes de santé publique. Une vingtaine d'années plus tard, ces Unités Sanitaires seront au nombre de soixante et treize, chacune ayant un territoire bien délimité; elles remplaceront progressivement pour les fins d'hygiène et de salubrité du milieu les bureaux de santé municipaux qui avaient auparavant cette responsabilité et seules les municipalités de Québec et du Montréal métropolitain maintiendront jusqu'à aujourd'hui leurs bureaux municipaux.

Les origines de l'éducation sanitaire organisée remontent donc à la fondation des Unités Sanitaires car, comme aimait à l'écrire celui qui fut longtemps le responsable de cette Division au Ministère de la Santé, le docteur Paul Claveau: "toute intervention des infirmières-hygiénistes, des inspecteurs sanitaires ou de l'officier médical de l'unité, que ce soit une clinique de vaccination, un dépistage radiologique de la tuberculose, une visite à domicile, etc., était une occasion privilégiée d'éduquer la population à une vie plus saine, plus hygiénique."(1) A mesure que le nombre d'Unités Sanitaires augmentait, c'était autant d'interventions éducatives qui se déroulaient, sur une base le plus souvent individuelle et la volonté de l'Etat de se doter d'un réseau organisé et efficace de santé publique où la prévention occupait une place prépondérante était manifeste à cette époque.

2) L'organisation et la croissance: 1940-1955.

C'est cependant en 1941 qu'un médecin hygiéniste de l'Unité Sanitaire

---

(1) Enoncé fréquent dans les rapports annuels du Ministère de la Santé.

de Shefford, le Docteur Jules Gilbert, suite à des constatations quant à la différence d'attitude face à l'hygiène des secteurs francophones et anglophones de la population qui semblait en partie attribuable à ce que les écoles véhiculaient comme habitudes (les écoles anglophones étant plus "hygiéniques"), commença à se préoccuper sérieusement de l'éducation des groupes et des collectivités au Québec. Grâce à un certain nombre de démarches personnelles, à l'appui financier d'une fondation américaine et à l'intérêt du sous-ministre de la santé du temps, il provoqua, après des séjours d'études et d'observation aux Etats-Unis, la création d'une "Division de l'Enseignement de l'Hygiène" au sein du Ministère de la Santé. Cette division, qui s'appella ensuite "Division de l'Education Sanitaire, en 1960, puis "Bureau d'Education Sanitaire" de 1964 à sa disparition en 1969, connut des années d'activité intense malgré sa taille réduite (elle n'eut jamais plus de deux officiers permanents: le Docteur Gilbert et un éducateur de formation, Monsieur Pierre De Léan).

La première préoccupation de la Division fut d'abord la formation professionnelle d'un certain nombre d'infirmières et de pédagogues qui séjournèrent aux Etats-Unis afin de se spécialiser, dans diverses universités, en éducation sanitaire des groupes et collectivités. Cela traduisit, dès l'origine, la volonté manifeste d'utiliser deux canaux privilégiés pour éduquer la population québécoise. Le premier était celui du réseau scolaire, par le biais de professeurs ayant reçu à l'Ecole Normale un enseignement approprié en hygiène dispensé par des pédagogues spécialement formés à cet effet. C'est très tôt que les attitudes et des habitudes saines se prennent et c'est à l'école qu'on peut les renforcer ou les faire acquérir si la famille n'a pas joué complètement son rôle à cet effet. L'école fut donc dès les débuts de la Division un secteur où l'on voulut absolument intervenir.

Le second canal fut celui des Unités Sanitaires où certaines infirmières triées sur le volet pour devenir "éducatrices-hygiénistes" animèrent le reste du personnel de l'Unité ainsi que la population en général après avoir reçu un enseignement spécialisé. Les premières personnes qui se rendirent étudier aux U.S.A. devinrent, pour quelques unes, professeurs à l'Ecole d'Hygiène. Cette école ouvrit ses portes en 1945 à l'Université de Montréal et permit de former ici, non seulement en éducation sanitaire mais encore dans toutes les disciplines reliées à la santé publique, le personnel des Unités Sanitaires. Plusieurs dizaines de personnes, médecins, infirmières, vétérinaires, inspecteurs sanitaires, enseignants, passaient chaque année à l'Ecole et y recevaient ensemble la formation professionnelle appropriée, formation qui durait généralement une année et qui conduisait à divers diplômes en Hygiène Publique. C'est dans ce contexte que les "éducatrices-hygiénistes" furent formées: elles suivirent le cours régulier de l'Ecole d'Hygiène puis reçurent une formation pratique supplémentaire sous forme de stage auprès d'éducatrices-hygiénistes déjà en poste et de cours spécialisés en nutrition.

Une deuxième préoccupation permanente de la Division de l'Enseignement de l'Hygiène a été de sélectionner des outils pédagogiques servant à l'éducation populaire (dépliants, films, etc.) dont la majeure partie était produite au niveau du gouvernement fédéral, et d'assurer leur diffusion dans la population via les Unités Sanitaires, les bureaux municipaux de santé et quelques autres organismes. La division produisait elle-même un certain nombre de dépliants, de livres, d'articles, de films, de messages télévisés, de séries radiophoniques, d'affiches, en collaboration tantôt avec l'entreprise privée (certaines compagnies d'assurances ou certaines entreprises pharmaceutiques ont toujours été très actives dans la production de matériel d'information), tantôt avec d'autres services du gouvernement québécois. Le

gros effort de production de la Division était cependant un kiosque à l'exposition provinciale de Québec où, à chaque année, un thème particulier était exploité.

La troisième activité d'importance était la coordination des activités d'éducation sanitaire au niveau du Québec, tâche qui comportait deux grands volets: d'une part, coordination des activités se déroulant sur le terrain, dans les Unités Sanitaires; d'autre part, coordination avec des organismes privés ou publics d'envergure provinciale, nationale ou internationale, avec d'autres divisions du Ministère de la Santé du Québec (les divisions des maladies vénériennes, de l'hygiène dentaire et de la nutrition, entre autres, attendu leur orientation marquée vers des stratégies d'éducation sanitaire), avec d'autres instances gouvernementales (en particulier l'Instruction Publique, quant au développement d'un enseignement de l'hygiène intégré au programme scolaire régulier et quant à la place à donner à l'éducation physique en milieu scolaire, etc., etc.). La section "relations extérieures" du rapport annuel de la Division est toujours très étoffée et la présence du Québec sur les plans national, nord-américain et international était alors bien réelle.

3) L'apogée: 1955-1960.

C'est cette période de cinq ans environ qui constitue ce qu'on peut appeler l'apogée de l'éducation sanitaire au Québec. La formation professionnelle allait bon train et l'Ecole d'Hygiène continuait à former ses enseignants et ses professionnels de la santé (ces derniers étant tous boursiers du Ministère de la Santé); les facultés de médecine commencaient à s'intéresser à l'éducation sanitaire et le Docteur Gilbert y était souvent demandé pour donner cours, causeries, conférences. Une vingtaine de profes-

seurs spécialement formés travaillaient dans les Ecoles Normales auprès des futurs enseignants.

Les activités "d'éducation populaire" de la Division avaient elles aussi atteint une maturité et un rythme de croisière qui seront inégalés par la suite: rencontres de divers organismes, revision et diffusion de matériel éducatif, publication d'innombrables articles, conférences, etc..

C'est aussi l'apogée des programmes d'éducation dans les Unités Sanitaires; trente-huit infirmières éducatrices-hygiénistes y consacrent la majeure partie de leur temps dans autant d'unités et la Division leur fournit un encadrement technique par l'intermédiaire d'un bulletin mensuel, Liaison, publié régulièrement depuis 1945. Des réunions regroupant toutes les éducatrices ont aussi lieu à tous les deux ou trois ans et le Docteur Gilbert passe souvent la majeure partie de ses vacances à sillonner le Québec en toutes directions afin de visiter les éducatrices dans leur propre milieu. Dans les Unités Sanitaires où il n'y a pas d'éducatrice -hygiéniste , l'éducation est généralement un sujet important; le passage suivant, écrit par un médecin qui a vécu cette époque, nous montre bien l'ardeur au travail et la foi dans l'éducation qui caractérisaient les interventions du personnel hygiéniste d'alors:

> "Après avoir vu autour de moi tant de dégâts causés par les préjugés et le manque de connaissances, j'entrepris une campagne d'éducation populaire. A la salle municipale, où j'avais accès à mon gré, je rassemblais des groupes de mamans pour leur parler d'hygiène, de puériculture, de maladies contagieuses, de premiers soins aux blessés, d'alimentation, etc.. Dans les paroisses où je devais passer la nuit, à cause de la distance, j'organisais des conférences le soir. Je ne refusais jamais une occasion de faire passer mon message".(1)

(1) Dumas, Albert; "La Médecine et mon temps",in Recherches Sociographiques, Vol. XVI, 1975, I, p. 31.

Le rayonnement international du Québec, en matière d'éducation pour la santé, atteignit alors un niveau inégalé et des délégations de plusieurs pays vinrent ici voir comment notre système d'éducation sanitaire fonctionnait. L'Organisation Mondiale de la Santé (OMS) envoyait en formation des francophones d'un peu partout au monde à l'Ecole d'Hygiène de Montréal, qui offrait même un diplôme de second cycle en éducation sanitaire.

Le Québec a cependant évolué depuis la création des Unités Sanitaires et pour ce secteur comme pour beaucoup d'autres au Québec, 1960 marque un changement radical.

4) Le ralentissement: 1960 - 1972

Une des personnes que nous avons interviewée a qualifié d'euphémisme la façon dont nous avons intitulé cette quatrième période; la débandade, la désagrégation, l'anéantissement auraient été, selon lui, des qualificatifs plus appropriés pour décrire ce qui est alors advenu des activités d'éducation sanitaire. Cette période devait en effet voir le Bureau d'éducation sanitaire du Ministère aboli en 1969, après avoir vivoté durant cinq ou six ans. L'Ecole d'Hygiène de Montréal disparaîtra elle aussi, les différents groupes auxquels elle s'adressait étant attirés dans d'autres facultés nouvellement crées (nursing, médecine vétérinaire, éducation, etc.) où les enseignements sur l'hygiène s'estomperont rapidement des programmes. Le réseau des Unités Sanitaires se détériorera considérablement et les enseignements d'hygiène faits aux futurs enseignants dans le cadre des Ecoles Normales disparaîtront avec celles-ci. D'une époque d'activité fébrile on en viendra à un vacuum quasi total.

Un grand nombre de faits précis (modification de la strucutre des Unités Sanitaires, départ du Docteur Gilbert en 1960, arrivée d'un nouveau

doyen à l'Ecole d'Hygiène de Montréal, etc.) peuvent être invoqués pour expliquer ces changements; ils sont la cristallisation d'une foule de phénomènes beaucoup plus profonds qui, de façon latente, modelaient la société québécoise depuis quelques années déjà et que nous allons tenter de faire ressortir.

Cette période de l'histoire du Québec a souvent été qualifiée de "révolution tranquille" en ce sens que la fin du règne de l'Union Nationale, marquée par la prise de pouvoir du gouvernement libéral de Jean Lesage en 1960, donnait libre cours à toute une série de changements que l'industrialisation et l'urbanisation, accélérées depuis la fin de la seconde guerre mondiale, avaient rendus imminents. Dans ce climat de mutations, la désaffection de la médecine préventive, à l'intérieur de laquelle prend place la désagrégation des interventions d'éducation sanitaire, peut se comprendre à la lumière de certains facteurs particulièrement importants.

Le premier est sans doute qu'une des raisons importantes de la création des Unités Sanitaires n'existait pratiquement plus: les grandes épidémies de maladies infectieuses qui décimaient le Québec au début du siècle étaient maintenant à toutes fins pratiques sous contrôle, en bonne partie d'ailleurs à cause de l'efficacité du travail de ces mêmes Unités Sanitaires et de l'amélioration générale du niveau de vie. Les Unités s'étaient par ailleurs ancrées dans une certaine forme de routine qui avait fait ses preuves par le passé, mais qui ne correspondait plus aux besoins du jour et n'avaient pas encore trouvé comment s'attaquer avec autant d'efficacité aux nouvelles épidémies: accidents, cancers, maladies cardiaques, etc...

Un second facteur important est l'évolution de la pratique médicale: au cours des années 1930, les Unités Sanitaires concurrençaient avantageusement

la pratique privée, car ces temps de crise économique étaient durs et les clients n'avaient pas souvent d'argent pour payer leur médecin [1]; l'Etat offrait un salaire au médecin intéressé à aller travailler dans les Unités Sanitaires et les conditions du marché étaient telles que la qualité du personnel recruté était excellente. A mesure que la pratique médicale a évolué, que la médecine générale pratiquée en cabinet privé ou sous forme de visites à domicile a cédé le pas à une médecine spécialisée pratiquée surtout à l'hôpital et rémunérée fort convenablement, le médecin hygiéniste, dont le salaire était maintenant loin d'être concurrentiel, a de plus en plus été dévalorisé par ses confrères. Il était petit salarié de l'Etat (et Gilles Dussault a fort bien montré comment c'est en réaction contre l'Etat que s'est élaborée une partie importante du discours public de la profession médicale)[2], voué à faire des vaccinations et autres pratiques routinières et sans gloire. C'était donc l'original ou parfois le cancre de la classe qui se consacrait maintenant à l'hygiène publique. Les différents groupes professionnels des Unités Sanitaires ont donc connu, à partir de 1960, un vieillissement progressif, ce qui les rendait encore moins attirants pour une relève de plus en plus rare.

Le troisième facteur, dans ce contexte général, c'est le changement de rôle de l'Etat. A partir de 1960, l'Etat québécois a progressivement délaissé les champs de l'hygiène et de la prévention, qui étaient auparavant ses seuls secteurs d'intervention, pour se lancer en 1961 dans un régime d'assurance-hospitalisation et, en 1970, dans un régime universel d'assurance-

(1) Dumas, Albert; Op. Cit., p. 21.

(2) Dussault, Gilles; La profession médicale au Québec, 1941-1971, ISSH, 1975.

maladie. L'accessibilité aux soins curatifs pour l'ensemble des citoyens, et le développement de l'infrastructure (hôpitaux, technologie, recherche biologique, etc.) requise pour les dispenser ont donc accaparé, pour différentes raisons de nature politique, économique et sociale, l'essentiel des énergies de nos gouvernements en ce qui concerne la santé. En même temps les sommes investies dans la prévention, et en particulier dans les Unités Sanitaires, se sont stabilisées puis ont décru. D'un organisme de services directs à la population axé sur la santé publique, le Ministère de la Santé est devenu le gestionnaire du paiement de services curatifs dispensés par les médecins et les établissements du réseau des Affaires sociales.

5) La relance: 1972 - ...

C'est cependant durant cette époque de grand ralentissement des activités préventives que l'Etat créait,au milieu des années 1960, une "Commission d'enquête sur la santé et les services sociaux" qui devait d'abord se pencher sur l'épineuse question de l'assurance-maladie et dont les travaux en vinrent à proposer une refonte complète du système de santé au Québec. En plus de fonder sa réforme sur une philosophie plus globale de la santé, la Commission Castonguay-Nepveu remettait en lumière la grande nécessité de la prévention (et de ses composantes éducatives, ajouterons-nous). Nous avons retenu 1972 comme moment de la relance parce que c'est l'année où, au niveau du nouveau Ministère des Affaires sociales, on créait à la Direction des Communications une section chargée de la production de matériel éducatif. C'est la même année que la règlementation de la loi sur la santé et les services sociaux prévoyait la création de trente-deux départements de santé communautaire qui, de concert avec les centres locaux de services communautaires, devaient intégrer dans leurs rangs ce qui restait de personnel dans les Unités Sanitai-

res et recevaient de façon privilégiée la mission de réorganiser les services préventifs au Québec.

C'est cependant en 1974 que la directrice de la production de matériel éducatif fut embauchée, au niveau du MAS, et que les D.S.C. commencèrent à s'organiser sur le terrain alors que le réseau des C.L.S.C., commencé dès 1971, se développait avec peine et sans toujours s'occuper de prévention.

De 1974 à 1977, alors que plusieurs autres secteurs du monde des Affaires sociales fonctionnaient au ralenti, celui des D.S.C. connaissait un développement important et recevait des allocations budgétaires lui permettant d'engager du personnel, de réactiver les programmes de prévention qui étaient dispensés par les Unités Sanitaires en les modernisant au besoin et d'en établir de nouveaux. Une nouvelle génération de professionnels de la santé, formés aux Etats-Unis pour la plupart, est venue animer d'un dynamisme nouveau ce secteur qui avait été délaissé.

Pour assurer la formation du personnel, les départements de santé publique des quatre facultés de médecine au Québec (Laval, Montréal, McGil et Sherbrooke) ont connu une activité de plus en plus fébrile; des certificats en santé communautaire ont vu le jour dans la quasi totalité des écoles de sciences infirmières de la province et de nouvelles catégories de professionnels (administrateurs, spécialistes des sciences sociales) se sont ajoutées aux professionnels de la santé habituels. La création d'un département de "Health Education" à l'éducation permanente de l'Université Concordia, à Montréal en 1972, la mise sur pied en 1976 d'un diplôme de maîtrise en Santé Communautaire à l'Université de Montréal, le nombre de plus en plus grand de personnes allant se spécialiser dans diverses écoles de santé publique à l'étranger, le nombre croissant de groupes universitaires intéressés à la re-

cherche en santé, voilà autant de signes qui laissent croire que la relance semble bien amorçée.

Il faut être optimiste et il est difficile de prévoir l'avenir, mais, malgré l'enthousiasme des deux ou trois dernières années, on sent déjà une stabilisation au niveau des budgets alloués pour le développement de la santé communautaire. De son côté, la section de production du matériel éducatif du MAS, malgré des dizaines de projets,(1) une équipe de quelques professionnels fort intéressés et plusieurs réalisations concrètes à son crédit, n'est pas dans une position structurelle qui lui permette, comme au temps de la "Division de l'Enseignement de l'Hygiène" d'avoir un pouvoir de décision sur les priorités en termes d'éducation, les programmes à mettre de l'avant, etc.. La nécessité absolue de trouver des solutions permettant de réduire les montants énormes que les services de santé coûtent à la collectivité québécoise crée un contexte où la prévention et l'éducation sanitaine seront inévitablement des solutions alternatives à explorer et l'avenir à moyen terme de ces approches semble donc prometteur même si, contrairement aux soins curatifs, elles devront toujours faire la preuve de leur rentabilité et justifier avec grande précision les maigres budgets qui leur sont alloués..

6) Conclusion:

Cette rétrospective historique nous a permis de voir que le développement de l'éducation sanitaire comme champ spécifique d'intervention est apparu aux environs de 1940 un peu partout dans le monde, que, au Québec

(1) Québec, Ministère des Affaires sociales; Esquisse préliminaire du programme de Communications, Editeur Officiel du Québec, 1975, 303 pages.

comme ailleurs, il a connu des hauts et des bas s'inscrivant dans l'évolution générale des stratégies préventives et qu'il semble voué, bien qu'il faille demeurer d'un optimisme modéré là-dessus, à un avenir prometteur.

BIBLIOGRAPHIE

1. BADGLEY, Robin; Health promotion evaluation in Canada , Toronto, Mimeo, juin 1975, 24 pages.

2. BAUER and HULL; Health education of the public, Saunders, London, 1942, 311 pages.

3. Canada, Ministère de la Santé et du Bien-Etre Social; Nouvelle perspective de la santé des Canadiens , Ottawa, 1974, 82 pages.

4. COSSETTE, Claude; MOTULSKY, Bernard; "Le Québécois et la consommation", in Cossette, Claude (ed): "Communication de masse, consommation de masse", Boréal Express, Sillery, 1976, pages 237-265.

5. DE GUISE, Jacques; "La communication persuasive" in Cossette, Claude (ed): "Communication de masse et consommation de masse", Boréal Express, Sillery, 1976, pp. 69-108.

6. DE LEAN, Pierre; L'éducation sanitaire et sociale en France, Ministère de la santé du Québec, 1966, 189 pages.

7. Département de sociologie, Université Laval; "Programme des études avancées 1975-76", Québec, septembre 1976, 7 pages.

8. DUMAS, Albert; "La médecine et mon temps" in Recherches sociographiques , Vol. 16, no. I, Québec, PUL, pages 21 à 43.

9. DUMONT, Fernand ; Le lieu de l'homme, la culture comme distance et mémoire , HMH, Montréal, 1971.

10. DUSSAULT, Gilles; "Les médecins du Québec 1940-1970" in Recherches sociographiques', 1975, Vol. XVI, No. 1, p. 69-85 (Numéro spécial consacré à la sociologie de la santé au Québec).

11. DUSSAULT, Gilles; 'La profession médicale au Québec, 1941-1971', Institut supérieur des sciences humaines, Québec, 1975, 133 pages. (Collection études sur le Québec, No. 2).

12. G.E.R.M. (Groupe d'étude pour une réforme de la Médecine), Journées de discussion sur l'éducation sanitaire , Ministère de la Santé Nationale, de la santé publique et des communications, Bruxelles, avril 1972, 215 pages.

13. GILBERT, Jules; 'L'éducation sanitaire, théorie et pratique', Masson et Cie, Paris, 1959, 310 pages.

14. GILBERT, Jules; 'L'éducation sanitaire', PUM, Montréal, 1963, 141 pages.

15. GILBERT, Jules; "The grandeur and decadence of Health Education" in Canadian Journal of Public Health , Vol. 58, août 1967, pages 355-358.

16. GREEN, Lawrence; "Evaluation and measurement: some dilemnas for health education" in American Journal of Public Health , Vol. 67, No. 2, février 1977, pages 155-161.

17. GRIGOPOL, Fédor; "Sur l'éducation sanitaire des asthmatiques en Roumanie", in Résumé des Conférences de la 9e Conférence Internationale d'Education pour la santé , UIES, Ottawa, 1976, p. 85.

18. ILLITCH, Ivan; 'Némésis médicale , Seuil, Paris, 1975, 212 pages.

19. JANIS I.L.; FESHBACH, S; "Effects of fear arousing communications" in Journal of Abnormal and social psychology , 1953, 48: 72-92.

20. JOHNSON, Rodrigue; "Les aspects éthiques de l'intervention de l'Etat dans la conduite dite privée", Ottawa, Santé et Bien-Etre Social Canada, janvier 1976, 51 pages (miméo).

21. KLEINEBERG, Otto; 'Psychologie sociale', PUF, Paris, 1967, 384 pages.

22. LALONDE, Marc; "Améliorer la condition humaine, quelle philosophie sera la nôtre" in Revue Internationale d'éducation pour la santé, XIX, 1976:3, p.146-148.

23. LIZIEWICZ, Adam; "Formulating a Quebec Health Education policy for the 1980's", in "Résumés des communications de la 9è conférence internationale d'éducation pour la santé", UIES, Ottawa, septembre 1976, p. 104.

24. LOZAN, Bruno; "Notre expérience sur l'éducation sanitaire antivénérienne dans la pratique de la médecine générale en Roumanie", in 'Résumé des communications de la 9è conférence internationale d'éducation pour la santé', UIES, Ottawa, 1976, p. 74

25. MEANS, Richard K.; A history of health education in the United States, Lea and Tebiger, Boston, 1962, 354 pages.

26. OMS; L'éducation sanitaire en URSS, Genève, 1963, 54 pages. (Cahiers de santé publique, No. 19).

27. OMS; L'éducation sanitaire en hygiène dentaire, Genève, 1970, 31 pages. (Rapports techniques No. 449).

28. OMS; "Planification et évaluation des services d'éducation sanitaire, Genève, 1969, 34 pages. (Rapports techniques, No. 409).

29. OMS - OPS; Rapport de la conférence sur la préparation post-universitaire du personnel de santé à l'éducation sanitaire, Genève, 1964, 53 pages. (Rapports techniques, No. 278).

30. OMS - UNESCO; Comité d'experts sur la préparation des enseignants à l'éducation sanitaire, Genève, 1960, 21 pages. (Rapports techniques, No. 193).

31. OMS; Rapport d'un comité d'experts de l'éducation sanitaire de la population, Genève, 1954, 46 pages. (Rapports techniques, No. 89).

32. OMS; 'Comité d'experts de la formation à donner au personnel de santé en matière d'éducation de la population', Genève, 1958, 45 pages. (Rapports techniques, No. 156).

33. PALKO, Michael E.; "Significant developments in Canadian Health Education" in Medical Services Journal , septembre 1967, p. 1060-1077.

34. PARSONS, Talcott; "Impact de la technologie sur la culture et sur l'apparition de nouveaux modes de comportement" in "Revue Internationale des Sciences Sociales , Vol. XXII, No.1, 1960, p. 660-667.

35. PINEAULT, Raynald; "Eléments et étapes d'élaboration d'un programme en santé communautaire" in L'Union Médicale du Canada , Tome 105, août 1976, pages 1208-1214.

36. Québec, Ministère des Affaires Sociales: "La participation", numéro spécial de '65 à l'heure' , Volume 4, no. 1, septembre 1976.

37. Québec, Ministère des Affaires Sociales: Esquisse préliminaire du programme de communications , Editeur Officiel du Québec, 1973, 303 pages.

38. Québec, Ministère des Affaires Sociales: "Cadre de référence pour l'adaptation des ressources aux priorités , septembre 1976, 35 pages (miméo).

39. Québec, Ministère de la santé et Ministère des Affaires Sociales, Rapports annuels"(de 1936 à 1976).

40. Québec, 'Rapport de la commission d'enquête sur la santé et le Bien-Etre Social , volume IV (La Santé), editeur officiel du Québec, 1970. (En particulier, tomeI: La situation actuelle, 212 pages et tome 2: Le régime de la santé, 209 pages).

41. Recherches sociographiques, Québec, PUL, XIV, 2, 1973. Numéro spécial sur "Le vécu".

42. RENAUD, Marc; "On the structural constraints to state intervention in health"" in International Journal of Health Services , Volume 5, No. 4, 1975, pages 559-571.

43. RENAUD, Marc ; The political economy of the Quebec state interventions in health: reform or revolution? . Thèse de doctorat non publiée, présentée à l'Université du Wisconsin, campus de Madison, U.S.A. en 1975.

44. ROCHON, Jean; "Measurements of quality of care, concepts and strategies for evaluation" in Annals of the Royal College of Physicians and Surgeons of Canada, 7: octobre 1974, pages 301-311.

45. SCHNEIDER, Robert E.; Methods and Materials of Health Education, W.B. Saunders et Cie, Philadelphie et Londres, 1964, 364 pages.

46. SIMMONS, Jeannette (Editeur): "Making Health Education work", supplément spécial d'octobre 1975 de l'American Journal of Public Health, 49 pages.

47. Task force on health promotion and consummer health education; Preventive medecine USA: health promotion and consumer health education, New-York, Prodist, 1976, 255 pages.

48. TURNER, C.E.; 'Principles of health education', DC Health et Cie., Boston 1939, 335 pages.

49. VILLEDIEU, Yannik; "Demain la Santé", les dossiers de Québec Science, Sillery, 1976, 302 pages.

Soc. Sci. & Med. Vol. 9, pp. 141–156. Pergamon Press 1975. Printed in Great Britain

METHODS AND EFFECTIVENESS OF HEALTH EDUCATION: THE PAST, PRESENT AND FUTURE OF SOCIAL SCIENTIFIC INVOLVEMENT*

N. DAVID RICHARDS
Department of Community Health, University of Nottingham, England

INTRODUCTION

The aims of this paper are (a) to focus attention upon Health Education as a key subject area lying fairly and squarely at the crossroads of social science and medicine and (b) to provide a selective review of social science concepts relevant to the practice and research of Health Education. Health Education is a wide-embracing term with a multitude of meanings conveying differing impressions to different people. Within the context of this paper it is taken to be a legitimate and dynamic area of concern for both social scientists and health professionals in their efforts not only to understand but also to bring about favourable changes in health behaviour and response to available health services. In this respect Health Education is viewed as an applied field for the translation of social and behavioural scientific study and research into action programmes which are designed to seek improvements both in knowledge and practice relating to health.

Following an opening section (1) in which an attempt is made to provide a working definition of the aims and objectives of Health Education, there follows a brief statement (2) outlining recent changes in the nature of disease, social and environmental conditions, the organization of health care, public health practice, and public opinion, mores and behaviour. Concomitantly there has been occurring a shift in the approach of Health Education dependent on a thorough understanding of health behaviour. At this stage it becomes relevant to introduce and comment upon certain key anthropological, sociological and psychological concepts as applied in Health Education practice and research (3).

The scene is thus set for an examination of some methods and techniques (4) which have been used in efforts to promote Health Education. This is in itself to suggest the need for an evaluation of both the efficiency and effectiveness of Health Education programmes (5). As will be clearly evident from sentiments expressed throughout this paper the current practice of Health Education, as a scientific, let alone social scientific endeavour, leaves much to be desired. Accordingly the final section of this position paper (6) underlines deficiencies in both present practice and research, and attempts a prognosis for the future.

This paper is then directed towards increasing knowledge about selected research findings and theory, and increasing awareness amongst both social scientists and health professionals about the kinds of practice and research opportunities and challenges in Health Education. It is not intended in any sense to be exhaustive in scope but merely attempts to draw together some major perspectives from the social sciences in an effort to stimulate a wider understanding of their value in the development of Health Education. With this aim in mind the paper does not attempt a comprehensive literature review but seeks instead to generate thoughts designed to facilitate future development in Health Education.

1. WHAT IS HEALTH EDUCATION?

Boundaries for discussion in the other main sections of this conference seem certain to be more limiting than those which are likely to suggest themselves in a section on Health Education. Just as the absolute boundaries of health are difficult to define, so too is it difficult to specify where Health Education begins and ends. Thus it perhaps comes as no surprise to discover that the term "Health Education" is clearly capable of many interpretations. Over the years there has certainly been no shortage of pronouncements as to what is meant by "Health Education". As a generic term, which may indeed have switched a lot of people off, it has been used to describe almost any effort to provide information and knowledge relating to health maintenance and promotion. This is to imply then that it has both informal and formal aspects.

From time immemorial there has clearly existed some form of Health Education which has become ensconced in the social and cultural milieu. Principally this has been developed about a knowledge base, i.e. the focus has been on providing information and knowledge, and has existed either in oral form and been transmitted, with amendment commensurate with contemporary development from person to person and group to group, or in written form perhaps in health manuals or "chat-books", such as *Aristotle's Works*, which became for several generations a source-book for the English working class on *inter alia* sex and conception. Whilst however recognizing the importance of informal transmission of empirical knowledge about health from generation to generation, and across families, classes, tribes and societies we are in this paper more concerned with its meaning in the formal sense implying planned and/or programmed efforts to bring about the promotion and continuation of health, and prevention of disease.

There would seem almost as many definitions of what constitutes Health Education as there are people discussing it: its objectives and aims therefore demand

* Paper prepared for the 4th International Conference on Social Science and Medicine, Elsinore, Denmark, 12–16 August 1974.

clarification. In particular this paper seeks to synthesize recent developments and to suggest that amidst a plethora of views in a number of countries there has seemingly developed a consensus of opinion that Health Education consists of more than the mere imparting of information. Implicit therefore has become some concept of modification of human behaviour towards healthier modes of living and away from deleterious practices. Emphasis has thus swung towards changing and modifying behaviour- inherent in such a task must be some deeper understanding not only of processes of education and change but also of the social, psychological, cultural, economic, political and organizational facets of life. Hence the derivation of the view of a Health Education specialist as a change agent in society, and his important need to understand more deeply the structure and forces of that society. At this point one may perhaps question where Health Education ceases and politics takes over.

Health Education specialists have developed because public health programmes could not achieve their goals without the understanding and support of the public. Thus many early efforts in the field were concentrated on the provision of information and it was assumed that if people were provided with information about health then they would change their behaviour to accord with recommended practices. Increasingly in recent years it has been realized that information and knowledge may not in themselves lead to recommended actions if these conflict with existing motives, attitudes, beliefs and values, and are not consonant with social and group norms.

Health Education like health itself is thus a multisided and multidisciplinary activity. Specifically it brings together the social and health sciences in an educational setting, wherein facts and findings from each of the disciplines are needed to reinforce and complement each other. It has grown up piece-meal around practice and has attempted to professionalize itself by efforts at building a core of theory and knowledge. Nonetheless solid research efforts have until recent years been scanty, and data have been built up impressionistically and often hardly upon a sound scientific base. In short there have been many unsupported deductions, and empiricism has ruled the day: hence the need (developed in the final section of this paper) to construct a framework of confirmed facts and practices out of which it will be possible to generate a knowledge base which will withstand rigorous scientific evaluation. Only by such efforts will it be possible to develop a definitive understanding of health behaviour.

Something more than removal of ignorance is therefore necessary—in essence the problem often appears simple whilst the solutions are neither easy nor quick. Merely to identify problems is insufficient: diagnosis that Health Educational effort is required becomes something of a platitude since effective methods of supplying that want may still prove elusive. Above all it is important to understand that Health Education does not exist in a vacuum-- hence Polgar's four fallacies of the "empty vessels", "separate capsule", "single pyramid" and "interchangeable faces" [1]. Of particular significance to Health Education is the fallacy of the "empty vessel": Health Education specialists proceed as if they were pouring their ideas and appeals into a void instead of an existing popular health culture [2]. Thus in advocating behavioural change it is important to remember the need to change a person's existing attitudes, beliefs, and practices. Indeed in arguing for cessation of such activities as cigarette smoking and sweet eating it becomes clear that replacement activities may need to follow.

A recent report from the World Health Organization has offered the following definition: "Health Education concerns all those experiences of an individual, group or community that influence beliefs, attitudes and behaviour with respect to health as well as the processes and efforts of producing change when this is necessary for optimal health" [3]. Health Education thus attempts not only to equip individuals, groups and communities with the wherewithal and means of making sound decisions but also to influence the outcome in a desired direction by such decisions.

The philosophical focus of this paper is in accord with the WHO definition, and Health Education is judged as a process which focuses on people, modifying their health behaviour where necessary and creating education opportunities and situations that foster and facilitate change. It has the difficult task of persuading people to forsake harmful yet sometimes enjoyable habits, and to take precautionary actions which may be regarded as troublesome. This is in essence to view the topic as being both the problem-oriented aspect of community and preventive medicine, and the action-oriented aspect of medical sociology, with its function being to bridge the gap between knowledge and practice and to encourage preventive health behaviour. As we have already suggested, the subject is multi-sided and multidisciplinary, and with no absolute monopoly of responsibility.

Since this paper is primarily directed at an audience of social scientists and health professionals their particular fields of study have been stressed. Equally important however will be a thorough understanding of Education and Learning Theory strategies. Furthermore the concepts of right and wrong are becoming blurred, as is a view of Health Education operating through an expert-client or master-servant relationship. It is a two-way process which involves working with and through people to achieve success—such a realization clearly follows upon reference to the numerous public health programmes which have failed for want of public appreciation and response. It is thus essential to adopt a two-angled stance to health behaviour and problems, and to view them as seen not only by health professionals but also by the general public. In this paper we largely beg the question of the ethics of Health Education as regards the right of a Health Education specialist to step in and imply the need for behavioural change. More specifically we direct ourselves to the need to provide a solid foundation on which to develop programmes which may perhaps be evolved in response to people-defined problems. The need to recognize any differences in professionally- and lay-defined problems is important—of paramount significance is however the need to have a scientific base for action programmes which may follow.

It is also important to consider the relationship between knowledge and practice. Their exact relationship is not known: for many however the inclination has been to suppose a change in knowledge would lead to an appropriate change in attitude, which in itself would facilitate a change in behaviour and eventually

in habits and customs. A person may believe in something but not practise accordingly, or he may adopt a certain practice whilst not believing in it. Divorce of practice from belief and vice versa have led to Festinger's propounding of a theory of cognitive dissonance—a topic we elaborate upon in Section 4. It is clear that people change their behaviour or fail to do so for a variety of reasons many of which we do not as yet begin to understand.

Health Education as an area for both practice and research evidences many deficiences—indeed it is startling to notice how many programmes have been devised without any clear stated objectives which are capable of evaluative research. Lack of objectives is of course also to question what is an absolute standard of acceptance of any planned programme: short of 100 per cent acceptance, what is it realistic to expect? Lack of precise statements of practical objectives is also to lead to "blunderbuss" effects—that is to say firing off shots in the vague direction in hope of scoring a direct hit. Much of the early social scientific effort was diagnostic in nature: in recent years with increasing sophistication of effort the targets have become more applied and theoretical, and the pendulum has begun to swing more in the direction of developing experimental and model programmes. Just as health is a dynamic concept, so too is Health Education—witness for example the changing definitions offered in recent technical reports on the subject from the WHO. Nevertheless one must note absence of cross-national guidelines and the paucity of replication of effort. In essence there are few answers to describe and explain what works where, how and why, and there exists a prime need for linkages between theory and practice.

It is important also to set out another distinction, namely that between information (what people think) and education (how people think). The ultimate aim of Health Education is practice resulting from changes in knowledge and behaviour. Hence the wisdom of John Ruskin's aphorism: "education is not that one knows more, but that one behaves differently".

Inherent in facilitating or impeding any change will be a host of cultural, social and psychological pressures which help mould behavioural patterns. Thus in order to be effective the planning, methods and strategies associated with Health Education must take into account the processes by which people acquire knowledge and change their behaviour, and factors which influence or act as barriers to such changes. The challenge is to understand how and when to intervene to maximum effect to produce both a desire for positive health behaviour, and also a willingness to adopt and maintain behaviour which will promote and encourage this desire. In essence then Health Edication is about people, their behaviour and processes of change, and as such it can be an important area for applied social scientific effort. The area might perhaps also appear to offer the promise of approaching Utopia, were the general public, and particularly those most at risk, to act upon the health messages and exhortations to action beamed at them in the name of Health Education. Nonetheless, to this commentator at least, it would seem that many would-be health educators seek to run before they have learned to walk—in the sense that in their haste to devise solutions to pressing problems they neglect to familiarize themselves with a thorough knowledge of social science findings which are applicable in action settings of health education practice.

In this section we have argued a role for Health Education, which implies the application of the techniques of education to health promotion and maintenance and within a social and cultural milieu, the broad frameworks of which are understood and appreciated. We continue by underlining the dynamic nature of both health and Health Education, and by examining the parameters of both an ever-changing need for and practice of Health Education itself.

2. THE CHANGING NATURE OF HEALTH PROBLEMS AND IMPLICITLY OF HEALTH EDUCATIONAL EFFORTS

Internationally the health care system may be said to be in a state of ferment. This is occasioned by changing patterns in the epidemiology of disease, social and environmental conditions, organization of health services, public health practices and public opinion and mores. Concomitantly, and hardly surprisingly there have also followed changes in approaches to Health Education, and in the techniques and strategies of action-oriented programmes.

During the bacteriological era there was comparatively little need for the incorporation of social and behavioural science into Health Education programmes. Efforts in disease control were highly successful (at least in Western civilizations) and infectious communicable and parasitic diseases began to disappear as the main causes of morbidity and mortality, only to be replaced by patterns of disease which do not respond to immunization, improvements in sanitary control, or in the face of compulsory legislative powers. These new conditions, the "diseases of civilization" (chronic degenerative diseases such as cancer, heart disease and arthritis) have assumed critical importance as reference to morbidity and mortality rates will show. Together with changes in age-patterns, dependency-ratios and the duration of disease they have occasioned a transfer from sudden to routine episodes of sickness. The association of behavioural patterns with disease patterns has led to the concept of disease being self-afflicted or socially induced, and a realization that a large and increasing proportion of contemporary morbidity is man-made. Research into the role of cigarette smoking in lung cancer, dietary factors and physical activity in heart disease, oral hygiene practices and periodontal and gingival conditions, and eating sweets and dental caries has furthered this concept. In the light of this Powles [4] has argued that technical medicine is not making much advance, and pointed to diminishing returns. Nonetheless, and despite the cognitive recognition of health information, people are still dying from ignorance or want of knowledge relating to the risks they are taking.

Changes in demographic, social and environmental conditions are leading to the emergence of new problem areas for Health Education. Increases in the amount of air pollution, numbers of motor-cars, use of food additives, breakdowns in social supports leading to stress and mental problems, and percentage of elderly people are all creating new challenges for public health practitioners.

The organization and structure of health services is subjected to change as pressures on obtaining services

start to shift in response to a changing epidemiological pattern. Changes in the quantity and quality of professional, semi-professional and allied staff taken together with the spiralling cost of services, and a growing emphasis on community care offer the prospect of further challenges for Health Education. Concurrently there are also evolving changes in public health practices, the most obvious of which is a shift, thanks to advances in curative medicine and the availability of such care, from disease cure to prevention and health maintenance. Adherence to recommended regimens is more a question of voluntary participation than legislative compulsion. Society is itself of course changing, together with its predominant public opinion, mores and behaviour. Enhanced powers and opportunities for adolescents, liberalism, wider sexual freedoms and more independence and career opportunity for women have in themselves helped to produce changes in disease-patterns, and society's response to them. Of particular note in this respect has been the awakening of a less subservient role for consumers of health services, and their ability to become involved in the mechanism of health delivery systems. Amidst all these changes there is also need for historical and comparative perspectives in understanding use of available health services. For example in researching reasons which might explain the reluctance of many people in Britain to visit dentists on a regular preventive basis, it is important to grasp the historical perspective, and to see how ingrained existing practices have become in cultural and traditional settings. Parallel studies in the field of medical sociology, and in the social history of medicine have much to offer in this respect.

The effects of these and allied changes during recent decades have been to bring about a new focus of emphasis for Health Education and to suggest scope for Health rather than Disease Education. Widening definitions of sickness, and the increased attention upon factors such as obesity, alcoholism, drug addiction and dependency, accidents and depression have led to wider concepts not only of treatment but also, and more importantly for Health Education, of prevention. Persuasive appeals rather than legal sanctions have assumed pivotal significance, as an individual's health has become largely a personal rather than public responsibility. Motivation for healthy living has therefore assumed crucial significance, and programmes of Health Education, at least in industrialized societies, are increasingly being directed at healthy people and particularly those with sociodemographic and behavioural characteristics which suggest they may be at greatest risk. To reiterate a point made earlier: promotion and maintenance of health has become a partnership of interests since a decline in the old authoritarianism, which often viewed the client or patient as subservient to the practitioner.

Control of communicable disease was dependent on the ability to change the physical environment—by contrast it would seem that control of chronic and man-made disease requires changing the psycho-social environment. Suchman [5–7] taking the epidemiological model of host, agent and environment, provides a useful link between preventive medicine and the social sciences. He sees man as the agent of many of his own diseases, and suggests that his health state may be more determined by what he does to himself than by what some germ or viral infection does to him. Through his own unhealthy attitudes and behaviour man (the host) becomes the agent of his own destruction rather than the innocent victim or accident of disease. Although medicine might say the causes of lung cancer and heart disease were chemical substances in cigarettes and cholesterol deposits in the blood, another explanation would be behavioural: smoking, eating and lack of physical exercise. Cigarettes, alcohol, drugs and cars are not in themselves dangerous until used incorrectly. To return to the epidemiological model: the host would be characterized by personal readiness factors, such as recognition of seriousness, acceptance of personal vulnerability, predisposition, motivation, ability to act and knowledge of the belief in the desired action. The agent would be represented by situational factors, such as effectiveness, pleasure and effort of action, previous experience, favourable environment in which action ensues, and attractiveness of action. The environment would be described in terms of social control factors, such as social pressures to act, social membership, incorporation in role performance and acceptability of action. Suchman's model, whilst not however specifying the relative weightings of such variables, introduces us to a wide range of factors explored in the social science literature, and to which we return in the following section of this paper. The model has particular relevance for Health Educational strategies and schemes of intervention, and Suchman himself used it in an analysis of acceptance or rejection of an accident prevention measure (wearing protective gloves) amongst Puerto Rican sugar cane cutters [6].

In the wider sense the object of Health Education will be (a) to decrease exposure of the host to disease-causing ways of life (smoking, drinking, stress) and susceptibility, (b) to increase supportive forces, help avoid disease and adopt health-promoting and -maintaining practices and to (c) decrease the frequency and virulence of exposure. In essence then we have called attention in this section to changes in disease patterns, approaches to medicine and health care, societal organization, and community response to health problems. This has led to a burgeoning of interest in determining the role of cultural, sociological and social psychological factors in pointing to success or failure in operation of Health Education programmes. It remains now to describe some of the social scientific concepts which have become of increasing significance in the successful practice of Health Education. Clearly more than the dissemination of knowledge is necessary in order to intervene successfully to promote and maintain health, and we can no longer ignore the parameters of social, psychological, cultural, economic and political power.

3. SOCIAL SCIENTIFIC CONCEPTS AND RESPONSE TO HEALTH AND DISEASE: LESSONS FOR HEALTH EDUCATION

It is apparent that health behaviour has cultural, social and psychological determinants, all of which are inter-related and which interact together. What have studies in these three areas suggested that is of immediate or practical relevance and application to those who seek to bring about amelioration in programmes of health prevention, promotion and maintenance?

486

(a) *Cultural studies*

The goal of Health Educational efforts lies not, as we have reiterated, in having individuals merely give back factual information but rather in demonstrating actions relevant to health promotion. As numerous anthropological and community development studies have shown, there however occur problems of resistance on account of a clash with cultural and symbolistic traditions. Hence a need, as Goodenough [8] notes, for an analysis of community forces which should be considered in introducing changes in beliefs and behaviour. The nature and influence of a conflict in cultural values, and thereby often a failure in health programme innovation and the introduction of unanticipated effects resulting from the integrated nature of cultural institutions, has been well-documented [1, 2, 8–18]. More specifically we can point to a failure to understand the relationship between traditional values, beliefs and behaviours and social, cultural, religious, and ethnic structures and systems. Paul's monograph [11] catlogues a series of intervention studies in widely different cultural settings, and provides overwhelming evidence of the need for successful innovation to take account of existing community organization, local customs, values, taboos and religious beliefs. He documents the failure of many programmes to recognize the social and cultural functions that medicine performs, and reports how public health practitioners at work in what are to them alien cultures often experience difficulties in convincing people of the value of such preventive health measures as immunization and decontamination of water supplies, since they fail to reinterpret these actions in terms of prevailing concepts in the local culture. In essence then it is a failure to weave scientifically-tested discoveries into the fabric of daily life within the given culture or community.

Reports from several investigations will serve to illustrate the point that what appears from outside as irrational belief and behaviour becomes intelligible when viewed from within. The frequently-cited study of Wellin [19, 20] relating to efforts to convince Peruvian housewives of the need to boil their drinking water supports the axiom that customs like men are not islands unto themselves. After two years' effort a resident hygiene worker visiting individual homes was able to persuade only eleven housewives to boil their water in addition to the fifteen who already did so. Water boiling, which was associated with being sick, was not a simple isolated action, and the meagre response could not be put down simply to apathy, ignorance or stubbornness. Social distance between health worker and client was also shown to be another important variable. Miner's study of attempts to prevent sleeping sickness in Nigeria [21] also illustrates the failure of a rational-authoritarian approach, and underlines the role of magical and religious ideas – in local terms sickness is caused by spirits not the tsetse-fly. Similarly Adeniyi [22] has demonstrated how efforts in Ibadan, Nigeria, to suppress a cholera epidemic with a control programme were neutralized by traditional beliefs. Some people believed that the epidemic was the result of the anger of gods, and sacrifices were publicly made to propitiate them and stop the epidemic. A considerable number resorted to native medicine because they believed cholera was similar to other diseases with symptoms of diarrhoea and vomiting. The literature is replete with similar reports relating to failures to implement preventive health programmes in various cultural settings. Four further examples serve to illustrate the importance of due attention being directed to this problem area: McDermott and his colleagues [14] have pointed to the advantages of trying to work with and through rather than against local Navajo traditional and indigenous practitioners - a suggestion echoed by Morgan [23] in Nigeria in the course of a smallpox and measles eradication programme. Barnes and Jenkins [24], reporting field experiences in a malaria eradication programme in Surinam, underline the importance of cultural barriers and rivalries between bush negroes and town people and the effect these have upon the malaria eradication programme. A more recent reminder has come from Baric and his colleagues [25], who studied poor response to a cancer screening programme in Southern India. They throw doubt on transferring uncritically to another area of the world Health Education methods which have proved successful in a totally different cultural setting. In the society they studied failure to respond was not so much as in western terms for fear of cancer but because women were not supposed to visit a doctor unless they were sick, and to do so openly was to be seen as a deviant from existing social norms. Thus if a woman came forward for a cancer test she tended to invent reasons for her visit and not to tell her friends about it or act as an opinion-leader.

In seeking to transmit new health practices several important lessions are to be learned. Firstly it is essential to recognize that perception of and response to a particular disease may differ considerably from a Western or professional definition. Secondly it is important to realize the role of natural and supernatural factors and superstition, and to understand the need to operate within a framework of existing values, perhaps by modifying this rather than seeking to introduce a completely new system. Above all cultural patterns may affect the manner in which a health education programme is perceived and received. Amongst the powerful determinants of behaviour will be the pressures exerted by groups to which the individual belongs, and the more integrated the individual is with these, the more he will behave in accordance with group norms. Hence to try to change individuals without group sanctions may be to invite structural and societal dysfunction.

Lest it should appear that the above comments refer only to Health Education efforts in developing countries, it is relevant to refer to the work of Zborowski [26, 27] and Zola [28] in drawing attention to differences in American society in perception of and response to pain and illness which may be put down to cultural differentiation, ethnicity and family socialization practices. Saunders' study of attitudes and behaviour relating to medical care amongst Spanish and Anglos in the south-west of the U.S.A. [29] points to diverging practices. The Anglos preferred modern medicine and hospitalization, whilst the Spanish folk favoured family care or support, which was more consistent with their cultural conceptions. Clark's study [30] of health in the Mexican-American culture also provides confirmation of sub-cultural differences within American society and demonstrates that these factors can affect not only attitudes towards pain and

illness, but also the way disease is presented to doctors, and the approaches to preventive measures. This then is to underscore sub-cultural differences in values, goals, language, technology, religion and life-styles, and to help explain that failure to respond to health campaigns may not be irrational but logical when seen in relation to cultural conflicts.

(b) *Sociological studies*

Diseases and illness patterns have for some long time been areas of sociological endeavour. Many medical sociological studies of health and sickness have focused on illness behaviour: and specifically on the sick role, reactions to pain, aches and symptoms, and the utilization of health services. Current emphases in preventive health have led to a widening of social scientific efforts from merely detailing who registered and legitimated themselves as sick, to the seeking of clues to the reasons why individuals adopt health behaviour, and in the absence of any immediate threat or experience of illness adopt measures to prevent disease and to maintain health. Kasl and Cobb [31], in a selective review of three major aspects of behaviour related to health and illness, health behaviour, illness behaviour and sick role behaviour, have provided a theoretical framework for understanding the pattern and sequence of interacting social and psychological forces on the individual as he reacts to his state of health. Health behaviour they define as "any activity undertaken by a person believing himself to be healthy, for the purpose of preventing disease or detecting it in an asymptomatic stage". Health Education is thus clearly associated with health behaviour, and its proponent will need to be aware of the reasons for adopting (or failing to adopt) appropriate behaviour.

To these three behaviours listed above Baric [32] has added a further dimension—at-risk behaviour, which he explains in terms of the following phases: a healthy person acquires certain information about a health threat, and validates it with respect to his own person, he then explores his social environment about the credibility and general opinion concerning the threat and the preventive action recommended, he seeks confirmation from health professionals and information as to whether the threat concerns him directly—if he receives the confirmation needed and decides to act upon it, he accepts his membership of a high-risk group and behaves accordingly. In this way he accepts an "at-risk" role. The implications for Health Education are important, and intervention points in this model need to be explored. It seems unlikely that we can hope to change everyone's behaviour: nonetheless, by reducing our target population to those at high risk, we may provide the necessary preconditions for concentrated effort and maximum success.

It should however be realized that individuals who are inclined to seek diagnosis and treatment when symptoms appear or to enter the sick role after diagnosis may not show a similar inclination to take preventive health actions. It is nevertheless important for those who seek to influence and condition individuals to behave in preventive fashion to have a thorough understanding of factors which facilitate and impede an individual's decision to seek health care, and adopt health-promoting procedures.

In order to comprehend reponse to exhortations to adopt health behaviour we need to have a broad vision of factors which may intervene between a person and his decision to seek care. In addition to the cultural factors outlined above, socio-demographic and sociological variables have been shown to affect outcome. Just as with cultural responses to illness, many noxious practices continue to persist not because people are unaware of their existence but because conflicting interests are at stake. At times there may be people who are willing to seek advice to prevent future health problems but are denied access to such assistance by bureaucratic structures—an example being the unwillingness of many health authorities for a variety of reasons to provide a birth control and family planning service. Mechanic [33] has catalogued some major factors which affect response to illness: (a) visibility and recognizability of symptoms; (b) the extent to which symptoms are seen as serious; (c) the extent to which symptoms disrupt normal activity; (d) the frequency and persistence of symptoms; (e) the tolerance threshold; (f) available information and cultural assumptions about the symptoms; (g) denial tendencies; (h) competing needs; (i) competing interpretations and rationalizations; and (j) availability of treatment and the economic and psychological cost involved. To these one might add the situational factors which make it more or less difficult to seek care (i.e. problems of transport and time) and faith in the medical care team to provide easy and lasting solutions. More recently Fabrega [34] has advocated a further theoretical analytical model to account for decisions that are made by a sick person in his career to patienthood.

Many studies have pointed to both a differential perception of symptoms and use of services. Koos [35] found in his now almost universally cited study of *Regionville* that upper class people more often reported themselves ill than did lower class persons, and were also more likely to seek care: by contrast lower class persons had more actual symptoms, reported less illness and made less use of health facilities. A large number of studies have been undertaken in recent years to document differing perceptions and reactions to illness, and illness behaviour itself—those of Baumann [36], Fred Davis [37] and Roth [38] offer significant descriptions relating to three different illness patterns. Another important contribution was Suchman's study of sociomedical variations among ethnic groups in New York City [39]—this revealed significant differences in knowledge about disease, attitudes to health services and in illness behaviour.

Medical sociologists have thus accumulated a stockpile of evidence to document the relationship of socio-economic and socio-demographic factors to knowledge of disease [40], use of medical, dental and allied health services [41-44], acceptance of preventive health services [45, 46], delay in seeking care [47] and use of self-medications [48, 49]. Equally the reasons for seeking care and the process or sequences in decision-making in relation to seeking care have been considered and theoretically formulated. Stoeckle, Zola and Davidson [50] have reviewed features in a patient's decision to seek medical aid, and Zola [51] has since developed the notion of a "critical incident" rather like Rosenstock's idea of a cue or trigger [45]. Zola conceives of five non-physiological triggers, each of which may assume differing degrees of importance in differ-

ent groups, and which may set off action: (i) occurrence of interpersonal crisis calling attention to symptoms; (ii) perceived interference with social life and valued activity; (iii) sanctioning and the social pressures of others; (iv) perceived threat and interference with work; and (v) temporalizing of symptomatology. Suchman [52] has conceptualized a five-stage model of decision-making points (in relation to illness behaviour) in the patient's career. Decisions that need to be taken are that (a) something is wrong (symptom experience), (b) one is sick and in need of care (assumption of sick role), (c) one actually should seek care (medical care contact), (d) one should transfer control to the doctor (dependent patient) and (e) after treatment one should relinquish the patient role (rehabilitation and recovery). Strategically one might postulate a similar model in terms of health behaviour involving (a) symptom perception, (b) assumption of at-risk role, (c) preventive visit or test, (d) co-operative patient and (e) symptom-free. Thus Health Education specialists might profit from the development of closer insights into the decisions an individual faces in relation to assumption of the health behaviour advocated in their campaigns.

Equally important is another paper by Suchman [53] in which he notes that, although public health practitioners have long been aware of the failure of people in low socio-economic groups and ethnic minorities to respond to preventive health campaigns, the reasons for such inactivity have not been precisely explained. Accordingly he advocates consideration of the structure of social groups to which an individual belongs, and the health orientation or value system to which he adheres. The form of social organization with which an individual was associated was found to be significantly related to his medical orientation. "Cosmopolitan" types of groups were more likely to hold a "scientific" health orientation, whilst "parochial" types adhered to a "popular" health orientation. In Health Education terms this is to suggest that the strong group identification and cohesion of "parochial-popular" groups provides a natural avenue for programmes which set group approval as being the appropriate, acceptable and desirable behaviour. By contrast the independent individualistic nature of members of "cosmopolitan-scientific" groups lends itself to the formulation of a programme which makes sense and appeals on rational lines. Furthermore when a person makes a decision to seek preventive care, he may be influenced not only by the structure and availability of services and his knowledge of health professionals and institutions but also by opinions amongst his social contacts—the social network which Freidson has termed "the lay referral system" [54, 55]. Thus people seeking care, and health practitioners may be viewed as two components of an interacting system—indeed there is need for a recognition of conflict between the professionally determined needs for care and the felt needs of the general public. Indeed it is essential for the Health Education specialist to underscore any differences in this respect, and to seek to explore means of compromising the two goals without losing sight of his behavioural targets.

Another revealing study from which Health Education specialists might draw insights is that of Banks and Keller [56] into the relationships between symptom experience and health actions. They focus on the role of prior experiences—a variable ignored by many other investigators—and show that respondents who report prior family experience with particular symptoms tend to express less inclination to seek prompt medical care. By way of explanation they suggest that knowledge from prior personal symptom experiences reduces anxiety and uncertainty, and leads to a diminution of motivation for immediate medical action. The significance of the gravity of the disease and value placed upon medical care may affect the issue, but certainly Health Education specialists may care to reflect upon this study's findings, and also upon those of Steele and McBroom [57] who questioned the multi-dimensional of health behaviour, and of Williams and Wechsler [58] who studied inter-relationships between preventive health practices, each of which may be affected by personal, social and situational factors. In another American study Green [59] confirmed, with certain exceptions, a pattern of colinearity of health behaviour variables and suggested a common source of influence and/or a mutually reinforcing influence of one behaviour on another—in particular socio-economic status appeared as the most important predictor of adoption of preventive health practices. More recently Coburn and Pope [60], using data from a sample of British Columbian working men, tested six hypotheses explaining the link between socio-economic status and preventive health behaviour. They found that separately none of the explanatory variables greatly accounted for the influence of socio-economic status and that education, age, income and social participation (in that order) provided the most parsimonious set for predicting preventive health behaviour, which was measured in terms of preventive medical examinations, preventive dental visits and polio vaccinations. The implications of a collinearity are to suggest a greater carry-over from programme to programme: further study is needed on this subject.

(c) *Social-psychological studies*

It is axiomatic that a health care system cannot do more for people than they themselves are willing to receive. Thus we need to know of people's opinions and attitudes in addition to their behaviour if we are to understand the relationship between, on the one hand, felt need and potential demand and, on the other, real demand. An unclear picture however exists of the exact relationship between beliefs and attitudes and either subsequent health or illness behaviour. Many health promotional efforts have been based on the premise that changes in knowledge and attitudes will produce significant changes in the direction of preventive health behaviour—in itself a questionable assumption.

Based on a series of research studies of response to programmes offering screening by chest X-ray for tuberculosis detection [61], immunization against Asian influenza [62], prophylaxis for rheumatic fever [63] and preventive dental care [64], there has been developed a paradigm known variously as the "Psychological Readiness Model", "Social Psychological Model" and "Health Belief Model" which attempts to explain the behaviour of people who on a voluntary basis avail themselves of preventive health services. It has been subject to a number of amendments—Rosenstock [45] provides a critique of its accomplishments and failings. Health behaviour is judged as an emer-

gent of several interacting beliefs which act as determinants of an individual's likelihood of acting. In essence these consist of a perceived threat, and the attractiveness and value of a perceived benefit. Thus it is contended that a person will engage in preventive health behaviour if (i) he believes himself vulnerable or susceptible, (ii) he believes contracting the particular condition will have serious consequences for him, (iii) he believes a given course of action will benefit him, (iv) he believes there are no barriers (in cost, time or pain) in adopting the requisite action and (v) he experiences a cue or trigger to set off the appropriate action. This latter incident may be dependent upon some instigating event—as Stadt and his colleagues [65] noted in relation to an increase in children's dental visits following the issue of dental health birthday cards as reminders—perception of bodily state, interpersonal action, the impact of mass media, or invitation to take part. Antonovsky and Kats [66] and Antonovsky [67] following applications of these ideas in Israel raise certain doubts about the model—the former study, in attempting to predict dental behaviour questions the need for a trigger or cue from outside and suggests that if a person holds the four beliefs he will set off his own cue to action. These authors also dispute that psychological readiness will lead to positive motivation to act—an individual may desire to act in the right direction but fear prevents him from translating his belief into action. Furthermore they call attention to the existence of threshold levels, and suggest that above certain beliefs there will be no gain in preventive behaviour.

Initial tests of the model were retrospective studies. Hochbaum [61] demonstrated that belief in susceptibility and in the benefits but not in severity were associated with seeking chest X-rays and Kegeles [64] showed those who believed themselves highly susceptible to dental problems made significantly more preventive dental visits, whereas neither beliefs in seriousness nor benefits of taking action were related to subsequent behaviour. In a further and prospective study Kegeles [68] followed up his earlier survey-group in an attempt to determine whether the beliefs recorded in the original data were associated with subsequent behaviour. Perceived seriousness and benefits did not predict behaviour, although perception of susceptibility did show a correlation with making subsequent preventive dental visits. The model has not been without its critics and two specific problem areas occur. It has assumed a causal relationship of beliefs and behaviour and also the beliefs preceded behaviour. In the retrospective studies the data about both beliefs and behaviour were collected at the same time, and beliefs were assumed to have existed before the behaviour—a questionable assumption in itself. Prospective studies are few in number, and it is therefore difficult to come to any firm conclusions in terms of predictions.

In a national study of health beliefs concerning cancer, tuberculosis, tooth decay and gum disease Kirscht and his colleagues [69] asked questions about perceived severity, susceptibility and benefits. Their study found little evidence of a general preventive orientation towards disease, and pointed to the need to plan different educational programmes to promote prevention of the various diseases. In cancer and tuberculosis education they suggested that the objective should be to increase beliefs in both susceptibility and the benefits of early detection. In the cases of gum disease the emphasis should be on increasing susceptibility and regular dental visits, and of tooth decay on severity and regular visits.

Like the anthropologists and sociologists, social psychologists have been building up a data base to help explain reasons for adoption or neglect of preventive health behaviour. Not surprisingly each discipline has often tended to ignore the others, and to fail to consider any possible interaction between cultural, sociological and personality factors. In the final paragraph of this section we note some of the recent efforts to synthesize and draw together the results of studies from these three fields, and point to the advantages for Health Education practice of a multi-sided search for explanation.

(d) *Multi-disciplinary studies*

Cultural influences, social pressures, socio-demographic factors and individual needs have, as we have described, all been used to help explain variation in response to campaigns to promote Health Education and preventive medicine. It seems apparent that any attempt by the Health Education specialist to change attitudes and behaviour must consider all three components. Young's valuable literature reviews [70–75] of research and studies relating to Health Education practice underscore this multi-disciplinary approach.

Several authors have attempted to co-ordinate these various viewpoints—hence Gibson's commentary [76] and McKinlay's overview of use of health services [42] which points to six analytically distinct research strategies: economic, socio-demographic, geographic, social-psychological, socio-cultural and organizational. Anderson [43] has suggested five similar approaches: socio-cultural, socio-demographic, social-psychological, organizational and social systems, and concluded that "what is sorely needed is careful, theoretically based attempts to explicate causal structures that incorporate major features of all these approaches". Anderson and Bartkus [77] with that need in mind have formulated a model consisting of multi-disciplinary variables to account for differential patterns of health and illness behaviour among members of a pre-paid medical group plan. Berkanovic and Reeder [78, 79] have concerned themselves with ethnic, economic and social psychological factors in use of health sources, and suggested a convergence of research and opinion leading to the conclusion that ability to pay is the major determinant of use of health services once symptoms are perceived as serious.

Synthesis of data from studies in these different fields should enable the Health Education specialist better to plan health promotion and maintenance efforts, which will be acted upon to a higher degree than hitherto. By creating an awareness of the role of cultural, sociological and social psychological factors the studies cited above and their numerous counterparts will offer the prospect of practical applications leading to improvements and increased sophistication of effort in all those many activities which go together to constitute Health Education.

4. METHODS AND TECHNIQUES IN HEALTH EDUCATION—SOME SELECTED APPROACHES

The broad canvas of the scope and potential of Health Educational efforts will have been apparent

throughout this paper. In this section we consider some of the methods and techniques which have been used in recent times. Reference to Young's compendia of recent research findings [70–75] will illustrate the wide range of effort. Clearly it is beyond the compass of this present paper to attempt to describe all of the strategies adopted for differing purposes in different countries. Instead several problem areas have been selected for comment: what follows in this section is thus intended as window-display of a much larger exhibition of products.

It is clear that efforts which have been undertaken under the general rubric of Health Education have varied in terms of sophistication, intensity, quantity, quality and effectiveness. A variety of tactics has been used to inform, educate, persuade and cajole individuals, groups and communities to adopt particular forms of health behaviour. Some of these have operated formally, others informally, and still others both formally and informally. Inputs have been varied, and whilst enthusiasm has usually and initially been high, much of what is referred to in every day terms as Health Education has been pragmatic, owing more to rule-of-thumb methods than scientific and objectively-tested techniques and strategies. Methods have varied from face-to-face intervention to use of mass media: some have worked, others have been less successful, but the reasons for success or failure amongst particular groups have often escaped explanation. Whilst decrying an apathetic public response, Health Education practitioners have frequently failed to ask themselves if the particular strategies used were appropriate, acceptable, and proven for the use to which they were put. In essence much Health Educational effort is dependent on communication—the purpose of which may be to inform, educate, motivate or re-assure. For programmes to succeed they must modify values and behaviours that are often deeply rooted and supported by social sanctions. It is idle to suppose that merely because individuals are provided with up-to-date information they will respond in the desired fashion. Health Educational efforts are only as successful as are the methods and techniques used to disseminate appropriate messages, and change agents are confronted with the problem of evolving a suitable strategy to initiate and sustain desirable behaviour change. Thus in this Section we consider techniques of communication and the role of compliancy, strategies of social change, the relationship between beliefs and behaviour, and the use of fear-arousal as a motivator of action. Again no attempt is made to be exhaustive in outlining available approaches, and questions of effectiveness are primarily left over until the next section. Whatever methods are used they have in common one focal objective—namely a behavioural change on the part of the recipient. Even when the overall objective is the same, different methods may need to be developed for different population groups, and what works in one setting may not do so in another. Above all it should be remembered that social scientists are only now beginning to tap the reasons for these differences of response, and hoping in experimental and applied settings to be able to influence response-patterns.

Health Education frequently begins with an effort to plug some gap or deficiency: people are not using facilities designed in one way or another to promote health and prevent disease, or they are acting in ignorance of or in opposition to medically-supported knowledge. Thus the need to communicate with them, and to attempt in some orderly fashion to intervene, and persuade them voluntarily to change their behaviour and beliefs. This is to suggest the need for a channel of communication—a new formal link, or an informal line which operates through existing social networks and the lay referral system. As Swinehart [80] suggests, we need to identify factors associated with receptivity to health communications, determine the relative importance of these factors, and develop ways of manipulating the most important determinants—all these areas need to be researched, and documented for further operational use. It would appear that lay individuals when presented with medical information recall best what they are told first, and what they consider most important [81]. Studies of communication in medicine and of the doctor–patient relationship [82–86] focus on features of the interaction process and show how this may facilitate or hinder the quality and quantity of care recived, and the adoption or neglect of preventive health behaviour. McIntosh [87] has provided a review of the role of formal and informal communication, information-seeking and control associated with one specific area, namely cancer. Thus communication may be incomplete: it may be satisfying to the communicator but leave the recipient unmoved. For Health Education communications to be truly effective they must be seen to offer more advantages than drawbacks. A variety of techniques of communication has been used in the field in efforts both to teach those who are healthy how to stay healthy, and those who are ill to get well. These range from informal dissemination and interpersonal communications using indigenous aides through to group techniques using opinion-leaders, or political, religious and professional pressures, and may or may not be aided by audio-visual materials and mass media. The literature is replete with studies of the effects of particular techniques in intervention schemes: one conclusion seems to stand out—namely that decisions are seldom made after only one approach, and multiple contacts are needed with a variety of sources of communications over time. Many are the reasons why people fail to respond to Health Educational propaganda—thus a variety of techniques may be necessary to obtain public support and recognition. Even when they have access to knowledge outlining correct behaviour, it seems certain from Patient Education and compliancy studies that many therapeutic processes and medical regimens are ignored [88–94]. The doctor's advice may clearly conflict with existing behaviour, and it may not be enough for Health Educational campaigns merely to prescribe correct action.

For many the techniques of Health Education have been described in terms of posters, pamphlets and flannelgraph displays. Methods and media are but tools and devices, and not therefore ends in themselves. Hence talks, group discussion, buzz groups, panel discussions and role-playing, together with use of such media as radio, television, films, slides, tape-recorders, wall-charts, flannelgraphs, posters, leaflets, exhibits, models, programmed instruction and teaching models, are but means to an end. Health Education does not end with these techniques – it starts there after a decision to implement a planned programme. Similarly a

host of factors may bear upon the method of approach - not the least of these will be the general characteristics of the change agent in building up either empathy or rapport. Other important factors include the use of local language and terminologies which are understood by the target population, the ability and technical competence of the innovator and the degree to which he is able to convince his audience, organizational affiliations, commitment to objectives, motivation of felt and latent needs, the social and cultural structure, timing of the intervention and a measure of the benefits of change in terms of both short- and long-range costs and rewards. In short the Health Education specialist has at his service an *armamentarium* of devices to help him in his cause: wrongly used they may however handicap considerably the attainment of programme objectives. A research priority would seem then to be a controlled examination of the value of particular approaches (such as mass campaigns).

For many people a change in knowledge and belief has been seen as a prerequisite to behaviour—hence the methodology of numerous attempts to promote health behaviour by attempting to change and modify attitudes. Nonetheless there are certainly limitations to an approach which focuses merely on changing attitudes, since (i) attitudes may have no counterpart in behaviour, (ii) it is just as reasonable to conclude that behaviour causes beliefs rather than vice-versa and (iii) it is important to have a measure of the consistency, persistence and reality of attitudes. Festinger [95] has noted the scarcity of data which show that changing attitudes leads directly to behavioural changes, and made a plea that the "problem needs concerted investigation". Rokeach [95] is another to have analyzed theory and research concerning the relationship between attitude and behaviour change. Festinger [95] interpreted existing evidence as suggesting that behaviour change did not necessarily accompany attitude change, but Greenwald [97] has questioned this for all circumstances, though agreeing that there is not automatic relationship between attitude and behaviour: in essence the occurrence of behaviour change is not dependent on the prior occurrence of attitude change and vice-versa. The debate on how much correspondence exists between attitudes and behaviour has continued, and according to Warner and Defleur [98] there are three predominant postulates: (a) a consistency (attitudes do act as guides to behaviour); (b) an independent variation (no consistency); and (c) a contingent consistency (interactional concepts may modify the relationship between attitude and behaviour). The evidence, they suggest, indicates that neither of the first two of these ideas is tenable. Festinger [99] has suggested that behaviour and attitudes may be related in terms of a reaction to offset "congnitive dissonance", which occurs when attitudes and behaviour are inconsistent with each other. If one acts in a way which is inconsistent with one's beliefs, then these beliefs may change to be more in line with one's actions. The inconsistency between the new behaviour and held belief may set up discomfort or dissonance, which if strong enough may prompt a change in attitude. At this stage the role of group pressures assumes significance.

If the objective is to bring about a change in voluntary preventive health behaviour the focus should then be on action and its determinants. The social psychological work of Festinger *et al.* [100] has suggested the importance of social pressure and support and prompted consideration of the role of others, of social adaptation through group norms and peer groups, and of conformity in bringing about health behaviour change. The suggestion is that an individual's perceptions, feelings and actions emanate from group membership. A paper by Freeman and Lambert [101] supports such an approach and questions the role of pressure groups and influence of voluntary associations as agents of social change. In the long run however it is clear that an individual is subject to a vast range of influences, some systematic and legitimate—others incidental and unreliable, and out of this nexus of influences he will select some and reject others. Equally it would seem that there are particular times at which behaviour change is more easily attainable: childhood and adolescence obviously are vital periods, hence a preponderance of studies of School Health Education [75].

Both external forces such as mass communications and internal forces such as family and social supports may influence individual behaviour. Emotional rather than rational factors may influence any decision: in a sense it would seem legitimate to use any force which succeeds in getting voluntary preventive behaviour —care should however to exercised that in solving one problem others are not being created as byproducts. Green [102], in a perceptive review in which he attempts to place educational strategies emphasizing change in attitudes and beliefs in a social perspective, notes the relevance of seven important strategies of communication: dissonance reduction, deviance reduction, reinforcement, foot-in-the-door, behaviour observability, participation and involvement, and feedback. By removing any conflict between beliefs and behaviour, and any problem associated with preventive behaviour being termed deviant, by insuring reinforcement, and by perhaps seeking at first small modifications rather than total change, by being seen to be acting, by getting individuals involved, and by constantly looking to see how the programme is judged and reasons for non-compliance, then intended or desired and actual behaviour may be brought closer together.

It remains now briefly to refer to fear-arousal as a method of inducing attitude and behaviour change. Evidence from the recent reviews of Higbee [103] and Leventhal [104] suggests that threat appeals have sometimes changed things, and on other occasions no significant changes took place. Many of the studies reviewed in these publications were carried out in laboratory settings, on young subjects and it is perhaps difficult to draw precise conclusions for everyday practice. As in so many areas of approach ambiguities exist: the options of a mild unfrightening approach, and a hard fear campaign await wide-scale empirical testing, and particularly outside of North America.

In this section we have not been able to offer many "best buys". From the range of tactics and methods available, it is however evident that efforts to promote health education, whether formal or informal, have a myriad of approaches, and there is a prime need to tailor the methods used to fit in with the frame of reference of the target group. Theoretical input has frequently been missing, and methods adopted have owed more to personal and untested ideas than to the find-

ings of scientifically and objectively based research. No doubt because of the enormous variety of inputs, so too has effectiveness also shown vast variation. Methods and materials are only of value if those to whom they are directed are learning and benefiting from them—in many cases it is indeed difficult to assess how effective they are, and how their value compares with other techniques used in attempts to solve the same problem. Indeed much of the research in this area seems technique- rather than problem-centred. The vital question is that of asking how best can information on any given topic be presented so to effect maximum change in attitudes and behaviour. Communication is after all an essential in some form for all methods and materials—behavioural responses to Health Education communications are dependent on an aggregate of cultural, social and personality factors.

5. EFFICIENCY AND EFFECTIVENESS OF HEALTH EDUCATIONAL EFFORTS

(a) *General comments*

Health Education will be as efficient and effective as the level and strength of inputs to its programmes. A measure of effectiveness is often hard to calculate precisely since objectives are seldom spelt out in full and more often than not are assumed. Thus the objective of screening programmes is assumed to be to get all people at risk screened: only then does one begin to ask exactly which groups this constitutes, and to question if a 100 per cent response is in fact ever possible. Nonetheless the issue remains that Health Educational efforts are conducted for reasons other than solving a conscience that some effort is necessary. More often than not success, however this is judged, seems to occur by change rather than as a result of a carefully designed programme. Not the least of the problems facing national agencies in these inflationary years is the question of cost-benefits—an analysis of Health Education in such terms is difficult despite the fact that one might reasonably ask what was the effect of spending so much money on a particular scheme. Indeed there might be occasions when uptake of preventive services could, perhaps only temporarily, swamp available facilities and resources, and perhaps be said to be overeffective. For the most part however many of the gains will be minor or steadily accumulative.

Effectiveness may depend on a series of factors which we now examine in some detail. Important will be the quality of both health educators and health services—in a sense the two will be interrelated. How good in terms of advancing knowledge and practising what is known will be their efforts? How far they will be encouraged and allowed to develop educational strategies will be dependent on the priority given to them, and the available budget. How adequately in terms of coverage, and how appropriate were the measures introduced? How efficient was the allocation of resources such as time, money and effort? It is indeed surprising how little evaluative effort of this nature has been conducted, and how seldom have programme objectives been clearly set out, and the rationale for behaviour change explained. In order for a high level of effectiveness to be achieved there must be (a) easy access to participation or involvement on the part of an individual in the target group (if indeed one is specified), (b) a high feasibility of putting into practice the measures advocated, (c) a ready acceptability that the new behaviour accords with the values of the society and culture and (d) clear understanding of the benefits to be derived from the new practice.

Measurement of change is a further important criterion—before and after measures will be necessary to document the case. Herein lies a difficulty particularly if we are measuring knowledge and attitudes rather than direct outcomes or visits for tests. This is of course to return to the rhetorical question as to at what level of uptake a preventive health campaign can be said to be effective—what is the hallmark of success in practical terms? Indeed how realistic is it to expect success with the knowledge and answers available to us? Efforts to discourage smoking go on at many levels —how much of our failure to have any large-scale and lasting effect as a result of many current strategies is due to our failure to understand what to do?

Health Education represents a long hard slog: just as we begin to get to grips with one health problem, another appears on the horizon. Are we learning the lessons of the past and can we say we are learning from our mistakes? Many of the questions to which we have no answers have been raised before. Do we spend enough of our efforts on research, and attempts to improve our techniques? Specifically we need studies of programmes in many areas to see what succeeded, and what failed, and the reasons for these results. Failure to adopt a planned approach to health education has led to a certain lack of continuity. All too often campaigns which were never properly designed or conducted are termed failures, and the exact reason for these disappointments never charted. As Suchman [105] reminds us there can be two categories of reasons for unsuccessful intervention—either it did not activate the causal process which should have culminated in the intended goals, or it may have set processes in motion but did not cause the desired effects. It is as well for us to note that in many problem areas there is a lack of a firm methodological basis on which to plan experimental interventions—thus Green and Figa-Talamanca [106] call attention to suggested designs for evaluation of Patient Education programmes.

Health Education specialists seldom choose problems—they find them left on their doorsteps and hasty solutions sought for them. Health Education can clearly promise no panacea—by critically examining and evaluating their efforts those who work in the field can however provide effective measures of the earliness, frequency, quality, range and persistence of response to their efforts. Programmes will only be effective if people do what is required of them. Inputs may professionally be said to be high and efficient but public involvement low—if this is the case it will not be because the public are ignorant or apathetic but because the action required of them has low priority or is out of keeping with their value system, i.e. the focus of the programme needs revision, and stating in different terms. Hence the need for constant evaluation of effort, and for feedback into programme implementation. Above all there is a prime need for good hard data using accurate and sensitive indices to tell us how well a programme is meeting the purposes for which it is established, and whether it should be continued, expanded, cut back, changed, reproduced or aban-

493

doned. Relatively few services rather than research programmes have been examined in this critical frame.

What clues are there for us to employ in seeking to make health education efforts highly effective? Public receptivity, participant reaction, costs, and availability of staff are all important elements in determining success. Amongst the barriers operating to hamper effectiveness are social, ethnic and cultural conflicts, language and jargon, personal and group values, group membership, the form and content of communications, cost of action or treatment, habits, customs and taboos, availability of services, lack of faith in medical services, denial of existence of problem, belief in vulnerability, attitudes towards mass medication, fear and anxiety levels, administration of facilities, lack of felt needs, and education. Failure of the measure on account of some of these factors could well mitigate against the problem being termed effective. Again what seems effective on the drawing-board may not be so in practice, and the need for pilot programmes before full-scale implementation and for model schemes cannot be overstated.

(b) *Studies of non-participants*

Studies comparing acceptors and non-acceptors of screening programmes help provide feedback and develop an understanding of how the programme is perceived. By examining the reasons for failure to participate in preventive schemes we can identify potential areas for restructuring educational approaches. What patients and clients think of facilities offered to them will affect the way in which non-participants continue to react, and help explain overall response-rates, which will be an important aspect of any measure of effectiveness.

A number of studies may be cited to show the value of this approach. Schonfield and his colleagues [107] in Cambridge, Massachusetts and Nickerson and Hochstrasser [108] in Kentucky have both studied non-participation in a tuberculin testing programme —the latter investigators concluding that the majority of the reasons for non-participation were health-related, while political, economic and religious factors played lesser roles. Ward [109] studied reasons for drop-out at a family-planning clinic in Sheffield, England—in her survey the focus was on failure not in a one-shot programme but in one which called for a series of visits. Fink and his associates [110] considered repetitive screenings for breast cancer, and studied the viewpoint of reluctant participants. These two latter surveys do of course underline the need for effecting behavioural change not as once-in-a-lifetime phenomenon but as part of a sustained and repeat-visit programme. Similar considerations apply in the field of screening for cervical cancer, where a number of studies have pointed to differences in socio-demographic factors, use of health services, and attitudes between acceptors and non-acceptors to an invitation to avail themselves of a test. In two Finnish studies both Kaupinnen and his fellow workers [111], and Vuori and his associates [112] underline the finding that some who fail to participate (and hence reduce response-rates and the apparent level of effectiveness) have perfectly legitimate reasons for their action. Similar conclusions were drawn from comparable studies undertaken in Malmö, Sweden by Bjerre [113] and in West Sussex, England by Richards and McEwan [114]. Understanding of reasons for non-participation permits an evaluation of the need for alternative educational strategies, or shifts in emphases in current problems.

(c) *The effects of innovation and diffusion*

A growing concern to maximize the effectiveness of health services has led to increasing attention being devoted to innovation and factors underlying implementation and acceptance of new programmes. Studies of innovation may profitably be focused on the organizations and factors which impede or facilitate change, and participants. Kaluzny and his colleagues [115] offer an overview of work in this field.

Effectiveness of a plan to promote some aspect of health may be judged by the speed with and degree to which there is manifest acceptance of the scheme. Vertinsky and others [116] have proposed a paradigm to try to understand some of the complexities involved in changing human behaviour. The important contributions of Rogers [117] and Katz [118] in the area of innovations suggest a number of interesting applications in the field of Health Education, particularly in the suggested sequence of stages in the adoption process of a new practice—awareness, interest, evaluation, trial and adoption. Also of value is a division of adopters into innovators, early adopters, early majority, late majority and laggards, together with a discussion of opinion-leaders and the "two-step flow of communication" [119]—the latter concept being questioned by Hingson and Lin [120] following experiences during a mass immunization scheme in El Salvador. The original "two-step" theory suggested information and influence flowed from mass media to opinion-leaders who relayed it to others less active—now it is believed that the formal leaders of health opinion may not be necessarily those who facilitiate "two-step flow". Clearly however the manner in which health communications are relayed through the formal and informal social system may exert an important influence on the apparent effectiveness of any proposed programme of behaviour change.

6. PROGNOSIS—SOME RANDOM THOUGHTS FOR FUTURE APPLICATION

In this paper we have tried to view Health Education with a wide-angle lens, and without specializing in any disease or national settings, to offer a broad picture of the possible scope of social scientific endeavour in the highly diverse field of Health Education. It is clear that there are presently considerably more questions to ask than there are answers available.

Studies of factors and strategies relevant to educational change and preventive health behaviour lie scattered about the literature of a number of disciplines. Little codification of these data has been attempted, nor indeed has the strength and practicality (in terms of direct use in programmes) of many of these findings been objectively judged. The theories and findings of separate investigations have all too seldom been linked together, or applied in other studies. Consequently there has hardly been a building-up brick by brick of a body of accepted knowledge, which has stood the test of time and cross-national experimentation. It is unclear how far we may generalize from local surveys, and from laboratory experiments. Chambers' Twen-

494

tieth Century Dictionary defines science as "knowledge ascertained by observation and experiment, critically tested, synthesized and brought under general principles". Can we therefore really say that there has already been a science of Health Education? How far indeed can we say that Health Education has a tested body of knowledge forming principles on which professional practice is dependent? There is then a need for linkages between theory and action and no substitute exists for scientific rigour and exactitude, well formulated theories of research, a sound tested body of knowledge, and a systematization of that knowledge. T.S. Eliot in *The Hollow Men* has written that

> Between the idea
> And the reality
> Between the motion
> And the act
> Lies the Shadow.

Health Education certainly has the idea and the motion, and is probably currently lying in the shadow, for much of what passes in many countries and in common parlance for Health Education consists of armchair philosophy, missionary zeal, and statements of personal beliefs which owe little to scientific and objective method. A less than adequate methodology, biased samples, lack of control groups and unreliable frames of measurement also characterize many research efforts in the area. Suchman's criticism may perhaps seem harsh but it does bite deeply: "the field of public health education has not made use of many of the current concepts and methods of public opinion and communications research . . . Unrepresentative samples, poorly constructed questionnaires, unskilled interviewing, and inadequate analysis are the rule rather than the exception" [121]. The ailing subject-area awaits therapy then at the hands of the social scientist and statistician, and technology is needed not merely to define problems but to also help provide answers.

Health Education needs to harness the unbridled enthusiasm of some of its most staunch advocates, join this with the objective and calculated approach of others, and amalgamate them both with the interest and involvement of the groups most at risk. It is time for Health Education practitioners to cease being preachers and teachers and to realize we are all in a learning situation. For this commentator there would appear six important priorities for future practice: Philosophy, Policy, Priorities, Planning, Problem-solving and Programmes and it is to be hoped that this position paper has served to illustrate the role social science can assume in each of these areas.

We conclude with reference to three particular areas for future growth and development. Firstly there can still be no substitute for the dissemination of up-to-date health information—the social climate and the frontiers of knowledge are constantly changing, but there is still a basic responsibility to see that the new facts are consistently disseminated. To take an example of a field in which it would be quite wrong to assume everyone knows the answers: a very small minority of British adults showed evidence of their understanding of periodontal needs. Comparison of an individual's own subjective and a dentist's objective assessment of treatment needs revealed a startling lack of congruity [122]. Secondly we point to the strategy of changing the environment when it is clear that response to propaganda is lacking. This is to suggest that we make it easier for the individual to respond—by offering a free service, by as far as possible making it available at hours convenient to the client, and by removing travel problems. Hence for example a do-it-yourself cytopipette test which can be used at home, and obviates the need for a clinic visit for a cervical smear [114], and the use of a mobile unit in an effort to take cytology services out into areas where response to traditional cervical smear campaigns has been shown to be low [123]. Thirdly there is a continuing need for comparative surveys to study the relationship between parallel techniques and programmes. The search for the one right way of influencing all people seems a nebulous approach—by understanding what works or fails in which groups we shall be progressing towards that total explanation. Watson [124] for example evaluated four different techniques in modifying knowledge, attitude and practice relating to smoking. In a study in five schools she compared didactic, group discussion, psychological persuasion and project approaches with a control group. In the field of dental health Myers and Downe [125] compared two different ways of influencing knowledge, attitude and practice and Richards and Taylor [126] are presently conducting a Dental Health Education programme which involves a comparison between teacher-instructed, class project, parent-involved and dental auxiliary-instructed groups of 6–8-yr-old children together with a control group.

To return to our basic tenet, Health Education is not forcing people to act positively—rather it is persuading them along certain lines to do what they actually want to do for themselves [127]. Thus participation by the public and self-reliance become key concepts wherein Health Education is viewed as a community affair in which health practitioners and the public combine together. George Bernard Shaw in *The Doctor's Dilemma* noted that "what the public wants is a cheap magic charm and a cheap pill or potion to cure all disease". Health Education presently has neither a cheap magic charm, pill nor potion to prevent all disease, although there are many who are hoping for these discoveries—all the greater then the challenge which confronts those who strive to effect a more scientific and successful development of Health Education. It is clear that this sampling of the literature is indicative of a wide approach to Health Education: what is needed now is a fusion of knowledge from the social, health and educational sciences, and subsequent incorporation in action programmes.

REFERENCES

1. Polgar S. Health action in cross-cultural perspective. In *Handbook of Medical Sociology* (edited by Freeman H. E., Levine S. and Reeder L. G.), pp. 347–419. Prentice-Hall, Englewood Cliffs, 1963.
2. Polgar S. Health and human behaviour: areas of interest common to the social and medical sciences. *Curr. Anthrop.* 3, 159, 1962.
3. World Health Organization. Research in health education: Report of a WHO scientific group. *Wld. Hlth. Org. tech. Rep. Ser.* 432, Geneva, 1969.
4. Powles J. On the limitations of modern medicine. *Sci. Med. & Man.* 1, 1, 1973.

495

5. Suchman E. A. *Sociology and the Field of Public Health*. Russell Sage, New York, 1963.
6. Suchman E. A. Preventive health behaviour: a model for research on community health campaigns. *J. Hlth. Soc. Behav.* 8, 197, 1967.
7. Suchman E. A. Health attitudes and behaviour. *Arch. Env. Hlth.* 20, 105, 1970.
8. Goodenough W. H. *Co-operation in Change*. Russell Sage, New York, 1963.
9. Spicer E. H. *Human Problems in Technological Change*. Russell Sage, New York, 1952.
10. Mead M. (editor) *Cultural Patterns and Technical Change*. UNESCO, Paris, 1953.
11. Paul B. D. (editor) *Health, Culture and Community*. Russell Sage, New York, 1955.
12. Loudon J. B. Social structure and health concepts among the Zulu. *Hlth. Educ. J.* 15, 90, 1957.
13. Galdston I. (editor) *Medicine and Anthropology*. Inter Universities, New York, 1959.
14. McDermott W. *et al.* Introducing modern medicine in a Navajo community. *Science* 131, 197, 280, 1960.
15. Cassel J. Social and cultural considerations in health innovations. *Ann. N.Y. Acad. Sci.* 107, 1739, 1963.
16. Paul B. D. Anthropological perspectives on medicine and public health. *Ann. Am. Acad. Pol. Soc. Sci.* 346, 34, 1963.
17. Niehoff A. H. (editor) *A Casebook of Social Change: Case Histories of Innovation*. Aldine, Chicago, 1966.
18. Read M. *Culture, Health and Disease*. Tavistock, London, 1966.
19. Wellin E. Water boiling in a Peruvian town. In *Health, Culture and Community* (edited by Paul B. D.), pp. 71–102. Russell Sage, New York, 1955.
20. Wellin E. Implications of local culture for public health. *Hum. Org.* 16, 16, 1958.
21. Miner H. Culture change under pressure: a Hausa case. *Hum. Org.* 19, 164, 1960.
22. Adeniyi J. D. Cholera control: problem of belief and attitudes. *Int. J. Health. Educ.* 15, 238, 1972.
23. Morgan R. Attitudes towards smallpox and measles in Nigeria. *Int. J. Hlth. Educ.* 12, 77, 1969.
24. Barnes S. T. and Jenkins C. D. Changing personal and social behaviour: experiences of health workers in a tribal society. *Soc. Sci. Med.* 6, 1, 1972.
25. Baric L., Stevens R. D., Karuppaiyan V. and Vadivell M. Cancer education in Kancheepuram: a case against transferring health educational approaches. *Int. J. Hlth. Educ.* 17, 32, 1974.
26. Zborowski M. Cultural components in response to pain. *J. Soc. Iss.* 8, 16, 1952.
27. Zborowski M. *People in Pain*. Jossey-Bass, San Francisco, 1969.
28. Zola I. K. Culture and symptoms—an analysis of patients' presenting complaints. *Am. Sociol. Rev.* 31, 615, 1966.
29. Saunders L. *Cultural Differences and Medical Care*. Russell Sage, New York, 1954.
30. Clark M. *Health in the Mexican-American Culture: A Community Study*. University of California Press, Berkeley, 1959.
31. Kasl S. V. and Cobb S. Health behaviour, illness behaviour and sick role behaviour. *Arch. Env. Hlth.* 12, 246 and 531, 1966.
32. Baric L. Recognition of the "at risk" role—a means to influence health behaviour. *Int. J. Hlth. Educ.* 12, 24, 1969.
33. Mechanic D. *Medical Sociology*. Free Press, New York, 1968.
34. Fabrega H. Toward a model of illness behaviour. *Med. Care* 11, 470, 1973.
35. Koos E. L. *The Health of Regionville*. Columbia University Press, New York, 1954.
36. Baumann B. Diversities of conceptions of health and physical fitness. *J. Hlth. Hum. Behav.* 2, 39, 1961
37. Davis F. *Passage Through Crisis*. Bobbs-Merrill, Indianapolis, 1963.
38. Roth J. *Timetables*. Bobbs-Merrill, Indianapolis, 1963.
39. Suchman E. A. Sociomedical variations among ethnic groups. *Am. Sociol. Rev.* 70, 319, 1964.
40. Samora J. *et al.* Medical vocabulary knowledge among hospital patients. *J. Hlth. Hum. Behav.* 2, 83, 1961.
41. Bice T. W. and White K. L. Factors related to the use of health services: an international comparative study. *Med. Care* 11, 104, 1970.
42. McKinlay J. B. Some approaches and problems in the study of the use of services—an overview. *J. Hlth. Soc. Behav.* 13, 113, 1972.
43. Anderson J. G. Health services utilization: framework and review. *Hlth. Serv. Res.* 8, 184, 1973.
44. Richards N. D. Utilization of dental services. In *Social Sciences and Dentistry* (edited by Richards N. D. and Cohen L. K.), pp. 209–240. International Dental Federation, London, 1971.
45. Rosenstock I. M. Why people use health services. *Milb. Mem. Fund Quart.* 44, 94, 1966.
46. Rosenstock I. M. Prevention of illness and maintenance of health. In *Poverty and Health: A Sociological Analysis* (edited by Kosa J., Antonovsky A. and Zola I. K.), pp. 168–190. Harvard University Press, Cambridge, 1969.
47. Blackwell B. The literature on delay in seeking medical care for chronic illness. *Hlth. Educ. Monog.* 16, 1973.
48. Wadsworth M. E. J., Butterfield W. J. H. and Blaney R. *Health and Sickness: The Choice of Treatment*. Tavistock, London, 1971.
49. Dunnell K. and Cartwright A. *Medicine Takers, Prescribers and Hoarders*. Routledge & Kegan Paul, London, 1972.
50. Stoeckle J. D., Zola I. K. and Davidson G. E. On going to see the doctor: the contributions of the patient to the decision to seek medical aid. *J. Chron. Dis.* 16, 975, 1963.
51. Zola I. K. Pathways to the doctor—from person to patient. *Soc. Sci. Med.* 7, 677, 1973.
52. Suchman E. A. Stages of illness and medical care. *J. Hlth. Hum. Behav.* 6, 114, 1965.
53. Suchman E. A. Social factors in illness behaviour. *Milb. Mem. Fund. Quart.* 47, 85, 1969.
54. Freidson E. Client control and medical practice. *Am. J. Sociol.* 65, 374, 1960.
55. Freidson E. *Patients' Views of Medical Practice*. Russell Sage, New York, 1961.
56. Banks F. R. and Keller M. D. Symptom experience and health action. *Med. Care* 9, 498, 1971.
57. Steele J. L. and McBroom W. H. Conceptual and empirical dimensions of health behaviour. *J. Hlth. Soc. Behav.* 13, 382, 1972.
58. Williams A. F. and Wechsler H. Interrelationship of preventive actions in health and other areas. *Pub. Hlth. Rep.* 87, 969, 1972.
59. Green L. W. Status identity and preventive health behaviour. *Pacif. Hlth. Educ. Rep.* 1, 1970.
60. Coburn D. and Pope C. R. Socioeconomic status and preventive health behaviour. *J. Hlth. Hum. Behav.* 15, 67, 1974.
61. Hochbaum G. Why people seek diagnostic X-rays. *Pub. Hlth. Rep.* 71, 377, 1956.
62. Leventhal H., Rosenstock I. M., Hochbaum G. M. and Carriger B. K. Epidemic impact on the general population in two cities. In *The Impact of Asian Influenza on Community Life: A Study in Five Cities*. U.S. Public Health Service Publication No. 766, Washington DC, 1960.
63. Heinzelman F. Determinants of prophylaxis behaviour with respect to rheumatic fever. *J. Hlth. Hum. Behav.* 3, 73, 1962.
64. Kegeles S. S. Some motives for seeking dental care. *J. Am. Dent. Assoc.* 67, 90, 1963.

65. Stadt Z. M., Blum H. L., Kent G. W., Fletcher E., Keys G. and Frost L. A. Direct mail dental motivation of parents of 3-year-old children. *Am. J. Pub. Hlth.* 53, 572, 1963.
66. Antonovsky A. and Kats R. The model dental patient—an empirical study of preventive health behaviour. *Soc. Sci. Med.* 4, 367, 1970.
67. Antonovsky A. A model to explain visits to the doctor with specific reference to the case of Israel. *J. Hlth. Soc. Behav.* 13, 446, 1972.
68. Kegeles S. S. Why people seek dental care: a test of conceptual formulation. *J. Hlth. Hum. Behav.* 4, 166, 1963.
69. Kirscht J. P., Haefner D. P., Kegeles S. S. and Rosenstock I. M. A national study of health beliefs. *J. Hlth. Hum. Behav.* 7, 248, 1966.
70. Young M. A. C. Review of research and studies related to health education practice (1961–1966): what people know, believe and do about health. *Hlth. Educ. Monog.* 23, 1967.
71. Young M. A. C. and Simmons J. J. Review of research and studies related to health education practice (1961–1966): psychosocial and cultural factors related to health education practice. *Hlth. Educ. Monog.* 24, 1967.
72. Young M. A. C. Review of research and studies related to health education practice (1961–1966): communication: methods and materials. *Hlth. Educ. Monog.* 25, 1967.
73. Young M. A. C. Review of research and studies related to health education practice (1961–1966): patient education. *Hlth. Educ. Monog.* 26, 1968.
74. Young M. A. C. Review of research and studies related to health education practice (1961–1966): programme planning and evaluation. *Hlth. Educ. Monog.* 27, 1968.
75. Young M. A. C. Review of research and studies related to health education practice (1961–1969): school health education. *Hlth. Educ. Monog.* 28, 1969.
76. Gibson G. Explanatory models and strategies for social change in health care behaviour. *Soc. Sci. Med.* 6, 635, 1972.
77. Anderson J. G. and Bartkus D. E. Choice of medical care: a behavioural model of health and illness behaviour. *J. Hlth. Soc. Behav.* 14, 348, 1973.
78. Berkanovic E. and Reeder L. G. Ethnic, economic and social psychological factors in the source of medical care. *Soc. Probs.* 21, 246, 1973.
79. Berkanovic E. and Reeder L. G. Can money buy the appropriate use of services? Some notes on the meaning of utilization data. *J. Hlth. Soc. Behav.* 15, 93, 1974.
80. Swinehart J. W. Voluntary exposure to health communications. *Am. J. Pub. Hlth.* 58, 1265, 1968.
81. Ley P. Primacy, related importance and the recall of medical statements. *J. Hlth. Soc. Behav.* 13, 311, 1972.
82. Bloom S. *The Doctor and his Patient.* Russell Sage, New York, 1963.
83. Cartwright A. *Patients and their Doctors.* Routledge & Kegan Paul, London, 1967.
84. Francis V., Korsch B. M. and Morris M. J. Gaps in doctor-patient communication. *New Engl. J. Med.* 280, 535, 1969.
85. Korsch B. M. and Negrete V. F. Doctor patient communication. *Sci. Am.* 227, 66, 1972.
86. Fletcher C. M. *Communication in Medicine.* Nuffield Provincial Hospitals Trust, London, 1972.
87. McIntosh J. Processes of communication, information-seeking and control associated with cancer. *Soc. Sci. Med.* 8, 167, 1974.
88. Davis M. S. and Eichorn R. L. Compliance with medical regimens: a panel study. *J. Hlth. Hum. Behav.* 4, 240, 1963.
89. Davis, M. S. Variations in patients' compliance with doctors' orders: analysis of congruence between survey responses and results of empirical investigations. *J. Med. Educ.* 41, 1037, 1966.
90. Davis M. S. Predicting non-compliant behaviour. *J. Hlth. Soc. Behav.* 8, 265, 1967.
91. Davis M. S. Variations in patients' compliance with doctors' advice: an empirical analysis of patterns of communication. *Am. J. Pub. Hlth.* 58, 274, 1968.
92. Davis M. S. Physiologic, psychological and demographic factors in patient compliance with doctors' orders. *Med. Care* 6, 115, 1968.
93. Marston M. V. Compliance with medical regimens. *Nurs. Res.* 19, 312, 1970.
94. Mitchell J. H. Compliance with medical regimens: an annotated bibliography. *Hlth. Educ. Monog.* 2 (New Series), 75, 1974.
95. Festinger L. Behavioural support for opinion change. *Pub. Opin. Quart.* 28, 404, 1964.
96. Rokeach M. Attitude change and behavioural change. *Pub. Opin. Quart.* 30, 529, 1966.
97. Greenwald A. G. Effects of prior commitment on behaviour change after a persuasive communication. *Pub. Opin. Quart.* 29, 595, 1965.
98. Warner L. G. and Defleur M. L. Attitude as an interactional concept: social constraint and social distance as intervening variables between attitudes and action. *Am. Sociol. Rev.* 34, 153, 1969.
99. Festinger L. *A Theory of Cognitive Dissonance.* Row Peterson, Evanston, 1957.
100. Festinger L., Schacter S. and Back K. *Social Pressures in Informal Groups.* Harper, New York, 1957.
101. Freeman H. E. and Lambert C. The influence of community groups on health matters. *Hum. Org.* 24, 353, 1965.
102. Green L. W. Should health education abandon attitude change strategies? Perspectives from recent research. *Hlth. Educ. Monog.* 30, 25, 1970.
103. Higbee K. L. Fifteen years of fear arousal: research on threat appeals 1953-1968. *Psychol. Bull.* 72, 426, 1969.
104. Leventhal H. Fear appeals and persuasion: the differentiation of a motivational construct. *Am. J. Publ. Hlth.* 61, 1208, 1971.
105. Suchman E. A. Evaluating educational programmes: a symposium. *Urban Rev.* 3, 163, 1969.
106. Green L. W. and Figa-Talamanca I. Suggested designs for evaluation of patient education programmes. *Hlth. Educ. Monog.* 2 (New Series), 54, 1974.
107. Schonfield J., Schmidt W. and Sternfeld L. Medical attitudes and practices of parents towards a mass tuberculin-testing programme. *Am. J. Publ. Hlth.* 53, 772, 1963.
108. Nickerson G. S. and Hochstrasser D. L. Factors affecting non-participation in a country-wide tuberculin testing programme in S. Appalachia. *Soc. Sci. Med.* 3, 575, 1970.
109. Ward A. W. Why family planning patients drop out. *J. Biosoc. Sci.* 3, 301, 1971.
110. Fink R., Shapiro S. and Roester R. Impact of efforts to increase participation in repetitive screenings for early breast cancer detection. *Am. J. Publ. Hlth.* 62, 328, 1972.
111. Kaupinnen M. A., Kauraniemi T., Koli T. and Voipio N. Response to the written invitation in a gynaecological mass screening by cytology arranged in Helsinki in 1966. *Acta Obst. Gyn. Scan.* 49 (Supplement 7), 1970.
112. Vuori H., Rimpelä M. and Grönroos M. Cytological screening programmes: the problem of non-participation. *Int. J. Hlth. Educ.* 15, 22, 1972.
113. Bierre B. Studies on population screening of the early carcinoma of the uterine cervix. *Acta Obst. Gyn. Scan.* 48 (Supplement 6), 1969.
114. Richards N. D. and McEwan P. J. M. Sociological factors affecting use of cervical screening tests. *Br. J. Prev. Soc. Med.* 27, 66, 1973.

115. Kaluzny A. D., Gentry J. T. and Glasser J. H. Innovation in health care organizations: review of research and plan of projected studies. *Hlth. Serv. Res.* 3, 316, 1968.
116. Vertinsky P., Vertinsky I. and Zaltman G. Health information diffusion: an integrated model. *Int. J. Hlth. Educ.* 15, (supplement) 1972.
117. Rogers E. M. *Diffusion of Innovations.* Free Press, Glencoe. 1962.
118. Katz E. The social itinerary of technical change: two studies in the diffusion of innovation. *Hum. Org.* 20, 70, 1961.
119. Katz E. The two-step flow of communications: an up-to-date report on a hypothesis. *Pub. Opin. Quart.* 21, 61, 1958.
120. Hingson R. and Lin N. Communication participation and individual receptivity to health innovations. *Int. J. Hlth. Educ.* 16, 5, 1973.
121. Suchman E. A. The survey method as applied to public health and medicine. In *Survey Research in the Social Sciences* (edited by Glock C. Y.) Russell Sage, New York, 1967.
122. Bulman J. S., Richards N. D., Slack G. L. and Willcocks A. J. *Demand and Need for Dental Care.* Oxford University Press, London, 1968.
123. Richards N. D., Hussin O. A. N. and Whitfield A. P. Evaluation of use of a mobile unit in screening for cervical cancer (in press).
124. Watson L. M. Cigarette smoking in school children. *Hlth. Bull.* 24, 5, 1966.
125. Myers S. E. and Downs R. A. Comparative findings in school systems with differing approaches to dental health education. *J. Sch. Hlth.* 38, 604, 1968.
126. Richards N. D. and Taylor I. School dental health education—a controlled study of experimental approaches (in preparation).
127. Richards N. D. Whither dental health education? A layman's viewpoint. *Comm. Hlth.* 5, 2, 1973.

Determining the Impact and Effectiveness of Health Education as it Relates to Federal Policy*

Lawrence W. Green, Dr.P.H.**
Johns Hopkins University
School of Hygiene and Public Health
Baltimore, Maryland

In response to a set of policy questions posed by the federal government, research evidence on the potentials of health education for contributing to resolution of national health issues was reviewed in detail. A synthesis of the findings concludes that greater investments in health education on selected problems for selected populations would yield a better cost-benefit return than comparable investments in tertiary medical care. In relation to selection and amount, high priority groups for specific education changes can be clearly identified, and the point of diminishing returns of expenditure can be identified; health education programs thus can be both focused and contained. Among the 36 specific recommendations drawn from the research evidence are a series of criteria for design of health education programs. A general recommendation for a major federal investment in health education research is outlined.

This analysis of research evidence on the potentials of health education to help resolve some of the national issues in health policy follows an outline suggested by a series of questions posed by the Department of Health, Education, and Welfare. The purpose of the analysis was not to evaluate individual programs or methods, but rather to assess broad classes of factors which might account for success or failure of health education. Thirty-six recommendations for federal policy are presented and documented in this paper in the order suggested by the questions posed, without concern for their priority. These recommendations are presented conservatively according to the data available to support them.

The major recommendations will be discussed in general terms prior to presentation of individual recommendations to reflect their logical relationship to each other in a federal strategy rather than to the evidence or the questions whence they came; these recommendations are expanded in this overview somewhat beyond the empirical data, on the basis of extrapolation from theory and experience in related fields.

*This paper is a revised, abbreviated version of a report prepared for the Office of the Assistant Secretary for Health, Deputy Assistant Secretary for Planning and Evaluation, U.S. Department of Health, Education, and Welfare; Contract No. SA-7974-75; Mona Mitnick, Project Director.

**Address requests for reprints to Dr. Green, Professor and Head, Division of Health Education, Department of Health Services Administration, Johns Hopkins University School of Hygiene & Public Health, 615 N. Wolfe St., Baltimore, Maryland 21205

SUMMARY OF MAJOR RECOMMENDATIONS

Health education is found to be highly interwoven with complementary and overlapping technologies of motivational and behavioral control such as mass communications and behavior modification. The federal government is faced with some expensive health problems for which there are obvious solutions or partial solutions if only the public would cooperate with changes in personal health practices. But the government is also constrained by the Constitution, and by ethical concerns, in the degree to which it can press for these changes without representation and informed consent from those citizens whose health-related behavior is in question. It is precisely the process of obtaining health behavior through participation and informed consent that defines the parameters of health education in relation to the overlapping technologies of communication, motivation, and behavior modification, and in conjunction with political and legal processes of legislation and code enforcement. Hence, health education is defined and recommended for governmental deployment as "any combination of learning opportunities designed to facilitate *voluntary* adaptations of behavior which will improve or maintain health."

The operational terms for government action at the program level are "combination" and "designed." How can learning opportunities be designed and combined to facilitate health maintenance most efficiently without coercing behavior? At the policy level, however, the issues are less with methods and design than with selecting those areas of health behavior in which behavioral adaptations can have the greatest payoff in cost-benefit which behavioral adapatations can have the greatest payoff in cost-benefit terms for individuals and for society.

The first part of the paper concerns itself with the question of methods and design of health education; but federal policy must be based primarily on cost-benefit analyses of alternative strategies, which is the subject of the last half of the paper.

In general, the latter section of the analysis concludes that greater investments in health education on selected problems for selected populations would have greater payoff in cost-benefit terms than comparable investments in construction, specialized training, or equipment for tertiary medical care. Indeed health education could help curtail the "upstream" sources of many "downstream" problems if adequately and systematically developed and supported, as has been demonstrated in immunization and family planning programs. Obviously, prevention and tertiary care are not mutually exclusive alternatives, but preventive and self-care improvements have not had sufficient governmental support to achieve even the cost containment

867

potentials they are capable of offering in relation to medical care.

Given the general conclusion that federal policy should shift some of the resources from future expenditure in medical care to increased current investments in health education, the questions of selection and amount remain critical in times of severely limited spending. Two conclusions are immediately possible from the analysis in relation to the questions of selection and amount:

(1) not everyone benefits equally from a given health behavior, and

(2) spending on health education need not be regarded as a bottomless pit.

Stated more positively and specifically, high risk groups and optimum age ranges can be clearly identified for most preventive and self-care actions which health education can facilitate. Hence, it is possible to target scarce educational resources on those problems and populations where impact and payoff will be greatest, as identified in recommendations 20-35. Furthermore, it is possible to identify the point of diminishing returns in health education expenditures in some types of programs. Additional research is needed on this, but the implication of existing evidence is that health education can be both focused and contained. Unfortunately, most contemporary efforts in health education have been diffuse and continuous, making it difficult to account for their effects or their costs.

A related conclusion drawn from the analysis is that there is a level of health education spending below which health education will have no effect. Like the evidence on the level of diminishing returns, the threshold level is unknown for most programs in health education, leaving expenditure policy to be defined by a large number of technical considerations, as outlined in recommendations 2-15, which must be applied case by case.

One of the most important generalizations that emerges from systematic comparisons of programs and experiments with positive results and those with negative findings (no effect) is that the greater the variety in educational methods used, the more likely the program or experiment will show positive results. This generalization applies both with individuals and with populations. At the individual level, variety in educational methods apparently helps to surround the learner with communications appealing to different senses and modes of learning which are mutually supportive. At the community level, variety in educational methods accommodates the fact that different media have different levels of exposure, attention, credibility, and response for different people in relation to different health problems.

These observations, and related analyses, lead to an emphasis in this report on a series of technical recommendations (#2-8) urging federal priority on programs utilizing multiple health education methods (#2); selected and combined on the basis of prior assessments of health beliefs, attitudes and behavior in the population (#3); with a clear specification of the behavioral outcomes expected for each phase of program development (#4), and evaluated with procedures adapted from standard social research and health services research methods (#6-8). These criteria for the identification of the most promising proposals for health education research or demonstration may be contrasted with proposals to test a *single* technique or method in isolation; based on *assumptions* about the health attitudes, beliefs and behavior in the population; or selected on the basis of convenience, convention, or professional preference; with vague outcomes to be evaluated with sophisticated but trivial measures. These characteristics of a majority of past programs and studies reviewed account for the largely pessimistic picture of health education portrayed by several state-of-the-art reviewers.

Because so many inadequate programs have been supported with inadequate resources, the cumulative effect on the literature related to health education is unfortunately discouraging. A general recommendation is hereby offered for a federal policy which would offset this effect: that *a five-year program of research and demonstration in health education be directed by the National Center for Health Services Research and/or the Bureau of Health Education in which developmental funds are restricted to a few projects with sufficient magnitude, continuity, and planning to assure the production of results which will answer the following questions in relation to selected health problems:*

—What is the threshold level of health education effort required to achieve initial, minimal changes in health behavior for different high-risk groups?

—What is the point of diminishing returns in health education efforts (e.g., number of home visits, frequency of group meetings, duration of patient counseling sessions) beyond which further investments do not achieve proportionate gains in behavioral change?

—At what point in time following an initial behavioral change is follow-up education most reinforcing in terms of preventing discontinuance of the behavior?

—Which sub-populations (e.g., age groups, socioeconomic groups) respond most favorably in terms of the foregoing criteria to each of the various combinations of educational methods used to achieve these behavioral changes?

If these four fundamental questions can be answered in relation to priority health problems, it will be possible to estimate the minimum and maximum expenditures on health education under National Health Insurance and to establish guidelines on the relative frequency and distribution of educational resources required for different populations.

This paper recommends a number of methods and procedures for obtaining answers to these questions (Recommendations 6-14), and offers an approach to the selection of health behaviors in which both potential effectiveness of health education and potential payoff should be high.

Finally, on the basis of what is known from previous research and experience, recommendations 16-36 identify action policies that can proceed while the research program seeks answers that will eventually improve their cost-effectiveness.

PURPOSE, ORGANIZATION, AND METHOD

The purpose of this study, conducted for the Office of the Deputy Assistant Secretary for Planning and Evaluation/Health (United States Department of Health, Education, and Welfare), was "to develop evaluation criteria to predict the success of various kinds of health education programs." The suggested evaluation criteria are organized here in response to the following specific questions posed by the government:

(1) what kinds of health education programs are most or least effective in achieving specific goals;
(2) for what reasons;
(3) for what health conditions or diagnoses;
(4) under what circumstances;
(5) on what populations and through what providers;
(6) through what kinds of funding mechanisms;
(7) at what costs?

The analysis that follows is based on reviews of the literature, including epidemiologic, demographic, administrative, economic, and behavioral studies of mortality reduction and disease control concepts, programs, activities and methods, with an emphasis on those related to voluntary changes in the health practices of the public. This paper will summarize the findings of these reviews without extensive bibliographic citations. Previously submitted working papers[1-3] and a report submitted in conjunction with Arthur D. Little Company (RFQ 42-75-OS, "Determining the Impact of Health Education Programs on Utilization of Health Services and Other Consumer and Provider Behavior") provide the fuller documentation of conclusions drawn here. Reference will be made more frequently to these and other recent reviews of the literature than to primary sources.

Recommendations for federal policy on health education are based on the findings, on the goals and health policy of the Department of Health, Education, and Welfare (*Forward Plan for Health* for Fiscal Years 1976-1980), and on the uses of research in setting national health policy. The emphasis is therefore on prevention, preparing for National Health Insurance, and quality assurance (the first three priorities), and on cost-effectiveness and cost-benefits of health education (related to the fourth and fifth priorities of the *Forward Plan*).

Priorities for health education in national health policy are analyzed in terms of evidence of effectiveness, potential benefits, target populations, and prospects for support under national health insurance. Recommendations will specify health education activities that need to be directed at maternity patients, preventive child care, preventive adult care, chronic and long-term care, and special settings such as schools and industry. The known efficacy and benefits of health education will be shown to follow this order from firm to tenuous evidence. Other, more tenuous but promising strategies are recommended for further evaluative research.

DEFINITIONS AND SCOPE

The boundaries of programs, activities, and methods that may be characterized as educational are vague. Most health education activities are embedded in other programs and many are not identified as health education. Indeed, those responsible for programs or studies sometimes disavow any association with health education in an attempt to distinguish their efforts as more innovative, modern, technological, behavioristic, client-centered or scientific than they perceive health education to be. Alternative labeling occurs even when the methods employed clearly derive their philosophical, technical and theoretical or scientific approaches from education, educational psychology, educational technology, or health education itself.

The alternative labels used for health education programs and activities reveal the scope and diversity of educational applications in health:

—*Motivation programs.* This term has been used especially in family planning to refer to the activities generally included in health education programs,[4] plus "incentive schemes" designed to appeal more directly to economic motives for family limitation. From a formal psychological standpoint, the term motivation is incorrectly applied when it refers to programmatic activities, because motivation is a construct intended to refer to the internal dynamics of behavior, not the external stimuli. The more correct designation for methods is "motive-arousing appeals."

—*Behavior modification.* This term has been extended from its original applications by behaviorist psychologists[5] to include a wider range of educational and political strategies for which the priority objectives are changes in behavior.[6] The original (purist) behavioral psychologists would disallow any educational method that is designed to achieve behavioral changes by means of changing knowledge or attitudes. But behavioral modification (operant conditioning) techniques qualify as educational methods so long as the subjects voluntarily submit to the conditioning procedures to achieve changes they desire in their own behavior.

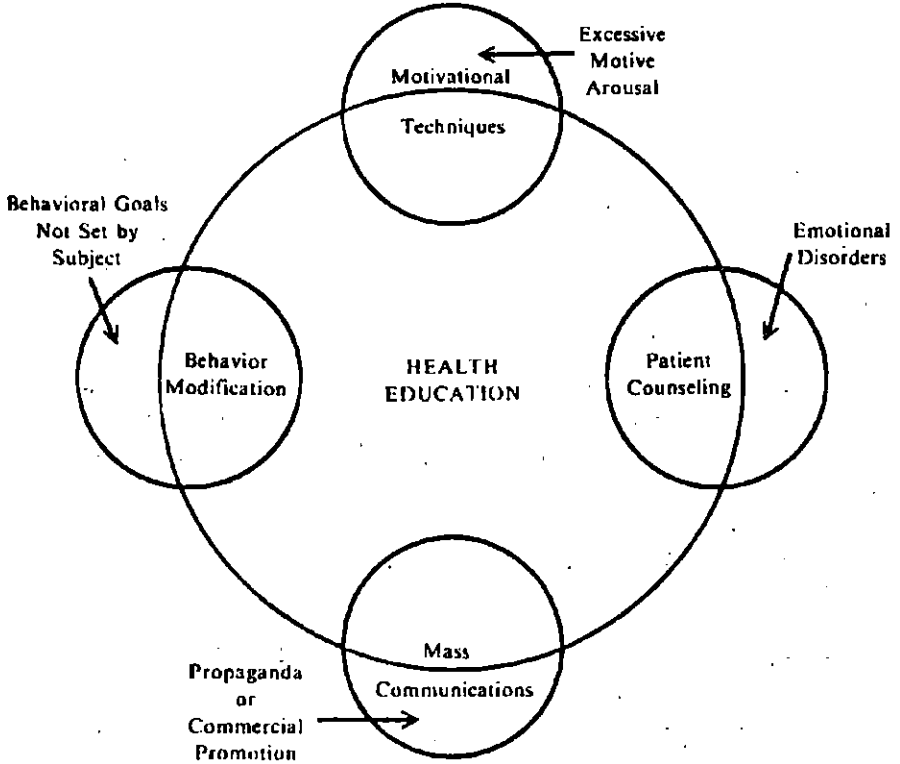
—*Health counseling.* This term and its variants (genetic counseling, diet counseling, patient counseling) constitute another subset of health education insofar as they represent an approach to voluntary change in health behavior. Most counseling methods of education have their theoretical and philosophical roots in ego psychology, which is at extreme variance with behaviorism. Counseling is outside the scope of health education when it is more psychotherapeutic than informational in its method and content.

—*Communications.* As an interdisciplinary concern of social psychologists, marketing and public opinion research, media specialists, health educators, and others, communications are applied and their effects on behavior are studied in every sphere of human endeavor.[7] Their applications in relation to health behavior are usually within the scope of health education programming,[8-10] except when they are used to advertise or promote products or causes inconsistent with the health needs of consumers.[11]

The foregoing examples of alternative labels used for health education

500

FIGURE 1
EXAMPLES OF OVERLAPPING TECHNOLOGIES FOR CHANGING HEALTH-RELATED BEHAVIOR



activities illustrate the kinds of commonalities and distinctions that need to be observed in assessing programs. Figure 1 provides an incomplete representation of these separate technologies in relation to health education, suggesting aspects or applications of each which do not qualify as educational. The defining characteristic of health education, as the term will be used here, is the voluntary participation of the consumer in determining his or her own health practices. This is not merely a philosophical tenet. *The overwhelming weight of evidence is that the durability of cognitive and behavioral changes are proportional to the degree of active rather than passive participation of the learner.* There are also practical and legal reasons for the federal government to adhere to this usage of health education to avoid public resistance or reaction to programs which might be perceived as propagandistic, manipulative, coercive, politically or commercially directed, threatening, or paternalistic.

Other forms and methods of health education which define its scope are community organization, in-service training, consultation, group work, computer-assisted instruction (CAI), other teaching machines and audiovisual methods, bibliotherapy, patient teaching, health fairs, exhibits, libraries, and conferences, and routine health provider-consumer interactions.

The scope of health education is defined as much by its expected outcomes as by its methods and forms. Section I outlines the process of health education in terms of the approximate relationships among the outcomes of interest for this analysis.[12]

This review of the scope and definition of health education suggests two recommendations for federal policy:

Recommendation 1. The Department of Health, Education, and Welfare should *retain the general use of the term "health education"* in preference over the narrower and in some instances more insidious alternatives currently used by specialized groups. The alternative terms should be used to refer to specific techniques or strategies of behavior change, but should be associated with health education only where it is clear that the patient or consumer is in control of the decision to change.

Recommendation 2. The Department of Health, Education, and Welfare should *encourage and support health education programs that coordinate two or more of the specific methods of health education in preference over programs employing only one technique or method.*

Because of the complexity of most health practices, no single educational input can generally achieve lasting changes in health behavior without supportive educational inputs directed at other factors influencing the behavior.

Behavioral scientists have made and continue to make significant contributions to the scientific base of health education through basic research on the processes of learning and behavioral change. The applied research of most behavioral scientists, however, tends to isolate single variables influencing behavior rather than multiple variables represented in health programs. This tendency of behavioral research to focus on unitary stimuli is understandable in terms of the goals of the behavioral scientist and the demands of experimental control. It is difficult, however, to apply the findings from tightly controlled, univariate behavioral studies to program planning in health education. As laboratory studies in psychology and survey research in sociology accumulate an increasingly lengthy inventory of determinants and correlates of health behavior, the practitioner and planner is further bewildered in deciding upon which factors to intervene. The foregoing recommendation is intended to encourage a shift in research resources from highly controlled studies of singular methods of health education to field studies or clinical trials of health education strategies combining several methods directed at multiple factors.

Randomized factorial experiments provide the ideal design for implementation of this recommendation. Where such randomization is not possible, quasi-experimental evaluations or demonstrations should be supported (see Recommendation 6) in which most recipients receive a comprehensive educational program and comparison groups receive everything but selected experimental methods.

I. WHAT KINDS OF HEALTH EDUCATION PROGRAMS ARE MOST OR LEAST EFFECTIVE IN ACHIEVING SPECIFIC GOALS?

This question required an assessment not only of the different kinds of programs partially identified in the foregoing discussion, but also of the variegated goals of health education. Several dimensions and criteria on which the goals and outcomes of health education can be measured are:

(a) *Immediate vs. intermediate vs. ultimate goals.* The immediate outcomes of health education programs should be changes (improvements) in selected predisposing, enabling, and reinforcing factors[12,13] which are known or expected to be causally related to the health behavior in question. The general components of health education programs that are found to be most effective with each of the three sets of immediate goals are identified as follows:

- — *Communication methods* directed at the consumers and the general public are most effective in achieving changes in *predisposing factors* such as knowledge, attitudes, beliefs, and norms.[7-14]
- — *Community organization* and organizational development methods directed at agencies and institutions are generally most effective in achieving changes in the *enabling factors* which facilitate or inhibit health practices, such as availability and accessibility of services, referral mechanisms, third party payment mechanisms, clinic hours, etc.[9-12,15-19]
- — *Staff development* methods directed at the health providers and other people who interact with the consumers are most effective in achieving changes in the *reinforcing factors* such as the attitudes and actions of clinic personnel toward walk-in patients, the rewards and punishments that parents and teachers provide children in response to health actions, or that employers provide employees in response to safety practices.[9-12,20-26]

Recommendation 3. In order to influence health practices, health education must effectively change the factors supporting the old behavior or blocking the new behavior; therefore, the Department of Health, Education, and Welfare should *support and encourage programs which include a systematic assessment of these factors prior to the selection of educational methods, in preference over programs which select educational methods on conventional or convenience bases.*

Intermediate goals of health education programs are the behavioral outcomes expected to occur if the immediate goals are achieved. If, for example, a program successfully increases public awareness of a health problem (predisposing factor) and improves the accessibility of services or resources to take action (enabling factors), and increases support for the behavior from parents or employers (reinforcing factors), it will be more likely to achieve changes in behavior than if only one or two of these immediate goals are achieved. Different emphases of programs are known to have varying effectiveness in achieving five different dimensions of behavior change as follows:

- — *Frequency.* If a health action must recur periodically and the goal of the program is to establish a pattern of regularity, then health education methods which provide cues at appropriate intervals tend to be more effective than one-shot or sporadic methods. Well-child visits, dental checkups, annual pap smears, and prenatal visits all have been shown to improve in frequency with a telephone or mailing system that reminds consumers, parents, or patients of their next appointment.[1,27-30]
- — *Persistence.* If the required health behavior is one that must be continued over a long period, as with medical regimens for hypertension or diabetes, then health education methods which provide social supports in the clinic and home environments are most effective. Several studies on compliance with medical regimens confirm the importance of the doctor-patient relationship and the patient's perception of support from family members, especially the spouse.[11,12] Persistence with diet and smoking cessation regimens show similar social influences and effectiveness of normative health education strategies.[10,11,31]
- — *Earliness.* The behavioral goal in many health education programs is primarily to reduce delay in seeking medical treatment for selected symptoms or to reduce delay in adopting a preventive practice such as immunizations, prenatal care or contraception. The former appears to be heavily influenced by family support and relationship with a health care provider and is therefore most effectively achieved through health education directed at reinforcing factors in the home and clinic environments, as in the preceding dimension of behavior.[14,15] There are site- and symptom-specific variations in this rule, however.[16] Preventive health actions, on the other hand, appear to be responsive in terms of reduced delay to less intensive health education programs. Mass media campaigns, for example, have demonstrated their effectiveness in recruiting new patients to family planning, immunization, venereal disease and various screening programs, but the patterns of clinic attendance in subsequent months and the characteristics of the patients recruited indicate that the media are triggering earlier action in those who were already motivated and who would have come to the clinics at later dates.[17,18] Thus, the mass media are useful in health education when the specific goals are to reduce delay in preventive health actions, but less effective when the goal is to reduce delay with illness or sick-role behavior.
- — *Quality.* When the behavioral goal of a health education program is to improve the appropriateness of health actions, i.e., when the consumer must choose between alternative products, sources of care or types of symptom response, the evidence suggests that prior experience with the product or source of care is most influential in determining the quality of the choice. Health education directed at reinforcing factors in the health

004

care system is most likely to influence appropriate utilization of health services.[10,39,40] Starting with failure to acknowledge and meet patient expectations on the first visit,[41,42] through failure to take family and work situation into account in prescribing therapeutic regimens,[43,44] several points in the health care system are identified where health education can strengthen the probability of appropriate health actions on subsequent occasions. Community organization methods are an essential component of health education when the quality of consumer choices is a function of the availability or accessibility of quality health products or services.[45] Similarly, communications with the public through food labeling,[46] counter-advertising of harmful substances, and public service announcements can increase awareness, arouse interest, and at least direct the choices of those already motivated to take action.

Range. When the behavioral goals of the health education program are highly complex because each individual is expected to adopt a range of different health practices, more intensive health education methods are required. Individualized counseling, home visits, and reinforcement of graduated steps in the behavioral repertoire may be required. Rehabilitation counseling, diet therapy and education of parents of disabled children are most effective when they include these several elements. Health practices which are routine for middle class housewives may be highly complex for rural poor homemakers and generally require more intensive and more long-term educational interventions for the latter.[47,48]

Another classification of behavioral goals sometimes lists utilization, preventive actions, consumption patterns, and compliance as different types of health behavior. As shown in the preceding paragraphs, these types of health practices vary according to the dimensions of behavior, so that there is less similarity between different types of compliance with a specific regimen and consumption of specific products where both the compliance and the consumption patterns are similar on dimensions of earliness, frequency, quality, range or persistence required.[10,12,35,38,48-50] The division of behavioral patterns into health behavior, illness behavior, and sick-role behavior is a theoretical convenience for medical sociologists[51,52] to which health education has had to add the "at-risk role."[53] But these distinctions have not proved as useful in guiding health education programs as has the assessment of behavioral goals in terms of earliness, frequency, quality, range, and persistence. This suggests a recommendation for federal policy in relation to health education programming and research.

Recommendation 4. In assessing the behavioral goals of health education programs, the Department of Health, Education, and Welfare *should seek specificity in terms of behavioral dimensions such as frequency, range, earliness, quality, and persistence* in preference over the broader theoretical classifications of health, illness and sick-role behavior.

The ultimate goals of health education programs consist of reductions in mortality, morbidity, fertility, and disability. These are usually measured as outcomes in community programs on an aggregate rather than individual scale, using vital statistics of incidence, prevalence, and distribution. When epidemiological measures are not accessible, or when the program is on a smaller scale, medical records, clinic or household interviews, and measures of symptoms or risk factors such as weight, blood pressure, blood or urine tests may be used to assess achievement of long-range goals beyond behavioral changes.

Evaluations of health education in terms of these longer-range goals have been rare for the following reasons:

- —Most health education programs have been too limited in time or resources to seek measures of these longer-range outcomes.[54]
- —Community health education programs, by the nature of their employment of mass media and community organization methods, preclude the availability of control groups unexposed to the program.[55]
- —Some of the most extensive health education campaigns have promoted health practices later found by epidemiologists to be unhealthful, such as the nutrition education programs of the 1940s and 1950s urging people to eat eggs, red meats, and shellfish.[10]
- —Health education programs are usually embedded in a larger program effort so that it is often impossible to unravel the effects attributable to health education from those attributable to other program components.[1]

To overcome these problems, investigators are sometimes tempted to go to methodological extremes which make their results entirely ungeneralizable. It is a mistake, for example, to attempt to isolate the health education component of a total health program and to test that component itself either in the community or in a classroom or psychological laboratory. The consistent failure of such programs to achieve the changes sought attests to the necessity of supporting health education with administrative and institutional arrangements conducive to the behavioral changes sought, and the importance of coordinating health education activities with other program components. These observations compel a further set of policy recommendations.

Recommendation 5. Health education programs or activities should not be supported in isolation from the agencies or institutions to which their consumers must relate. Programs intended to increase utilization of specific health services, for example, should ordinarily be conducted in cooperation with those services rather than apart from them.

Recommendation 6. The Department of Health, Education, and Welfare should *encourage and support evaluations of health education in which the experimental groups receive health education within the context of a total program and the control groups receive everything except health education,* in preference over evaluation in which experimental groups receive only health education and the control groups nothing.

603

The latter recommendation is fundamental to the external validity or generalizability of findings from evaluative research on health education. The findings from highly controlled but isolated educational interventions are usually negative because health education alone is usually insufficient. They are of limited applicability even when they are positive because the isolation of health education experiences is unnatural as are some of the other conditions of experimental control. The best way to avoid these anomalies of evaluative research is to design the evaluation of health education within the context of an on-going, multi-method health educational program as suggested in Recommendations 2 and 6.

Recommendation 7. The Department of Health, Education, and Welfare should *use behavioral changes as the primary criteria of health education impact*, in preference over criteria of changes in knowledge and attitudes and in preference over the more demanding but less realistic criteria of changes in morbidity, mortality, fertility, or disability rates.

Achieving changes in knowledge, attitudes, and other predisposing factors remains an important step in diffusion and behavioral change processes, but there is little need for additional evidence on the effectiveness of different health education methods in achieving these changes. Increasingly, the important problems for health education are in achieving behavioral changes in populations where high levels of knowledge and often positive attitudes already exist.[19,51,52,58]

Achieving changes in mortality, morbidity, fertility, and disability, on the other hand, while more important ultimately than changes in behavior, is often dependent on non-behavioral factors beyond the direct influence of health education programs. *The judgment of the effectiveness of health education should not have to depend entirely on criteria, measures of which may not be available for many years and may be overwhelmingly influenced by other circumstances over which health education could have no control.* Moreover, the epidemiological, pharmacological, and medical evidence linking the behavior to the vital outcomes is sometimes tenuous, as noted earlier in relation to the nutrition education on eggs and red meats in the 1940s. Today, there is considerable debate in the biomedical and health care literature on the efficacy of some of the procedures and practices on which new intensive health education programs are already underway. Health education can only be held responsible for the behavioral changes, not for the failure to achieve changes in vital rates if the behaviors are successfully changed.

(b) *Time Frame for Goals.* The foregoing classification of health education goals into immediate (improvements in predisposing, enabling, and reinforcing factors), intermediate (behavioral changes), and ultimate outcomes (reductions in morbidity, etc.) is concerned with causal priorities. The first must be achieved before the second can be achieved because the first causes the second, and so on from second to third. There is another time dimension on which health education goals may be framed and the effectiveness of programs judged. This temporal dimension concerns how soon the goal is to be achieved. Some health education goals can be achieved quickly, even reductions in morbidity; while others, including some changes in predisposing factors, require months, years, or decades to materialize.

Health education programs vary in their capacity to achieve different rates of change, but these variations cannot be analyzed without considering two other dimensions of goal statements: *extensity* (how many people are expected to change), and *target populations* (which groups or individuals in the population are expected to change).

Thus, goal statements in health education, when formally stated as program objectives, contain four elements:

(1) *Who* is to undergo or achieve (2) *how much* of (3) *what kind* of change (4) by *when*?

For example, a health education program goal might be stated as follows: By 1980, the national percentage of former smokers among men in the 21-24 year age group will increase from 20% to 40%.

The time over which the change is expected to occur is very largely a function of (1) the level at which the program begins (the baseline level) and (2) the amount of change expected. The first consideration can be analyzed independently, as a function of timing; the second is a more complex issue to be analyzed in connection with population, cost, and other considerations in later sections.

The relative *rates* of change to be expected from community health education programs can be stated simply: early in the introduction of a new idea or a new program the rate of change (goal achievement) will be slow; as the program progresses the rate of change in the population will speed up until it has included the majority of potential consumers; then the rate of change will become slower as the program attempts to reach the more socially isolated, poorer, and more resistant subgroups. This pattern of rate change in program achievement has been observed consistently across a wide range of program goals where the unit of change was percentage of the population adopting a new practice or innovation.[7,8,10,12,57-61]

The baseline level of prior adoption of the health practice in a community at the time a program is introduced determines the rate of goal achievement to be expected. If the health education program is introduced before anyone in the community has had an opportunity to adopt the practice recommended (zero baseline), it will tend to start slowly. Most programs, however, start in populations where at least some of the people have already adopted the recommended practice and many others are predisposed to adopt but need some information or cue to act. In these circumstances, there will tend to be a temporary leap in the rate of adoption or change. This is sometimes referred to as "skimming the cream off the top" of the community, where "the cream" is analogous to the highly motivated who have risen to the top and are awaiting the program.

Programs that attempt to change attitudes and behavior in the "hard-to-

204

reach" groups, where the baseline level of acceptance for the larger community is high but recalcitrants remain, will have slow rates of change. Efforts will have to be more intensive and the educational cost per new acceptor will be high, but the benefits per new acceptor will also be high because these tend to be high-risk groups.

Recommendation 8. The Department should use criteria in assessing the relative impact of a given type of health education program which *gives greater weight to changes when baseline levels are higher than to equal changes from lower baseline levels*, in preference over criteria which give equal weight to every percentile point of change.

One index of relative impact that can be applied to data from different health education programs to make the comparisons conform to the criterion in Recommendation 8 is the Effectiveness Index (EI):

$$EI = \frac{P_2 - P_1}{100 - P_1} \times 100$$

Where P_1 is the percentage of the population who have previously adopted the recommended health practice at the baseline, P_2 is the corresponding percentage at a later date (after at least part of the program). The denominator of this criterion represents potential change and the numerator actual change.[62] This measure will tend to favor programs reaching the poor.

2. FOR WHAT REASONS?

The explanations for the variable levels of achievement and rates of change for health education programs can be classified according to the following sources of variation in addition to baseline levels and amounts of change expected: (1) the nature of the behavioral changes sought and the symptoms or diagnoses to which they relate, and the circumstances surrounding the behavior or the health problem; (2) the characteristics of the population exposed to the program, and the source (provider) of the educational communication; and (3) the level of spending on the educational components of the program.

3. FOR WHAT PROBLEMS?

Three facts divert analysis of health education away from health problems. (1) The prevention or reduction of many health problems (e.g., heart disease) requires more than one health behavior. (2) Many health practices (e.g., diet) relate to several health problems, conditions, diagnoses, and symptoms. (3) Many health practices are motivated by beliefs, values, fears, and aspirations unrelated to the specific health problems affected by the behavior.

Given these considerations, it has often been more productive to assess health practices and develop preventive communications in terms of specific populations and their *perceptions* of the health practices rather than the health problems themselves. The analysis will focus here on behavioral problems and will return to specific health problems in the final sections.

Recommendation 9. The Department should attempt *to support health education innovations which do not fit conveniently under categorical funding programs* by channeling some of the categorical funds into health education programs directed at changing practices, beliefs, and attitudes associated with several health problems.

It is sometimes difficult to find demonstration and evaluation funding for school health education, health exhibits, mass media programs, health education resource centers, and generalized patient education and community organization programs because these do not relate exclusively to the concerns of one of the Institutes, Bureaus, or Divisions. The Bureau of Health Education has functioned effectively as a conduit for the mixing of some categorical funds in this way, but lacks sufficient staff to manage many more contracts than it now manages. The National Center for Health Services Research has recently given encouragement to behavioral research having more general health education applications. The recently established National Center for Health Education holds some promise of developing and coordinating federal and private resources for non-categorical health education.

Among the perceived attributes of practices that may influence the rate and extent of adoption of a new health behavior or an innovation are its relative *advantage* over the practice it supercedes; its *compatibility* with the existing values, past experiences, and needs of the receivers; its *complexity*; its *trialability* in terms of being divisible into less complex parts, and its *observability* to other people who may reinforce or disapprove of the behavior.[7]

National surveys on smoking between 1964 and 1970 indicate that the diffusion of smoking cessation has been responsive to these variables, as shown by attitudes and practices in three surveys.[63] Schwartz and Dubitzsky[61] compared the expressed willingness of smokers to try ten different smoking cessation methods. They found that smokers were not willing to use methods most frequently offered by physicians, psychologists, and public health workers. The preferred methods were the more individualized (low observability) methods.

It is bewildering to note in contrast the apparent popularity of group methods for weight control[64] and for exercise programs.[65] Thus, there is evidently an interaction effect between the characteristics of the innovation and the social circumstances surrounding its introduction or promotion.

The acceptability of different health practices is reflected partially in the relative levels of recall knowledge in the public of specific heart-risk factors. The Harris survey[66] of U.S. adults revealed major differences in the salience of specific risk reduction possibilities available to them. A recent community survey in Calgary, Alberta serves as a rough comparison. Mackie found that 53% of her Canadian sample could mention overweight and 42% mentioned smoking as a risk factor, but only 12% mentioned cholesterol. Insufficient exercise was mentioned by 25% of the Canadian respondents.[67] The Harris

S1 S1

FIGURE 2

RELATIVE ACCEPTABILITY OF VARIOUS PREVENTIVE ACTIONS
PUBLIC PERCEPTIONS OF ATTRIBUTES OF THE HEALTH BEHAVIOR
(+ = positive, 0 = neutral or ambivalent, - = negative)

Health Behavior Innovation	Relevance	Social Approval	Advantage	Complexity	Compatibility with Values, Experiences, and Needs	Divisibility or Trialability	Observability
1. Smoking cessation	+	+	+	-	-	+	+
2. Weight control	+	+	+	-	-	+	+
3. Blood pressure control	+	0		-	-	+	-
Medication	+	0		-	-	+	-
Low Sodium Diet	+	-	-	-	-	+	-
4. Low Cholesterol diet	+	0	+	-	-	+	-
5. Exercise	+	+	-	-	+	+	+
6. Preventive medical examinations	+	0	-	+	+	-	+
7. Immunizations	+	+	+	+	+	+	0
8. Prenatal care	+	+	0	+	+	+	+
9. Breast self-examination	+	0	-	-	-	+	-
10. Well-child care	0	+	-	0	+	+	+
11. Seat belt use	-	0	-	0	-	+	0
12. Teeth care	+	+	+	+	-	+	0
13. Defensive driver training	0	-	0	-	+	+	0
14. Family planning	+	+	+	-	0	+	-

survey asked Americans what they thought were the best ways to prevent heart trouble and got almost a reverse order of risk-factor salience: 44% mentioned proper diet, 34% exercise, 17% smoking, and 16% mentioned weight. The U.S.-Canadian differences could be a function of dissemination of information or possibly only differences in coding the two surveys.

Figure 2 scores various health practices plus, minus, or neutral on each of the seven attributes according to their presumed perception in the general population. The attribute which would seem to be most consistently troublesome is complexity. Most of the risk reduction practices that we recommend to the public are not very simple to understand or to implement. If one were to predict from this heuristic chart the health practice least likely to be followed, assuming that all of the perceived attributes are equally weighted, one might predict poorest results with blood pressure control, especially the diet aspects.[10] National data indicate that approximately 50% of those diagnosed with high blood pressure are under treatment, and only 50% of those under treatment for hypertension are under adequate blood pressure control.[66]

The medical and health education literature on compliance with medical regimens can be regarded as documentation of one aspect of *discontinuance* of adopted health practices.[68-70] Other aspects of discontinuance, including clinic dropouts and broken appointments,[28-30,71-73] and rejection after adoption of intra-uterine devices and oral contraceptives[74] are even better documented with life-table techniques. The following analysis will concentrate on the less organized literature related to discontinuance of exercise and diet for illustrative purposes. Smoking cessation dropouts tend to follow similar patterns.

Discontinuance of exercise programs appears to vary widely depending primarily on the rigor of the program. This supports the generalization that the more difficult exercise innovations would have a slower rate of adoption as well as a higher rate of discontinuance. Thus, Mann had a 41% dropout rate over 6 months in a very demanding program of exercise frequency (4 or 5 sessions per week) and intensity,[75] whereas Oja et al lost only 9% over 18 months with 3 sessions per week and with intensity of exercise individually prescribed at 60 to 70% of the subjects' maximal capacity.[65] Between these two extremes are reports of shorter programs with dropout rates roughly corresponding to effort and duration: 14% in an 8- to 10-week program consisting primarily of running at or near maximum capacity of subjects;[76] 27% in a 12-week program of jogging.[77]

Recommendation 10. The Department of Health, Education, and Welfare should *support applied research on the development and evaluation of preventive measures which simplify the requirements for consumer effort, time, technical knowledge, and are adaptable to individual variations in commitment.*

Data on the characteristics of participants in community programs for weight reduction and their discontinuance of dieting practices are limited.[78]

909

Within the past 15 years, traditional clinical approaches to weight reduction have been replaced with more encouraging results by behavior modification procedures,[79,80] self-help groups,[64,78,81-84] and combinations of the two. These more effective modern approaches still require continued study of the persisting problem of discontinuance.

The self-help group approaches represented by TOPS (Take Off Pounds Sensibly) and Weight Watchers have apparently minimized the dropout problem by providing graduated social rewards over time. Yet, the clubs continue to have "a large but undetermined turnover in membership."[64] The average duration of membership in 22 chapters of TOPS studied by Stunkard in West Philadelphia was 16.5 months, with an average weight loss of 15 pounds. Levitz and Stunkard compared these results very favorably with group efforts under medical auspices.[85] There are few evaluations of health education or self-help programs that trace participants for more than a few months after exposure to the program, partly because of the fiscal realities of grants and contracts.

The premature termination percentages for participants in behavior modification experiments for control of obesity have ranged from 0 to 83, with higher dropout rates for female college students than for other groups.[80] The longest program was that of Jeffrey et al, which lasted 18 weeks and lost an average of one percent of its adopters per week.[86] In general, the discontinuance rates are impressively low, but long-term follow-up of 6 months or more has not been included in enough studies to determine the extent of discontinuance of self-management innovations after the experimental phase of weight control, or of exercise and smoking cessation programs.

Recommendation 11. The Department of Health, Education, and Welfare should *seek ways of funding longer term follow-up evaluations of health education without necessarily providing continuous support between program and follow-up.* But duration of program funding should not be unduly shortened just to make evaluation possible.

Ways of implementing Recommendation 11 might include grants or contracts with two separate payments; one during the year of the program and one during the year of the follow-up; funding separate investigators to conduct the program and the follow-up evaluation; soliciting proposals for follow-up of experimental and control groups from major programs conducted in the past. The caution to be sounded is that program design, quality, or duration should not suffer for the sake of evaluation; or there will be nothing worth evaluating.

4. ON WHAT POPULATIONS AND THROUGH WHAT KINDS OF PROVIDERS?

Generalizations concerning the receivers and providers of health education come largely from the literature on diffusion and communication effects. Combined with social structural effects, the variables in this class probably account for most of the remaining variance in health behavior. Many of the communication variables cannot be separated from the structural effects to which they are related.

Paraprofessional and Indigenous Workers

The earlier research on the relative effectiveness of different health communication sources has given way to process interpretations related to the social distance and identification between patient and communicator.[7,11,19,47,48,57,59,87-90] This perspective, combined with cost-control considerations,[1] has supported the trend in developing new categories of indigenous health workers, health education aides, etc. to improve communications with poor patients and ethnic and socioeconomic groups that differ from those of most professionals.[29,91-93] Grosser, for example, evaluated the effectiveness of "local residents" employed in a neighborhood health center as mediators between professionals and low-income patients. He found that they were more effective than professionals in interpreting and predicting the attitudes and behavior of the patients, but not significantly more effective in changing their attitudes and behavior. His interpretation was that local residents begin to identify more with the professionals than with the population from which they were drawn, thereby defeating the purpose for which they were hired and trained.[94]

Other explanations for the failure of indigenous workers to prove as effective as patient educators as might have been expected are that they are not given sufficient training[95] or support by professional workers,[96] they are not perceived as sufficiently expert by patients,[89] or they have become socioeconomically distant as perceived by patients because of their new income, status, and agency identification.[48] Holder demonstrated with maternity patients in an urban hospital that *racial* similarity alone between communicator and patient was inadequate to influence subsequent behavior, but that other dimensions of similarity were important independent of race.[59] Simons et al question the similarity recommendation on grounds that "certain dissimilarities have positive effects on attitude change."[97]

Recommendation 12. The Department of Health, Education, and Welfare should support and encourage the use of indigenous health workers to *supplement rather than replace the communications from professionals.*

Communicator similarity has been found important as a determinant of influence on behavior relevant to health maintenance.[98-100] Mazen identified physical condition (pregnant vs. non-pregnant), race, shared experiences, and practical knowledge as key dimensions of similarity affecting behavioral influence and communicator effectiveness.[101] Similar results have been obtained in family planning.[102]

Grouping of Patients or Consumers

Similarity in physical condition or health problems of participants has been found an effective element in group discussion approaches to health

507

maintenance behavior among alcoholics, obese adults, neighbors meeting on accident prevention, presurgical patients in a general hospital, ambulatory hypertensive patients, parents of children with rheumatic fever, ulcer patients, adult cardiacs, parents in child health conferences, adult asthmatics, congestive heart patients, and, more frequently, diabetics.[103-105] The effectiveness of bringing together individuals who share common health problems appears to be based not only in the tendency to trust and conform to the judgment of others who have the same problem, but also in the quality of pertinent, relevant, and understandable discussion and mutual reinforcement that occurs among participants in such groups. If such groups of patients can also be organized with some socioeconomic homogeneity, the discussion of solutions to common problems is also made more relevant and credible. Self-help groups have followed this principle.[104]

Numerous studies have established the efficacy of group discussion as a powerful educational method to improve preventive health practices, increase visits to physicians, clinic, or mobile units for health screening or early diagnosis, and improve patient compliance with medical regimens.[105] Appropriate utilization of health services and facilities and return visits for medical care appear to be improved by group discussion as an educational strategy with ambulatory patients, such as hypertensives.[1,9-12,41,42,56,78-86,104]

Provider-Patient Communication

In 100 pre-1970 studies on compliance reviewed by Mitchell,[70] the only variables, apart from seriousness of the illness, that showed consistently positive associations with compliance were those indicating a good "doctor-patient relationship" and those indicating explicit attempts to provide patient education. Numerous, more recent compliance studies (not included in Marston's 1970[69] or Mitchell's review) support the need for improved relationships and communication between doctors and patient.[106-117]

Recommendation 13: The Department should *support the development and demonstration of teaching materials and continuing education programs to improve the communication and educational skills of health workers.*

Channels Other than Interpersonal

Communication channels are usually assessed on the basis of the types of adopters (e.g., early vs. late) or the stages of the adoption process (e.g., awareness, interest, trial, evaluation, adoption) with which they work best. In general, mass media are relatively more effective with early adopters and in creating awareness and interest, whereas more interpersonalized media and channels are more effective with late adopters and with behavioral change.

Until recently, the most widely quoted statement on "the role of mass media in public health" was Griffiths' and Knutson's (1960) article with that title. These authors built a persuasive case against the mass media as having anything but an information and support function, and that function primarily with people other than "those populations with which public health is primarily concerned" (p. 521). The main theme of their article was that mass media primarily reach opinion leaders and should "be employed in such a way that these opinion leaders are more effectively utilized in the communication networks they informally serve" (p. 522).[118]

This conclusion is quite consistent with current diffusion theory, but it must be called into question today not for its validity but for its limited view of modern media. When Griffiths and Knutson wrote their article, television was relatively unavailable to high-risk segments of the population. In fact, the latest of the three references cited by Griffiths and Knutson to document their claim of limited exposure to the media was dated 1950, when less than 10% of the U.S. homes had television. Swinehart noted that it was during the 1950s that "quality" magazines had their greatest circulation growth, and the percentage of homes with television had jumped from less than 10 to nearly 95% in 1963.[119]

It is still generally accepted in the diffusion literature as well as the marketing and health education literature that the effects of the mass media are limited primarily to the awareness stage of the adoption process and can be expected to achieve but minor shifts in behavior which is already motivated by other sources of influence.[7-12,31-38,52-61,120-123] It is also true that the urban poor continue to rely very much on interpersonal rather than mass channels of information.[124] Finally, it is still true that television is too expensive and inaccessible to most health agencies and institutions to do them much good beyond the information function.[125]

Competition for television time is not only formidable in economic terms, but the pervasiveness of erroneous and at times dangerous health information in the mass media is perhaps impossible to control without total censorship on advertisements as well as dramatic scripts. Manoff notes that nearly one billion dollars was spent on television by food advertisers in 1972, and nearly one third of all foods advertised can be classified as frivolous at best (candy, snacks, soft drinks).[126]

Nutrition education by other means continues to show increments in knowledge and sustained awareness of certain problems, but changes in behavior tend to be proportional to the intensity and continuity of the effort.[127-131] The majority of nutrition education efforts described in these and earlier articles are of brief duration and are still more didactic than experiential in their approach. The school lunch and Head Start mothers programs[132] and the outreach programs such as the Expanded Food and Nutrition Education Program (EFNEP)[47,133] appear to have greater potential than the one-shot "classes" or clinical approaches. In EFNEP, homemakers are visited monthly by nutrition aides who show them how to prepare new recipes. This kind of outreach program with home-centered health education and continuity has been found cost-effective in family planning as well, so long as the number, frequency, or duration of visits do not exceed their points of diminishing returns.[1,4,8,59-61,93,102]

At least a small residual of nonadopters and even of people who disavow

any awareness of innovations for which there have been saturation health programs will usually remain. In a study of employees of a company in Chicago, Stamler found that of the men examined annually for 10 years and told repeatedly that they were hypertensive, and referred to their private doctors, 6% still disavowed knowledge of their elevated blood pressure.[134] Thus, no matter how targeted the communication, relevant the message, and personalized the medium, there will be other constraints on knowledge, persuasion, and adoption that preclude a complete result in a given population. Health education requires more than communication.[1-14,56]

5. AT WHAT PROGRAM COSTS?

Previous reviews of the literature suggest the cost-benefit advantages of health education initially on the basis of cumulative evidence of relationships between specific health education strategies and improvement in specific indices of health or behavior.[2,12] By extrapolation from these more or less consistent relationships, cost-benefit estimates can be made on the trade-offs between investments in health education and alternative strategies related to specific diseases in terms of projected savings in primary health care, bed-days of inpatient care, work-days lost, long-term home or institutional care, and rehabilitation.

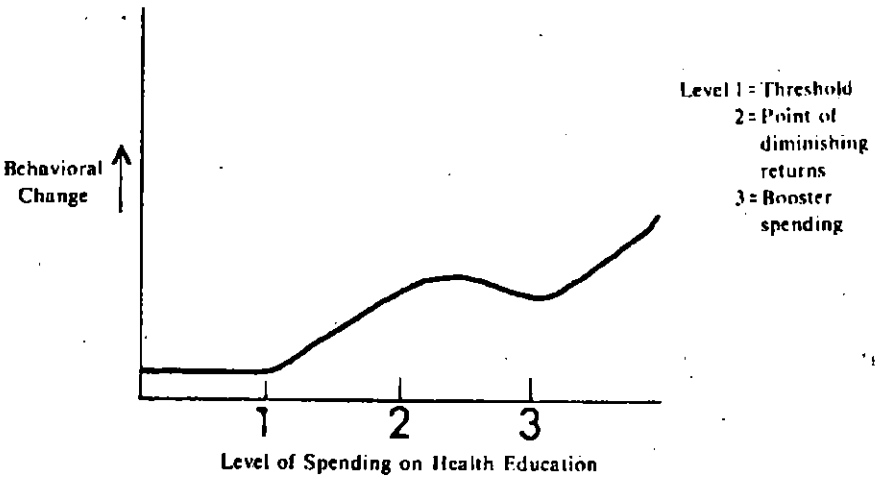
Where the trade-offs cannot be shown to be cost-effective for investments in health education, alternative policy considerations are suggested by the analysis, such as the necessity of more coercive preventive measures to achieve sudden changes in health behavior. Where, for example, the cumulative evidence does not indicate the probability of a sufficiently high success rate for health education directed at seat-belt usage to effect a reduction in emergency room visits, then alternative strategies, such as legislation directed at drivers, manufacturers, or highway engineers, are suggested.

Before recommending program areas and target populations in which health education should receive priority support, some final guidelines are suggested here on how to assess the optimum level of expenditure for specific programs or purposes.

The minimum that should be spent on health education for a specific purpose is the amount required to achieve that purpose. While this appears tautological, it is a tragic fact of much health education funding that the budget has been below the minimum required to obtain a desired effect. A threshold level of input is required before a difference in behavior will be perceptible.

Cost-benefit estimates suggest whether anything should be spent on a given educational program. If the ratio of potential benefits to costs is not greater than one, then nothing should be spent, strictly speaking. Cost-effectiveness ratios, while still in short supply, will enable policy-makers to compare the costs of two or more methods in achieving the same outcome. Then, depending on how much of that outcome is needed, the threshold level of spending might be directly obtained from the cost-effectiveness measure by

FIGURE 3
RELATIVE LEVELS OF EXPENDITURE ON HEALTH EDUCATION AND THEIR THEORETICAL CONSEQUENCES



multiplication. The range within which decisions on variable amounts of spending should be made is the range between the threshold level and the point of diminishing returns, shown as levels 1 and 2 in Figure 3.

It is seldom the case that educational effects are permanent. Research on behavioral change, dropouts, discontinuance of health practices, learning curves, and memory curves predictably yields a pattern of forgetting and backsliding. Reinforcement is known to be as important to education as booster shots are to sustained immunization. Thus, after reaching the saturation level of spending, a period of time should elapse before introducing an additional expenditure on education for the same population. At the point when the behavioral changes achieved begin to deteriorate, booster spending on reinforcement or new educational methods may be necessary, shown as level 3 in Figure 3.

Recommendation 14. The Department should *review proposals for health education spending to determine whether they have given consideration to threshold, saturation, and booster levels of spending.*

Economy of Scale: Methods Increasing Cost-Effectiveness

By aggregating, standardizing, and formalizing health education, economies of scale can be achieved at the expense of some effectiveness at the individual level. The cost per unit of production per patient educated, or per message disseminated can be reduced through aggregation. The goal in educational innovation is to design standardized materials, methods, and procedures that can be produced or implemented on a larger scale (therefore at lower unit prices) without sacrificing a proportionate amount of

305

effectiveness inherent in more individualized, personalized and flexible methods and procedures.[1]

The Timing of Investments

There are optimum times in the life history of a given health problem or program when spending for specific kinds of educational inputs will minimize risk and maximize payoff. As educational methods become more aggregated they tend to minimize risks while usually also minimizing payoff. More individualized methods increase the risks of loss because they are more expensive, but they also increase the possibility of benefits to the extent that they are usually more effective. The timing of investments in one educational method vs. another should follow the same decision rules governing priority setting in other areas of administration.

Recommendation 15. In a new program, the people most likely to respond first are those who are already motivated to adopt the recommended health practice. *During this early phase, low unit-cost measures such as written materials (pamphlets, etc.) are effective enough. As the program moves through the target population to increasingly "hard to reach" and high-risk groups, more expensive educational methods such as counseling sessions and home visits should be supported.*

The extra cost to reach one high-risk patient will be offset in most cost-benefit computations by the greater benefit accruing from behavioral change in a high-risk as contrasted with a low-risk patient. In economic terms, the marginal utility of behavioral change is greater in high-risk than in low-risk populations. This will be discussed in a final set of criteria: problem areas and target populations most likely to benefit from health education.

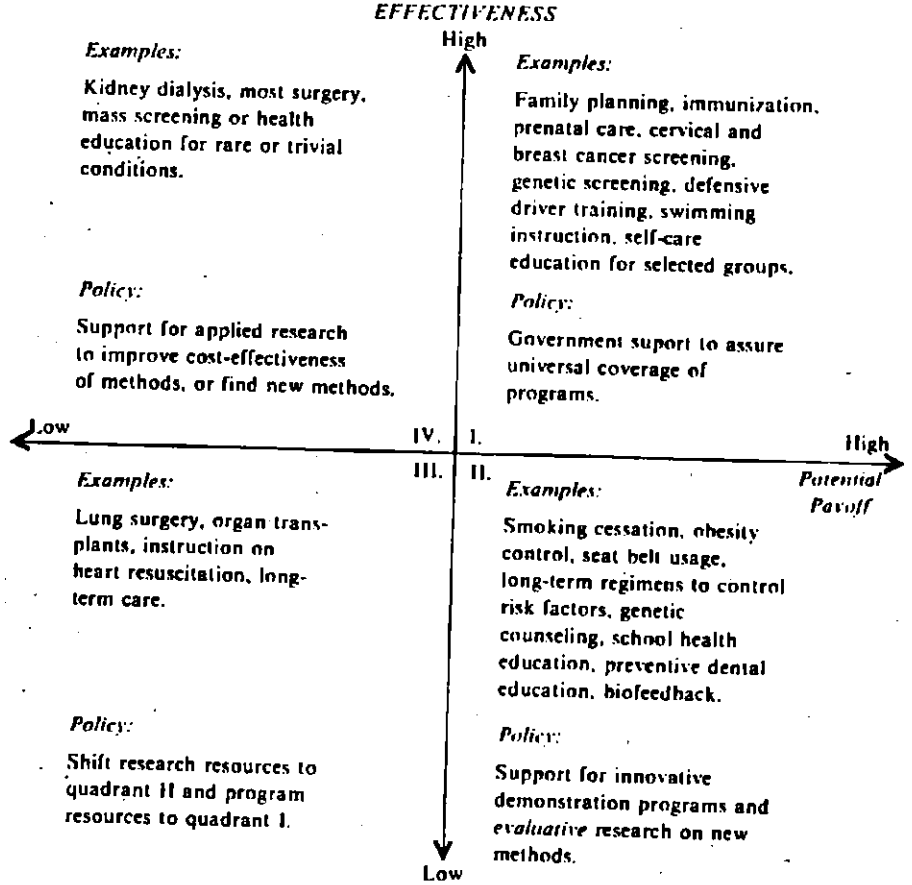
CRITERIA FOR POLICY RECOMMENDATIONS ON PROGRAM AREAS

Medical and health education strategies can be classified according to their known effectiveness, based on the foregoing review, and their potential payoff in cost-benefit terms. In this two-dimensional scheme, specific strategies can be placed in one of four policy categories, as shown in Figure 4. In quadrant I, where effectiveness is high and payoff is high, there is clear justification for major investments in programs to assure universal coverage. Examples include health education in relation to immunizations, family planning, cervical and breast cancer screening, and patient education in relation to acute conditions for which the therapeutic regimen is brief and relatively uncomplicated.

In quadrant II, where potential payoff is high but effectiveness is still low—as with seat belt usage, smoking cessation, and patient education with chronic conditions *where the regimens are long-term, complicated, painful, or otherwise troublesome*—the policy recommendation is for evaluative research on health education methods to improve effectiveness and behavioral research to understand the barriers and motives to change.

FIGURE 4

POLICY IMPLICATIONS FOR HEALTH STRATEGIES CLASSIFIED ACCORDING TO EFFECTIVENESS AND COST-BENEFIT POTENTIAL



Quandrant IV includes strategies for which effective means exist but payoff is low because the conditions are rare, the means are too expensive, or the problems are trivial. They include many surgical procedures, kidney dialysis, and mass screening or health education for diseases or conditions with very low incidence rates or very minor consequences. The policy recommendations for these problems is continued or increased investments in applied research to improve the cost-effectiveness of the methods.

The methods in quadrant III are those for which both effectiveness and potential payoff are low. They include procedures such as surgery for lung cancer and organ transplants and biofeedback where some spectacular successes have been recorded, but the overall cost-effectiveness and cost-benefit ratios are poor. The policy recommendation for these is to shift some of the public funds now being used in support of tertiary care facilities to programs in quadrant I and evaluative research in quadrant II.

Primary Prevention

The great potential payoffs are in preventing exposure and vulnerability to disease and injury agents. Primary prevention has had its successes and its disappointments, but it has seldom had investments proportional to its potential. Compared with the medical specialties, prevention has been—relative to other parts of the health care system—undervalued, unglamorous, and capable of helping save costs in other parts of the health care system if given the opportunity.

The inventory of primary prevention measures available to the public is well known, but most of these measures require behavioral changes and cooperation on the part of private citizens. Some of these can be legislated, but the majority depend on voluntary change, which depends heavily on adequate investments in health education.

Secondary Prevention

Earlier diagnosis and treatment of many diseases and the earlier identification of risk factors that can be controlled are the main hopes for containing the major causes of mortality and disability in this country. These include earlier diagnosis and treatment of neoplasms, alcoholism, streptococcal infections, tuberculosis, asthma, irregularities in the functioning of selected organs (heart, hearing defects, amblyopia, kidney), and other conditions for which there are medical or surgical technologies to spare the victims at least some of the suffering they would otherwise experience with later treatment. In addition, there are conditions not necessarily regarded as diseases themselves, the early identification of which can lead to treatment plans or reorganization of life styles or environment that will reduce the otherwise high probability for these people that they would experience a particular kind of health problem. These include the coronary-risk factors of cigarette smoking, hypercholesteremia, and hypertension; the accident-risk factors of alcoholism, epilepsy, and visual defects; the genetic-risk factors of sickle-cell trait or Tay-Sachs; the lung-risk factors of smoking, industrial or housing exposures to coal dust, asbestos, mite dust, pollens, exhausts or radiation; the allergenic sensitivity to specific substances, and so forth.

The strategies for achieving secondary prevention in this regard, apart from the therapeutic procedures themselves, are *multiphasic screening; periodic medical and dental examinations; mass screening for specific conditions* (including genetic screening); *public education regarding early recognition of symptoms* (including breast self-examination, cancer "danger signs," indicators of a "drinking problem," and early onset signs of acute attacks—heart, asthma, use of a fever thermometer) and the importance of *seeking medical care early for diagnosis* of these symptoms; patient education following diagnosis and prescription to improve *compliance with therapeutic regimens; genetic counseling;* and continuing reinforcement education for the adoption of risk-reduction behavior.

This is not intended as a complete inventory, but rather as a broad classification of strategies. Furthermore, the order in which these strategies are listed does not imply priority. The temporal precedence of screening over patient education for compliance does not necessarily mean that H.E.W. should first emphasize screening and then shift its efforts to patient education. In fact, there is some justification for concern with the wisdom of the National High Blood Pressure Education Program's extensive screening activities, for example, which have the potential of flooding the clinics with hypertensives before the current record of 50% compliance among those under treatment can be improved upon through patient education research and evaluation. Indeed, the reduction of the 50% of those not currently under treatment would likely increase the percentage of treated but not under control, because they are probably less motivated than those now under treatment.

The issue for policy consideration is whether any or all of these strategies are worthy of shifting medical resources from their current downstream emphasis where they have at least a known effectiveness. The main advocates for the shift upstream have been public health and preventive medicine who have argued more from logic than from data that earlier diagnosis and treatment must make a difference in outcomes when organized on a large-scale basis. Unfortunately, the data do not consistently uphold the logic. But the data are strong enough to recommend health education efforts in relation to selective screening programs for selected (high-risk) groups.

The proliferation of specialized screening programs runs the risk of leading into all of the same problems that specialization and fragmentation of treatment has had. The arguments for comprehensive primary care should apply equally to screening whenever possible. Health education could then be based in screening centers with a single, rather than multiple, recruitment effort. Mobile units should concentrate particularly on industrial settings to reach blue-collar, high-risk groups least likely to seek periodic examinations.

Patient Education

One of the objections of health professionals to mass screening concerns the observation that large proportions of patients screened into treatment do not adhere to their therapeutic or risk-reduction regimens.

The research on patient compliance with medical regimens already reviewed strongly points to problems of doctor-patient and nurse-patient relationships and gaps in communication as major sources of noncompliance. Starting with failure to acknowledge and meet patient expectation on the first visit, through failure to communicate therapeutic information effectively and persuasively, failure to take family or work situation into account, and inability of the practitioner to predict or recognize the noncompliant patient, several points for educational intervention were identified by the research literature. Obviously, the education must include staff development and consultation on communication needs and skills. There are, of course, alternatives to face-to-face communication.[135,136]

Simonds has suggested as a policy goal for this decade that hospitals "should take a major role in the provision of services for health promotion and health maintenence, including the provision of health education to all who use the hospital or are within its geographical area."[137] Economic justifications for such roles for hospitals currently apply primarily to those serving prepaid plans and tax-based hospitals rather than to proprietary hospitals. New provisions for reimbursement of health education services under Medicare, Blue Cross, and potentially, National Health Insurance, make health education economically feasible even for hospitals that have depended on keeping patients horizontal.

Reports previously submitted or recently published reviewed the cost-benefit status of patient education and concluded that the potential benefits are much greater than the costs of supporting health education in medical care settings.[2,66,67] Apart from the direct administrative benefits to the health care system and the direct benefits for the patient, health education also offers indirect benefits by improving the outcomes of medical and surgical procedures. Major emphasis should be given to the development of professional and material resources for patient education programs in anticipation of national health insurance legislation. The following four recommendations are endorsed from those of the National Conference on Preventive Medicine of the Fogarty International Center and the American College of Preventive Medicine.[138]

Recommendation 16. The Department should *"require, as a condition of federal, state, and local funding, that all hospitals, health maintenance organizations, neighborhood health centers, and other community health facilities—both public and private—provide...in regulations to be promulgated by the Secretary of DHEW, consumer health education as an integral part of their service responsibilities to their individual users and to their service constituencies.*

Recommendation 17. The Department should "*...encourage all third-party payors—public and private—to reimburse providers for the net costs of patient education.*"

The following additional recommendations of the National Conference on Preventive Medicine are consistent with the foregoing analysis and findings:

Recommendation 18. The Department should *"encourage coordinated, comprehensive school health education programs through federal research and development funds and grants in aid to states."* The federal support should be directed to "evaluating school health education curricula, training teachers in health education, and encouraging state education agencies to take a more active role in promoting school health education."

Recommendation 19. The Department should continue to *"provide federal support for research and development in consumer health education techniques, methodologies, and programs, and their evaluation."*

Setting Priorities

With the preceding analysis of the state-of-the-art and general recommendations as background, the remaining task is to focus health education on those target problems and populations that will yield the greatest improvements in the nation's health. For each health problem (whether defined in terms of death, disease, or disability), the priority relative to other health problems varies for different populations.

Hence, the first general recommendation to be offered is the *endorsement of a method for assigning different priorities to the same health problems in different populations,* in lieu of a general priority ranking for the nation. Following the recommendation of this technique for allocation decisions within population groups, a series of recommendations for high priority populations will be offered on the basis of anticipated health education coverage in National Health Insurance.

An Index for Assigning Program Priorities Within Populations

The two technical problems facing the policy analyst in deciding which problems to tackle first in a given population are: (1) How should the various available measures of the problems (deaths, bed-days, outpatient days) be weighted? (2) How much improvement is possible in each problem, regardless of its extensity relative to other problems? The administrator obviously must take other (non-technical) considerations into account, but these two at least would provide the basis for a scientific recommendation. The first is a question of relative *need* for improvement. The second is a question of relative *potential* for improvement. The "need" question corresponds roughly to the "potential payoff" of "benefit" dimension of Figure 4 discussed in the previous section. The "potential for improvement" question corresponds to the "effectiveness" dimension of Figure 4.

A method has been developed for the Indian Health Service[139] to combine both of these dimensions in a single index (Q) which uses national averages (age- and sex-adjusted) as the standard of "potential improvement" for a given population. Thus,

$$Q = \frac{M_i}{M_n} DP + \frac{274A + 91.3B}{N}$$

where

M_i = age- and sex-adjusted mortality rate for the target population for a given cause of death,

M_n = age- and sex-adjusted mortality rate for the reference population, i.e., total U.S. population,

D = crude mortality rate (per 100,000 population) for the target population,

P = years of life lost because of premature death for the target population,

A = hospital days for the target population,

B = outpatient visits for the target population,

274 and 91.3 = constants used to convert A and B to years per 100,000 population; 3 outpatient visits are equated in time to one day.

The Q index takes into account the major considerations that go into both professional judgment and community values about the relative urgency of different health problems. If a community has an excess of deaths, hospital days, or outpatient visits for a given disease or condition relative to the national average after adjusting for age and sex distributions, that community will have a higher Q rating for that disease or condition. This allows the national averages to serve as a benchmark for the state-of-the-art, including medical and health technology, health care facilities, self-care, and life styles. The application of this index should therefore lead to the focusing of health education resources on those problems for which selected populations have suffered disproportionately.

Table 1 illustrates how the Q index resolves major discrepancies in the ranking of health problems by hospital days, outpatient visits, and mortality.

Recommended Priorities for Health Education Goals

The following recommendations for target populations and behavioral goals to be achieved through health education are based on assessments[140] of the most productive areas in which payment under National Health Insurance (NHI) might be made:

Recommendation 20. Payments for health education services should be made under national health insurance as part of (a) maternity care, (b) preventive child care, (c) adult preventive care, and (d) chronic illness care, in that order of priority.

Recommendation 21. Community health education efforts should seek to increase prenatal visits to an average of one per month through the 28th week of pregnancy, two per month from 28 to 36 weeks of pregnancy and one per week after the 36th week of pregnancy.

Recommendation 22. During pregnancy, instruction should be offered on nutrition, family planning, self-care, delivery, and parenthood: up to six hours of group instruction each, for mother and father.

Recommendation 23. During intrapartum care, through the fifth postpartum day, hospitals should provide reimbursable bedside instruction to mothers up to six hours total on self-care, infant care, and family planning.

Recommendation 24. At least one post-partum home visit by a public health nurse or other qualified health provider within the first two weeks after discharge from the hospital should be provided in order to assist with implementation of home care and to provide additional education for parenthood.

Recommendation 25. Community health education should encourage and NHI should provide for one office or clinical visit each for the mother and infant to their respective physicians during the first eight weeks after birth for routine examination, advisement, laboratory procedures, and follow-up family planning service.

Recommendation 26. Community health education for preventive child care should encourage, and NHI should support, seven visits during the first year of life, three visits during the second year of life, and one visit each year between ages 2 and 5.

Recommendation 27. Health education in preventive child care settings should include counseling and anticipatory guidance regarding nutrition, accidents, hygiene, and child development according to standards to be established by the Secretary of HEW.

Recommendation 28. Community and school health education should encourage, and NHI should reimburse total costs of, no more than one visit per year to an office or clinic for purposes of maintaining immunizations, or

TABLE 1

CLASSES OF DISEASE RANKED BY TRADITIONAL HEALTH DATA AND BY THE Q INDEX (IN THE INDIAN AND ALASKAN NATIVE POPULATION)[139]

Classes of Disease[1]	Rank Order			
	Inpatients	Outpatients	Deaths	Q
Infective and parasitic diseases	1	2	9	2
Neoplasms	12	14	5	15
Endocrine, nutritional, and metabolic diseases	9	7	10	9
Diseases of the blood and blood-forming organs	17	15	14	17
Mental disorders	13	13	13	11
Diseases of the nervous system and sense organs	6	3	6	6
Diseases of the circulatory system	8	9	2	12
Diseases of the respiratory system	2	1	3	3
Diseases of the digestive system	4	5	4	4
Diseases of the genito-urinary system	7	10	11	10
Complications of pregnancy, childbirth, and the puerperium	5	8	15	7
Diseases of the skin and subcutaneous tissue	11	6	16	13
Diseases of the musculoskeletal system and connective tissue	10	11	17	14
Congenital anomalies	14	16	12	16
Certain causes of perinatal morbidity and mortality	16	17	8	8
Symptoms and ill-defined conditions	15	12	7	5
Accidents, poisonings, and violence	3	4	1	1

[1] *International Classification of Diseases, 8th revision. Adapted for use in the United States. PHS Publication No. 1693. U.S. Government Printing Office, Washington, D.C., 1968.*

513

for physical development evaluations, and counseling regarding sexual development, alcoholism and drug abuse; and including five consultation visits when requests for those visits originate from a school health service for reasons of unsatisfactory progress in school.

Recommendation 29. Community and school health education should encourage, and NHI should support, (a) *routine dental services up to age 13 including oral examination, X-ray, prophylaxis, topical fluoride, counseling and education; and restorations;* and (b) developmental vision care and audiometry services up to age 18.

Recommendation 30. Community, college, and industrial health education programs should encourage, and NHI should support, adult preventive care including at least:

(a) 2 routine physicals between ages 18-30,
(b) 3 routine physicals between ages 31-40,
(c) 4 routine physicals between ages 41-50,
(d) 5 routine physicals between ages 51-60,
(e) annual physicals over age 60.

Recommendation 31. In addition to the above preventive services, all women in reproductive ages should be encouraged through health education to receive annual breast and pelvic examinations, a Papanicolaou smear and a G.C. culture; and should receive in the annual visit health education and family planning counseling.

Recommendation 32. Health and nutrition education should be provided by NHI as part of reimbursed home health services for both the patient and the family, as one means of reducing the length of stay and the probability of returning for further institutional care.

Recommendation 33. Patient education supported by NHI on both an individual and small group basis should be geared to meet the following objectives:

(a) development of appropriate self-management techniques and increased compliance with prescribed regimens;
(b) assist the patient or health consumer to utilize benefits appropriately;
(c) reduction of risks leading to health problems, minimization of the effects of an illness, and avoidance of recurrences; and
(d) orientation of the individual to seek early screening or treatment.

Recommendation 34. Within schools, health education and safety education (including driver education) should be carefully geared to the age-specific and sex-specific interests of students and should be taught only by teachers with specific preparation in the topics.

Recommendation 35. Within industry, alcohol and drug abuse education should be integrated with reimbursable, non-punitive programs of alcoholism counseling and rehabilitation.

Recommendation 36. Community health education and referral services should be designed to minimize the barriers to seeking care by individuals in need of crisis intervention for suicides, drug and other health problems having social stigma or opprobrium associated with them.

These priorities, in approximately the foregoing order, should yield the maximum cost-benefit advantage to society and to consumers currently obtainable from health education.

The earlier recommendations for development, evaluation, and demonstration in less effective but high payoff areas such as smoking cessation, obesity control, diet therapy, genetic counseling, and long-term hypertension and diabetes management will lead hopefully to a further strengthening of health education as a central feature of health policy.

ACKNOWLEDGEMENTS

Connie Kansler assisted in several stages of this project. Sambasiva Reddi assisted with economic analyses. Sigrid Deeds, Ruth Faden, Pearl German, and Virginia Wang provided helpful suggestions. Mona Mitnick guided the entire effort with encouragement and critical reviews.

REFERENCES

1. Green LW: Evaluation and measurement: Some dilemmas for health education. Am J Public Health 67:155-161, February 1977.
2. Green LW: The potential of health education includes cost-effectiveness. Hospitals, JAHA 50:57-61, May 1, 1976.
3. Green LW: Methods available to evaluate the health education components of preventive health programs. In American College of Preventive Medicine Preventive Medicine USA. New York, Prodist, 1977. Also in Somers AR, Health Promotion and Consumer Health Education. New York, Academic Publications, 1976.
4. Deeds S: A Guidebook for Family Planning Education. DHEW Publication no (HSA) 74-16002. Rockville, Md, Bureau of Community Health Services, 1973.
5. Kazdin AE: Behavior Modification in Applied Settings. Homewood, Ill, Dorsey Press, 1975.
6. Ubell E: Health behavior change: A political model. Prev Med 1:209-221, 1972.
7. Rogers EM: Communication of Innovations. New York, Free Press, 1971.
8. Rogers EM: Communication Strategies in Family Planning. New York, Free Press, 1973.
9. Mico PR, Ross HS: Health Education and Behavioral Science. Oakland, Third Party Associates, 1975.
10. Green LW: The application of diffusion theory to cardiovascular risk behavior in the public. In Enelow A, Henderson JB (eds): Applying Behavioral Sciences to Cardiovascular Risk. New York, American Heart Association, 1975.
11. Green LW: Should health education abandon attitude change strategies? Perspectives from recent research. Health Educ Monogr 1(30):25-48, 1970.
12. Green LW: Toward cost-benefit evaluations of health education: Some concepts, methods and examples. Health Educ Monogr 2 (suppl):34-64, 1974.
13. Andersen R: A Behavioral Model of Families' Use of Health Services. University of Chicago Center for Health Administration Studies, Research Series No. 25, 1968.
14. Becker MH (ed): The Health Belief Model and Personal Health Behavior. Health Educ Monogr 2:324-473, 1974.
15. Kerri JN: Anthropological studies of voluntary associations and voluntary action: A review. J Vol Action Res 3:10-25, 1974.
16. Marris P, Rein M: Dilemmas of Social Reform: Poverty and Community Action in the United States. Chicago, Aldine Publishing, 1973.
17. Grosser CF: New Direction in Community Organization: From Enabling to Advocacy. New York, Praeger Press, 1973.

18. Wang VL, et al: An approach to consumer-patient activation in health maintenance. Public Health Rep 90:449-454, 1975.
19. Warren RL: Community change: Some lessons from the recent past. Community Develop J 9:2-10, 1974.
20. Anthony WA: Human relations training and rehabilitation counseling: Further implications. Rehabil Coun Bull 17:171-175, 1974.
21. Canfield RE: The physician as a teacher of patients. J Med Educ 48(12, pt 2):79-87, 1973.
22. Carrera MA, Rosenberg G: Inservice education in human sexuality for social work practitioners. Clin Soc Work J 1:261-267, 1973.
23. Levine DM, Bonito AJ: Impact of clinical training on attitudes of medical students: Self-perpetuating barrier to change in the system? Br J Med Educ 8:13-16, 1974.
24. McKinlay JB: On the professional regulation of change. Sociol Rev Monogr No 20:61-84, 1973.
25. Weiss RL, Swearingen RV: Chairside Psychology in Patient Education: A Self-Instruction Course. US Department of Health, Education, and Welfare, National Institutes of Health. San Francisco, Dental Health Center, 1973.
26. Wolle JM: Multidisciplinary teams develop programming for patient education. Health Serv Rep 89:8-12, 1974.
27. Becker MH, Drachman RH, Kirscht JP: A new approach to explaining sick-role behavior in low income populations. Am J Public Health 64:205-216, 1974.
28. Finnerty FA Jr, Mattie EC, Finnerty FA III: Hypertension in the inner city: I. Analysis of clinic dropouts. Circulation 47:73-75, 1973.
29. Fletcher SW, Appel FA, Bourgeois MA: Management of hypertension: Effect of improving patient compliance for follow-up care. JAMA 233(3):242-244, June 21, 1975.
30. Shroeder SA: Lowering broken appointment rates at a medical clinic. Med Care 11:75-78, 1973.
31. Becker MH, Green LW: A family approach to compliance with medical regimens: A selective review of the literature. Int J Health Educ 18:2-11, 1975.
32. Mullen PF: Health education for heart patients in crisis. Health Serv Rep 88:669-675, 1973.
33. Green LW, Green PF: Intervening in social systems to make smoking education more effective. In Steinfeld J, et al (eds): Proceedings of the Third World Conference on Smoking and Health. Washington, US Government Printing Office, 1977.
34. Antonovsky A, Hartman H: Delay in the detection of cancer: A review of the literature. Health Educ Monogr 2:98-128, 1974.
35. Green LW, Roberts BJ: The research literature on why women delay in seeking medical care for breast symptoms. Health Educ Monogr 2:129-177, 1974.
36. Green LW: Site- and symptom-specific factors in secondary prevention of cancer. In Cullen J, Fox B, Isom R (eds): Cancer: The Behavioral Dimensions. New York, Raven Press, 1976.
37. Atwater JB: Adapting the venereal disease clinic to today's problem. Am J Public Health 64:433-437, May 1974.
38. Udry R, et al: Can mass media advertising increase contraceptive use? Fam Plann Perspect 4:37-44, July 1972.
39. Kalmer H: Member-Professional Agreement Related to Use of a Prepaid Adult Ambulatory Care Setting. Paper presented at the American Public Health Association, Chicago, November 1975.
40. Stimson GV: Obeying doctor's orders: A view from the other side. Soc Sci Med 8;97-104, February 1974.
41. Geertsen HR, Gray RM, Ward JR: Patient non-compliance within the context of seeking medical care for arthritis. J Chronic Dis 26:689-698, 1973.
42. Zola IK: Pathways to the doctor – from person to patient. Soc Sci Med 7:677-689, 1973.
43. Oakes TW, Ward JR, Gray RM, et al: Family expectations and arthritis. Patient compliance to a hand resting splint regimen. J Chronic Dis 22:757-764, 1970.
44. Porter AMW: Drug defaulting in general practice. Br Med J 1(5638):218-222, 1969.

45. Cauffman JG, Lloyd JS, Lyon ML, et al: Health information and referral services within Los Angeles County. Am J Public Health 63:872-877, 1973.
46. Littlefield JL: Nutritional labeling revisited. Regulatory considerations. Paper presented at Food Update XIII, Scottsdale, Arizona, 1974.
47. Wang VL, Ephross PH, Green LW: The point of diminishing returns in nutrition education through home visits by aides: An evaluation of EFNEP. Health Educ Monogr 3:70-88, 1975.
48. Green LW: Status Identity and Preventive Health Behavior. Pacific Health Education Reports No. 1. Berkeley, University of California, 1970.
49. Steele J, McBroom WH: Conceptual and empirical dimensions of health behavior. J Health Soc Behav 13:382-392, 1972.
50. Williams AF, Wechsler H: Interrelationship of preventive actions in health and other areas. Health Serv Rep 87:969-976, 1972.
51. Kasl SV, Cobb S: Health behavior, illness behavior and sick role behavior. Arch Environ Med 12:246-266; 531-541, 1966.
52. Kasl SV: The health belief model and behavior related to chronic illness. Health Educ Monogr 2:433-454, 1974.
53. Baric L: Recognition of the "at-risk" role: A means to influence health behavior. Int J Health Educ 12:24-34, 1969.
54. Hazell JW, Hodges FB, Cunningham GC: Intermediate benefit analysis--Spencer's dilemma and school health services. Am J Public Health 62:560-565, 1972.
55. Green LW, Figã-Talamanca I: Suggested designs for evaluation of patient education programs. Health Educ Monogr 2:54-71, 1974.
56. Green LW, Levine DM, Deeds S: Clinical trials of health education for hypertensive outpatients: Design and baseline data. Prev Med 4:417-425, 1975.
57. Burt R: Differential impact of social integration on participation in the diffusion of innovations. Soc Sci Res 2:1-20, 1973.
58. Feldman JJ: The Dissemination of Health Information: A Case Study in Adult Learning Chicago, Aldine Publishing Co, 1966.
59. Green LW: Identifying and overcoming barriers to the diffusion of knowledge about family planning. Adv Fertil Control 5:21-29, 1970.
60. Lin N, Hingson R: Diffusion of family planning innovations: Theoretical and practical issues. Stud Fam Plann 5:189-194, 1974.
61. Vertinsky P, Vertinsky I, Zaltman G: Health information diffusion: An integrated model. Int J Health Educ 15(suppl), 1972.
62. National Clearinghouse for Smoking and Health: Adult Use of Tobacco, 1970. DHEW Publication no (HSM) 73-8727. Atlanta, US Department of Health, Education, and Welfare, Center for Disease Control, 1973.
63. Schwartz JL, Dubitzky M: Expressed willingness of smokers to try ten smoking withdrawal methods. Public Health Rep 82:855-861, 1967.
64. Stunkard AJ: The success of TOPS, a self-help group. Postgrad Med 51:143-147, 1972.
65. Oja P, Teräslinna P, Partanen T, et al: Feasibility of an 18 months' physical training program for middle-aged men and its effect on physical fitness. Am J Public Health 64:459-464, 1974.
66. Harris L & Associates: The Public and High Blood Pressure: A Survey Conducted for the National Heart and Lung Institute. DHEW Publication no (NIH) 74-356. Bethesda, Md, National Institutes of Health, 1973.
67. Mackie M: Lay perception of heart disease in an Alberta community. Can J Public Health 64:444-454, 1973.
68. Davis MS: Variations in patients' compliance with doctor's advice: An empirical analysis of patterns of communication. Am J Public Health 58:274-288, 1968.
69. Marston M: Compliance with medical regimens: A review of the literature. Nurse Res 19:312-323, 1970.
70. Mitchell JH: Compliance with medical regimens: An annotated bibliography. Health

5/13

Educ Monogr 2(1):75-87, 1974.
71. Finnerty FA, Shaw L, Himmelsbach C: Hypertension in the inner city, II: Detection and follow-up. Circulation 47:76-78, 1973.
72. Stokes J III, Payne G, Cooper T: Hypertension control—the challenge of patient education. N Engl J Med 289:1369, 1973.
73. Tagliacozzo DM, Ima K, Lashof JC: Influencing the chronically ill: The role of prescriptions in premature terminations of outpatient care. Med Care 11:21-29, 1973.
74. Tapham RJ, Mauldin PW: National family planning programs: Review and evaluation. Stud Fam Plann 3:29-52, 1972.
75. Mann RA: The behavior-therapeutic use of contingency contracting to control an adult behavior problem: Weight control. J Appl Behav Anal 5:99-109, 1972.
76. Kilbom A, Hartley LH, Saltin B, et al: Physical training in sedentary middle-aged and older men. I. Medical evaluation. Scand J Clin Lab Invest 24:315-322, 1969.
77. Harris WE, Bowerman W, Fadden RBM, et al: Jogging: An adult exercise program. JAMA 201:133-135, 1967.
78. Dwyer JT, Feldman JJ, Mayer J: The social psychology of dieting. J Health Soc Behav 11:269-287, 1970.
79. Abramson EE: A review of behavioral approaches to weight control. Behav Res Ther 11:547-555, 1973.
80. Hall SM, Hall RG: Outcome and methodological considerations in behavioral treatment of obesity. Behav Ther 5:352-364, May 1974.
81. Cornacchia A: A layman's view of group therapy in weight control. Can J Public Health 58:505-507, 1967.
82. Kornhaber A: Group treatment of obesity. GP 38:116-120, 1968.
83. Stunkard AJ, Levine H, Fox S: The management of obesity: Patient self-help and medical treatment. Arch Intern Med 125:1067-1072, 1970.
84. Wollersheim JP: Effectiveness of group therapy based upon learning principles in the treatment of overweight women. J Abnorm Psychol 76:462-474, 1970.
85. Levitz LS, Stunkard AJ: A therapeutic coalition for obesity: Behavior modification and patient self-help. Am J Psychiatr 131:423-427, 1974.
86. Jeffrey DB, Christensen FR, Pappas JP: A case study report of a behavioral modification weight reduction group: Treatment and follow-up. In Research and Development Report No. 33. Salt Lake City, University of Utah Counseling Center, 1972.
87. Berkanovic E, Reeder I: Ethnic, economic, and social psychological factors in the source of medical care. Soc Prob 21:246-259, 1973.
88. Bossé R, Rose CL: Age and interpersonal factors in smoking cessation. J Health Soc Behav 14:381-387, 1973.
89. Holder I: Some theoretical and practical considerations in influencing health behavior. Health Educ Monogr 1(30):49-69, 1970.
90. Holder L: Effects of source, message, and audience characteristics on health behavior compliance. Health Serv Rep 87:343-350, 1972.
91. Golden A, Carlson D, Harris B: Non-physician family health teams for health maintenance organizations. Am J Public Health 63:732-736, 1973.
92. Jordan HA, Levitz LS: Behavior modification in a self-help group. J Am Diet Assoc 62:27-29, 1973.
93. Westheimer RK, et al: Use of paraprofessionals to motivate women to return for post-partum check-ups. Public Health Rep 85:625-635, 1970.
94. Grosser CF: Local residents as mediators between middle-class professional workers and lower-class clients. Soc Serv Rev 40:56-63, 1966.
95. Hoff W: The importance of training for effective performance. Public Health Rep 85:760-766, 1970.
96. D'Onofrio CN: Aides—pain or panacea. Public Health Rep 85:788-801, 1970.
97. Simons HW, et al: Similarity, credibility, and attitude change: A review and a theory. Psychol Bull 73:1-16, 1970.

98. Berscheid E: Opinion change and communicator-communicatee dissimilarity. J Pers Soc Psychol 4:670-681, 1966.
99. Janis IL, Hoffman D: Facilitating effects of daily contact between partners who make a decision to cut down on smoking. J Pers Soc Psychol 17:25-35, 1970.
100. Mendelsohn GA, Rankin NO: Client-counselor compatibility and the outcome of counseling. J Abnorm Psychol 74:157-163, 1969.
101. Mazen FR: Communicator Similarity and Acceptance of Recommendations: A Study of Attitudes and Behavior of Pregnant Women. Doctoral dissertation. New Haven, Yale University, 1967.
102. Fisher AA: The Characteristics of Family Planning Opinion Leaders and Their Influence on the Contraceptive Behavior of Others. Doctor of Science thesis. Baltimore, Johns Hopkins University School of Hygiene and Public Health, 1974.
103. Rosenberg SG: Patient education leads to better care for heart patients. HSMHA Health Rep 86:793-802, 1971.
104. Green LW, Werlin SH, Schauffler HH, et al: Research and demonstration issues in self-care: Measuring the decline of medicocentrism. Health Educ Monogr 5(2):161-184, 1977.
105. Young MAC: Review of research and studies related to health education communication: Methods and materials. Health Educ Monogr 1(25):63-70, 1967.
106. Aiken L, Aiken JL: A systematic approach to the evaluation of interpersonal relationships. Am J Nurs 73:863-867, 1973.
107. Alpert JJ: Slave patients and free physicians. N Engl J Med 284:667-668, 1971.
108. Anderson FP, Rowe DS, Dean VC, et al: An approach to the problem of noncompliance in a pediatric outpatient clinic. Am J Dis Child 122:142-143, 1971.
109. Bernstein L, Headlee R, Jackson B: Changes in acceptance of others resulting from a course in the physician-patient relationship. Br J Med Educ 4:65-66, 1970.
110. Blackwell B: Drug therapy—patient compliance. N Engl J Med 289:249-252, 1973.
111. Boyle CM: Differences between patients' and doctors' interpretation of some common medical terms. Br Med J 2:286-289, 1970.
112. Caldwell JR, Cobb S, Dowling MD, et al: The dropout problem in antihypertensive treatment: A pilot study of social and emotional factors influencing a patient's ability to follow antihypertensive treatment. J Chronic Dis 22:579-592, 1970.
113. Charney E: Patient-doctor communication. Pediatr Clin North Am 19:263-279, 1972.
114. Dodge JS: What patients should be told: Patients' and nurses' beliefs. Am J Nurs 72:1852-1854, 1972.
115. Heagarty MC, Robertson LS: Slave doctors and free doctors—a participant observer study of the physician-patient relation in a low-income comprehensive care program. N Engl J Med 284:636-641, 1971.
116. Hulka BS, Kupper LL, Cassell JC, et al: A method for measuring physician's awareness of patients' concerns. HSMHA Health Rep 86:741-751, 1971.
117. Inui TS, Yourtee EL, Williamson JW: Improved outcomes in hypertension after physician tutorials: A controlled trial. Ann Intern Med 84:646-651, 1976.
118. Griffiths W, Knutson AL: The role of mass media in public health. Am J Public Health 50:515-523, 1960.
119. Swinehart J: Voluntary exposure to health communications. Am J Public Health 58:1265-1275, 1968.
120. Hingson R: Obtaining optimal attendance at mass immunization programs. Health Serv Rep 89:53-64, 1974.
121. Lin N: Information flow, influence flow, and the decision-making process. Journalism Q 47:33-40, 61, 1971.
122. Lin N: The Study of Human Communication. Indianapolis, The Bobbs-Merrill Co, Inc, 1973.
123. Lin N, Burt RS: Roles of Differential Information Channels in the Process of Innovation Diffusion. Albany, International Center for Research State University of New York, Monograph No. 001, 1973.

124. Greenberg BS, Dervin B: Use of Mass Media by the Urban Poor. New York, Praeger Press, 1970.
125. Kalba K: Communicable medicine: Cable television and health services. Socio-Econ Plann Sci 7:611-632, 1973.
126. Manoff RK: Potential uses of mass media in nutrition programs. J Nutr Educ 5 12, 3-4, 1973.
127. Bauman HE: What does the consumer know about nutrition? JAMA 225.61-62, 1973.
128. Garner JL, Mulcahy MJ: MIC project: Another kind of dietetics outreach program. Hospitals 47:96-97, 1973.
129. Goldberg JP: Some community nutrition services in a Boston program. J Am Diet Assoc 62:537-539, 1973.
130. Neumann AK, Neumann CG, Ifekwunigwe AE: Evaluation of small-scale nutrition programs. Am J Clin Nutr 26:445-452, 1973.
131. Nichaman M, Collins G: Nutrition programs in state health agencies. Nutr Rev 32:65-69, 1974.
132. Zimmerman RR, Munro N: Changing Head Start mother's food attitudes and practices. J Nutr Educ 4:66-68, 1972.
133. Marketing Economics Division: Impact of the Expanded Food and Nutrition Education Program on Low-Income Families: An In-Depth Analysis. Washington DC, US Department of Agriculture, Economic Research Service, Report no 220, 1972.
134. Stamler J: High Blood Pressure in the U.S. An Overview of the Problem and the Challenge. Proceedings of the National Conference on High Blood Pressure Education. DHEW Publication no (NIH) 73-486. Rockville, National Heart and Lung Institute, 1973.
135. Bartlett MH, Johnston A, Meyer TC: Dial access library - patient information service. N Engl J Med 288;994-998, 1973.
136. Skiff A: Experiences with methods for patient teaching from a Public Health Service hospital. Health Educ Monogr 2:48-53, 1974.
137. Simonds SK: Health education as social policy. Health Educ Monogr 2(suppl):1-10, 1974
138. Somers AR: Conference recommendations. The National Conference on Preventive Medicine: Summaries and Recommendations. In Preventive Medicine USA. New York, Prodist, 1977.
139. Miller JE: An indicator to aid management in assigning program priorities. Public Health Rep 85:725-730, 1970.
140. American Public Health Association: Proposed Preventive Benefits to be Covered on a First Dollar Basis Under National Health Insurance. Washington, DC, July 16, 1974, processed.

EDUCATION SANITAIRE, 1983-4 : 7-10, 1981

Analyse critique des perspectives canadiennes en promotion de la santé

Document de travail présenté par Ronald Labonté et Susan Penfold, M.D., F.R.C.P. (C)*

Introduction

La promotion de la santé est devenue en très peu de temps le concept d'avant-garde en matière de programmes de santé, englobant même le domaine de l'éducation sanitaire, qui avait vu le jour bien avant. L'objectif visé: "prévenir les maladies tant physiques que mentales par la promotion de l'hygiène personnelle et l'élimination des risques auxquels l'individu s'expose délibérément".[1] Bien que les spécialistes de la santé se plaignent souvent de ce que les programmes de promotion de la santé sont reléqués au second plan dans les prévisions budgétaires, il reste que ces programmes se sont beaucoup fait remarquer aux niveaux tant provincial que fédéral, au cours des cinq dernières années. En tentant de déterminer quels devraient être les objectifs prioritaires des programmes de promotion de la santé au cours des années 1980, nous soulèverons et analyserons dans notre article deux questions très simples, à savoir: d'une part, si les programmes actuels en la matière sont efficaces, et d'autre part (et il s'agit là d'une question plus pertinente pour les fins de notre thèse), si les modes de vie des Canadiens sont essentiellement à l'origine du piètre état de santé d'un grand nombre de ces derniers.

Les programmes actuels de promotion de la santé sont-ils efficaces?

La réponse est "oui", mais avec certaines réserves. Si l'on se reporte aux définitions actuelles en matière de promotion de la santé, la plupart des programmes arrivent parfois à avoir un impact, du moins sur certaines personnes. Voyons de plus près certains exemples:

– Les campagnes anti-tabac menées ces dix dernières années ont eu des répercussions très certaines sur la consommation de tabac. S'il est vrai que le nombre d'adolescents (de jeunes filles surtout) qui commencent à fumer ne cesse de croître, il est également vrai que le nombre de fumeurs adultes ayant trouvé la volonté nécessaire pour abandonner la cigarette est encore plus impressionnant.[2]

– Les programmes de promotion et d'éducation axés sur une saine alimentation ont provoqué une prise de conscience manifeste chez les consommateurs. Il suffit pour s'en convaincre de voir à quel point ces messieurs de *Madison Avenue*, s'évertuent à proclamer la supériorité de leurs produits par rapport à ceux de leurs concurrents du point de vue de la valeur nutritive.

– Une énorme campagne de promotion de la santé menée en Colombie-Britannique et axée sur les maladies vénériennes a été couronnée d'un succès tel que, malgré la confusion qui règne dans l'esprit des gens, maintenant que la maladie n'est plus désignée sous le nom de "maladie vénérienne" mais par l'expression "maladie transmise par voies sexuelles", le nombre de malades qui se sont fait soigner a atteint un chiffre record. La campagne a dû être mise en veilleuse jusqu'à ce qu'on trouve des installations plus grandes et davantage de ressources humaines pour répondre aux besoins.

– Les programmes de promotion de la bonne forme physique ont aussi donné des résultats intéressants. Ils ont réussi à inciter les constructeurs de luxueux immeubles en copropriété à installer des courts de racquetball et des salles de gymnastique. Autre constatation: les magazines de sport occupent maintenant une place de choix dans les kiosques à journaux. De toutes parts, les annonces de programmes de conditionnement physique foisonnent. Les magasins d'articles de sport ont de moins en moins l'aspect d'une salle de casiers et rappellent de plus en plus la boutique en haute mode.[3]

Nous admettons avoir omis de citer toute une gamme d'études pour étayer notre allégation selon laquelle certains programmes de promotion de la santé méritent tous nos éloges. Sans vouloir nier la valeur de telles recherches et études d'évaluation, nous devons cependant admettre que *toute* information sur les moyens à prendre pour être en meilleure santé, surtout si elle fuse de partout – autobus, journaux, radio, télévision, affiches, brochures, voire chèques d'allocations familiales – doit inévitablement influer *dans une certaine mesure* sur le comportement des gens. Il ne faut pas oublier que notre comportement est largement façonné par le type et la quantité d'information qui nous est transmise.

S'il est un fait que nous avons fait des progrès, aussi menus soient-ils, pourquoi alors avoir convoqué une autre conférence sur la promotion de la santé dans l'idée d'examiner les objectifs prioritaires? Pourquoi, aussi, est-ce que nous, éducateurs sanitaires, remettons sans cesse en question notre discipline, nos objectifs, nos méthodes, bref, notre *raison d'être*?[4] La raison est fort simple: parce que, comme nous sommes tous prêts à l'admettre, il se trouve que les programmes de promotion de la santé n'ont aucun effet sur la plupart des gens, la plupart du temps, et que par conséquent, nos nobles objectifs sont loin d'être atteints. Cette constatation est difficile à accepter pour ceux dont le gagne-pain, le statut professionnel et la satisfaction personnelle dépendent des résultats concrets des programmes qu'ils ont lancés et mis en oeuvre.

Pourtant, il n'est pas *impossible* d'adopter des mesures vraiment efficaces qui auront un impact réel et considérable sur la santé de *tous* les Canadiens. Pour cela, il faut entre autres mettre l'accent sur le contexte social dans lequel les individus choisissent leurs modes de vie, élément dont font abstraction la plupart des programmes actuels de promotion de la santé.

Le principe actuel: C'est la victime qui a tort!

La principale raison pour laquelle les programmes de promotion de la santé sont loin de produire les effets bénéfiques escomptés (amélioration des normes de santé, réduction du budget des soins,[5] plus grande sensibilisation de la population à la question de la santé), c'est qu'en dépit du fait qu'ils atteignent la masse[6] en ayant recours à tous les médias d'information, ce ne sont que des individus isolés qui sont visés. Par ailleurs, en faisant de la santé et de la maladie une question qui dépend d'abord et avant tout du mode de vie de chaque individu, ces programmes ont faussé le débat en passant complètement sous silence une foule de facteurs autrement plus déterminants sur l'état de santé, à savoir: la pauvreté l'inégalité des sexes, le racisme, les risques professionnels et la

*Le présent article est la version revue d'une étude présentée par le docteur Penfold (anciennement Stephenson) lors de la *Première conférence canadienne sur la promotion de la santé: Mandats et Actions pour les années 80*, qui s'est tenue à Ottawa les 19 et 20 janvier 1981. Mme Sue Penfold est professeur agrégé à la division de la psychiatrie infantile de l'Université de la Colombie-Britannique, à Vancouver. M. Ron Labonté, qui est également de Vancouver, est conseiller en éducation sanitaire, auprès du ministère de la santé de la Colombie-Britannique. Les délégués à la conférence n'auront que la nouvelle version est plus riche en données que l'ancienne, mais de l'avis des auteurs, le fond de la thèse demeure inchangé.

dégradation de l'environnement par la pollution industrielle. N'est-ce pas là une constatation éloquente du point de vue politique? Nous allons sous peu décrire brièvement les contre-coups de ces "agents pathogènes d'origine sociale". Pour l'instant, nous nous appliquerons à montrer comment nous parvenons si bien maintenant à rejeter les torts sur la victime.

En 1974, lorsque Marc Lalonde faisait paraître son fameux document intitulé: *Nouvelle perspective de la santé des Canadiens*,[7] la promotion de la santé se vit doter d'une brillante philosophie et d'une méthodologie nouvelle – la "conception globale de la santé" – selon lesquelles la santé était considérée comme étant tributaire de facteurs tels que les habitudes de vie, l'environnement et l'organisation des soins de santé. Malheureusement, le savant concept ne fut jamais mis en pratique, et par la suite, aussi bien M. Lalonde que les spécialistes de l'éducation sanitaire et de la promotion de la santé ont eu tendance à considérer les divers facteurs qui influent sur la santé comme des phénomènes isolés, entre lesquels il n'y a pour ainsi dire aucun rapport. Le rapport de M. Lalonde a eu pour effet de donner l'impression que la santé est surtout fonction du mode de vie choisi par l'individu, et l'état de santé déplorable des Canadiens, à l'heure actuelle, est attribué à toute une litanie de risques auxquels l'individu s'expose délibérément.[8] En effet, on a tendance à penser que tout irait pour le mieux dans le meilleur des mondes, si seulement les éducateurs sanitaires parvenaient à convaincre les gens de cesser de fumer, de consommer moins d'alcool, de s'alimenter sainement, de faire un peu d'exercice, de s'exercer à combattre le stress, de cesser d'ingurgiter des pilules, de boucler leur ceinture de sécurité, d'être prudents au volant, et de mettre de l'ordre dans leur vie sexuelle.

Encore une fois, lorsqu'on insiste tant sur la responsabilité individuelle dans les programmes de promotion de la santé, on contribue à jeter le voile sur le contexte social dans lequel s'exercent les choix personnels. Prenons le cas d'une annonce de *Dialogue sur l'alcool* parue récemment à la télévision, mettant en scène une femme dans un bar qui décrit une entrevue qu'elle vient d'avoir avec un employeur, ses dix années de métier de femme au foyer qu'elle a offertes à ce dernier, et le verre qu'elle a dû prendre avant l'entrevue pour combattre son angoisse. Il est vrai qu'il est important de mettre les gens en garde contre le danger d'avoir recours à l'alcool pour faire face à une situation de stress, mais il reste que la situation évoquée dans cette annonce cache une réalité encore plus sombre. En effet, le sexisme de la société canadienne fait que la femme doit toujours assumer l'entière responsabilité d'élever et d'éduquer ses enfants, qu'elle n'a pas toutes les possibilités de poursuivre des études et de faire carrière qui sont données à l'homme, et que son salaire, quand elle réussit à se trouver un emploi, est en moyenne la moitié de celui de l'homme.[9] Présenter la femme qui surmonte le stress par l'alcool sans pour autant adopter une attitude critique face à l'infériorité de la condition sociale de la femme équivaut à tenir les femmes individuellement responsables de l'inégalité sociale même qui de toute évidence les pousse dans une certaine mesure à boire.

Ainsi, subtilement, souvent involontairement, mais toujours de façon désobligeante, les responsables de programmes de promotion de la santé (où c'est la victime qui est incriminée) font peser uniquement sur l'individu le poids du piètre état de santé ou des mauvaises habitudes des gens. Si l'on pousse plus loin ce raisonnement implicite (sur lequel, ne nous le cachons pas, reposent la plupart des programmes de promotion de la santé), l'on doit forcément en déduire qu'en fait les gens *optent* pour un mauvais état de santé. Dans un même ordre d'idées, on peut être porté à conclure que sans l'intervention des spécialistes et les nombreuses initiatives (promotion et éducation) destinées à réconforter, à prendre en main, à motiver, et en quelque sorte, à "changer" ces individus qui se détruisent sans réagir, les Canadiens choisiraient de plus en plus la maladie plutôt que la santé. L'individu est donc perçu comme "l'anti-héros", qui agit en dépit du bon sens et de la nature holistique de l'homme, et qui à certains égards est tout simplement irresponsable. Notre société, qui perpétue les inégalités fondées sur le sexe et les classes sociales, qui tolère les risques que peuvent représenter pour la santé le milieu de travail et l'environnement, ainsi que le stress très répandu, provoqué par ces problèmes d'origine sociale, fait figure de "meilleur des mondes possible" et est dégagée de toute responsabilité face à la myriade de problèmes de santé auxquels nous devons maintenant faire face.

Certains trouveront peut-être notre accusation un peu hardie. Qu'il leur suffise cependant d'examiner certaines des contradictions inhérentes aux programmes de promotion de la santé, la plus flagrante (et par conséquent, la moins aberrante) étant celle qui a trait à l'alcool et au tabac. On constatera à cet égard que d'un côté, le gouvernement accorde des subventions aux industries et tire un revenu des taxes appliquées à ces produits, tandis que de l'autre côté, il finance des programmes qui visent à réduire la consommation de ces mêmes produits. Ces deux réalités sont soigneusement masquées dans un contexte économique qui ne remet jamais en question le "droit" qu'a un secteur de l'industrie de tirer des profits de produits qui entraînent la maladie. La crise de l'emploi qui sévit actuellement au Canada est vite invoquée par ceux qui tentent de justifier le maintien en vie d'une industrie pathogène comme celle du tabac.[9] C'est ce même type de raisonnement fallacieux que tiennent d'autres secteurs de l'industrie qui soutiennent que les avantages à tirer de la réduction des risques professionnels et de la lutte contre la pollution doivent être comparés (par qui?) aux coûts (c'est-à-dire, la baisse des profits pour le secteur concerné). Bref, d'une part, il y a une foule de programmes et de politiques de l'État qui nous incitent à la production et à la consommation, et d'autre part, il y en a d'autres qui nous disent que ce genre d'habitudes met en danger notre santé et traduit une attitude irréfléchie chez l'individu.

Depuis quelques siècles, l'interdiction proprement dite de l'alcool et du tabac est jugée incompatible avec le caractère libéral et démocratique de notre société. Nous, spécialistes de la santé, sommes tous au courant du dilemme dans lequel nous a enfermés cette situation économique et politique. Ce que nous avons parfois tendance à oublier c'est que les gens que nous essayons d'atteindre avec nos messages et dont nous essayons de changer le comportement ne sont pas plus dupes que nous, et que cette politique ambivalente pourrait amener beaucoup d'entre eux à rejeter nos avances, considérant que nous ne faisons que mettre la charrue devant les boeufs. Cela ne revient pas à dire que nous devrions nous abstenir de dire la vérité au sujet de la cigarette et de l'alcool, mais plutôt que nos "messages" devraient être davantage axés sur la critique sociale. En d'autres termes, nous devons nous montrer plus honnêtes à propos du contexte politique et économique qui entoure les industries du tabac et de l'alcool et être moins empressés de faire peser sur l'individu tout le fardeau des maladies "rattachées aux habitudes de vie."

Ce que nous avons décrit jusqu'ici est une situation qui sans doute saute aux yeux de la plupart des éducateurs sanitaires et des autres spécialistes de la santé. Ce qui est moins évident ce sont les contraintes sociales qui s'exercent sur les domaines que la plupart des gens considèrent comme la substance des programmes actuels de promotion de la santé, à savoir: l'alimentation, le conditionnement physique et la lutte contre le stress. A première vue, ces préoccupations ne semblent dissimuler aucune réalité équivoque et n'apparaissent aucunement contaminées par des contradictions d'ordre politique, ce qui explique qu'au cours des dernières années elles soient devenues la nouvelle tendance en matière de programmes de santé.[10] Or, on retrouve là aussi les mêmes fausses conceptions et une tendance à dissimuler les mêmes questions évoquées plus haut au sujet de l'alcool et de la cigarette.

L'alimentation et la nutrition

Prenons le cas d'une personne qui est attablée pour manger. Logiquement, on doit présumer qu'elle a choisi les aliments qui se trouvent dans son assiette. En fait, il y a une réalité complexe qui est sous-jacente à ce choix: l'industrie alimentaire, l'omniprésence de la publicité relative aux aliments sans valeur nutritive, la relative accessibilité des aliments sains par opposition aux produits traités, les dangers ignorés des additifs alimentaires, le phénomène social très répandu que sont devenus les restaurants-minute. Ajoutons à cela que la majorité des produits alimentaires qui font l'objet d'annonces publicitaires à la télévision ont peu de valeur nutritive, sinon aucune. La plupart du temps ils sont très riches en sucre et en d'autres produits qui pourraient nuire à la santé. Les études l'ont clairement montré: l'aptitude des gens à faire la distinction entre ce qui est vrai et ce que leur racontent les industriels en matière d'alimentation est fonction de leur niveau d'instruction et de leur revenu annuel.[11] Aussi ne doit-on pas s'étonner de constater que plus le niveau de revenu est élevé, plus la consommation des fruits et légumes est grande, et la consommation de féculents réduite. Autrement dit, il est tout simplement plus probable qu'une personne venant d'un milieu aisé ait poursuivi des études universitaires, qu'elle puisse comparer les prix des aliments et se procurer les produits chers et hors-saison, que son voisin qui est économiquement faible.

Voici une réalité qui risque d'étonner la plupart des éducateurs sanitaires: 1 657 017 enfants canadiens vivent dans des conditions économiques tellement défavorables qu'elles peuvent occasionner des problèmes de santé causés surtout par les carences nutritionnelles;[12] et plus de 27 pour cent de tous les Canadiens adultes vivent sur le seuil de pauvreté ou au-dessous de celui-ci.[13] Ce serait faire preuve de cynisme ou de myopie que de ne pas tenir compte de l'impact de la pauvreté et du manque d'instruction sur les choix des aliments.

En outre, il ne faut jamais perdre de vue que les choix alimentaires du consommateur, abstraction faite du revenu ou du niveau d'instruction de ce dernier, sont manipulés par des industriels de l'alimentation qui brassent des milliards en chiffre d'affaires, et qui cherchent non pas à promouvoir la santé et la saine alimentation, mais d'abord et avant tout à réaliser des profits. Il importe de le souligner: le traitement des aliments, dont les bienfaits nutritionnels sont loin d'être prouvés, représente le plus important secteur de l'industrie manufacturière au Canada, avec une main-d'oeuvre de plus de 200 000 travailleurs.[14] Le conglomérat de l'alimentation est un chef-d'oeuvre d'intégration. En effet, des sociétés géantes exercent un monopole sur les produits à partir de l'exploitation agricole jusque chez le détaillant. Le docteur Cornerly, ex-président de l'*American Public Health Association* n'exagérait sans doute pas lorsqu'il affirmait que "les déplorables habitudes alimentaires des Américains étaient principalement attribuables à la politique des conglomérats de l'alimentation".[15] Pauvreté, manque d'instruction, promotion fallacieuse des aliments sans valeur nutritive, sociétés monopolistiques – voilà la trame du contexte social dans lequel la plupart des Canadiens doivent exercer leurs choix alimentaires.

Le manque d'exercice

On a tendance à associer le manque d'exercice au fait que l'individu ne se livre pas à des activités physiques pendant ses heures de loisir, du moins pas suffisamment. Par conséquent, les programmes de promotion de la santé finissent souvent par s'en prendre aux "affreux paresseux" qui, après une dure journée de travail sédentaire, rentrent chez eux, s'affalent devant leur poste de télévision, n'ayant pas la force de volonté nécessaire pour aller faire du jogging ou se rendre au court de racquetball. La vérité c'est que la mauvaise forme physique de bien des Canadiens ne s'explique pas simplement par le fait que les individus ne se livrent à aucune activité physique pendant leur temps libre – c'est un phénomène propre à toutes les sociétés hautement industrialisées qui, en l'espace d'un siècle seulement, ont presque éliminé l'activité physique du vécu quotidien du travailleur.

On a constaté par exemple que parmi les ouvriers agricoles, il y a environ quatre fois moins de cas de coronaropathies que parmi les fonctionnaires sédentaires.[16] Au lieu de mettre uniquement l'accent sur l'activité physique pendant les périodes de loisir, les responsables de programmes de promotion de la santé feraient bien de préconiser l'intégration de l'activité physique dans la législation relative aux normes du travail.

Comment les choses se passent-elles actuellement? Prenons le cas de PARTICIPaction, le plus vaste programme de promotion de la santé axé sur la forme physique jamais mis en œuvre jusqu'ici. Cet organisme vend son programme aux sociétés en invoquant qu'il s'agit d'un programme qui n'empiète aucunement sur la journée de travail, qui ne nécessite aucune installation particulière, dont le volet "exercices physiques" n'est nullement organisé, et qui peut être administré par n'importe quel membre du personnel."[17]

Pourtant, étant donné que c'est l'organisation même du travail qui veut que la plupart des gens soient inactifs sur le plan physique pendant à peu près la moitié de leur temps de veille, n'est-il pas raisonnable de laisser entendre qu'il appartient en partie à cette même organisation d'offrir de nouveau aux individus la possibilité de combler cette lacune? Ainsi, les employeurs pourraient prévoir pendant les heures de travail rémunérées de tous leurs employés des séances d'exercices et des pauses-gymnastique, de sorte qu'ils "mériteraient" certains des avantages financiers dont ils bénéficieraient du fait d'avoir des employés en forme. De plus, on pourrait inclure dans les normes de construction l'installation de centres de conditionnement physique, disposition qui vaudrait dans le cas de la construction de tout nouvel immeuble à bureaux ou toute nouvelle usine.

Il faut se rendre à l'évidence: les programmes de promotion de la santé par le conditionnement physique véhiculent des conceptions propres à la classe moyenne sur la question (sa tendance à "pratiquer la politique de l'autruche", à refuser de voir les disparités sur les plans économique, sexuel et culturel parmi les Canadiens). Par exemple, ces programmes partent du principe que chacun devrait pratiquer le jogging (sans tenir compte du fait que les souliers de course coûtent cher) ou la natation (bien que les économiquement faibles n'aient pas les moyens d'utiliser les installations financées directement ou indirectement par l'État) ou encore, faire des excursions les fins de semaine (en dépit du fait qu'un grand nombre de gens, surtout les mères célibataires, ont du mal à trouver l'argent nécessaire pour s'évader de la ville, et encore plus pour s'acheter l'équipement approprié.)

Une fois de plus, l'inégalité sexuelle qui règne dans notre société fait qu'il est difficile pour les femmes de prendre part régulièrement à un programme d'exercice physique. Encore de nos jours, c'est surtout à la femme qu'il revient de prendre soin des enfants. Même lorsqu'elle travaille en dehors du foyer,[18] c'est encore elle qui doit assumer la plus grosse part de responsabilités en ce qui concerne le ménage et les enfants.[9] Et quand elle trouve du travail, c'est généralement dans les secteurs les moins bien rémunérés de l'économie, de sorte que ses pauses-repas durent à peine plus d'une demi-heure. Enfin, d'ordinaire, ce n'est pas elle qui gère le budget familial, ce qui fait que quand bien même elle le voudrait, elle aurait du mal à se payer le temps (par exemple en payant les services d'une gardienne) et l'équipement nécessaires pour pouvoir participer au moindre programme de conditionnement physique.

Lutte contre le stress

Depuis quelque temps, la lutte contre le stress et la médecine psychosomatique comptent parmi les principaux axes des programmes de promotion de la santé, sans doute parce qu'on s'est aperçu qu'il existe un lien indéniable entre le stress et les coronaropathies, lesquelles sont les principales causes de décès observés dans notre société d'abondance. Il reste que ces programmes aussi mettent l'accent sur le mode de vie de l'individu, et partent du principe qu'il y a divers choix de comportements qui s'offrent à lui. Pourtant, fait intéressant à noter, d'après une étude menée en 1971 aux États-Unis par le ministère de la Santé, de l'Éducation et du Bien-être:

> . . . le principal déterminant de la longévité est la satisfaction au travail, le deuxième en importance étant l'attitude générale de l'individu face à la vie . . . D'autres facteurs jouent également: l'alimentation, l'activité physique, les soins médicaux et le patrimoine héréditaire. Les recherches menées sur la question tendent cependant à montrer que ces facteurs n'interviennent que dans une proportion de 25 pour cent dans le cas de cardiopathies. Il semblerait par ailleurs que l'utilité du travail, les conditions de travail et d'autres facteurs sociaux peuvent, dans une large mesure, rendre compte des facteurs de risque qui, dans 75 pour cent des cas demeurent inconnus.[19]

Contrairement au stéréotype très répandu selon lequel le cadre soumis au stress serait le meilleur candidat aux maladies cardiaques (et par conséquent celui qui aurait le plus besoin de cours axés sur la lutte contre le stress et d'énormément de loisirs), on a constaté que la fréquence des coronaropathies augmente à mesure que l'on descend les échelons de la hiérarchie, de sorte que l'incidence des maladies de ce genre est quatre fois plus élevée chez les travailleurs manuels que chez les cadres.[20]

Ainsi que nous l'avons déjà soutenu, l'accès à la forme physique dépend de la situation financière, de l'appartenance sexuelle et du bagage culturel de l'individu. Pour que ce dernier puisse profiter des activités de loisir et trouver un dérivatif au stress, il faut qu'il en ait les moyens. Par voie de conséquence, le gagne-petit a bien entendu beaucoup moins facilement accès aux activités récréatives que l'homme privilégié qui est grassement rémunéré. A cela on peut ajouter divers types de stress rattachés au travail, qui déteignent sur la vie familiale, et que l'on retrouve uniquement dans les couches sociales défavorisées, à savoir: les difficultés financières, la crainte du chômage, l'insécurité devant la retraite, la mécanisation du travail, etc.

Il existe un autre facteur qui explique sans doute encore mieux cet écart constaté en ce qui concerne l'incidence des coronaropathies. Il s'agit du conditionnement psychologique auquel est soumis l'ouvrier. En effet, le travailleur qui se trouve aux échelons inférieurs de la hiérarchie doit généralement obéir à des ordres. Son travail est routinier, ne laisse pour ainsi dire aucune place à l'apport individuel autre qu'à l'application mécanique d'une technique donnée. Ce qui est encore plus démoralisant pour le travailleur c'est de savoir que fatalement il n'est indispensable que dans la mesure où il est disposé à se conformer aux ordres de ses supérieurs. Autrement dit, sa position est diamétralement opposée à celle du cadre qui dans bien des cas se voit, lui, récompensé pour les raisons mêmes qui sont à l'origine du renvoi d'un employé subalterne: le sens de l'initiative, l'autonomie dans le travail, le choix de sa propre méthode de travail, etc. Dans ces conditions, on comprend facilement pourquoi les employés subalternes pourraient être moins portés que les cadres à *choisir* des moyens de se libérer du stress. Il y a non seulement les obstacles d'ordre financier et le manque de temps qui jouent contre les gagne-petit, mais encore leur expérience au travail les conditionne à *ne pas* faire de choix personnel!

Pour une philosophie qui mettrait l'accent sur la responsabilité sociale

Il ressort des exemples précités que nous devrions nous défier de toute activité de promotion qui présente les problèmes sociaux (par exemple la santé) comme étant essentiellement individuels et appelant une solution individuelle. Soulignons-le une fois de plus, les divers types de programmes de promotion de la santé partent tous implicitement du principe que la santé et la maladie sont des phénomènes exclusivement individuels. Pour notre part, nous ne voyons pas les choses du même oeil. En ce qui nous concerne, la maladie est un phénomène vécu par l'individu, mais qui à l'origine est moins déterminé par les caractéristiques de ce dernier (mode de vie, faiblesses congénitales, etc.) que par les structures et le contexte sociaux. Ici, le terme-clé est "social". Il faudrait que dorénavant on cesse de percevoir les problèmes de santé comme étant la conséquence des risques auxquels s'expose délibérément l'individu et auxquels ce dernier doit remédier en changeant d'habitudes de vie et de comportement. Il importerait par contre de bien se rendre compte que la santé et, son revers, la maladie sont des phénomènes profondément déterminés par le contexte social. D'où la nécessité des changements sociaux par le biais de l'action collective.

L'action collective n'est ni plus ni moins que le regroupement des initiatives d'individus qui se rendent compte que leurs problèmes de santé ont pour dénominateur commun le contexte social, et que la meilleure façon de les régler est de recourir aux initiatives de groupe. En effet, cette forme d'action offre divers avantages (elle permet aux intéressés de s'encourager mutuellement, d'obtenir l'avis des uns et des autres, de s'entraider et de s'informer, d'instaurer le dialogue pour faire une analyse critique de leur démarche et d'avoir la force de frappe que confèrent les démarches de groupe) qu'on ne retrouve pas dans le cadre des programmes actuels de promotion de la santé axés sur l'effort individuel plutôt que collectif. Nous étudierons un peu plus loin de quelle façon on peut transformer l'éducation sanitaire pour faire en sorte qu'elle devienne une expérience de groupe comme celle que nous venons de décrire.

Jusqu'ici, nous avons tenté de montrer que même dans le domaine des habitudes d'hygiène qui, comme on le sait, sont par définition tout à fait personnelles, on ne peut faire abstraction de l'influence primordiale exercée par le contexte social. Il importe maintenant de pousser plus loin notre raisonnement de façon à tenir compte du vaste champ des maladies rattachées au contexte socio-économique et à l'environnement, où le "choix personnel" ne sert qu'à justifier l'attitude irresponsable des autorités publiques, des industriels, du corps social et médical.

Disparité économique

Un des auteurs, ayant tenté d'obtenir de Statistique Canada des données précises sur le rapport entre le revenu et la santé, a reçu la réponse suivante: "Nous n'avons pas les statistiques que vous cherchez."[21] Autre fait révélateur: la publication intitulée *Indicateurs sanitaires* qui regroupe les données ayant trait à la santé ne fournit aucun renseignement sur le lien qui existe entre l'état de santé et le revenu, la catégorie sociale, la profession, etc. Le fait que ce genre de renseignements ne soit pas recueilli montre bien qu'on ne considère pas le revenu et la catégorie sociale comme des variables significatives pour ce qui est de la santé, et explique pourquoi il est difficile de prouver le contraire.

Il existe cependant d'autres statistiques (la plupart d'origine américaine) qui mettent en évidence la primauté du niveau de revenu comme variable explicative en ce qui concerne l'état de santé:

— Les taux de mortalité en milieu urbain sont de 50 à 100 pour cent supérieurs dans les quartiers pauvres.[23]

— Les familles à faible revenu sont quatre fois et demie plus susceptibles de se déclarer en mauvais état de santé que les familles plus aisées.[23]
— Il existe un rapport certain entre la piètre santé de certains enfants et la pauvreté.[24,25,2]
— Les groupes à faible revenu ont des modes de vie qui les exposent davantage à la maladie que les groupes mieux nantis.[25]

Ces constatations font ressortir le fait que les "sous-utilisateurs des services de santé ont tendance à être victimes de l'inégalité sociale sous ses formes les plus diverses"[27] et que ces programmes sont conçus pour la classe moyenne.
— Il existe une corrélation positive entre le taux de mortalité infantile (souvent utilisé comme un important indicateur de l'état de santé général d'une population) et les faibles revenus.[28] La même étude a mis en évidence une corrélation entre les décès attribuables à toutes les causes et la pauvreté.
— L'hypertension (qui représente un facteur de risque dans les maladies du coeur et des reins, les accidents cérébrovasculaires et cardiovasculaires et qui peut entraîner la mort) est deux fois plus courante dans les groupes économiquement faibles.[29]
— On constate qu'il existe une corrélation positive entre la pauvreté et l'incidence accrue des maladies mentales.[30,31]
— Les taux de mortinatalité et de mortalité au Canada sont supérieurs dans les provinces où les niveaux de revenus sont plus faibles.[32]
— En Colombie-Britannique, le taux de mortinatalité est presque deux fois plus élevé chez les Indiens que chez les non Indiens. La différence la plus frappante entre ces deux groupes est celle du degré relatif de pauvreté.[33]

Inégalité des sexes

Plus les gens sont pauvres, moins ils disposent de conditions économiques suffisantes pour être en bon état de santé. La pauvreté est plus accentuée parmi les femmes que chez les hommes. Soixante pour cent de la population adulte canadienne vivant dans la pauvreté est composée de femmes.[34] La situation est encore plus désolante lorsque la femme représente l'unique source de revenu. La femme a trois fois plus de risques d'être pauvre lorsqu'elle vit seule que quand elle vit avec un homme, et les familles dont le chef est une femme sont cinq fois plus susceptibles de se situer dans la catégorie des économiquement faibles que celles dont le chef est de sexe masculin.[35] Le cliché selon lequel la différence entre l'assistée sociale et la femme qui ne l'est pas ne tient qu'à la présence d'un homme est malheureusement vrai.

Il existe cependant d'autres facteurs qui font que les inégalités fondées sur le sexe représentent un grave risque pour la santé:
— Près de la moitié de toutes les Américaines, à un moment ou à un autre, se voient prescrire des psychotropes — soit plus de deux fois le nombre d'hommes qui utilisent ces médicaments.[36]
— Soixante-dix pour cent des utilisateurs de tranquillisants mineurs et 72 p. cent des utilisateurs d'antidépresseurs sont des femmes.[36]
— Au Canada, plus des deux tiers des psychotropes sont prescrits aux femmes, et 15 à 20 p. cent de ces dernières en absorbent à un moment donné.[37]
— Pendant la période comprise entre 1971 et 1977, les Canadiennes représentaient respectivement 65, 69 et 70 p. cent des personnes souffrant d'effets indésirables par suite de l'absorption d'analgésiques, d'antipyrétiques, de sédatifs, d'hypnotiques et de psychotropes.[38]

Il a été amplement démontré que le traitement particulier que le milieu médical réserve aux femmes est lié aux hypothèses sexistes concernant le rôle social de la femme, aux idées patriarcales que reçoivent les médecins pendant leur formation au sujet de la santé et des problèmes de santé de la femme, hypothèses confortées par la publicité relative aux produits pharmaceutiques, laquelle tend à "suggérer fortement que la femme qui se montre incapable de se conformer au rôle qui lui est traditionnellement dévolu, qui n'est pas à même d'assumer sa condition de femme et les responsabilités qui en découlent, doit être soignée avec des médicaments."[39]

Les inégalités sexuelles et les attitudes patriarcales sont également à l'origine des risques auxquels sont exposées les femmes dans des domaines comme la contraception. La "pilule", par exemple, a été introduite sur le marché après seulement cinq années de recherche, dont la validité, sur les plans déontologique, méthodologique et scientifique, est actuellement mise en doute.[36,2] De l'avis d'un chercheur réputé, l'utilisation persistante de la contraception orale constitue "la plus formidable expérience jamais faite avec des cobayes humains."[40]

Il y a également eu l'affaire de la thalidomide, fléau qui, selon certains, risque de resurgir avec l'utilisation du Bendectin.[41] Certains médecins et certaines cliniques de l'Amérique du Nord persistent à utiliser comme contraceptif post-coïtal le diéthylstilbestrol, produit cancérigène très connu.[41] L'obstétrique et la gynécologie demeurent des disciplines où la prééminence masculine est très nette,[42] et l'intervention des sages-femmes est toujours perçue par certains professionnels de la santé comme pouvant compromettre la santé des mères,[43] et ce, en dépit de l'ample documentation qui prouve le contraire.[44,45,46] Compte tenu de l'ampleur des risques que font peser sur les femmes les valeurs et attitudes conditionnées des membres masculins du corps médical, un auteur a été amené à recommander que "l'on interdise aux étudiants en médecine de sexe masculin l'accès à l'obstétrique et à la gynécologie."[47]

Risques professionnels

M. Arnold Miller, qui prenait la parole en qualité de président de l'*United Mine Workers of America*, a déclaré une fois que "la maladie professionnelle infeste notre pays comme la peste."[48] Il suffit, pour s'en convaincre, de considérer ce qui suit:
— En 1977, 1 036 000 travailleurs canadiens ont subi des accidents sur le lieu du travail, accidents ayant entraîné plus de 900 décès tandis que la morbidité a été établie à plus de 500 000 cas.[49] Cette estimation est jugée prudente compte tenu du nombre considérable de "risques professionnels dont les manifestations ne sont pas aussi soudaines, mais qui entraînent tout autant l'incapacité et la mort (et parce que l'estimation repose uniquement sur les accidents qui sont couverts par la loi et qui sont signalés à la commission des accidents du travail)".[49]
— Aux États-Unis, plus de 128 000 travailleurs meurent chaque année à la suite d'accidents liés au travail ou de maladies d'origine professionnelle.[50,51]
— Plus de 50 000 décès signalés chaque année en Amérique du Nord sont attribués à des cancers reliés à l'amiante.[52]
— Chaque année, on découvre des milliers de nouveaux produits chimiques industriels. Or, rares sont ceux qui font l'objet d'une analyse sérieuse destinée à déceler les risques auxquels peuvent s'exposer les travailleurs appelés à manipuler ces produits. Encore plus rares sont ceux dont l'usage est restreint d'une façon quelconque. Et même lorsque ce genre de mesures existent, elles n'ont la plupart du temps aucunement servi à protéger convenablement les travailleurs contre les dangers que présentent ces produits pour la santé.[53]

Souvent, le travailleur se trouve impuissant à changer tout seul ces conditions de travail. La plupart ne sont toujours pas syndiqués et risquent d'être renvoyés s'ils se plaignent à ce sujet, surtout lorsque leur revendication implique une modification en profondeur des conditions de travail.

Des dizaines de milliers de travailleurs canadiens travaillant à l'usine doivent supporter le bruit assourdissant des machines, respirer de l'air chargé de polluants pouvant avoir un effet toxique, tout en manipulant des substances sur lesquelles on les renseigne à peine, et à propos desquelles on

est encore mous bien renseigné. Nous, spécialistes de l'éducation sanitaire, ferons bien de nous interroger sur l'impact que pourraient avoir nos messages légèrement moralisateurs et édulcorés sur les dangers de la cigarette et l'alcool ainsi que nos allègres exhortations à l'activité physique et à l'adoption d'un mode de vie sain, lorsqu'on songe au vécu quotidien de la plupart des ouvriers.

Risques associés à l'environnement

Le docteur Samuel Epstein, qui a été au service de l'*U.S. Environmental Protection Agency*, vient de faire paraître une étude très détaillée de 628 pages sur le cancer aux États-Unis.[54] L'une des conclusions auxquelles il aboutit est que les facteurs environnementaux sont à l'origine d'une bonne partie (70 à 90 p. cent) de tous les cancers et que l'incidence accrue de ce genre de maladies, récemment constatée, est attribuable à la pollution industrielle.[55] En outre, il fait ressortir très nettement le lien entre le tabac et le cancer du poumon, mais à la différence de la plupart des programmes de promotion de la santé, il n'hésite pas à s'attaquer à une question beaucoup plus pertinente et gênante: celle des risques de cancer auxquels sont exposés les ouvriers sur le lieu de travail.

Nous ne cherchons pas ici à nier la valeur des programmes de promotion de la santé qui attirent l'attention sur les effets néfastes de la cigarette. Cependant, nous tenons à souligner qu'ils ne font que toucher à la pointe de cet iceberg qu'est le cancer, et que ce silence que les responsables de ces programmes observent sur les graves risques que représentent pour la santé les polluants atmosphériques (ce qui n'a manifestement rien à voir avec le choix personnel du consommateur) doit être contesté. Elles ne cessent de s'accumuler, les preuves selon lesquelles, avant la fin de notre millénaire, la pollution de l'environnement par les déchets industriels sera sans doute la "maladie" la plus dangereuse que notre planète ait jamais connue.

Notre état de santé est compromis par d'autres facteurs reliés à l'environnement, à savoir:

— la pollution de l'eau de boisson, et tous les risques qui y sont associés mais qui ne sont pas encore déterminés;
— la pollution de l'air, qui peut provoquer diverses maladies respiratoires;
— l'amenuisement de la couche d'ozone, qui accroît les risques de cancer de la peau;
— les pluies acides qui compromettent les récoltes, tant au Canada qu'aux États-Unis;
— la pollution des océans, qui fait diminuer la faune aquatique (poissons et crustacés);
— les engrais cancérigènes et les additifs alimentaires qui polluent toute la production alimentaire.

Ce sont tous là des facteurs déterminants du point de vue de la santé. A l'instar de tous les agents pathogènes d'origine sociale, ils sont profondément ancrés dans la structure même de notre société industrielle, et l'on ne peut espérer y remédier si l'ensemble du corps social ne s'engage pas à apporter des changements radicaux à notre façon de fonctionner en tant que société. Selon nous, il est essentiel qu'au cours des années 1980 la promotion de la santé commence à tenir compte du phénomène très réel des agents pathogènes d'origine sociale plutôt que de persister à ne s'en tenir qu'à une analyse qui fait de la santé et de la maladie une question strictement individuelle. La condition de la femme, de la personne démunie, de l'ouvrier travaillant à l'usine, de celui qui vit dans une ville polluée ou à proximité d'une usine ou d'un cours d'eau pollué, l'expérience quotidienne de tous ces gens-là: voilà la réalité socio-politique de la santé et de la maladie. En persistant à taire cette réalité, nous nous condamnons, dans le meilleur des cas, à l'anachronisme, voués à tenir encore plus de conférences où nous continuerons de nous poser la sempiternelle question: "Pourquoi nos programmes n'ont-ils pas de *véritable* impact?" Au pire, nous nous ferons les complices de ce qu'on pourrait qualifier d'une des plus vastes tromperies collectives de l'histoire.

Éveiller l'esprit critique par l'éducation: un défi à relever

Les accusations que nous venons de porter sont graves et sévères. Toutefois, nous pensons vraiment qu'elles s'imposent. Par ailleurs, nous estimons que se contenter de fustiger le milieu de la santé ne vaut guère mieux que tenir les Canadiens personnellement responsables de leurs piètres habitudes de vie. C'est pourquoi, en guise de conclusion, nous donnerons aux éducateurs sanitaires et aux promoteurs de la santé une idée de la façon dont ils peuvent procéder pour relever le défi que nous leur avons lancé.

D'abord et avant tout, il s'agit de redéfinir le concept de la responsabilité individuelle, de l'étendre de façon à ce que les individus se sentent une obligation non seulement à l'égard de leurs propres habitudes en matière de santé, mais aussi face à des comportements susceptibles de compromettre la santé des autres. Ainsi, le conseil de direction et les actionnaires d'une usine polluante doivent assumer la responsabilité des troubles respiratoires des gens qui vivent sous le vent des cheminées, au même titre que les fumeurs que l'on rend responsables de polluer leurs propres poumons. Dans une même optique, nous devons, dans nos messages, faire le lien entre les choix personnels que l'individu peut réalistement faire en matière de santé et les choix sociaux devant être assumés sur une base collective. Ce changement d'orientation implique la démarche suivante:

> a) faire le point sur les divers aspects de la réalité (par exemple, les problèmes de santé, de logement, etc.); b) remonter à la source des problèmes; c) tenir compte des implications et des conséquences de ces questions, et enfin, d) mettre au point un plan d'action en vue de s'attaquer aux problèmes cernés collectivement.[56]

Cette réorientation des programmes ne peut qu'éveiller le public et le rendre en mesure de jeter un regard critique sur le milieu dans lequel il évolue, et d'avoir une vision globale et non pas fragmentaire des choses. Les programmes actuels de promotion de la santé insistent sur la dimension très restreinte du comportement individuel en matière de santé, dissociant cette réalité de son contexte social. De même, ils assimilent les problèmes sociaux tels que la pauvreté et la pollution aux questions politiques d'actualité, sans rapport aucun avec la santé des individus. Au cours des années 1980, les programmes de promotion de la santé devront tenter de rapprocher ces deux pôles de façon à ce que la responsabilité individuelle devienne synonyme de responsabilité sociale. Autrement dit, ils doivent comprendre d'une part des initiatives collectives qui viseraient à modifier les aspects de notre environnement social qui compromettent notre santé, et, d'autre part, des initiatives individuelles dont l'objet est de changer les mauvaises habitudes personnelles.

Le rôle de l'éducateur sanitaire et du promoteur de la santé devra lui aussi être totalement repensé. Nous, spécialistes de la santé, devrons descendre de notre piédestal d'où nous essayons de déterminer quels sont les meilleurs moyens de promouvoir la santé. Travailler auprès des gens nous apprend beaucoup plus que consommer passivement des messages véhiculés par les médias ou écouter les discours savants des spécialistes d'un domaine. Pour que les activités d'éducation et de promotion soient efficaces, il faut qu'elles ne soient pas fondées sur les rapports hiérarchiques. En d'autres termes, nous devons cesser de travailler *pour* les gens (cesser de chercher des moyens de les "motiver"), et nous mettre à travailler *avec* eux. La lutte qu'ils mènent pour accéder à un mode de vie sain est aussi notre combat. Il est de notre devoir de *répondre* à leurs besoins, et non pas de les *déterminer*.

En raison de la nature même du métier que nous exerçons, nous avons le privilège d'avoir acquis, par nos études et par notre expérience professionnelle, des connaissances dans le domaine de la santé que possèdent très peu de Canadiens. C'est un privilège que nous devons partager de façon à amener les gens à redécouvrir leur pouvoir individuel et leur force de frappe collective, et à reprendre en main leur propre santé en étant conscients que c'est le contexte social, économique et le sexisme qui sont en grande partie responsables de notre piètre état de santé.

Nous n'entendons pas par là qu'il faille abandonner nos efforts actuels. Là où nous voulons en venir, par contre, c'est qu'il importe de changer nos priorités de façon à ce que les risques d'origine sociale et les risques attribués aux habitudes de vie soient placés sur un même plan.

Plusieurs des points que nous avons soulevés dans notre article sont des questions sur lesquelles on ne s'est pas penché jusqu'ici, mais qui doivent être posées.

Si nous ne perdons pas de vue le fait que les gens *souhaitent* être en santé plutôt que malportants, que ce désir est contrarié par un ensemble complexe de structures et de rapports de pouvoir dans la société, et que notre rôle doit consister à oeuvrer *auprès* des gens, à les aider à résoudre les paradoxes sociaux et personnels qui les touchent de près, alors, l'éducation et la promotion sanitaires peuvent encore être appelées à jouer un rôle important et efficace, et nous pourrons peut-être faire des bonds en avant dans le sens de la réalisation de notre objectif: veiller à l'état de santé de tous les Canadiens.

RÉFÉRENCES ET RENVOIS

1. Éconcés de politique (Promotion de la santé) de l'Association canadienne d'hygiène publique, *Canadian Journal of Public Health*, vol. 65, no 2, mars/avril 1974, p. 140.

2. Samuel Epstein, *The Politics of Cancer*, Anchor Books, New York, 1979, pp. 152-154. Comme on pouvait s'y attendre, les industriels du tabac sont en train de créer de nouveaux marchés dans le pays du Tiers-Monde en fabriquant et en vendant des cigarettes qui sont trois à cinq fois plus fortes que les marques qui se trouvent actuellement sur le marché nord-américain. Ainsi, même si tous les Canadiens et tous les Américains devenaient des non-fumeurs, l'industrie du tabac ne connaîtrait aucun ralentissement d'activité et la question morale de savoir si l'on devrait permettre à une industrie nettement pathogène de poursuivre ses activités resterait posée.

3. L'exploitation par les hommes d'affaires de ce nouvel engouement pour le conditionnement physique témoigne d'un problème qui est à la base des programmes de promotion de la santé; à savoir une tendance, propre à la classe moyenne, à tenir pour acquis que tout le monde nage plus ou moins dans l'opulence. Les programmes de promotion axés sur le conditionnement physique sont conçus par des professionnels et sont malheureusement suivis surtout par des professionnels.

4. Il est important de se demander pourquoi la responsabilité du maintien et de la promotion de la santé qui appartenait pendant des siècles aux mères, aux familles et à l'ensemble de la collectivité, a été usurpée par des spécialistes au cours des trente dernières années.

5. Fait étonnant, la plupart des économistes considèrent que l'argent consacré aux quelques mesures de prévention n'entraîne: a aucune réduction des dépenses au titre des soins curatifs. Les auteurs de cet article perçoivent les programmes de promotion de la santé et les principes qui les sous-tendent, ainsi que les économies tant vantées, qu'ils doivent permettre de réaliser comme une mesure d'adaptation prise par les autorités publiques en pleine période de crise économique pour tenter de justifier une politique visant à maintenir, voire à diminuer, le niveau des crédits affectés à des services sociaux et de santé qui ne sont nullement rentables. En investissant des sommes minimes dans des campagnes spectaculaires qui rejettent sur l'individu la responsabilité de son mauvais état de santé, le gouvernement a trouvé un argument pour justifier les coupures budgétaires au chapitre des services de santé. Dans ce sens-là, on peut établir un parallèle entre la politique en matière de promotion de la santé et la tendance au "retour à l'essentiel," observée dans le monde de l'éducation ainsi que l'attitude, propre au secteur de l'assistance publique, consistant à perpétuer un certain paupérisme. Ce point est étudié de façon plus détaillée dans la version intégrale de notre étude, qui sera publiée dans le courant de l'année par le Bureau de la Région de l'Ouest, Direction générale de la promotion de la santé, ministère de la Santé nationale et du Bien-être social.

6. Paulo Freire, *Education for Critical Consciousness*, Continuum Publishing, New York, 1980, p. 8. Une société "massifiée" est celle où les gens ... ont été transformés par l'élite en une masse non pensante et facile à manipuler."

7. Marc Lalonde, *Nouvelle perspective de la santé des Canadiens*, Santé et Bien-être social Canada, Ottawa, avril 1974.

8. Ibid., p. 16. Le document de M. Lalonde n'est pas le seul où l'on retrouve le concept des "risques auxquels l'individu s'expose délibérément". Les États-Unis et la Grande-Bretagne ont aussi fait paraître des ouvrages analogues où les problèmes de santé sont définis dans une perspective individualiste, tout aussi restreinte. Il s'agit de: *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention*, U.S. Department of Health, Education and Welfare, Washington, 1979, et *Prevention and Health: Everybody's Business*, Department of Health and Social Security, Londres, 1976.

9. M. Meissner, et autres "No Exit for Wives: Sexual Division of Labour and the Cumulation of Household Demands", *Review of Canadian Sociology and Anthropology*, Vol. 12, 1975, pp. 424-439.

10. Ces points sont devenus les principaux thèmes de la conférence intitulée "The Shifting Medical Paradigm: From Disease Prevention to Health Promotion", co-parrainée par le bureau de la Région de l'Ouest de la Direction générale de la promotion de la santé, ministère de la Santé nationale et du Bien-être social, par le ministère de la santé de la Colombie-Britannique et par le *Vancouver City Health Department*, et qui a eu lieu les 20 et 21 mars 1980 à Vancouver.

11. Bureau des sciences de la nutrition, La nutrition et la santé au Canada, Document hors-série n° 5, Santé et bien-être social Canada, Ottawa, janvier 1979, p. 40.

12. Conseil national du bien-être social, *Les enfants pauvres*, Rapport du Conseil national du bien-être social sur les enfants pauvres au Canada, Ottawa, mars 1975, p. 41.

13. Statistique Canada, *Répartition du revenu au Canada selon la taille du revenu*, (13-207), Ottawa, 1978, Tableau 64.

14. *La nutrition et la Santé au Canada*, p. 56 (voir n° 11).

15. Vicente Navarro, *Medicine under Capitalism*, Prodist, New York, 1976, p. 127.

16. David Montgomery, "L'activité physique et les maladies coronaires", *Actualités – Prévention*, vol. 2, n° 3, automne 1980, p. 2.

17. PARTICIPaction, *"Faites et gestes"*, pochette documentaire, Toronto, 1980.

18. Statistique Canada, *Population active féminine du Canada*, (98-804E), Ottawa, 1980. Soixante-deux pour cent de toutes les femmes faisant partie de la population active sont mariées. Quarante-deux pour cent de toutes les femmes mariées travaillent.

19. U.S. Department of Health, Education and Welfare, *Work in America*, Rapport d'un groupe d'étude spécial présenté au Secrétaire d'État à la santé, à l'éducation et au bien-être social, Cambridge, Mass., M.I.T. Press, 1971.

20. Lawrence E. Hinkle, et autres, "Occupation, Education and Coronary Heart Disease," paru dans *Dynamics of Wellness*, publié par William Sechrist et Edward Hart, Wadsworth Publishing, Belmont, California, 1970.

21. Échange de lettres avec John Davis, directeur adjoint, Division de la santé, Statistique Canada, 22 août 1980.

22. Barbara Ouellet, *Indicateurs Sanitaires*, Santé et Bien-être social Canada, Ottawa, 1979.

23. U.S. Public Health Service, *Selected Vital and Health Statistics in Poverty and Non-poverty Areas of 19 Large Cities*, Série 21, n° 26, novembre 1975, p. 7.

24. R. Davie et autres, *From Birth to Seven: Report of the National Child Development Study*, Longman, Londres 1972.

25. Joint Commission on the Mental Health of Children, *Crisis in Child Development*, Harper and Row, New York, 1970.

26. Neil Holtzman, "Prevention: rhetoric and reality," *International Journal of Health Services*, Vol. 9, n° 1, 1979, p. 32.

27. Alex Scott-Samuel, "Why don't they want our health services?" *Lancet*, Vol. 1, n° 8165, 23 février 1980, pp. 412-415.

28. J.M. Chapman et A. Coulson, "Community Diagnosis: An analysis of indicators of health and disease in a metropolitan area," *International Journal of Epidemiology*, Vol. 1, 1972, pp. 75-81.

29. J. Eyer, "Hypertension as a disease of modern society," *International Journal of Health Services*, Vol. 5, N° 4, 1955, pp. 539-558.

30. R. Jay Turner, "Jobs and Schizophrenia," *Social Policy*, Vol. 8, N° 1, 1977.

31. Ramsay Liem et Joan Liem, "Social Class and Mental Illness Reconsidered," *Journal of Health and Social Behaviour*, Vol. 19, N° 2, juin 1978, pp. 139-156.

32. Statistique Canada, *La statistique et l'état civil*, vol. III (84-206), Ottawa, 1976, Tableau 4.

33. Santé et Bien-être social Canada, Direction générale des services médicaux, Rapport annuel de la direction générale des services médicaux, Région de l'Ouest, Santé et Bien-être social Canada, Ottawa, 1979.

34. Conseil national du bien-être social, *La femme et la pauvreté*, Ottawa, octobre 1979, p. 7.

35. Ibid., p. 14.

36. Gena Corea, *The Hidden Malpractice*, Jove Publications, New York, 1979, p. 92.

37. Susan Stephenson et Gillian Walker, *"Psychotropic Drugs and Women"*, *Feminist Quarterly*, vol 2, n° 1, printemps de 1980, pp. 20-38.

38. Statistique Canada, *Perspectives Canada III*, Ottawa 1980, Tableau 5.17., p. 58.

Voir "Références et renvois", à la page 14

"Références", suite à la page 10 .

39. Robert Seidenberg, "Images of Health, Illness and Women in Drug Advertising", Journal of Drug Issues, vol. 4, n° 2, été de 1974, pp. 264-273.

40. Samuel Epstein, *The Politics of Cancer*, (voir n° 2) p. 258.

41. Mark Dowie et Carolyn Marshall, "The Benedictin Cover-up", *Mother Jones*, novembre 1980 .

42. Diana Scully, *Men Who Control Women's Health*, Houghton Mifflin, Boston, 1980.

43. A.F. Mandeville, "Parapros threaten care", *B.C. Medical Journal*, Vol. 22, N° 9, septembre 1980, p. 397. M. Mandeville est l'actuel président de la *B.C. Medical Association*.

44. Gena Corea, *The Hidden Malpractice*, (Voir N° 36) chapitres 11 et 12.

45. John Hopkins University, "Traditional Midwives and Family Planning," *Population Reports*, Série J, N° 22, Vol. 8, N° 3, mai 1980, pp. J439-J488.

46. Barbara Ehrenreich et Deidre English, *Witches, Midwives and Nurses*, Billings and Sons, Londres, 1973.

47. Barbara Seaman, citée dans l'ouvrage de Gena Corea, *The Hidden Malpractice* (Voir N° 36) p. 14.

48. Arnold Miller, "The Wages of Neglect: Death and Disease in the American Work Place," *American Journal of Public Health*, Vol. 65 N° 11, novembre 1975, pp. 1217-1220.

49. David Brusegard, "Santé," *Perspectives Canada III*, Statistique Canada, Ottawa, 1980, p. 48.

50. D.P. Discher, *et autres, Pilot Study for the Development of an Occupational Disease Surveillance Method*, National Institute for Occupational Safety and Health, U.S. Department of Health, Education and Welfare, U.S. Government Printing Office, Washington D.C., 1975, pp. 75-162.

51. J.B. Gordon, A. Ackman et M.L. Brooks, *Industrial Safety Statistics: A Reexamination: A Critical Report Prepared for the U.S. Department of Labour*, Praeger Publishers, New York, 1971.

52. Samuel Epstein, *The Politics of Cancer* (Voir N° 2) p. 80.

53. *Ibid.*, pp. 76-125.

54. Samuel Epstein, *The Politics of Cancer*, Anchor Books, New York, 1979.

55. *Ibid.*, p. 17.

56. Meridith Minkler et Kathleen Cox, "Creating critical consciousness in health," *International Journal of Health Services*, Vol. 10, N° 2, 1980, pp. 311-322.

Article on Health and Social Policy

HEALTH EDUCATION: CAN THE REFORMERS BE REFORMED?

E. Richard Brown and Glen Elgin Margo

Health educators have created a new professional role that emphasizes the changing of individuals rather than their social conditions. The article shows how historical roots, ideological perspectives, and structural constraints have combined to create an ambiguous, generally conservative role for the health education profession.

Epidemiological evidence is presented that contradicts many implicit notions of disease etiology that underlie health education approaches. Finally, the authors suggest an "ecological" model of health education that takes account of the multiple causes of disease and is committed to progressive social change.

Over the past 30 or 40 years, health educators, fighting for their own turf in the crowded health field, have developed several educational and social functions into a professional role.

Nevertheless, health education's role remains ambiguously defined, and the profession is still struggling for recognition and financial support. Health education has received verbal encouragement—for example, from the President's Committee on Health Education (1) and from the National Health Planning and Resources Development Act (PL 93-641), which requires that health education be included in health services planning—but little money has come its way.[1]

Professional leaders in academic and policy-making jobs define roles that often have little to do with the real functions of health educators in the field. These spokespersons for the profession conceive of health education as being both a part of the health care delivery system and a force for social change, but structural and ideological forces anchor the profession's practice more in "the system" than among the forces for change (3). In its frustrated drive for professional status and authority, health education has thus become a conservative agent in the very arena in which it claims expertise and leadership as an agent of progressive social change.

In this article, we will show how historical roots, ideological perspectives, and structural constraints have combined to create an ambiguous, generally conservative

An earlier version of this paper, written by the first author as a policy paper, received financial support from the California State Department of Health.

[1] Of the $128 billion spent in the U.S. on health in 1976, only 2–2.5 percent went to preventive health care services, and only one-half of one percent (0.5 percent) was spent on health education (2).

International Journal of Health Services, Volume 8, Number 1, 1978

© 1978, Baywood Publishing Co., Inc.

role for the health education profession. We will also show that epidemiological evidence contradicts many implicit notions of disease etiology that underlie major health education approaches. Finally, we will suggest an "ecological" model of health education, one that takes account of the multiple causes of disease and is committed to progressive social change.

HEALTH EDUCATION: CONTRADICTIONS IN THE PROFESSION

There is a delusional gap between what most health educators *do* and what the profession's norms dictate they *should* do. A similar disparity exists between norms and practice in most professions, but the marginality of health educators' functions in the medical system and their low professional status threaten to drown the profession in a sea of ambiguity. While the profession's national leadership worries about appropriate role definitions (1, 3, 4), most health educators find themselves in structures that define their roles for them. In essence, they are behavior-change technicians.

Health education has been traditionally directed at encouraging people to change their behavior to more healthful "life-styles" (5). Such programs generally take their essential content from physicians and translate that content into behavior-change strategies—compliance approaches, in particular. If physicians or epidemiologists determine that smoking is harmful to health, health educators apply their skills to educating people to the importance of stopping the habit, and then develop techniques to help them stop. Dorothy Nyswander, an early leader in the profession, articulated this perspective nearly 30 years ago (6):

> Health education is a process of growth in an individual by means of which he alters his behavior or changes his attitudes toward health practices as a result of new experiences he has had.

Health educators usually trace their professional beginnings to the school health movement (7). By 1925, major health programs were under way in big city school districts; included in these programs were provisions for school nurses, hygiene education, and physical examinations for students. Over the next decade, the programs were emulated throughout the country.

However, the functions and perspectives of health educators predate the school health movement. They were developed from about 1870 to 1930 by physicians, the emerging social work professions, and the broader charity organization movement (8). Physician control of the whole health care system increased as doctors were put in charge of newly created public health boards (endowed with police powers), and of hospitals and outpatient clinics which were rapidly multiplying throughout the country. They defined the objectives and the preferred methods for the entire health system.

The medical model became *the* paradigm and perspective through which all social and health workers were trained to see disease and social problems. The clinical medical model centered on professional intervention in individuals' lives to relieve or correct physical, emotional, or social "maladaptation." Even in public health, most practitioners simply saw community-wide problems as maladaptation of groups to

society, correctable through professional intervention in the community. Seldom did public health professionals consider social conditions and social structure as the primary causes of social and health problems. Only when touched by growing political movements, for example, during the Progressive Era, the Great Depression, and the more recent civil rights movements, has the public health movement viewed advocates of major social change as more than a fringe to be tolerated.

New occupations were created out of functions that emanated from the medical model. In the first decade of this century, "medical social workers" were trained to go into poor and working-class communities to teach principles of healthful and moral living, train mothers in techniques of personal hygiene, and channel individuals to voluntary social service and relief organizations (9). As the medical care and public health enterprise grew in the 1930s and 1940s, hospitals and health agencies became more bureaucratic and more technologically based. Old roles were divided and new roles added. Medical social work became focused on casework assistance to physicians and/or patients. Public health nurses trained community people in hygienic practices, and health education was largely done by publicists hired to write leaflets, pamphlets, and press releases explaining correct health practices for schools and public consumption.

Thus, parallel to the school health movement, community health educators developed their own territories as the profession emerged from hospitals, health departments, and voluntary agencies. Community and school health educators in this period shared a perspective and methodology related to personal and community health. Their role now became focused on developing more effective methods of changing people's values and beliefs as a means of affecting their behavior. Some of the major tools employed by health educators were described by the National Health Assembly (10):

> To *propagandize* is to spread particular systematized doctrines; to *advertise* is to proclaim desirable qualities in order to arouse a desire to purchase or invest; to *campaign* is to conduct a series of operations to bring about a desired result; to work at *public relations* is to seek to win public goodwill for an organization; to *inform* is to give news or factual data; to *publicize* is to spread information designed, or "slanted," to advance special interest; to *interpret* is to bring out the real meaning; to *teach* or to *educate* is to cause, or facilitate, learning; and to *learn* is to find out about, and to gain knowledge and understanding.

By the 1940s, health educators had firmly established their role as teachers in a campaign to educate the public on practical steps to reduce the spread of disease. In addition, a second role had grown up to link this preventive role with "appropriate" use of medical services (11, 12). When a new service was funded for a target population, generally among the poor or a particular ethnic minority, health educators were often included to publicize the existence of the service, to teach people how to use it, and to overcome the reluctance of people to "take advantage" of the service.

Although some people are seen as not making enough use of services, others are accused, especially today, of "over-utilizing" them. Physicians and medical care administrators often argue that their services are over-utilized by people who run to the doctor with health problems they could deal with at home or go to the wrong facility for other problems. If physicians or administrators determine that persons

with flu symptoms should treat themselves at home instead of going to their doctor, health educators develop educational stratagems to encourage people with such symptoms to follow this policy. For the people who flock to hospital emergency rooms because they do not have a private physician, health educators may develop ways of re-directing them to clinics and private practice physicians.

ECLECTIC THEORY PRODUCES A NEW ROLE

Underlying this behavior-change function was a body of theory on which a new, higher status role was claimed. Health education theory is drawn essentially from psychology, sociology, and social psychology, corresponding to the training of the early personnel in the field. From psychoanalytic theory, early health educators drew their notions of socialization and social relations. "Civilization" was defined as the collection of patterns by which people are socialized to satisfy inborn emotional and physical needs; and "community" was a collection of socialized individuals (13). Psychoanalysis was incorporated into social work and health education during the 1920s period of economic growth. As Cox and Garvin observed, "many believed that the social environment offered so many opportunities and was otherwise so benign that any problems must be the result of individual failures" (14, p. 47). Psychoanalytic theory thus provided health educators with some justification for an individual-oriented approach to disease prevention and treatment, but it offered a very weak basis for understanding the *structure* of social relations in society.

Health education also took many cues from psychology's learning theory in developing methods for inculcating desired health attitudes and behavior. But this application was, of course, only a rough approximation. The precision of laboratory methods was difficult to transfer to social situations in which people met conflicting reinforcement patterns and had their own rational and irrational methods of mediating their actions.

Social psychology, however, provided health educators with a significant new handle on changing behavior, and an expanded professional role. Small-group dynamics and Lewin's field theory (15) gave them the technique and theory for an ostensibly effective method of using group processes to change the behavior of individuals. Field theory provided a more sophisticated psychology to explain the interaction of individuals in groups (15). Group dynamics developed methods by which professional group leaders could substantially control group processes and help determine their outcomes (16). While the contributions from social psychology were useful in building a theory for health education's behavior-changing practice, they kept health educators' attention and efforts focused on the individual and not on the system.

These theories from social psychology were added to sociological concepts of community derived from the Chicago School and the development of community organization as a social work method. Together they formed health education's community organization theory. "Community organization" is often claimed as a technique of social change, but its historical genesis in the early charity organizations, settlement houses, urban leagues, community councils, and United Fund movements assured it a conservative, social adjustment role (14). In addition to their casework

role, social workers became technicians who planned and implemented social changes approved by funding agencies or individuals who were often far removed from the "target" communities. Planned social change in the United States has, as Galper (17, p. 112) noted, isolated "change processes from the arena of social and economic conflict" and assured its confinement to the institutionalized, less disruptive structures and dynamics of the society.

Health educators applied these community organization concepts of planned change to the health field and developed a new role for themselves. Community organization theory provided an ideological justification for administrative techniques masquerading as methods of social change. For example, health educators tend not to regard social class differences as sources of conflict and instead treat them as influences on behavior. The importance of class, claims DuBois (18), "lies not in class *per se*, but in class as an index to cultural goals and institutional roles." For DuBois, "social class is a dependent, not an independent variable." Community health education usually ends up containing conflicting interests within a community and cooling them out.

In this role, health educators attempt to apply group process skills to facilitate change broader than that at the individual level. Health educators see themselves helping to formulate health and medical care policy, and, through community organization techniques combined with planning and evaluation skills, to bring together the concerned parties and mediate a solution to health care problems. On this role, health education hangs its professional hat. As mediators of change among *all* parties, health educators believe they are subordinate to none of them. Most broadly, this perspective adds to individual health-behavior change the objective of obtaining "active participation of all individuals concerned with a problem in order to explore all facets of the problem and to develop realistic solutions to it" (11). Within this role, however, health educators see themselves as advocates of consumer participation in health planning and of the needs of the poor for health services (5, 11), a position that sets them apart from physicians and administrators.

STRUCTURAL CONSTRAINTS ON THE PROFESSION

The mediation role has been the focus of much tension between the policy-oriented, usually academically based leadership of the profession and other health professionals and policy makers. Health educators are constrained in this role by contradictions between their own conceptions of their role and the structures in which most health educators work.

Health educators conceive of their role as mediating differences and conflicting interests among health care providers and consumers. As Richard McGrail, an official of the American Cancer Society and a participant in the President's Committee on Health Education, put it, "Learning and behavior are facilitated through the two-way communication of information, knowledge, values, and attitudes" (1). This means teaching health providers how to relate to the needs of the poor, ethnic minorities, and other consumers, and it means helping consumers understand how to use health services and participate in their planning structures (11). This mediating role, however, exists side by side with the profession's conception of itself as advocate of consumers

in general, and the poor, minority groups, the aged, and other "medically underserved" groups in particular (5, 11, 19, 20). There is, of course, a logical and practical conflict between simultaneous roles as mediator and advocate. The conflict inevitably gets resolved in the structures in which health education is planned and those in which health educators work.

In concrete attempts to develop consumer participation in health planning, health educators act as facilitators of co-optation more frequently than they increase the power and influence of consumers. Health educators are usually employed by an institution whose material interests may differ from those of consumers. They are asked to involve consumers in a planning process in which, as Grossman has observed, "people are encouraged to set their own goals within a framework of predetermined goals" (19). The question before planning committees is not, for example, whether an occupational health program for an industrial community should expose violations of the Occupational Safety and Health Administration's regulations, for such a policy would be vetoed by industry members of the committee. Rather, the policy is framed in terms acceptable to industry, one that makes the workers' health subordinate to the closed process of a labor-management occupational health and safety committee. The community member is involved, in the words of Grossman, "to change *his* behavior or to add his weight to the achievement of the already determined outcome" (19). If by some chance things were to get out of hand, the effect would be little different; those with more power in the situation can effectively veto the decisions of others or simply ignore the outcome of the process.

HA!

Furthermore, despite the lofty ideals of the profession's leaders, most jobs open to health educators are in public health departments and medical care institutions in which health educators are subordinate to other health professionals and administrators (21). The bold prescriptions of university-based health educators that "health education must concern itself with the broader social issues that are obstacles to good health" (22) are lost in the mundane jobs that require health educators to focus their energies on changing people's life-styles or their use of health services.

CHANGING INDIVIDUALS AND SOCIETY

It seems clear that most of the profession's leadership realizes this conflict because few of its recognized spokespersons give more than lip-service to social change efforts. Even recent public pronouncements on the field focus mainly on individual behavioral change. As a 1966 statement defined it (23),

> Health education is a process which effects changes in the health practices of people and in the knowledge and attitudes related to such changes. Education is an internal process for the individual concerned. . . . Education thus places responsibility on the individual and is essentially different from a compliance approach. It involves motivation, communication, and decision-making.

In the extreme, however, this process orientation concludes that "the real index of health education is behavior," and the effectiveness of health education is measured in "changing behavior, not in reducing morbidity and mortality statistics alone" (24). Health educators have made their mark by obediently concerning themselves less with *what* should be changed than with *how to do it*.

Following this usually exclusive concern with process, few health educators show much concern for enabling people to develop rational insight into their lives—an understanding on which to base personal and community decisions. While some policy reports consider health education "a process with intellectual, psychological, and social dimensions relating to activities which increase the abilities of people to make informed decisions affecting their personal, family, and community well-being" (3), this is usually considered fulfilled by facilitating health-promoting behavior and discouraging health-damaging life-styles. "Harnessing recommended health behaviors and practices to values which have real meaning to adults requires imaginative approaches," writes William Griffiths, a prominent member of the profession. "The coronary patient may be persuaded to follow a prescribed regimen, not because he is concerned about himself but because of his great desire to join his companions for the opening of the hunting season" (22). If a man's own health is not sufficient motivation for him to follow recommended practices, then find another incentive that will be effective in changing his behavior.

This same line of reasoning has led to the incentive sterilization programs in India through which men and women are given cash, radios, or other popular consumer goods if they submit to vasectomy or tubal ligation. It has also led to drug and alcohol treatment programs, like the Department of Labor's "Employee Health Program," designed to "stabilize work behavior" and provide other cost-effective benefits without regard for the broader social causes of addiction or the frustrated personal and occupational needs of the addict (25). Behavior change is the important objective, not helping people understand why they have little regard for their bodies or helping them change personal and social conditions that encourage attitudes perceived by health professionals to be careless, irrational, or both.

Certainly not all efforts to change health-related behaviors are manipulative or reactionary. Some of this focus on teaching health-promoting behavior is motivated by a desire to free people from dependency on doctors, to enable people to do things for themselves that probably could not be done any better (if as well) by health professionals. Ivan Illich and his growing following within and outside the health professions raise these issues when they advocate self-care (26, 27). Many policy makers are clearly concerned about saving medical care dollars by reducing utilization of physician services. Encouraging health-promoting behavior change can be both well-meaning and useful. But much of the emphasis of this perspective is implicitly reactionary (28).

By focusing on behavior itself, health educators do not deal with the social relations and structures that may underlie and contribute to the behavior patterns they find objectionable and the diseases they wish to prevent. The alienation of people from their bodies, stressful working conditions, and inadequate nutrition due to insufficient income or inadequate food supply, are ignored. The ultimate absurdity of this position is represented by a study of lead poisoning among children in Cleveland's black ghettos (29). By focusing blame for the children's lead-paint eating on Southern blacks' "permissive socialization of oral behavior," the researchers took the heat off the slumlords.

Health education programs often rely on techniques that manipulate behavior rather than facilitating individuals' and groups' abilities to influence and control

their physical, social, and economic environments. Such programs implicitly, and often explicitly, reinforce the notion that, say, the individual poor person is to blame for his/her poverty (30). Persuading smokers to kick the habit is certainly in their interest, but when smoking behavior is a response to stressful conditions and no one helps smokers examine or change those conditions that create the "need" to smoke, then anti-smoking campaigns become "blaming the victim" crusades. As C. Arden Miller, a past president of the American Public Health Association, recently said, "For the vast majority of people in our society, the life circumstances leading to poor health are not adopted as a matter of personal choice, but are thrust upon people by the social and economic circumstances into which they are born" (31).

Whether it is manipulative or not, the practice of displacing people's attention and energy from changing the social context of behavior to changing individual behavior leaves a major problem area untouched. As Dorothy Nyswander regretfully concluded at the end of her career, "My efforts were expended in working on the symptoms of closed societies; the basic conditions giving rise to the symptoms were untouched." This founder of the health education profession was compelled to ask herself, "Have I not actually helped to maintain the status quo in these situations? Have I not taught people to accept those gifts approved by the establishment which would make life more bearable but which would not threaten the power of the establishment itself?" (32).

We need to increase *both* the ability of people to take maximum power over their *individual* lives *and* their ability to change the *social* relations and structures in which they live and work. This is an important and feasible role for health educators. The relevance of this dual emphasis can best be seen by contrasting the prevailing assumptions in the *practice* of health education with the substantial evidence that it is misdirected.

AN ECOLOGICAL PERSPECTIVE ON HEALTH AND DISEASE

Outside the writings and goals of health education policy leaders, the realities of health education programs must be confronted. Health education practice generally rests on three assumptions: (a) the most important variable in health status is individual behavior or life-style; (b) people are responsible both for seeking medical care when sick and for preventing the onset of disease in the first place; and (c) health behavior can be substantially changed without changing the social relations within which a person lives and works.

Against this perspective stands a growing preponderance of epidemiological evidence that draws our attention to the social bases of health and disease. Our broadened perspective should take into account the relationship between all environmental factors—physical, social, and economic—and health. Without attempting to review all the evidence, let us just summarize some of the major findings and their implications for health education.

Historical epidemiological evidence overwhelmingly demonstrates that improvements in nutrition, sanitation, and housing and working conditions rather than medical science or medical care brought down high mortality rates and raised life expectancy over the past two centuries (33, 34). McKeown (34) convincingly argues

that the increase in food supplies beginning in the 18th century and environmental sanitation in the 19th century greatly reduced mortality rates. These measures were well under way by the middle of the 1880s, before either the concept of specific causes of disease or the germ theory was widely accepted. Specific preventive and therapeutic medical measures gradually introduced in the 20th century slightly accelerated the already substantial decline in mortality and improved physical health (34).

The contribution of medical science to improved health has been greatly overrated. "In reality," as Dubos has observed, "the monstrous specter of infection had become but an enfeebled shadow of its former self by the time serums, vaccines, and drugs became available to combat microbes" (35). Tuberculosis, the Great White Plague of the 19th century, is a prominent example of the decline of infectious diseases. In the middle of the 19th century, TB was killing 500 people per 100,000 each year. By 1900, its toll fell to 200 out of 100,000, and by 1967, the U.S. rate had dropped to 3 deaths per 100,000. This tremendous rate of decline was only slightly affected by the introduction of collapse therapy in the 1930s and chemotherapy in the 1950s. It was improved housing and nutrition, brought about by increased productivity and social reforms, that were responsible for the decline of TB (36).

Medical science has contributed drugs that are effective in treating bacterial infections, thus reducing both death rates and days of hospitalization. It has also created the potential, through immunization, of eliminating *some* diseases. Smallpox has already been eradicated, and, with effective immunization campaigns, polio and measles could be eliminated as well. These and some other technological and curative interventions notwithstanding, medicine's impact on overall death rates and most other measures of health status has been substantially smaller than the effect of social and environmental factors.

In combination with improvement in general living conditions and public sanitation, improved nutrition has probably contributed more to increased life expectancies than any other factor. Its contribution has been greatest for children. Improved living conditions have reduced the spread of infectious diseases, and better diets have strengthened children's bodies and enabled them to fight formerly major killers. The infectious diseases that took the heaviest toll of children for example, scarlet fever, diphtheria, whooping cough, and measles declined over the last century due to these environmental improvements (33). With rising living standards, infant mortality in the U.S. fell from about 150 per 1,000 live births at the turn of the century to 16.5 in 1974. In areas where malnutrition is widespread, infectious diseases still take a heavy toll of children under five (37).

Most of the increased average life expectancy at birth in industrialized countries has been due to a sharp decline in infants' and young children's mortality from infectious disease. Longevity for men in the U.S. who reach 50 years of age has not increased for half a century. The major killers of adults are not infectious diseases, but heart disease, stroke, cancer, and accidents. The mortality rates due to these causes of death are not responsive to simple increases in the standard of living.

In fact, some health professionals consider these contemporary major killers diseases of affluence. Fuchs, for example, argues that these deaths are related to our "life-styles"—to the food we eat, our lack of exercise, our smoking, our overindulgence in alcohol, our reckless driving, and our violent social relations in which people kill

each other (38). Attributing these causes of death to "life-styles," however, implies a substantial degree of individual choice. There is a strong tendency to blame industrial accidents on "behavioral problems of workers" (39) and to conclude that health status will improve when people take "individual action" to stop smoking, drink less alcohol, eat better, and get more exercise (40). Epidemiological evidence contradicts these simplistic conclusions.

The physical environment plays a central role in contemporary disease and death rates. The role of air pollution in the etiology of disease is widely known. Even "normal" levels of air pollution have been associated with adverse clinical disease rates. The effects of pollutants on children (in whom smoking and occupational dust exposures are assumed to be minimal) include temporary deterioration of pulmonary function and increased frequency of lower respiratory tract infections. Air pollution is also associated with lung, stomach, and other forms of cancer, as well as chronic bronchitis and asthma (36, 41). More broadly, we know that cancer (in industrialized countries, the second leading cause of death after heart disease) is strongly influenced by environmental factors. At least 75 percent of all cancer is caused by harmful effects of the environment (42). And the president of the National Blue Cross Association has estimated that "31 percent of workers' health problems are caused by factors in their environment" (43).

Social conditions outside individual choice are likewise central in determining who will live and who will die. Despite increasing standards of living throughout the society, great social and economic disparities pervade advanced capitalist countries and are associated with similar disparities in morbidity and mortality. Within the U.S., social class, race, education, and occupation are very much related to rates of sickness and death (44-46). A study of causes and rates of mortality in Los Angeles County found that low birth weight, infant mortality, and deaths from all causes were positively correlated with being black and negatively correlated with income (47). Similarly, the probability of being disabled (temporarily or permanently) is negatively related to income and education, but positively related to being black (48, 49). Epidemiological evidence forces the conclusion that poverty, as Winkelstein put it, "remains among the most powerful determinants of altered health status and clinical disease today. It may well be that elimination of poverty in and of itself would drastically alter the health status of the population in a favorable direction" (50).

Undoubtedly, individual behavior works hand in glove with the physical environment and social conditions in the etiology of many important diseases. An effective treatment program must deal with them all. Should health educators focus single-mindedly on changing individual behavior when the physical environment, social class, and social conditions are such important variables in the etiology of disease? The case of essential hypertension[2] demonstrates the error of ignoring the social environment. It also provides an example of how efforts to change social conditions that contribute to disease can be incorporated into treatment and prevention programs.

It is sobering to realize that although some 25 million Americans are believed to have high blood pressure, few of the abundant epidemiological findings concerning

[2] "Essential hypertension" is high blood pressure not definitely attributable to any known cause (such as kidney disease). It comprises about 90 percent of all hypertension

the condition have found their way into treatment and prevention programs. A very brief review of these findings will make the point.

High blood pressure is a major contributing factor in heart disease, stroke, kidney disease, and other cardiovascular diseases and deaths (51-53). It is generally agreed that along the blood pressure gradient, the higher one's blood pressure, the more likely it is the person will suffer from one of the diseases for which hypertension is a known risk factor[3] (52).

Hypertension is not distributed evenly among all population groups. About 30 percent of all U.S. blacks have high blood pressure, twice the prevalence rate for whites (51). The differences between blacks and whites are *not* believed to be due primarily to genetic factors. Liberians and other West African population groups from which many of the American slaves were captured have lower blood pressure levels than U.S. blacks. Furthermore, when socioeconomic differences between blacks and whites are considered, much of the difference between their blood pressure levels disappears (51, 52, 54).

Overall data on race, social class, and other studies bearing on heritable factors strongly suggest that environment is a powerful influence on blood pressure. More specifically, it appears that inheritance and environment work together in essential hypertension. As epidemiologist Michael Marmot summed up the issue, "It seems likely that certain people are genetically more susceptible to developing high blood pressure than others, but for the high blood pressure to develop, exposure to environmental factors is also important" (52).

Similarly, salt consumption and obesity, often heralded as major factors in the etiology of hypertension, play complex, but limited, roles. Excessive salt consumption is now generally believed to facilitate the development of hypertension when other causally related factors are present. Obesity is causally related to hypertension, but not universally. It, too, seems to work in conjunction with other factors (51).

Clearly, hypertension is caused by a multiplicity of factors, but one broad environmental factor - social and psychological stress - seems very important and yet is usually ignored. Many studies demonstrate that when traditional peoples' cultures are disrupted by "modernization" or migration to industrially more developed areas, their blood pressures increase. Other studies demonstrate that migration from rural to urban areas generally is accompanied by increases in blood pressure and coronary heart disease (51, 55).

Although differing methodologies limit comparability among studies, there is substantial evidence that persons in piecework and other assembly-line type work (56) and high-pressure jobs of all kinds (57, 58) have higher blood pressures than persons in less pressured jobs. People whose lives are integrated around pressured careers and unrelenting competition are good candidates for high blood pressure and coronary heart disease (59). Such widespread stressful factors in our society as unemployment (60) and feelings of powerlessness (61) have been associated with increased rates of hypertension.

Treatment and prevention strategies should be geared both to epidemiological

[3]Hypertension is usually considered to include diastolic blood pressures of 95 mm Hg or above; 90-94 mm Hg is considered "borderline" hypertension.

evidence concerning the causes of hypertension and to clinical effectiveness in lowering blood pressures. In reality, it is the quick, technical, clinical "fix" that is universally applied whether in the doctor's office, high blood pressure screening and treatment programs, or massive federal and state campaigns. Anti-hypertensive drugs are by far the preferred treatment despite the fact that the most conclusive study to date has demonstrated their value mainly for older males or those with highly elevated blood pressures (62, 63). Quite appropriately, changes in diet and exercise programs are frequently recommended, along with weight reduction for obese patients. Stress-reducing techniques are just coming into vogue on the fringes of the medical profession: biofeedback and meditation may be useful in reducing blood pressures in some cases of borderline or moderate hypertension (64-67).

While anti-hypertensive drugs may be the easiest first line of attack by overly busy physicians, one would expect major state and national programs to develop broader lines of treatment. Yet, like its counterparts throughout the nation, the 1976 *California State Plan for High Blood Pressure Control* (53) relies mainly on pharmaceutical treatment, giving secondary emphasis to altering diet and other health-related behaviors, and meditation and biofeedback to relieve stress.

Drug treatment is immediately effective for sharply elevated blood pressure. When non-drug therapies are combined with the drug regimen, they may permit reducing the amount of medication used and, therefore, reduce the probable side effects and the high cost that anti-hypertensive drugs represent (53). While relaxation techniques provide a degree of adaptation to stress, they leave the stressful conditions unchanged. The real challenge in the treatment *and* prevention of high blood pressure is to *change the immediate social conditions that contribute to stress*. Although this might be difficult, to avoid it would be to ignore a fundamental cause of hypertension.

We will illustrate how this might be done after briefly describing an "ecological" model of health education.

AN ECOLOGICAL MODEL OF HEALTH EDUCATION

It is one thing to know what needs to be done to improve health and it's quite another to *do* it. Can health educators develop a work role in which *as health professionals* they help change social conditions that make people sick, literally? We believe they can *not* as technicians drawing plans for social change, but as advocate-facilitators for constituencies that need them.

Health educators could provide the means for people to examine the social bases of their work and lives for unhealthful conditions *and* help them learn tools to change those conditions. This is a substantially different role than putting complete responsibility on the individual to change unhealthful behavior, although it definitely retains a strong emphasis on the responsibility of individuals to participate and take the lead in alleviating such conditions.

This ecological perspective in health education recognizes that individuals act within a social context that limits their individual options and the impact of their individual actions. Milio (68) argues that health education should not just provide consumers "with information or knowledge in order to make them aware of the costs and benefits to their health to be derived from particular behaviors." Health educators

must also be concerned "to provide ready access to health-promoting options," she continues. "For example, a local effort at conveying more knowledge about healthful diets is not likely to result in changes of eating patterns unless it is accompanied by a combination of healthful, low cost, readily available foods - changes which require effort beyond the individual or small-group methods, and extend to the community public and private organizational structure."

Health educators should not only explain to people what behaviors or conditions would be beneficial to their health, they should also help them act to remove health-damaging conditions and substitute health-promoting ones. Because both kinds of conditions are usually outside any individual and often beyond the power of any individual to change, the health educator must be a kind of community organizer as well. In other words, the role of health educator should be not only to set up educational programs to change individuals' health-related behavior or even to help community members define their needs; it should *also* be to help people challenge health-damaging conditions in their communities and workplaces and to work toward health-promoting conditions.

As facilitators, health educators would help individuals and groups identify health problems in their communities *and*, if desired by the community people with whom the health educator works, teach them organizational and political skills necessary to challenge and alleviate those problems. For example, in many poor communities supermarkets are known to raise their prices on days that people receive their welfare checks. A resourceful health educator concerned with the community's nutrition might encourage community people to form car pools to shop at stores outside the barrio or ghetto, enabling them to boycott the offending stores until they end their discriminatory pricing practices. The health educator would not need to lead the demonstration; she or he needs to be a resource person who can teach effective organizing skills.

Two Hypothetical Programs

Two model programs following the dictates of the ecological perspective will illustrate the priorities we have in mind, the breadth of methods used, and the kinds of relationships between health professionals and others that are envisioned.

The first model we will describe is occupation-based. It might be sponsored by a union, it might be located in a large work setting, or it might be found in a community dominated by one or two industries that produce relatively large numbers of occupational health problems. Whatever the setting, the program would have to be geared to the specific health needs of that set of workers and the conditions in which they work and live.

The optimal program would concern itself with all occupational health problems found among a given set of workers. There may be contamination of the air that workers breathe, causing respiratory problems or cancers; excessive heat or cold, contact with toxic chemicals, or radioactivity may be other problems. There may be loud noise, speed ups, tension among workers and between them and supervisors, all of which may produce stress, leading to accidents, hypertension, and general wear and tear of bodies and minds

With the help of an industrial hygienist, discussions with workers themselves, surveys of physicians in the community, any of several manuals available (69-72), or any creative method, the health educator and the workers should determine the health problems facing that particular group. Following standard health education methods, the group members should then decide what problems they want to tackle, set priorities for themselves, and develop a program for achieving their goals.

For a hypothetical illustration, let us pick an automobile assembly plant. A health educator is hired by the local union to develop an occupational health program in that plant. Under provisions of a contract negotiated with the company during recent collective bargaining, the union runs the health education program and the company pays for an agreed-upon number of hours for workers to be released from regular plant jobs to work with the occupational health and safety committee, while continuing to receive their regular wages.[4]

At a meeting of the workers' committee during which occupational health literature and survey forms[5] drawn from similar work settings are reviewed, the health educator and the workers decide to evaluate the conditions in the plant. With the results of their survey, they conclude that excessive noise and the pace of work are the plant's major health hazards. The literature, informal discussions with groups of fellow workers, and systematic observations in a few locations in the plant persuade the committee that loss of hearing and stress-induced diseases especially hypertension are serious health consequences of these conditions. They decide on a four-part program of action, including education, screening, treatment, and prevention; in addition, they establish evaluation criteria and procedures.

Defining specific program objectives, the committee's goals are: (a) to survey plant conditions and develop systematic procedures in an ongoing program to identify and monitor all health hazards in the plant; (b) to train at least two workers from each work area or unit to identify and measure health hazards on their jobs; (c) to educate workers in the plant to understand how hearing loss and hypertension may result from working conditions; (d) to screen all workers for hearing loss and hypertension; (e) to enroll in treatment programs at least 90 percent of all workers found to have hypertension and refer workers with hearing loss to hearing specialists; (f) to abate all such hazards found in the initial plant survey; and (g) to establish procedures for continued abatement of subsequently identified hazards. The committee then sets specific criteria of success and designs methods of measuring the achievement of goals for each objective.

The health and safety committee develops a program to educate plant workers to dangers in their workplace. The health educator and the committee together train selected assembly workers to identify and measure hazards in their work areas with environmental monitoring equipment provided by the company,[6] and to educate

[4] The United Auto Workers nationwide contract provides paid release time for some workers in each auto plant to perform health and safety work under the union's direction. The International Association of Machinists, Local 1781, has a similar contract arrangement with United Airlines at San Francisco International Airport.

[5] For sample forms, see reference 72, pp. I-18 to I-20.

[6] The United Auto Workers contract requires the company to provide union health and safety committees with noise, carbon monoxide, and air flow testing equipment.

other workers to the effects of such hazards. A one-day training course prepares these workers to survey conditions in all work areas of the plant.[7] Involving workers in educating each other builds interest and a committed following within the plant, and it helps more workers intimately understand the problems facing them and enables them to lead efforts to deal with these problems.

The committee also develops programs to screen all workers for selected health problems and treat those with correctable health problems that may result from plant conditions. With funding obtained from a (hypothetical) hypertension control program of the U.S. Public Health Service, the committee hires public health nurses to measure the blood pressure of all plant workers. The union pays for audio screening done at the same time as hypertension screening. Workers who are found to have hearing loss are informed by the screening team whether their loss is potentially compensable under workmen's compensation laws. Workers with blood pressure higher than is considered medically satisfactory are provided several forms of treatment. Anti-hypertensive drugs are known to be effective in reducing blood pressure, but they do not deal with the stressful conditions that are known to be associated with hypertension. Therefore, the treatment program combines drug therapy under medical supervision (provided by the union health plan) with relaxation exercises *and* problem-solving groups (73, 74) designed to identify stressful situations in the lives of affected workers.

These groups help their members examine their work and living conditions for sources of stress and help them formulate goals and methods of making their lives more satisfying. Such methods may include working with fellow workers to reduce stressful conditions in their workplace or trying to develop more satisfying relationships with a person's family and friends. The stress-reduction goals are developed by each person with the help of the group and the health educator leading the group. The group also helps each person deal with obstacles to achieving her/his goals.

In the same collective bargaining contract that provided release time for health and safety committee members, the company had agreed to provide two hours a week paid release time for each hypertensive employee to participate in the treatment program. Different combinations of therapies—drugs, diet, weight reduction, exercise, and problem-solving groups—are used with the consent of the workers to determine which combinations are most effective in different cases. Measurements of changes in blood pressures are compared with types of treatment to determine the most effective therapeutic programs. As conclusions are drawn from this formative evaluation, the programs are modified to take advantage of important findings.

Finally, health and safety committee members develop a prevention program to remove health hazards from the plant or at least to alleviate them. Part of this prevention aspect is accomplished by having already trained workers identify hazards in their work areas and by providing them with equipment to measure environmental problems. Identifying the problems is only the beginning, however. The self-education of workers and the screening program in which they have seen large numbers of health problems attributed to working conditions has turned the formerly lethargic occupa-

[7]Similar training programs are now conducted by the Labor Occupational Health Program at the University of California, Berkeley. The Oil, Chemical and Atomic Workers Union has taught some of its members how to use monitoring equipment.

tional health and safety committee into a more dynamic body. With technical assistance and encouragement from the health educator, the committee identifies not only excessively noisy or dangerous machinery as hazards, but also begins to define the speed of the assembly line and the organization of work itself as potential health hazards that increase risk of hypertension for all exposed workers.

The committee, therefore, attempts to file grievances for abatement of all hazards, and specifically insists that the assembly line be slowed to a speed it and line workers consider tolerable. For some workers, meeting increased line speed is the only interesting challenge left in assembly-line work. However, because speed-up may be health damaging, the committee concerns itself with different ways to organize production for example, team assembly to make the jobs more interesting and less damaging. It also develops a proposal to have worker representatives survey their work areas at least twice a year, and another proposal to require the company to abate all identified hazards within 30 days of any report.

Quite naturally, the company, whose profits depend substantially on the productivity of its assembly lines, does not want to consider assembly speed a grievable issue. It is willing, plant managers report, to institute reasonable design modifications to abate machine noise and dangers, but line speed is a management prerogative it will not relinquish.

Because of the grassroots education the occupational health and safety committee has conducted and because of committee leadership provided by concerned rank-and-file workers, the issue is not laid to rest on the grievance committee table. The committee presses its demands, builds support among workers throughout the plant, and insists that assembly line speed, related working conditions, and the organization of work itself become primary issues in the next collective bargaining negotiations. The membership of the union local is prepared to strike over these issues that affect their well-being and longevity, prerequisites to their being able to enjoy the modest financial benefits the union is also winning.

In this optimistic hypothetical case, the health educator performs a crucial and delicate role. As a health professional, the educator provides technical information and resources that would be much more difficult for the occupational health and safety committee members to obtain without this expertise. The health educator also provides knowledge of and experience with educational techniques, group process, health institutions, grant proposal project design, and the like that facilitate the committee's work every step of the way. At no time, however, does the health educator become the director of the operation or the spokesperson for the workers. The educator sees her/his role as encouraging, facilitating, and teaching self-organization of workers to change unhealthful working conditions. Every phase of the occupational health program from worker-run educationals and screening, to problem-solving groups in the treatment program, to the prevention program itself is designed to contribute to the activation and organization of plant employees to protect their own health.

A second illustration will suggest a way of applying the ecological model of health education to general community settings. Again, the focus of the program would depend, of course, on the particular health problems in each community. The community constituency for a program might be an ethnic or racial minority group;

it might be poor, working-class, or middle-class people; or it might be gay people, children, senior citizens, mothers, women in general, physically disabled people, or a cross section of the community.

Ideally, any program would be developed around a broad range of community health needs, but it might be focused more narrowly on one or two problems because a sponsoring community organization is mainly interested in one issue or a funding agency is willing to provide a categorical grant to deal with one health problem. Thus, for illustration, let us say that a health educator, hired by the county health department to develop a program to improve the nutritional status of all residents in a poor and working-class area of a large city, must approach this monumental problem.

Since this health educator does not have a ready-made organization to work with, she or he must go into the community and talk with community organizations of all kinds - women's organizations and bowling leagues, senior citizen centers and youth organizations, ethnic groups and neighborhood associations. Approaches for making contact with persons not using community organizations must also be developed. Finally, meeting with collected representatives of organizations and concerned individuals, the health educator and the community residents discuss the community's varied nutritional problems with a nutritionist as consultant. After reviewing evidence of local problems, including residents' informal observations and data gathered by the local health department, the group concludes that three significant problems all need immediate attention: many older people lack either the financial or personal resources to obtain adequate meals (75, 76); many pregnant women get inadequate prenatal care and inadequate diets (77); and people in general but especially teenagers and women consume diets that are too high in refined starch, sugar, and fat (78, 79).

The group designates itself the "Community United for Better Nutrition." With the participation of the health educator, the group develops programs linked together around the common issue, but with separate components and working committees to deal with each population group. Without detailing the process of each group, we can summarize the essentials of the programs they develop. The senior citizen group develops a proposal for meals-on-wheels and a congregate dining program, creating several neighborhood-based centers which provide nutritious meals catering to local tastes, conduct nutrition education, *and* make meals social occasions which help reduce the isolation of the elderly. The group submits its plans to the city council and the county board of supervisors for funding. Having been turned down by the two bodies' respective health and welfare committees, the senior citizen group appeals to the wider Community United group for support. The health educator urges all members to join in solidarity with the older persons' committee, reminding the members that each group may need to call upon the others for support from time to time.

The group that is concerned about maternal health, learning from the experience of the senior citizens, involves union women, church groups, women's organizations, and neighborhood associations in developing a proposal for free maternal and child health clinics to be opened in several locations in the spread-out community. They also propose that the Women, Infant and Children (WIC) program be implemented more aggressively and comprehensively. Then, with the support of these groups and the broader Community United for Better Nutrition, the group presents its proposals to

the board of supervisors for revenue-sharing funds. They insist that the program must be governed by a board of community people from the nutrition organization and the other interested groups, and that sufficient funds be provided for outreach and education programs.

Finally, obesity and poor eating habits in general present a difficult problem to the group working on the issue. This problem is not simply a lack of facilities and programs. The group, which includes obese people themselves as well as others more generally concerned with the issue, concludes that people eat the wrong foods for different and often overlapping reasons. Some people eat too much highly refined starch, fatty foods, and too many sweets because they don't know they are bad for them; others eat them because they don't have enough money for an adequate diet and come to depend on these foods for energy to get them through the day; and others eat snacks obsessively because they feel helpless to deal more constructively with the deep dissatisfactions they experience in their lives. Contributing to all these factors, supermarkets and food processors encourage excessive consumption of potato chips, starchy snacks, candy, cookies, cakes, sodas, and sugar-laden cereals because they provide a profitable line of products to a receptive market that may be enlarged with advertising. To be effective, the program must deal with each of these problems.

With the health educator's help, the group designs educational programs to explain the consequences of obesity and teach proper nutrition in schools, community centers, organizations, and other places where people gather regularly. These educationals reach out to overweight people and to those with no obvious weight problem. They involve people in small-group discussions about their eating habits, why they eat what they do, how they may use food in ways that parallel tobacco or alcohol use, how adult eating behavior provides models to children, why supermarkets and the food processing industry encourage consumption of poor foods by their advertising, and so on. Ongoing problem-solving groups are set up in schools, community organization centers, and large workplaces. They are designed to help people change dietary habits that are injurious to their health and to develop more healthful and satisfying eating habits. These groups encourage their members to examine their personal lives for sources of dissatisfaction, stress, and other factors that may contribute to their compulsive or otherwise destructive food consumption. Group members are given support to change those conditions within their control, and to work together to overcome shared problems. The groups are led by community people, public health nurses, and health educators who receive special training in group process and problem-solving perspectives and techniques.

Finally, direct action is needed on two fronts to deal with causes beyond the scope of individual change. With the health educator playing a facilitating role, the group develops broad community support for a more accessible and liberal food stamp program to provide adequate diets for people without adequate incomes, and takes their demands to the county board of supervisors.

With the consciousness of community members already raised by educational programs on nutrition, demands are put to supermarkets in the community to remove candy display counters around checkstands and reduce the amount of sugar-starch food items they carry, making healthful or at least less harmful foods more readily available. Uncooperative stores find themselves boycotted by many of their regular

customers (80). The group also encourages efforts to launch a cooperative market that will better serve the community's nutritional and financial needs. They also build support within schools and the community to improve school lunches and remove "junk" food from school vending machines (81).

As in the first illustration, the health educator acts as technical adviser and process facilitator to the organization. She or he is neither spokesperson nor director of the community's campaign. Some efforts will succeed and others will fail. But in each illustration, the health educator helps people articulate their health needs, develop programs of individual and social change to meet their needs, and learn from their successes and their failures.

Discussion

In both illustrations, some health problems were behavioral, or situations over which the affected individual could exercise considerable control. Therefore, educational programs and supportive problem-solving groups were provided to encourage and help people make informed changes in their lives for example, changing their diets and eating habits or dealing with stressful relationships in their personal lives. In each case, however, there were factors outside the control of any individual in the "target" population inadequate incomes and profit-oriented marketing of food in one case, and noisy machinery and assembly line speed-ups in the other. All these problems were the active concern of the health educator who presumably had both training and experience adequate to deal with them.

Clearly, the health educators in this model do not sell their loyalties to whoever pays their salaries. Since the mediator and advocate roles often conflict, and since the mediator role is most often a fiction that obscures the person's co-optive function, the ecological model requires the health educator to choose the advocate role. Although the health educator can continue to be the advocate of a hospital, employer, or government agency's institutional interests, we are suggesting a change from the status quo. We believe that health educators, individually and as a profession, must choose to work for the interests of communities, individuals, workers, and consumer groups. It should be clear that, in many respects, we are trying to give substance to the avowed prescriptions of many academic professional leaders (11, 22). Yet this role is seldom seriously considered within the profession.

We make four assumptions in developing this role. First, we agree with the profession itself that health educators bring to their work a greater knowledge of health and health care than most lay persons. Although this knowledge is not itself taught in most health education training programs, the tools for acquiring it are supposed to be provided. Full-time work in a particular field, of course, also imparts more extensive factual knowledge, and, hopefully, greater understanding of principles and theory than the average lay person usually possesses. In dealing with the ever-increasing complexities of health problems and medical care institutions, having this knowledge or knowing how to get it is a distinct advantage.

Second, we assume that people dispossessed from power over institutions that daily influence their lives should learn where they have collectively forgotten how to determine what is in their best interest and how to struggle successfully to achieve it.

We believe the health educator can effectively bring together such persons around their common interests that relate to health, use their process skills to facilitate the group in deciding what goals are in their interests, and be a resource in developing methods to realize those goals.

Third, we believe that the process skills of health educators can be effective in confrontations between their community constituency and people in authority or power over them by clarifying the issues that divide the group. The objective of this clarification is *not* to smooth out differences, but to sharpen the dialogue about issues and facilitate a resolution of differences a resolution that genuinely meets the needs of the community. The demarcation between facilitating changes sought by the community and co opting the community is both fine and often faint, but we believe that health professionals who self consciously define their roles in the way we have suggested and who are held accountable by their constituency can work with a community on behalf of its interests.

Finally, health educators cannot operate on a free lance basis. They need an institutional base that will pay their salaries and provide them with autonomy needed to be effective advocate/facilitators. Ideally, funds should be secured by the constituent group to whom the educator would be accountable. This would be least compromising for the educator and provide the most direct means of assuring her/his accountability to the group.

However, it is possible to use a health education position paid for by another agency, but applied to the needs of such a constituency. For example, Health Systems Agencies, under their responsibility to develop health education programs in their areas, might fund health educators to work with community groups. County health departments often employ health educators, and some might be persuaded to support them in roles we have suggested. Even when job titles or supervisors do not officially support such a role, committed and enterprising health educators could use their positions in this way. Unsanctioned roles are risky but may be very effective. They require that the educator develop strong community support that will defend him/her in case of a confrontation with bureau heads, politicians, representatives of powerful medical institutions, or other financial interests. Above all, risk taking requires that the health professional will ultimately have to place social goals ahead of personal desires for professional advancement and financial security.

CONCLUSION

Our conclusion may seem like common sense to many readers, but sadly it is viewed by most health professionals as heretical. Essentially, we argue that health professionals should redefine the scope and methods of disease prevention and even treatment to make them consistent with our knowledge about causes of disease and death. It is not the responsibility of "someone else" to deal with social and economic factors while physicians and other health professionals emphasize "early detection," sanitation, immunizations, clinical care, and compliance.

Ours is hardly a new concept. It has been around since Virchow and other 19th-century scientists and physicians joined struggles for social reform. We do not expect

every health professional to abandon his/her profession to lead a movement for socialism as Salvador Allende did or to fight guerilla wars as Che Guevara did.

We do not argue for throwing out proven medical therapies or demonstrably effective social techniques. But we do suggest evaluating each set of methods in relation to the larger physical, social, and economic contexts which propagate disease. It is not social methods that must be integrated into medical work, for that leads inevitably to the ameliorative casework approach. Rather, *medicine and the other health professions must integrate their methods into the broader struggles to alter the social and physical environment.*

It may well be, as we believe, that a society based on profit making and the private control of productive resources will always breed more disease than is otherwise necessary. When profit is the primary goal in organizing production and other economic and social relations, those who control these resources will tend to have a careless attitude toward workers' lives and the physical environment. It may be necessary to replace private control and accumulation of wealth with societal ownership and democratic control of all these resources in order to eliminate these constant assaults on our health. If socialism is a pre-condition to establishing a healthful environment, then health professions must not back off from the challenge. However, socialism is not precisely the issue we have tried to address in this paper.

Instead, we are emphasizing that the difficult task of redefining professional roles must begin. Health education is an easy mark because of its present broad and ambiguous role definition. We hope our effort will stimulate women and men in other health professions to incorporate measures responsive to the clear-cut findings of social epidemiology

REFERENCES

1. *Report of the President's Committee on Health Education*. Department of Health, Education, and Welfare, Washington, D.C., 1972.
2. John E. Fogarty International Center for Advanced Study in the Health Sciences, National Institutes of Health, and American College of Preventive Medicine. *Preventive Medicine U.S.A.. Health Promotion and Consumer Health Education*, pp. 67 Prodist, New York, 1976.
3. New definitions. Report of the 1972-1973 Joint Committee on Health Education Terminology *Health Education Monographs* No. 33 63-70, 1973.
4. D'Onofrio, C. N. What Is Health Education? Paper presented at the National Health Education Workshop sponsored by the Extension Committee on Organization and Policy and the American Medical Association, Roanoke, Va., March 1975
5. Making health education work. *Am. J. Public Health* 65 (supplement), 1975.
6. Nyswander, D. B. Evaluation of health education practice *Public Health News* 30, 1949.
7. Nyswander, D. B. Education and public health. *Annals of the American Academy of Political and Social Science* 302: 34-43, 1955.
8. *Proceedings of the National Conference of Charities and Correction*, annual issues, 1882-1916.
9. Cabot, R. C. *Social Work: Essays on the Meeting-Ground of Doctor and Social Worker.* Arno Press, New York, 1972 (originally published, 1919).
10. *Planning for Health*. National Health Assembly, Federal Security Agency, Washington, D.C., 1949.
11. Mazelis, S., Allen, N., and Frost, L. Health education in delivery of health services. In *State Plan for California Comprehensive Health Planning*, Appendix B. State Department of Health, Sacramento, undated.

12. Mico, P. R., and Ross, H. S. *Health Education and Behavioral Science*. Third Party Associates, Oakland, Calif., 1975.
13. Freud, S. *The Basic Writings of Sigmund Freud*, translated and edited by A. A. Brill. Modern Library, New York, 1938.
14. Cox, F. M., and Garvin, C. Community organization practice 1865-1973. In *Strategies of Community Organization. A Book of Readings*, by F. M. Cox, J. L. Erlich, J. Rothman, and J. E. Tropman, Ed. 2., pp. 39-58. Peacock Publishers, Itasca, Ill., 1974.
15. Lewin, K. *Field Theory in Social Science*, edited by D. Cartwright. Harper and Brothers, New York, 1951.
16. Cartwright, D., and Zander, A. *Group Dynamics: Research and Theory*. Row and Peterson, Evanston, Ill., 1968.
17. Galper, J. H. *The Politics of Social Services*. Prentice-Hall, Englewood Cliffs, N.J., 1975.
18. Dubois, C. The public health worker as an agent of socio-cultural change. *Health Education Monographs* No. 5: 3-19, 1959.
19. Grossman, J. Health for what? Change, conflict and the search for purpose. *Pacific Health Education Reports* 2: 51-66, 1971.
20. Darity, W. A., and Wilson, W. J. An awareness of the economic class structure in health education. In *New Directions in Health Education*, edited by D. A. Read. Macmillan, New York, 1971.
21. Richards, N. D. Methods and effectiveness of health education: The past, present and future of social scientific involvement. *Soc. Sci. Med.* 9: 141-156, 1975.
22. Griffiths, W. Health education definitions, problems, and philosophies. *Health Education Monographs* No. 31: 7-11, 1972.
23. Professional preparation in health education in schools of public health. *Health Education Monographs* No. 21, 1966.
24. Keyes, L. L. Health education in perspective: An overview. *Health Education Monographs* No. 31: 13-17, 1972.
25. Schramm, C. J. Measuring the return on program costs: Evaluation of a multi-employer alcoholism treatment program. *Am. J. Public Health* 67: 50-51, 1977.
26. Illich, I. *Medical Nemesis: The Expropriation of Health*. Pantheon Books, New York, 1976.
27. Levin, L. S., Katz, A. H., and Holst, E. *Self-Care: Lay Initiatives in Health*. Prodist, New York, 1976.
28. Navarro, V. The industrialization of fetishism or the fetishism of industrialization. *Soc. Sci. Med.* 9: 351-363, 1975.
29. Chatterjee, P., and Gettman, L. H. Lead poisoning: Subculture as a facilitating agent? *Am. J. Clin. Nutr.* 25: 324-330, 1972.
30. Ryan, W. *Blaming the Victim*. Pantheon Books, New York, 1971.
31. Miller, C. A. Societal change and public health: A rediscovery. *Am. J. Public Health* 66: 54-60, 1976.
32. Nyswander, D. B. The open society: Its implications for health educators. *Health Education Monographs* No. 22: 1-15, 1967.
33. Powles, J. On the limitations of modern medicine. *Sci. Med. Man* 1: 1-30, 1973.
34. McKeown, T. A conceptual background for research and development in medicine. *Int. J. Health Serv.* 3: 11-28, 1971.
35. Dubos, R. *Mirage of Health: Utopias, Progress, and Biological Change*. Anchor Books, Doubleday and Company, Garden City, N.Y., 1959.
36. Winkelstein, W., and French, F. F. The role of ecology in the design of a health care system. *California Medicine* 113: 7-12, 1970.
37. Bodenheimer, T. S. The political economy of malnutrition: Generalization from two Central American case studies. *Archivos Latino-Americanos de Nutricion* 22: 495-505, 1972.
38. Fuchs, V. R. *Who Shall Live? Health, Economics, and Social Choice*. Basic Books, New York, 1974.
39. Draft of President Ford's Economic Report to Congress. Cited in "Ford-termed cool to Ukey health and safety agencies," by D. Burnham. *New York Times*, January 16, 1976.
40. Cooper, T., Assistant Secretary of Health, Department of Health, Education, and Welfare. Cited in "The Nation's Health." *Time*, January 26, 1976.
41. Ellington, G. A. Health effects from air pollution. In *Environmental Problems in Medicine*, edited by W. F. McKee, pp. 114-124. Charles C. Thomas, Springfield, Ill., 1974.
42. Hueper, W. C. Environmental cancer risks in industrial economy. In *Industrial Pollution*, edited by N. I. Sax, pp. 118-149. Van Nostrand Reinhold, New York, 1974.

43. *Blue Cross Association Consumer Report*, March 1976.
44. Syme, S. L., and Berkman, L. F. Social class, susceptibility and sickness. *Am. J. Epidemiol.* 104: 1 8, 1976.
45. Nagi, M. H., and Stockwell, E. G. Socioeconomic differentials in mortality by cause of death. *Health Services Reports*, 88: 449-456, 1973.
46. Antonovsky, A. Social class, life expectancy, and overall mortality. *Milbank Mem. Fund Q.* 45: 31-73, 1967.
47. Chapman, J. M., and Coulson, A. Community diagnosis: An analysis of indicators of health and disease in a metropolitan area. *Int. J. Epidemiol.* 1: 75-81, 1972.
48. Luft, H. S. The Probability of Disability: The Influence of Age, Race, Sex, Education and Income. Paper presented at Annual Meeting of the American Public Health Association, Chicago, November 17, 1975.
49. Socioeconomic differentials in morbidity. *Metropolitan Life Insurance Company Statistics Bulletin*, pp. 10-12, 1972.
50. Winkelstein, W. Epidemiological considerations underlying the allocation of health and disease care resources. *Int. J. Epidemiol.* 1: 69 74, 1972.
51. Eyer, J. Hypertension as a disease of modern society. *Int J. Health Serv.* 5(4): 539-558, 1975.
52. Marmot, M. Epidemiology of hypertension. In *Hypertension Assessment and Management A Training Manual for Extended Nursing Roles*, edited by J. Fisher. Drew Medical-Dental Center, East Palo Alto, Calif., 1976.
53. *California State Plan for High Blood Pressure Control, 1976*. Chronic Disease Control Section, California State Department of Health, Berkeley, 1976.
54. Syme, S. L., Oakes, T., and Friedman, G. Social class and racial differences in blood pressure. *Am. J. Public Health* 64: 619-620, 1974.
55. Syme, S. L., Hyman, M. M., and Enterline, P. E. Cultural mobility and the occurrence of coronary heart disease. *Journal of Health and Human Behavior* 6:178-190, 1965.
56. Levi, L. *Stress and Distress in Response to Psychosocial Stimuli*. Pergamon Press, New York, 1972.
57. Friedman, M., Rosenman, R., and Carroll, V. Changes in serum cholesterol and blood clotting time in men subjected to cyclic variation of occupational stress. *Circulation* 17: 852-861, 1958.
58. Russek, H., and Zohman, B. Relative significance of heredity, diet, and occupational stress in coronary heart disease of young adults. *American Journal of Medical Science* 235: 266-277, 1958.
59. Friedman, M., and Rosenman, R. *Type A Behavior and Your Heart*. Knopf, New York, 1974.
60. Kasl, S., and Cobb, S. Blood pressure changes in men undergoing job loss: A preliminary report. *Psychosom. Med.* 32: 19-38, 1970.
61. Naditch, M. P. Locus of control, discontent, and hypertension. *Social Psychiatry* 9: 111-117, 1974.
62. Veteran's Administration Cooperative Study Group on Anti-Hypertensive Agents. Effects of treatment on morbidity in hypertension: Results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. *JAMA* 202: 1028 1034, 1967
63. Veteran's Administration Cooperative Study Group on Anti-Hypertensive Agents. Effects of treatment on morbidity in hypertension II Results in patients with diastolic blood pressures averaging 90 through 114 mm Hg. *JAMA* 213 1143 1152, 1970.
64. Benson, H., Rosner, B. A., Marzetta, B. R., and Klemchuk, H. P. Decreased blood pressure in borderline hypertensive subjects who practice meditation. *J. Chronic Dis.* 27: 163 169, 1974.
65. Patel, C. 12 Month follow-up of yoga and bio feedback in the management of hypertension. *Lancet* 1: 62-64, 1975.
66. Stone, R., and DeLeo, J. Psychotherapeutic control of hypertension. *New Engl. J. Med.* 294: 80-84, 1976.
67. Segal, J. Biofeedback as a medical treatment. *JAMA* 232: 179-180, 1975.
68. Miller, N. A framework for prevention Changing health damaging to health generating life patterns *Am. J Public Health* 66 435 439, 1976
69. Stellman, J. M., and Daum, S. M. *Work Is Dangerous to Your Health*. Vintage Books, New York, 1973.
70. *How to Use OSHA*. Urban Planning Aid, Inc., Cambridge, Mass., 1975.
71. Fowler, R. *Seven Steps to Hazard Identification*. Labor Occupational Health Program, University of California, Berkeley, 1976

72. Brecko, A., with Brunt, M. *Working for Your Life: A Woman's Guide to Job Health Hazards.* Labor Occupational Health Program and Public Citizen's Health Research Group, University of California, Berkeley, 1976.
73. Steiner, C. Radical psychiatry and movement groups. In *The Radical Therapist*, edited by J. Agel, pp. 18-26. Ballantine Books, New York, 1971.
74. Kinnes, R. How to be a radical therapist. In *The Radical Therapist*, edited by J. Agel, pp. 27-44. Ballantine Books, New York, 1971.
75. Position paper on nutrition and aging. *J. Am. Diet. Assoc.* 57: 448, 1970.
76. *Proceedings of the Sessions on Nutrition and Aging at the 25th Annual Scientific Meeting of the Gerontological Society, 1972.* American Journal of Clinical Nutrition, Bethesda, Md., 1973.
77. Committee on Maternal Nutrition, Food and Nutrition Board, National Research Council. *Maternal Nutrition and the Course of Pregnancy.* National Academy of Sciences, Washington, D.C., 1970.
78. Robinson, C. H. Nutrition education: What comes next? *J. Am. Diet. Assoc.* 69: 126-132, 1976.
79. *Preliminary Findings of the First Health and Nutrition Examination Survey, United States, 1971-1972: Dietary Intake and Biochemical Findings.* Department of Health, Education, and Welfare, National Center for Health Statistics, January 1974.
80. Shoemaker, C. Influencing supermarket policies. *Nutrition Action* 3-5, July 1976.
81. Doerschuk, B. West Virginia schools ban junk foods. *Nutrition Action* 10-11, April 1976.

Manuscript submitted for publication, October 7, 1976

Direct reprint requests to:

Dr. E. Richard Brown
1707 Buena Avenue
Berkeley, California 94703

LECTURE POUR MERCREDI LE 15 JUIN 1983

DE: Foster & M, Anderson B.D. Medical Anthropology N.Y.: John Wiley & Sons, 1978

chapter 13

Anthropology and Medicine in a Changing World I: Lessons from the Past

In the next two chapters we discuss international public health, an area in which anthropologists and health personnel have had a fairly long, and in many ways mutually satisfying, relationship. In the present chapter we describe some of the things that anthropologists have learned about the dynamics of introducing scientific medical practices into societies that previously have been dependent largely or entirely on traditional services. Our orientation is essentially the past, and our definition of "problem" is narrowly that of health personnel engaged in international work.

Although in Chapter 14 we continue with some of these same topics, our orientation changes to the present and future. We are particularly concerned with contemporary trends that seem to us to be not fully appreciated, and with their implications for health planning and policy. Among these trends are, on the one hand, the growing success of scientific medicine in winning acceptance and, on the other hand (and perhaps anomalously), the rapid growth of alternate or alternative medical systems, particularly in urban areas. Thus the problem (as viewed by formal health personnel) is no longer that of scientific versus traditional medicine. Instead, the question has become one of roles, or possible roles, for scientific, traditional, *and* alternative medical systems in

155

meeting health needs as defined not only by health personnel but also as perceived by the consumers, the patients.

THE PUBLIC HEALTH "PROBLEM"

In its ability to prevent or cure infectious diseases, to eliminate the scourges of the past, such as malaria and smallpox, to prevent infant mortality, and to mend through surgery, scientific medicine has no peers. Almost all peoples who have been exposed to scientific medicine, whatever their cultural levels, acknowledge its superiority, at least with respect to some illnesses. Most would like greater access to scientific medicine. In spite of such instances as the Indian government's support of Ayurvedic academies, the health care programs of all of the countries of the world are based largely on scientific medicine.

In view of this widespread, if not unanimous, appreciation for scientific medicine, one wonders why there should be problems, beyond those of trained personnel and money, in universalizing modern medicine. Yet almost always when scientific medical programs, and particularly preventive programs, have been introduced into areas where indigenous systems are yet strong, there have been great resistances. The initial high hopes of international medical specialists time after time have been dashed, as they have found to their astonishment that seemingly technologically sound programs have failed completely or achieved only partial success. Westerners are highly ethnocentric; we assume that the superiority of our civilization over technologically less complex societies is obvious, and we assume that, given the chance, peoples in "less fortunate" countries would like to adopt our ways. Western medical personnel have been, if anything, even more ethnocentric about the superiority of scientific medicine, finding it difficult to believe — despite evidence to the contrary — that given the opportunity, all peoples will not accept it in its totality.

As a consequence, for nearly 50 years, international public health programs and their medical specialists operated largely on the assumption that better health depended on the design and execution of scientifically sound programs, in which the people to be benefited would be enthusiastic participants. In 1916, after having demonstrated in the southern United States how hookworm could be controlled and ultimately eliminated, the Rockefeller Foundation turned to Ceylon for a massive demonstration of how disease control methods could be applied in other parts of the world. The *technical* part of the program appears to have been sound: a census, sanitary surveys to locate sources of infection, microscopic examination of feces and blood samples of everyone in the trial area, treatment of infected persons, and latrine-installation campaigns. But 6 years later, when the program was brought to an end, hookworm was still widespread, and even today it is endemic. The clinical superiority of Western hookworm control and treatment proved inadequate to the elimination of the

disease. Far from winning the gratitude of tea plantation workers, active opposition was incurred (Philips 1955).

Many years later, in 1942, when the Institute of Inter-American Affairs (a forerunner of the Agency for International Development) began public health programs in cooperation with Latin American governments, the same ethnocentrism prevailed. Better health for Latin Americans, it was assumed, would result from the adoption of U.S. practices, especially the dichtomy between clinical, curative, private sector medicine, and preventive, public health, public sector medicine. That uniquely American institution, the public health center, offering pregnancy and postnatal services to mothers and infants, immunization programs, communicable disease control, dental services, environmental sanitation services, and similar activities, was adopted as the keystone of the program. To the surprise of the Americans, public acceptance of the new programs was far from enthusiastic. Many eligible people genuinely in need of the services failed to come at all; others, after initial visits, skipped appointments or dropped out. Utilization of available services was far below planned capacity (cf. Foster 1952 for detailed discussion of these programs).

With the post-World War II establishment of the World Health Organization and the development of a vast international health bureaucracy designed to bring modern medicine to all of the world's countries, similar mistakes have been made. Regardless of country of origin, specialists trained in scientific medicine, when first exposed to cross-cultural health problems, have found it difficult to believe that the "obvious superiority" of their methods is not alone sufficient to win total acceptance.

During recent years, fortunately, we have learned a good deal about what happens when people who previously have relied only on traditional medicine are afforded the option of seeking scientific help. We know the cultural, social, psychological, and economic barriers that have often made them reluctant to try the new; and we also know the kinds of experiences that make them full or partial converts to scientific medicine. A part of the resistance to change, we have discovered, lies in the social and cultural forms of the people themselves. More recently we have also realized that both the quality of scientific medicine and the institutional forms set up to deliver it have not always been characterized by the excellence that wins confidence.

Nevertheless, for all of the problems and shortcomings that have been described for scientific health programs and services, the fact remains that scientific medicine wins converts every day, and traditional medicine, however effective it may be in some areas, is losing its hold. Villagers who, 25 years ago (as in Tzintzuntzan) rarely if ever consulted a physician, today routinely seek his services, turning to traditional curers only for minor problems or (in an interesting reversal of an earlier pattern) if the physician does not bring the immediate cure they have come to expect of him. In the pages that follow we describe what anthropologists and medical specialists have learned about the complex process of bringing scientific medicine to the world's peoples. Some

of the "problems" discussed are problems only in an historical sense; today they have largely disappeared. Other "problems" (e.g., the frequent poor quality of health services) require solutions before optimum use is made of scientific medicine.

RESISTANCES IN COMMUNITIES RECEIVING NEW HEALTH SERVICES

In their value and belief systems, in their social structures, and in their cognitive processes, tribal and peasant peoples display forms that sometimes inhibit their acceptance of scientific medicine. All peoples are ethnocentric; they are attached to their traditional ways and beliefs, and they assume that these ways are the equal to, and probably better, than the ways of others. With respect to that complex of beliefs and values associated with food, health, and illness, this seems particularly true. Health beliefs, body imagery, and concepts of illness are parts of a wider world view. And, just as this wider world view is rarely questioned, so too are individual elements that make up its totality accepted unquestionably. If ghosts, witches, and forest sprites are a part of the supernatural environment of a people, the lecture of a scientifically trained health educator is hardly going to convince them that these are not causes of disease. People whose basic view of health and illness is based on humoral pathology, on the hot-cold equilibrium that they have been taught to strive for since infancy, are no more likely to abandon this interpretation of the health environment than are educated Americans to give up the germ theory.

In a clinical sense it is easy to argue (as we have done in Chapter 7) that scientific medicine is superior to non-Western medicine. But, as Young points out, it is incorrect to say that non-Western medicine does not "work," at least when viewed through the lens of its believers. In defining "work," Young makes the important distinction between resulting in what people *hope* will happen, and producing what people *expect* may happen. The two meanings sometimes are congruent: treatment, as hoped and expected, restores the patient to health. Sometimes it does not. But treatment has "worked" in the sense that it has met the expectations of family and friends as to what should be done, and it has perhaps produced expected clinical evidence that the diagnosis has been correct (Young 1976:7). Traditional medicine has played a powerful positive role in maintinaing the psychological and physical well-being of those who believe in it.

1. The "adversary model"

In short, the health and illness views of every people are a part of their innermost being, not to be cast aside lightly until overwhelming evidence indicates that there are better explanations. It is not, to quote Polgar, a question of the "empty vessels" model, the erroneous assumption that the "new wine" of scientific health information and practice can be poured into minds essentially devoid of illness beliefs (Polgar 1963:411). In trying to understand the changing health behavior of tribal and peasant peoples and in trying to help change traditional views, anthropologists have made use of an adversary model, the "conflict" between primitive and folk medical systems, on the one hand, and scientific medicine, on the other. We have tended to see traditional medicine as the defender, under attack by scientific medicine, which little by little forces the traditional medicine to give ground, contesting every inch of the way. As a general rule, we can say that when scientific medicine first becomes available to people whose previous experience has been with traditional systems, therapeutic practices that are interpreted to be in direct conflict with traditional beliefs about health and causes of disease will encounter the most resistance. Richard Adams has given us a classic example from a Guatemalan Indian village, where blood is viewed as a finite good, as a "nonrenewable" or "nonregenerative" substance. Blood is seen as the source of strength, but when it is lost through injury or disease, a person's resistance to disease weakens forever. When medical doctors on a health team wished to take blood samples from children for a health survey, they met great resistance. Villagers simply could not understand how outsiders who obviously were weakening their children could claim to be bettering their health (R. Adams 1955:446-447). The view that blood is nonregenerative is common in much of the world. In Latin America it seems to be one of the reasons why blood banks and blood transfusions are less successful than in the United States: people are most reluctant to part with their precious blood.

In societies in which belief in witchcraft is strong, people are careful to conceal their excreta and other bodily wastes so that evil doers cannot find materials on which to work contagious magic. To such people, the thought of advertising the presence of excreta in a sanitary latrine, advocated by public health workers, is the height of folly. In a pilot project in social medicine among the Zulu of Durban, South Africa, projects such as home gardening and compost pits were modestly successful, but pit latrine programs less so. "Adults and children are found to micturate in the immediate vicinity of their huts, but defecation, except in the case of infants and young children, always took place some distance from the homestead. Bushes were preferred, but as long as a place was found that offered protection from the public gaze it was considered satisfactory. Modesty, and perhaps even more important, the desire to conceal the identity of the persons defaecating, demanded this protection" (S. and E. Kark 1962:26). In response to questions as to why pit privies were not used, answers showed clearly the fear of witchcraft. "People do not want to defaecate in one place because they still think that someone can bewitch them through their faeces. Because of this belief it is hard to get them to build pit privies" (*Ibid.*).

2. A cognitive dichotomy

Physicians unacquainted with the popular medical beliefs of their patients often are puzzled by the apparently erratic decisions sometimes made in the face of illness. In Vellore, Tamil Nadu, India, for example, it was found that young children suffering from diarrhea and severe dehydration were sometimes brought to the hospital, but other times they were not, or they were brought too late to save them. Research revealed that there are two distinct types of diarrhea, *bedhi* and *dosham. Bedhi* is viewed as a "natural" diarrhea caused by an excess of heat in the body, commonly caused by eating foods classed as "hot" according to the Ayurvedic system. To go to the hospital is viewed as an appropriate decision in cases of *bedhi. Dosham* diarrhea, in contrast, has its origin in views about ritual pollution; it occurs most commonly when a mother feeds her infant after seeing a woman who has suffered a miscarriage. Children who are believed to be suffering from *dosham* are treated with elaborate purification rites that are not, of course, a part of the treatment received in the hospital; hence, they are rarely taken to physicians (Lozoff et al 1975).

Diarrhea in Vellore is a modern illustration of a cognitive dichotomy first noted more than a generation ago. In the early 1950s anthropologists studying public health programs in Latin America discovered that traditional peoples who had access to scientific medical services felt that illnesses with certain etiologies could be cured or prevented by physicians so that, in spite of initial distrust, they sought such services. Illnesses with other etiologies, however, were felt best treated by home remedies or by *curanderos,* since it was believed that physicians did not recognize, and hence could not be expected to cure, such cases. Illnesses with magical or emotional etiologies, such as the evil eye or *susto* (fright), were the ones most frequently judged to be unsuitable for treatment by the modern physician (Foster 1952). Infectious diseases quickly relieved by antibiotics or surgery fell into the other category. Erasmus found in Ecuador that even *curanderos* admitted that the physician was best qualified to treat things such as diphtheria, tuberculosis, venereal disease, and appendicitis, and that they had "nothing but praise for the yaws campaign and a complete conviction that yaws was a disease which the doctor alone was capable of curing" (Erasmus 1952:416, 417).

This basic dichotomy between illnesses the physician can cure and those he cannot cure has subsequently been found in widely separated parts of the world. Thus, in North India Gould found that "critical incapacitating dysfunction" (i.e., usually infectious diseases) were taken to the physician, while "chronic nonincapacitating dysfunctions" (i.e., degenerative, recurrent diseases) tended to be treated in the village (Gould 1957).

A second type of cognitive dichotomy, that between "our" diseases and "their" diseases, has also been noted. Madsen quotes a Mexican-American *curandero* from south Texas who makes clear this distinction. "I do not know why God has afflicted only us with the evil eye, *empacho,* fallen fontanel, fright, and many other diseases. But God is merciful. He sends us these afflictions but he also taught us how to cure them . . . . They come to us and only we understand them. Only we of *La Raza* can cure them. . . . God sends disease to punish us but he also lets us learn to repent and undo the suffering. His great concern must reflect a very great love for *La Raza.* You punish most the erring children whom you love best. Therefore, it is not strange that the Anglos never suffer these divinely sent afflictions" (Madsen 1964:79).

The Pima Indians of Arizona have a similar belief. The *ká:cim múmkidag,* or "staying sicknesses," are "Indian sicknesses that afflict only Pimans, are diagnosed only by Piman shamans, and are cured by native ritual curers" (Bahr et al. 1974:19). These are the sicknesses associated with the moral order; beliefs about their etiology encourage correct behavior. Staying sicknesses are caused by the "strength" of sacred or dangerous objects whose "way" has been violated by an evil or thoughtless human act. "In trespassing against a dangerous object's 'way,' the patient is said to have committed an impropriety against rules which were set down for Pimans at the time of creation. The *ká:cim* sicknesses were intended specifically for Pimans as a race of chosen People" (*Ibid.,* 21).

The Scot missionary-physician Barker found the same pattern among his Zulu patients: a strong belief that their diseases were African diseases, thus hidden from foreign eyes. "After all," says Barker, "it is human enough to believe our problems to be unique. We speak of *our* gall-bladders, *our* stomaches, *our* appendices, as if we owned them like so many lap-dogs to be pampered and fed" (Barker 1959:88). During the early years of his practice, most diseases that patients recognized in themselves were considered to be peculiarly African. "You wouldn't have been able to help," women would tell him, in commenting on the deaths of their children. "The child had one of the sicknesses of the people." The usual symptoms of these "people's diseases" were diarrhea, vomitting, coughing, high fever, which "indicated nothing very mysterious yet the belief persisted that such symptoms were an exclusively African privilege" (*Ibid.,* 88-89). Little by little Barker and his physician wife gained the confidence of his patients, who brought more and more "African" diseases to him for treatment. However, this did not indicate a weakening of belief about their exclusiveness; instead, it was a testimonial to Barker for his skills in having learned to treat African diseases!

3. Resistance to hospitalization

Sometimes people resist hospitalization, not only because hospitals have been perceived historically as places people go to die, but because hospital practices often conflict with traditional patient care. Many tribal and peasant peoples ritually dispose of the placenta following childbirth, burying it beneath the ashes of the hearth or placing it in a running stream. In Tzintzuntzan the

placenta beneath the three hearth stones symbolically ties the child to its ancestral home. When first offered the opportunity to deliver in hospitals or clinics, a number of mothers refused, because they were worried about the placenta disposal. In some hospitals whose clients share these beliefs, the placenta is delivered to the relatives of the mother for ritual disposal.

Other hospital procedures may be frightening. Clark tells of an elderly Mexican-American woman in a California hospital who, when past the acute phase of her illness, was forced by a well-meaning nurse to take a shower. Returning to her bed, she was struck by "bad air"; to the dismay of the hospital staff, she then insisted on going home where she could receive "proper care" (Clark 1959a:202). For this woman, as in Tzintzuntzan, bathing was not to be taken lightly; it was safe only when she was fully recovered. Clark also tells how postpartum food taboos discourage Mexican-American women from accepting hospitalization for delivery. The postpartum period among village Mexicans is a dangerous time for mother and child; the mother's diet is strictly limited, emphasizing chicken and excluding things such as fruit juices and certain vegetables, which are too "cold" to follow immediately after the heat of pregnancy and childbirth. Some mothers are worried that in hospitals they will be asked to eat foods that they believe are dangerous to them (*Ibid.*, 167, 227).

4. Differing perceptions of role behavior

When scientific medicine is first made available to people who previously have relied on traditional methods, one of the most important initial resistances stems from differing expectations between patient and physician about appropriate role behavior. These differences, and the frequent mutual bewilderment that results, have been partially explored in the section on the "Therapeutic Interview" in Chapter 6. Other examples further illustrate the problem. Clark, for example, in discussing health problems of Mexican-Americans in California, points out how Anglo doctors are taught that impersonal objectivity, professional "efficiency," an authoritarian manner (which connotes confidence?), and a right to "lecture" patients on what they should do are appropriate behavior for their roles. Mexican-American patients' expectations, however, are quite different. They routinely visit doctors accompanied by family members who, it is assumed, participate in all deliberations; this situation is reassuring to the patient, but hardly contributes to "efficient" use of often cramped space and busy schedules of doctors. Since Americans assume that at least some sickness is the result of the patient's failure to take routine precautions, physicians may suggest that a condition has resulted from failure to exercise proper precautions. Mexican-Americans, however, do not feel guilt or responsibility for illness. "A medical worker who implies that a sick person is at fault and is somehow responsible for his condition may find his statements received with indignation or hostility. To the patient and his family such a view is unjust or even malicious" (Clark 1959a:230). A leg ulcer responds to the penicillin, and not to the hand that injects it; the patient, however, responds to the person who injects, and even if the results are seen to be beneficial, an unpleasant therapeutic encounter will discourage future visits.

In this context Alland's comments on the Abron of the Ivory Coast are pertinent. "Confidence in Western drugs seems greater than confidence in doctors. . . . There is little ritual associated with Western medical treatment. The paraphernalia of the examination room is seldom seen, and examinations are usually cursory; thus the doctor often appears to be an unnecessary adjunct to the distribution of medicine" (Alland 1964:720). In other societies traditional peoples accept Western medicines, but view their dispensers with distaste. In North India Gould found that modern medical practitioners "are rarely able, because rarely inclined, to establish the kind of rapport that would win acceptance of *themselves* in the village" (Gould 1965:202). In developing countries physicians often feel that they are degrading themselves and their science if they listen sympathetically to recitals of magical or supernatural causes of illness. Sometimes they heap scorn on traditional beliefs, thus alienating patients who would like to turn to them because of their obvious clinical skills.

Most medical anthropologists believe that therapy in intercultural or interclass situations is most effective when the physician has an understanding of the medical beliefs and treatment expectations that the patient brings to consultation. Smithsonian Institution anthropologists studying public health programs in Latin America in the early 1950s stressed the importance of knowing folk medicine. "If public health personnel are acquainted with prevailing concepts of folk medicine, in many cases these beliefs can aid rather than hinder the doctor — the good or the useful can be separated from the bad and the useless, and programs planned with this in mind" (Foster 1952:10). The anthropologists pointed out that an understanding of folk medicine would aid health personnel in phrasing curative and preventive measures so that they would be intelligible to patients accustomed only to traditional medicine, and that confidence in the new would be instilled in patients if it were apparent to them that health personnel understood *their* beliefs (*Ibid.*).

More recently anthropologists working in intercultural and interethnic settings in the United States have come to the same conclusion. Harwood, speaking of Puerto Ricans in New York, writes that "To communicate effectively with a patient about his illness or treatment regimen, a physician must know something about how the patient conceives of disease, its etiology, and therapeutics in general. . . . In order to treat patients of a different sociocultural background effectively, the physician must . . . develop a special understanding of their medical beliefs and practices" (Harwood 1971:1153). Snow voices similar thoughts with respect to American minority ethnic groups. "It is important that the physician know what the patient has been using to combat the illness — if it is harmless, it might be left in the treatment plan and the

physician's own suggestions added" (Snow 1974:94-95). And Wintrob, a psychiatrist-anthropologist, writing of American blacks specifically and minority groups in general, believes that "Sensitivity to the culturally determined beliefs and practices of patients whose ethnic and social class background is different from that of the treatment personnel is essential in the assessment of those patients' illnesses, whether medical or psychological. And in evaluating the data one must take account of folk beliefs and folk medicine, no less than interpersonal stresses and intrapsychic factors" (Wintrob 1973:325).

5. Preventive medicine and the maintenance concept

Whatever the specific etiological views of traditional peoples or their image of the therapeutic interview, they share one major premise: illness is manifest through pain or discomfort. They find it difficult to believe — faintly ridiculous, in fact — that unknown to a victim, a serious illness can appear and develop slowly, making itself known only when it is too late to help the sufferer. When a person feels well, when his body functions normally, that is all one can ask. The time to seek medical aid is when one does not feel well, when one's body is not functioning normally. Since much of preventive medicine is based on the philosophy of taking action before illness appears, through immunization that prevents it, or through early detection that increases the likelihood of successful treatment, traditional peoples are not the best candidates for this branch of medicine.

In this premise about health, traditional peoples — probably all preindustrial peoples — reflect a much wider world view: maintenance, Western style, is little valued. Europeans and Americans have had two centuries in which to learn the importance of maintenance. Routine but elaborate care of machinery is basic in industry. In our private lives we have learned that automobiles must be serviced at regular intervals, that houses last longer if periodically painted, and that clothing can profitably be mended. Although we sometimes ignore or delay such prescriptions, there is concensus that we do so at our peril. We have aphorisms that, from childhood, impress on us the importance of maintenance: a stitch in time saves nine; never put off until tomorrow what you can do today; an ounce of prevention is worth a pound of cure; and so forth. With this basic premise about the importance of maintenance deeply engrained in us, we are good candidates for the proposition that maintenance of the human body (i.e., preventive medicine) is a sound idea. Preventive medicine is consistent with our wider world view, and we take it for granted that it is a wise policy.

In traditional and developing societies just the opposite is true; the maintenance of machinery in new factories is always a struggle, house roofs are repaired after they begin to leak, highways are rebuilt more often than they are repaired and, in some areas such as Africa south of the Sahara, the mending of worn clothing is conspicuous for its absence. Lacking basic agreement that maintenance is essential to a smoothly running society, it is not surprising that preindustrial peoples are less receptive to maintenance of the human body through preventive medical activities than are people whose ancestors began the practice generations ago. Major forms of preventive medicine are not consistent with the wider world view of traditional peoples, and they will not accept it with the same relative alacrity with which Europeans and Americans have done so.

RESISTANCES IN SCIENTIFIC MEDICAL BUREAUCRACIES

When international public health workers first began to think in cultural terms, to realize that indigenous peoples do not necessarily perceive the benefits of Western medicine in the same ways that they did, they jumped to the conclusion that "cultural" and "social" barriers such as those just described were rooted in the ways of the target group members. Anthropologists working in early international public health programs shared this view. We believed that if the social and cultural forms of recipient peoples were studied, then it would be possible to design health programs that could clearly be understood by the people concerned. We assumed that people failed to understand Western medicine, and that if it could be presented to them in such a way that they could perceive its obvious advantages, they would be happy to accept it. Posparturient Mexican mothers fear many of the foods normally given to American women? Find out the foods that they feel are safe and construct a nourishing diet around these foods. Salvadorean village women believe that their husband's shirt, inside out, under the mattress, facilitates delivery? By all means allow them to have the shirt in the hospital. Physician's instructions are not phrased in ways that conform to, or build on, traditional interpretations of illness? Proper rephrasing is a reasonable goal that does little or no violence to the integrity of scientific medicine.

Public health and other medical personnal generally have been receptive to what anthropologists have told them about the customs and beliefs of the people with whom they work. After all, the problem is identified as being "out there," among the patients. Physicians, like most other people, prefer to believe that shortcomings in their performances are due to outside factors. Unquestionably public health and medical programs designed with an awareness of cultural factors, and the beliefs and attitudes of the target groups, have been more successful than those based only on the assumption that scientific medicine is best, and only a people's stupidity prevents their acceptance of it.

Yet as time has gone on we have slowly come to accept a disquieting thought: *at least as many of the resistances encountered in the promotion of scientific medicine are rooted in the medical profession and in health bureaucracies as in the target peoples.* Medicine's clinical assumptions, the statuses of its practitioners, its bureaucratic organization, the hours of its services — all of

these, and many more factors of its own creation, inhibit the acceptance of scientific medicine.[1]

1. Erroneous planning assumptions

Much international planning continues to be based on the assumption that the ways that have worked best in Western countries are the model to be followed in developing countries. Rifkin lists things such as "the belief that increased facilities constitute improved health. . . . ; that highly trained manpower is the only method by which to deliver health care; that medical care is exclusive of social factors including health education and preventive activities; and that disease and hospital-based systems are, without reservation, the most appropriate to meet health needs" (Rifkin 1973:249). Similarly, Bryant asks questions about medical schools in developing countries, giving the case of the University of Dakar in Senegal. Many basic facilities are excellent. Yet French influence, which accounts for much of this basic excellence, also raises disquieting questions; for example, the size and composition of the entering class is determined by the same rigorous examination system used in France. Consequently, in the 1964 entering class, only three of the 28 students were Senegalese. Moreover, the entire medical curriculum is planned in France and conforms to the same plan determined by the Ministry of Education for French medical schools. Bryant asks: "Is a curriculum designed for an advanced, industrialized country appropriate for a far less advanced, nonindustrialized country? Are the attitudes, skills, and concepts needed by a physician to direct the health affairs of a sector of Senegal sufficiently different from those required for France (or the United States or England) to warrant designing an educational program more specifically for that situation?" (Bryant 1969:63-64).

Bryant obviously believes that the answer lies in developing curricula adapted to the local needs. To illustrate his point, he describes a day in the life of a Western-educated Senegalese physician in charge of a hospital in which he, aided by a nurse and auxillary personnel, and by seven dispensaries, each with a nurse, serves 100,000 people! The doctor sees perhaps 150 patients every day. Five or six patients are admitted weekly to the hospital, and about ten women come for delivery. There is no midwife, no X ray, no running water, no facilities for simple surgery, and insufficient medicine. Time must be taken occasionally to visit the dispensaries, where services are rudimentary in the extreme. It seems clear that to maximize his utility, a Senegalese physician in a rural area must have many skills that are unnecessary in urban Europe.

Bryant believes that the "health team," led by a physician and composed of a variety of health workers especially trained for specific tasks and to work together as a unit, is the answer to health problems in developing countries. The major problem, he says, is that the physician simultaneously wishes to retain his traditional role as primary diagnostician and therapeutist, a role that obviously cannot be combined well with team direction. Instead, says Bryant,

"He must fill a role in which he manages limited resources to meet the comprehensive health needs of large numbers of people rather than serving as personal physician for a few. But there is widespread reluctance to accept this concept — it is in fundamental conflict with much of current medical education and professional thought as to what the physician's role should be" (*Ibid.*, 141-142).

A part of the greatness of the good physician, says Bryant, is his acceptance of responsibility to give unstintingly of himself to those who need his help. But this is also the basis for the physician's reluctance to share his activities with others, because it is intertwined with the feeling that no one other than a physician can provide that help. "A curious side of this concept is the value the physician places on the particular acts of diagnosis and prescription of treatment. Physicians are anxious to use every level of health worker in furthering a health program . . . but the words 'diagnose' and 'prescribe' evoke the strongest feelings of professional possessiveness" (*Ibid.*). The concept that the physician must attend personally to his patients actually determines the form of most health services, says Bryant, and it can obstruct efforts to change the design of health systems. "Thus while logic tells us that the physician's role should be determined by the health needs of the entire population, implementation of this logic is obstructed by the insistence of the medical profession that only physicians can evaluate and treat the sick. This stand of the medical profession has a paralyzing effect on the design and implementation of health services and is one of the most serious obstacles to the effective use of limited health resources" (*Ibid.*, 143).[2]

2. Clinical versus preventive medicine

Still another example of the imposition of inappropriate Western models on developing countries is the American distinction between clinical, curative, private sector medicine, and preventive, public health, public sector medicine. Considering American history, emphasis on private enterprise, numerous medical schools, and large numbers of patients who can pay for medical services, this dual system seemed reasonable for many years. It developed as a function of the reality of the total American cultural and economic pattern. But it was not God-given, inherent in nature, an answer for all peoples in all times. In most developing countries relatively few people have the means to consult private physicians. The curative needs of most of them must be met through government clinics, free or subsidized, through social security benefits, and in other ways in which curative and preventive services are combined.

Yet, more than 30 years ago, when the first cooperative programs with American aid began in Latin America, they were based on the assumption that health centers with maternal and child health and preventive services such as environmental sanitation and immunization were what was most important.

Whatever the merit of a health program based on preventive medicine, the average Latin American is primarily interested in doctors and nurses because he hopes they can cure his, or his children's, ills. When people for whom health centers were designed discovered the limited range of services and the relative neglect of curing activities, they were strongly critical. Criticism was particularly bitter when, in some instances, it was discovered that sick children would be examined by a doctor only if they had previously been registered in the so-called "well baby clinic" and were under health center "control."

Preventive medicine undoubtedly has contributed more than curative medicine in promoting longevity and high health standards. Certainly in the developing world major improvements in health levels will owe more to public health measures than to clinical services. But the prevention and the cure of illness are part of the same process and, through successful curing, people can be led to appreciate the importance of contemporary preventive measures. To assume that American institutional models of health care hold the solution to health problems of other countries is the height of ethnocentrism. Fortunately, the health services of most developing countries now realize this, and *their* personnel, not Americans or Europeans, are in the vanguard in the development of "social" or "community" medicine.

3. Personal priorities of health personnel

A third major problem in health programs based on scientific medicine is the frequent assumption of health personnel that their *personal* priorities are also those of the target group. Time after time we find that medical personnel fail to realize that they confuse their personal values in research and health goals with those of the people they hope to serve. Moreover, when we consider the total spectrum of problems faced by everyone, it is clear that health problems "compete" for priority ratings with economic problems, family problems, and a host of other kinds of problems. Busy Americans, thoroughly convinced of the importance of an annual medical examination, often postpone the visit to the doctor because they place a higher priority on other activities that compete for precious time.

A public health physician notes the frequency of cancer of the cervix among American women, and he knows that with regular Pap smears and early surgery, a high percentage of deaths can be avoided. He wonders at the apathy of most women, who do not share his consuming passion for control of cancer of the cervix, and he despairs of reaching his goal: annual examinations for all women. However, if one makes a list of immediate priorities for the average housewife — grocery bills, children not doing well in school, teenage truancy and perhaps brushes with the law on marijuana, a husband whose ardor has cooled, or a mother-in-law who does not hesitate to give advice — it is clear that an annual checkup for something that in any event is not apt to strike is of very low priority.

Even within the list of individual health priorities, the physician's own personal priority may be far down the list. In developing countries, public health personnel, not unreasonably, have stressed preventive measures. But without exception, on the priority list of the average person preventive measures fall below curative measures, i.e., immediate attention to their health needs as *they* perceive them. In the Rockefeller Foundation antihookworm campaign in Ceylon this became clear: "Some villagers were irritated by the concentration on hookworm disease in view of their other overwhelming medical needs. . . . The villagers were more interested in having their wounds and abcesses dressed and their miscellaneous acute illnesses attended than continuing in the dull routines of anti-hookworm work" (Philips 1955:289). Despite a home-office warning not to scatter their energies by engaging in such activities, field directors found they had to treat all kinds of complaints in order to gain support for the hookworm work.

The story of an attempt to bring pure drinking water to the village of Arenal, in Peru's Ica Valley, illustrates the same point. The regional health department, concerned with developing economical and technologically simple ways to achieve a protected water supply, offered to provide a concrete cover for the village well and add a hand pump. A health department sanitary engineer made the simple installation and turned the maintenance over to the villagers. A year later the village had returned to the use of a rope and pail to draw water. The problems of maintenance had proved insuperable. Various reasons account for the failure. Perhaps the overriding problem was that of priorities. "To be sure, Arenal could conceivably have used its meager funds for pump repairs and even for paying a stipend to a well-custodian. Had it done so, however, this would have meant diverting available resources from the important town celebrations (the patron saint's fiesta, the Independence Day celebration, and others). Much as Arenal favored the new hand pump over the former rope-and-bucket arrangement, the town fiestas were even more important than the pump" (Wellin 1966:123-124).

A similar story comes from a tribal group in Orissa state in India, where the National Malaria Eradication Program met with scant success. "They are pre-occupied with other aspects of their daily life, related to agriculture, construction of roads, and drainages which they view as problems. Though they mentioned 'fever' in interviews, they did not consider it a problem to worry about and, therefore, were not motivated to cooperate" (Dhilon and Kar 1963:20-21).

It is clear that unless the members of a target group share or come to share similar health priorities with the specialists who wish to help them, interest and cooperation in specific projects will be difficult to achieve.

4. Erroneous assumptions about decision making

A continuing problem in the doctor-patient and health educator-client rela-

tionship in the developing world is the tendency of medical people to assume that the patient or potential patient makes the decision about what kind of medical help to seek. In fact, medical decisions in the traditional world usually are group decisions, and matters such as status, rank, age, sex, and traditional roles may be involved. To illustrate, today in Indonesia family planning is a major government activity, with services provided through the regular health clinic network. In spite of apparent technical excellence, the program has been only moderately successful. Among the reasons for a low rate of acceptance of family planning has been the assumption, in the face of all evidence to the contrary, that it is the woman who decides to take this step. Hence, family planning education and propaganda are largely directed toward the wife, although she is the person whose desires are least important in the family. Major decisions are made by the husband and, on matters such as children and grandchildren, the views of grandparents are also highly important.

The same erroneous assumption about who makes decisions often has characterized maternal and child health services in new health centers in developing countries. Mothers and mothers-to-be seem like the obvious educational targets. But the older women in traditional families often consider themselves to be the repositories of wisdom on pregnancy, childbirth, and infant care, and their role as experts on these matters is widely accepted. Not surprisingly, to many of them a new health center may be seen as a direct competitor, an intrusive government activity that threatens their status and power within their families; therefore they oppose it vigorously. In families where the elderly women's roles are accompanied by real power they, as well as the young mothers, must be won over to the new system.

5. Deficiencies in health services

Still other factors discourage the acceptance of scientific medicine and preventive programs. Sometimes the reaction is negative because, in fact, the scientific medicine is not very good, its mode of delivery does not meet the needs of the people, or it is accompanied by side effects that are viewed adversely by the peoples concerned. For example, one of the most successful of all preventive measures (if we discount possible long-term consequences) has been residual spraying with DDT to eliminate the *Anopheles* mosquito vector that spreads malaria. Large parts of the world have been rendered completely free of malaria as a consequence of this activity. Yet opposition to spraying often has been intense. In many places cats have died in large numbers; they rub against DDT-sprayed walls, lick their fur, ingest DDT, and die. Since cats play the role of mousers in traditional societies, rat and mice populations increased, causing damage to food supplies. In tribal areas of Orissa state in India there were additional complaints: the spray gave a bad odor to the unventilated rooms in which people sleep; bedbug populations were believed to explode after spraying; and whitish DDT deposits disfigured colorfully painted house walls (Dhilon and Kar 1963:22-23).

WHO-UNICEF personnel found similar resistances in an antimalaria program among the Bush Negroes of Surinam; not only cats, but dogs and chickens also died from the effects of DDT and dieldrin. Furthermore, the cockroach problem increased drastically after spraying; the insects bit sleeping children and ate stored rice, cassava, and bananas. The head of one antimalaria team sprayed a box of cockroaches with insecticide, confidently expecting them to be dead by morning. The cockroaches, however, had become DDT-resistant and were alive and active when the box was opened. This demonstration did nothing to increase the villagers' confidence in the program's effectiveness. Finally, the Malaria Eradication Program workers, all Surinam nationals, were comparatively well paid by local standards. Consequently, they were both envied and feared by the local men; their relative affluence made them attractive to some married women who preferred them to their own less affluent husbands (Barnes and Jenkins 1972).

On other occasions it has been noted that vaccines have lost their potency because of inadequate refrigeration. When children fall ill of the diseases against which they were presumably protected, parents lose confidence not only in immunization but also in other modern medical procedures. Health personnel, too, may be lax about their duties. In the Indian villages that he and his colleagues studied, Banerji found general (but not universal) acceptance of smallpox vaccination. However, "The number of children who are left unvaccinated due to lapses of the parents appear to be a very small fraction of those who remain unvaccinated due to the lapses of the vaccinators and their supervisors" (Banerji 1974:11).

Other bureaucratic practices that at first thought appear to have no bearing on health services utilization can be viewed as deficiencies in those services. Family planning in Indonesia appears to have suffered because of pill procurement practices: the sizes, shapes, and colors change from time to time as the government, with U.S. support, gives succeeding contracts to the lowest bidder. The unexplained changes in what seems like a standard product make village women distrustful of the program.

These examples illustrate the kinds of factors that have often inhibited the acceptance of modern medicine among traditional peoples. They are not intended to convey the impression that traditional medical systems stand largely unchanged, fiercely defended by their supporters, that representatives of scientific medicine have learned nothing about the delivery of health care to traditional peoples. Just the opposite is true; in spite of the historic and continuing problems that characterize the introduction of scientific medicine to traditional peoples, more and more people are turning each day to the trained physician instead of to the indigenous curer. In turn, physicians who practice among patients from social or ethnic backgrounds quite different from their own are far more aware than a generation ago of the implications of this gap in

expectations. New therapeutic role expectations can be, and are being, learned both by physicians and patients; in many instances well-trained doctors show remarkable sympathy and understanding in dealing with patients from lower socioeconomic classes or different cultural backgrounds. Increasing numbers phrase their questioning and instructions to conform, as nearly as possible, to local concepts and terminologies.

In a recent study in rural Thailand, Goldschmidt points out that "Experienced doctors agree upon [the necessity of] trying to use vocabulary which is adapted to their patients' level of understanding," even though it distresses them that some of the words they must use are "impolite." They have learned, however, that "polite words may easily lead to mistakes confusing patients" (Goldschmidt 1972:11). Goldschmidt also found that many of the Thai doctors he observed showed great understanding of the sociocultural problems of their poor patients, and that they often went to considerable lengths to try to help them with these problems (*Ibid.*, 16).

Presently in Tzintzuntzan, villagers often express distrust for physicians in the abstract while at the same time they describe satisfactory encounters with them, often praising individual doctors for their skill and for the kindnesses they have shown to their patients.

6. Professional role conflicts

Cohen has called attention to an ethical and policy dilemma that, although of a somewhat different order from the "bureaucratic barriers" just described, can nonetheless be considered as a part of the wider problem of providing good medical care to all those who need it. Large numbers of illegal aliens live and work in the United States: 2 million, and perhaps many more. They are poor, they speak little or no English and, in self-defense, they maintain a low profile, avoiding agencies and services where their illegal status might be detected. Many of these people are hard workers; they have come to the United States to better themselves economically. They are lawbreakers, but they are not criminals.

Their health problems are at least as great as those of low-socioeconomic-class Americans but, in addition to the barriers that often prevent the latter from seeking medical help, the illegal aliens are also deterred by fear of exposure and deportation. The situation, says Cohen, "provokes fundamental questions regarding the expression of altruism and forms and functions of giving towards such groups" (Cohen 1973:184). Americans subscribe more and more (although far from universally) to an ideology of equal access to health services for all people. Yet they wonder if they should be taxed to provide such services for people who have violated American laws by coming to this country.

Cohen describes how many health professionals cope with the ethical dilemma that faces them. They recognize the moral obligation of the health

professional to aid those in need; they are also circumscribed in their actions by agency regulations that distinguish between eligible and ineligible clients. Cohen believes that the low-profile behavior of illegal immigrants tends to protect "the health care giver from potential conflict between the ideals of his calling, and the societal norms embodied in agency guidelines. The hidden status of these potential clients appears to have relieved health personnel from accountability for the full exercise of professional ideals. It contributed to the avoidance of conflict with the regulations which govern the health systems under consideration" (*Ibid.*, 188).

Since aliens tend to limit their contacts with public health agencies, health personnel are relieved in considerable measure of the compulsion to practice the "Good Samaritan principle." Professionals who express frustration because follow-up home visits often are fruitless because of the high mobility of illegal aliens are, in fact, relieved of the responsibility of justifying their (legally) questionable actions. These are ethical and policy problems, and answers satisfactory to all concerned are not easily found.

NOTES

[1] See Foster (1976a) for a discussion of this point in the historical context of international health planning.

The tendency to assign problems encountered in health programs to the target peoples instead of looking for possible problems in health bureaucracies is not limited to international health programs. In a study of an underutilized neighborhood health clinic in a large eastern U.S. city, Jones found that "the health practice of many people in the area could be explained in part by the poor treatment people received in public health clinics — a common complaint" (D. Jones 1976:225). The problem initially was seen by health personnel as rooted in attitudes shared by the target population. But, "as it turned out, part of the problem was traced to the attitudes of health professionals toward the poor" (*Ibid.*, 226). And part of the problem simply resulted from lack of thought of health professionals about how the health world looks to the poor. The clinic was located in what was known as a "fancy" hospital; many people, not believing that there could be a free outpatient clinic in the hospital, made no attempt to patronize it. Others tried unsuccessfully; there were no signs directing patients to the part of the hospital containing the clinic, and some receptionists either did not know of the clinic or were reluctant to speak of it (*Ibid.* 226).

[2] Many observers in health roles, physicians included, may feel that this picture, while undoubtedly true in an earlier era, is overdrawn for today. More and more, medical doctors are not only willing, but anxious, to delegate much

560

of the diagnosis and treatment of patients with both acute and run-of-the-mill medical problems to other medically qualified personnel. In the discussion of changing nursing roles (Chapter 11) we saw how clinical nurse specialists have assumed responsibilities in intensive care units that would have been exercised only by physicians in an earlier period. "Physician assistant" programs, too, have been established at a number of medical schools, and the employment of men and women in this new role is gaining increasing acceptance by physicians who are hard pressed to keep up with the demands on their time. Medical practice, like nursing, is undergoing rapid change, and this will unquestionably be reflected in planning for health needs in developing countries.

525

DE LANDY D. Culture, Disease and Healing
N.York : MacMillan, 1977
562

VII Public Health and Preventive Medicine

Prevention of Illness and Social Control

The prevention of illness and the containment of disease are part of every medical system. But much more is involved than questions of sanitation and private and public cleanliness and robustness. For notions of contagion are bound up with religion and world view and with perceptions of the powers and intentions of one's neighbors and friends, not to mention strangers. Though intent may not be subject to control or change, behavior to some extent is, and, therefore, the medical system, especially with regard to contagion and sanitation, is directly hooked into local systems of social organization and social control. In the ultimate sense, public health as a separate field of study and practice is in fact the attempt by a society to control the behavior of its members for what the reigning groups and belief systems define as the welfare of the community as a whole. In preindustrial societies this usually includes also the control, or at least the effort to influence, the spiritual and cosmic forces that may attack the health and well-being of any mortal.

In some instances these societies evolve pragmatic means of sanitation that, whatever the intended religious or magical purpose, may in fact function to protect health and ward off disease. That is, through a long process of trial-and-error a society learns to invent at least some measures that will aid its quest for health and freedom from sickness. But empirical devices are not to be fully depended upon and appeal to spiritual beings must also be effected through ritual, prescribed ways of behaving and not behaving, and the use of fetishes, amulets, and talismans to ward off all forms of evil that may bring disease or other misfortune (Sigerist 1951:143–146).

Malevolent beings not only may be superhuman but human as well, and so the maintenance of health and control of disease also include systems of sorcery and countersorcery. Furthermore, in the sense that preindustrial societies place heavy emphasis upon individual responsibility and culpability in regard to one's own behavior, the maintenance of health requires the observance of a usually complex system of taboos, and for many illnesses it will be assumed that the individual is himself the culprit because he has transgressed a social, spiritual, or moral law.

As the papers in this and other sections show, such beliefs and practices, to be comprehended, must be seen within the context of the social structure, religion, and other institutions and processes of a society. Recently, especially with the

surge of interest by anthropologists in the functions and significance of symbolic behavior, this topic has undergone rather interesting analysis and theoretical discussion. One such study, *Purity and Danger: An Analysis of Concepts of Pollution and Taboo* by Mary Douglas (1970), provides some fascinating insights and hypotheses. Needless to say, this aspect of what has come to be called symbolic anthropology is of direct significance to medical anthropology.

In an essay that is too intensively organized and lengthy to be anthologized, Charles Hughes' (1963) "Public Health in Non-Literate Societies" makes a signal contribution to this area of study by drawing the topics of public health and preventive medicine into the purview of anthropology. Hughes finds that in a basic sense all health is public, that is, that no person is ever isolated completely from society so that his health necessarily is a cause of concern to others, and each ailing person, no matter how rare his disease, is a member of a class of persons suffering from the same disability. His review of the cross-cultural data impels Hughes (1963:166–174) to make four generalizations:

1. In any society, practices relevant to public health are generally neither wholly "magical" nor wholly "empirical" or scientific; they are a mixture of both orientations, and often may simply be habitual. . . .
2. The relations between the state of public health of any group and its way of life are reciprocal. . . .
3. In "primitive" societies the health of the public is a continual and encompassing concern of the entire group, which recognizes few bounds in protecting itself from disease or the threat of disease. . . .
4. A people's health is a function of a total life situation.

Hughes believes that the behaviors of preindustrial societies that have a direct bearing on the health of the community may be subsumed under the following categories: magico-religious prevention, empirical prevention; personal hygiene; clothing and adornment; physical culture; cosmetic and nutritional practices; occupational health, housing, and settlement patterns; handling of contagion, and other health-relevant practices. These propositions are profusely illustrated with ethnographic data and the reader is urged to consult the original of this important work.

Alland (1970:128–133) takes a rather dim view of preventive medicine in preindustrial societies, but though he feels it is not very systematic or effective, he nevertheless states that ". . . Much (certainly not all) of what appears to be irrational behavior will turn out to fit the prevailing ecological conditions and demonstrate that much of behavior is either health-oriented or produces maximization [of resources] in some other way." Alland even goes on to conclude that ". . . The most effective behaviors in relation to disease actually fall under the rubric of preventive medicine as it has been defined here rather than under native medical theory of disease, treatment, or disease prevention." But this is because Alland would probably define medical system more narrowly than I would. In other words, I would contend that in whatever segment of culture certain measures may result in the improvement of health and avoidance or control of disease, these should be included as parts of the system of public health and preventive medicine, and, therefore, as a part of the medical system. That such measures may also be seen as religious, economic, political, and so on in no way gainsays this point.

25 The Role of Beliefs and Customs in Sanitation Programs

Benjamin D. Paul

Although health professionals see health as a value of the highest priority and public health and sanitation measures are viewed as the most important obligation of a community to its citizens, health is not a value of equal priority among the people of a society or the power structure that governs them. In this paper Benjamin Paul indicates why public health workers—in this case sanitation engineers, but the same lessons apply to all health workers—must "case" the culture of the community with which they intend to deal. As in the assessment of Paul's theoretical contributions to medical anthropology in the earlier paper by Wellin, Paul treats culture as a system of interlinked beliefs and customs. Introducing sanitation programs is an exercise in culture change, and Paul shows why the health worker must identify and understand the linkages of that sector of a culture he or she wishes to change instead of simply presuming that the people should be able to grasp the "rationality" of the beliefs and practices being introduced. As Paul says in his casebook of public reactions to health programs (Paul 1955),

> If you wish to help a community improve its health, you must learn to think like the people of that community. Before asking a group of people to assume new health habits, it is wise to ascertain the existing habits, how these habits are linked to one another, what functions they perform, and what they mean to those who practice them.

Polgar (1963) describes four fallacies that commonly afflict public health programs: (1) the fallacy of the empty vessels (the subject populations do not have established health customs and are empty vessels waiting to be filled with whatever health program is being advocated); (2) the fallacy of the separate capsule (health beliefs and practices comprise a bounded, separate capsule of behavior and cognition apart from the remainder of culture); (3) the fallacy of the single pyramid (the communication structure of a society is organized within the social units as a single pyramid so that information and behaviors poured in at the top will trickle down to all levels); and (4) the fallacy of the interchangeable faces (all clients are alike). All of these errors derive from a lack of understanding of the nature of culture and social organization and their involvement in public health practices. Incidentally, the high valuation placed on health not only is not necessarily shared by the members of another society but may not be shared by members of the health worker's own society, either. Here I differ a bit from Paul's statement in this paper: "An engineer can construct health facilities in his home area without worrying too much about the cultural characteristics of the people who will use the facilities." Members of our own society still place the pleasures of smoking, eating, easy living, and many other practices far above their valuation of health. The billions of dollars spent on advertising demonstrate that most Americans have to be persuaded to be concerned about their health (or have their anxiety level needlessly raised over what may not really be a critical health problem at all).

Reprinted from Benjamin D. Paul, 1958, "The Role of Beliefs and Customs in Sanitation Programs," *American Journal of Public Health*, *48*: 1502–1506, with permission of the author and the American Public Health Association.

Man is a biological and social animal; he is also a cultural animal. He is cultural in that he runs his life and regulates his society not by blind instincts or detached reason alone, but rather by a set of ideas and skills transmitted socially from one generation to the next and held in common by the members of his particular social group. Culture is a blueprint for social living. Man resides in a double environment—an outer layer of climate, terrain and resources, and an inner layer of culture that mediates between man and the world around him. By applying knowledge which comes to him as part of his cultural heritage, man transforms his physical environment to enhance his comfort and improve his health. He also interprets his environment, assigning significance and value to its various features in accordance with the dictates of his particular culture. Among other things, culture acts as a selective device for perceiving and understanding the outer world. Since cultures vary from group to group, interpretations of the physical environment vary correspondingly.

Ordinarily people are unaware that culture influences their thoughts and acts. They assume their way is *the* way or the "natural" way. Interacting with others in their own society who share their cultural assumptions, they can ignore culture as a determinant of behavior; as a common denominator, it seems to cancel out. An engineer can construct health facilities in his home area without worrying too much about the cultural characteristics of the people who will use the facilities. Sharing their habits and beliefs, he has in effect taken them into account. But in another country with another culture, his assumptions and those of the residents may not match so well. In

parts of Latin America maternity patients of moderate means expect a private hospital room with an adjoining alcove to accommodate a servant or kinswoman who comes along to attend the patient around the clock. In parts of rural India the hospital should be built with a series of separate cooking stalls where the patient's family can prepare the meals, in view of cultural prohibitions against the handling of food by members of other castes. And of course the effect of cultural differences looms even larger where sanitation has to be built directly into the habit systems of people, rather than into structures and plants that serve the people.

Anyone familiar with the operation of technical assistance programs knows about the kind of behavioral differences I have mentioned. Unfortunately, however, it is easy to misconstrue these observed differences. Three kinds of misinterpretation are common. The first is to suppose that "they" have more odd beliefs and habits, while we have less of them. We tend to see them as captives of blind custom and ourselves as relatively free from cultural peculiarities. The fact is that all men are creatures of their culture with its inevitable admixture of rational and nonrational elements. Cultures differ and rates of cultural change differ, but peoples do not differ appreciably in the degree to which their actions are influenced by their respective cultures. We are quick to apply the term "superstition" or the epithet "uncouth custom" to the other fellow's manner of thinking or behaving. We may be repelled by the custom of eating domesticated dogs and yet impatient with people who would rather go hungry than eat their cattle. Americans take offense at the odor of night soil in the settlements of Korea and other parts of eastern Asia; a Korean gentleman on his first visit to New York was asked by a friend how he liked the great city, whereupon he replied: "Oh, very well, but the smells are so bad!" (Moose 1911). Measured by the standards of one culture the manifestations of another are bound to appear more or less arbitrary or bizarre. We need to realize that we have culture, too, and that our ways can seem as peculiar to others as theirs do to us.

Even allowing that our behavior as well as theirs bears the stamp of cultural conditioning, a second facile assumption is that our ways and ideas are more advanced than theirs, that they have yet to catch up with us. The trouble with this assumption is that it represents a partial truth: Some aspects of culture, notably scientific knowledge and technical skills, are indeed subject to measurement and relative ranking. But knowledge is not wisdom, and many aspects of culture, including language, esthetics, moral codes, and religious values, lie beyond objective rating for want of a culture-free standard of measurement. It is a mistake and an insult to imply, as we inadvertently do at times, that because some areas of the world are technically underdeveloped their people or their cultures are in general underdeveloped.

A third and particularly common shortcoming in our understanding of cultural differences is a tendency to view customs and beliefs as isolated elements rather than as parts of a system or pattern. The linkages between the parts of culture may be loose or tight and the connections are not always apparent upon first inspection, but it frequently turns out that people cling to a particular practice or belief not merely because it is familiar and traditional but because it is linked to other elements of the culture. Conversely, a change effected in one area of the culture may bring with it unexpected changes in other areas or may result in awkward dislocations, as the following illustration will indicate.

On the island of Palau in the western Pacific the pattern of living calls for frequent and large gatherings of people to celebrate or solemnize certain social events. In the old days, Palauan houses were large enough to hold many people. There were no partitions, and it was possible for each man attending a feast to receive his food in the order of rank and to sit in such a way as not to cause offense by turning his back on anyone. Since the last war, most Palauans live in small two-or three-room houses built in the Japanese or American style. They try to maintain the old customs but they have their troubles. Visitors overcrowd the small house and sit packed together on the floor. They must suffer the insult of having to look at a neighbor's back and must take their food in any order they can. The Palauans are incessant betel chewers—and spitters. The old houses had several doors and numerous floor cracks to accommodate this habit. The new buildings, especially the Quonset huts now being created for chiefs' dwellings and council chambers, have caused a minor crisis. The two Quonset doors are premium locations; knotholes in the plywood floors are too scarce to provide relief for the majority of chewers. Tin cans are coming in as spittoons, but these are in scarce supply (Barnett 1951:91).

Housing customs and hospitality customs, once closely linked in Palau, are now in strained relationship. It should be stated parenthetically

that social or cultural strains are not necessarily good or bad in themselves; depending on the case, they can lead to increased cultural disorganization or to an eventual reorganization of the sociocultural system on a new basis.

In some instances people strive to prevent cultural strain by resisting environmental and sanitary improvements. In rural India, fecal contamination of food and water by direct contact or contact through flies and rodents constitutes a difficult problem. The source of the trouble is the custom of defecating in the open fields. Use of latrines would go far toward solving the problem. Public health engineers and others working in India have devised special types of latrine adapted to the local squatting posture and designed to meet varied soil, climatic, and water supply conditions. Numerous latrines have in fact been installed, but follow-up studies reveal that only a small proportion are used regularly. Women in particular tend to avoid the latrines. Every morning and afternoon women go in groups to the field, not only to relieve themselves but also to take time off from busy domestic routines, to gossip and exchange advice about husbands and mothers-in-law, and to bathe with water from tanks located in the field. The linked habits of going to the fields for social gatherings and for toilet and bathing activities meet a strongly felt need for community living and relaxation from daily toil. In the women's view, defecation customs are usefully linked to other customs. In the view of sanitation specialists these customs are, harmfully linked to a cycle of contamination and intestinal disease. To disrupt the contamination cycle the women are urged to use the new latrines. They shy away from following this advice, partly because doing so would disrupt an ensemble of customs they prefer to keep intact, and partly because their culture has given them little basis for comprehending the connection between feces and enteric diseases.

I began by saying that culture mediates between man and his material environment. In an article analyzing the outcome of a rural sanitation program in a small Peruvian town, the author explains how perceptions of so common an environmental element as water are culturally screened:

> A trained health worker can perceive "contamination" in water because his perceptions are linked to certain scientific understandings which permit him to view water in a specially conditioned way. The Peruvian townsman also views water in a specially conditioned way. Between him and the water he observes, his culture "filters out" bacteria and "filters in" cold, hot or other qualities that are as meaningful to him as they are meaningless to the outsider.

An important part of the local culture is a complex system of hot and cold distinctions. Many things in nature, including foods, liquids, medicines, body states, and illnesses are classified as essentially "hot" or "cold" or something in between, irrespective of actual temperature. Sick people should avoid foods that are very cold, such as pork. "Raw" water is cold and fit for well persons; "cooked" water is hot and fit for the sick. The times of day when water can be boiled are hedged in by limitations of time and fuel and further restricted by "hot" and "cold" considerations. Water is consumed mainly around noon. Water boiled later in the day and standing overnight becomes dangerously "cold" and must be reboiled in the morning. So it is useless to boil it at any time other than the morning in the first place. The patient efforts of a local hygiene worker to persuade housewives to decontaminate their drinking water by boiling it met with only limited success in the face of these cultural convictions (Wellin 1955).

It is interesting to note that the hot-cold idea system now widely current in Latin America apparently goes back many centuries to the humoral theory of disease expounded by Hippocrates and Galen and transmitted by Arabs to Spain and by Spaniards to the New World, where it retained a place in formal medical teaching until the 18th century (Foster 1953). Folk theories of medicine in contemporary rural India and in other parts of Asia indicate that the humoral theory spread in that direction, too, if indeed it did not have its origin somewhere in Asia. In the course of its travels the humoral theory underwent modification, so that its present form in Asia is not identical to the one in Latin America. It is remarkable that cultural complexes such as the hot-cold idea system should persist, however altered, through such long periods of time.

Objectively viewed, the cosmos and all its contents are morally neutral; nothing is good or bad in itself; it simply is. But man clothes his cosmos in a moral cloak. He evaluates it, holding some things to be good and others evil. Values, the fundamental bases for choosing between alternative courses of action, are a central part of any group's culture. Values differ, but these differences are less apparent than differences in language, dress, posture, rules of etiquette, or other overt

features of the culture. Because values usually remain below the level of awareness, we are particularly apt to impose our own values upon others on the innocent assumption that we are merely helping them achieve better health.

Members of our own middle class tend to make a virtue of tidiness, apart from its possible bearing on sanitation. Cleanliness is both a health measure and a cultural value. This distinction can be appreciated if we glance back through history to see the shifting value assigned to bathing and cleanliness from the time of the ancient Greeks. Such a review also illustrates the connectedness of the parts of culture (Kroeber 1948:600–602; Sigerist 1941:69–72; Varron 1939–1940: 213–214).

Although they built no great baths, the Greeks valued athletic sports and despised the Persians for their false modesty in keeping the body covered. The Romans, taking over much of the Greek cult of the body, constructed enormous public baths where men of leisure spent hours daily. The early Christians set themselves against the established pagan religion and also against many of the attitudes and amenites inherent in Roman culture. Baths were construed as instruments of paganism and vice, as devices for softening the body rather than saving the soul. Before long, even minimum cleanliness by current standards was seen as the road to ruin. The ascetic saint was indifferent to filth; attention to personal cleanliness, especially on the part of a man, incurred the suspicion that one might not be too good a Christian.

Bathing occupied an important place, however, in the lives of Europeans during medieval times. As the vessel of the soul, the body needed to be preserved. The monastery of the early Middle Ages had its bathroom for friars and pilgrims. By the 13th century, public bathhouses had come into use in the cities, providing both steam and water baths along with haircuts and minor surgery. But the presence of food and drink, girls and music increasingly converted the bathhouses into places of amusement and eventually earned the opposition of the clergy. Moreover, the bathhouses became centers of infection when syphilis began to plague Europe at the end of the 15th century. Municipal bathing disappeared from the urban scene, private houses lacked baths, and the entire custom of bathing was condemned for reasons of morality and health.

Interest in bodily cleanliness was revived in the 18th century with the growth of enlightenment, the increase in comforts, the refinement of social manners, and the rise of the bourgeoisie. The lead in this direction was taken in countries where the new wealthy middle class became especially influential; hence the scrubbing of Dutch doorsteps and the proverbial Englishman with his portable bath. Today, in the United States, prosperity, democracy, and frequency of bathing have become linked values. Americans say that cleanliness is next to godliness, an indication that bathing and cleanliness are affect-laden values in contemporary middle-class culture as well as a means to better health. Yet even in the United States bathing is neither as old nor as general as people now assume. Ackerknecht reminds us that President Fillmore was as much attacked for buying a bathtub for the White House in 1851 as was Harry Truman in our time for his balcony.

We might have more success in exporting our technical means for improving the world's health if we could manage to divest these means of the values and other cultural trappings that accompany their use in the American scene. It might then be easier to fit our technical means into foreign cultural contexts. To do this we need to become skilled in perceiving our own cultural contours and those of the country we strive to help. This is one of the reasons why instruction in cultural anthropology and other social sciences is rapidly being introduced into schools of public health.

26 Social and Cultural Implications of Food and Food Habits

John Cassel

This chapter illustrates, through a descriptive analysis of a program for changing health practices among Zulu peoples in South Africa, many of the concepts and principles described in the preceding paper by Paul and in my introduction to this section. John Cassel, a public health physician and epidemiologist, exhibits a fine appreciation for the contributions to be made by social science knowledge to public health problems and by the use of social scientists on teams dealing with these problems and attempting to introduce cultural

Reprinted with abridgments from John Cassel, 1957, "Social and Cultural Implications of Food and Food Habits," *American Journal of Public Health*, 47:732–740, with permission of the author and the American Public Health Association.

changes. This paper is concerned only with food but similar experiences have been found in attempting to change other categories of cultural practices and values. (See the case studies in Paul [1955] in which much of the present paper first appeared in extended form.) Without in any way underestimating the factor of poverty in the malnutrition and undernutrition of these Bantu people, it is nevertheless clear that ingrained food customs (or "habits" as the author says, though I would reserve this term for individual practices, not group ones), as well as many other attitudes tightly wrapped in deeply held beliefs, some ancient and some, as the reader will see, only thought by the people to be of long tradition, obtruded in the path of healthful nutrition. The success of the program ultimately is reflected in the changes that took place after a dozen years in infant mortality and the near disappearance of pellagra and kwashiorkor.

. . . It is no coincidence that one of the major areas of interest common to both the health professions and the social sciences should center round the topic of food. From the dawn of medical history the role of food in health and disease has been under investigation by health workers, and its significance increases as further advances in nutritional knowledge are made. For social scientists a study of food ways and the system of attitudes, beliefs, and practices surrounding food may constitute an important technic in unraveling the complexities of the over-all culture pattern of a community (Du Bois 1941). Health workers are, in addition, now learning that food habits are among the oldest and most deeply entrenched aspects of many cultures, and cannot therefore be easily changed, or if changed, can produce a further series of unexpected and often unwelcome reactions. . . .

This paper is an attempt to illustrate by means of a case study some of these social and cultural factors of significance in programs designed to change food habits; an attempt will be made to derive certain general principles from this illustration; and finally some of the specific contributions that social scientists could make in such programs will be indicated.

Before considering these factors, however, it is desirable to elaborate somewhat on the degree to which we as health workers are "culture bound." The type of training the majority of us receive makes it difficult for us to see any merit in points of view or patterns of behavior different from our own. Food patterns for example which differ from our concepts of "good" practices are automatically labeled "bad." Attitudes and beliefs differing from ours are regarded as "illogical," "misinformed" or "wrong." The people who hold these different concepts are regarded as "ignorant" or perhaps "superstitious," "childlike" or plain "stupid."

Unfortunately there are still too many of us who are convinced that our own particular set of beliefs, attitudes, and practices is the only correct way of life and one that should be emulated by people of all cultures and all social classes. Such a philosophy on our part presumes that only we as professional health workers know what is good for all people. Furthermore it is frequently our fond belief that our nutritional education is being executed in a "knowledge vacuum" (to use the words of Edward Wellin 1955) as far as the recipients are concerned; that because the population we are serving knows nothing of our nutritional concepts, they therefore have no concepts whatever about nutrition. It is evident that documentation of the relevance of social and cultural factors to health programs is unlikely to lead to any effective application of these concepts as long as such attitudes are maintained by the majority of health workers.

THE CASE STUDY

In 1940 the Pholela Health Centre was established on a "native reserve" in the Union of South Africa. Health conditions among the Zulu tribesmen to be served by this center were extremely poor. The crude mortality rate was 38 per 1,000 population and the infant mortality rate 276 per 1,000 live births. Poor environmental sanitation and communicable disease represented two major health problems, but of greater significance, perhaps, was the extensive malnutrition that existed. Eighty percent of the people exhibited marked stigmata of this malnutrition and evidence of gross nutritional failure in the form of pellagra or kwashiorkor was common.

The Pholela Health Centre, charged with the responsibility of providing comprehensive medical care and health services to the population, was organized on the basis of a number of multidisciplined teams. Each team, consisting of a family physician, a family nurse, and a community health educator (health assistant), was responsible for the health of a number of families living in a defined geographical area. The major health problems of the area were defined by the teams and broadly classified into "unfelt" and "felt" needs. Health programs to meet these needs were then initiated on two planes in which the promotive-preventive and curative aspects of the services could be integrated.

Basically the function of the health assistants was to develop community health education programs whereby the "unfelt" needs could be made apparent to the community, and members of the community stimulated to change aspects of their way of life to meet these needs. Simultaneously curative and preventive services to individuals within their family units was the responsibility of the doctor and nurse on the team. Each individual was offered complete periodic health examinations, followed by a family conference to discuss the health problems of that family, and any illness in any family member would similarly be treated by his family doctor and nurse. At the clinical sessions the community health education programs were reinforced in the specific advice given the individual patients.

The case study to be presented here is a description of one program which was designed to change the food habits of mothers and infants. Dietary surveys revealed that the existing diet consisted principally of a single staple, corn, prepared in numerous ways, supplemented by dried beans, negligible amounts of milk and occasionally meat and wild greens. Potatoes and pumpkins were eaten seasonally, and millet (sorghum) was fermented to brew beer which was consumed in large quantities. Whenever funds allowed, sugar and white bread were also bought. Even though this was a rural agricultural community, poor farming methods and poverty of the soil made it impossible for the vast majority of the people to raise adequate food supplies. Consequently a large percentage of the food consisted of refined corn meal purchased with money sent home by the migrant laborers. Furthermore, during the best month of the year, if all the milk had been equally distributed, only one-twentieth pint per head per day would have been available. Having determined the existing food patterns an analysis was then made of the major factors responsible for the inadequacy of the diet. These included the extreme poverty of the people, the extensive soil erosion, certain traditions as to which foods were customary, and inefficient use of available resources. In addition prevailing cooking methods frequently destroyed a large fraction of what nutrients were present.

As anticipated, early attempts at group meetings and home visits to demonstrate the causal relationship between a poor diet and a low standard of health encountered considerable skepticism. People maintained that their diet had always been the diet of their people. On this same diet, they maintained, their ancestors had been virile and healthy and consequently there could be no possible relationship between the present diet and ill health.

It was difficult to refute this point without having any reliable information about the health of their ancestors.

A search of the available literature, however, revealed that the present diet had not always been the traditional diet of the Zulus and other Bantu-speaking tribes. Prior to the arrival of the whites, the indigenous cereal had been millet, corn having been brought into the country by the early white settlers. Because of its greater yield, corn had gradually supplanted millet as the staple cereal, millet being reserved solely for brewing. Furthermore, historically the Zulus were a roving pastoral people owning large herds of cattle; milk and meat had played a prominent part in their diet. So important was milk as an article of food that no meal was considered complete unless milk was included. The relatively fertile nature of the land at that time and the extensive wild game resulted in further additions of meat to the diet and a plentiful supply of wild greens. Roots and berries gathered from the forests were also extensively used.

These facts were presented to the community at the regular small group meetings initiated by the health assistants and discussion invited. In particular, confirmation was sought from the older members of the tribe, and in most instances such confirmation was received. This endorsement by the prestigeful and usually most conservative segment of the population insured that interest in the topic of food would remain high. Realization that the modern diet was not as traditional as had been supposed assisted in reducing resistance to the changes that were later advocated.

At subsequent meetings concepts of digestion and the functions of different foodstuffs were discussed. It was widely, though not universally, held that in some manner food entered the blood stream, but what happened to it thereafter was unspecified. Greatest interest was aroused by considering how a fetus was nourished in utero; discussion on this point would sometimes continue for hours. Many women were of the opinion that there must be a breast in the uterus from which the fetus suckled, but others indicated that calves in utero must be receiving nourishment in the same way, and no one had ever seen a breast in a cow's uterus. Some women maintained that the fetus was nourished by the placenta as the Zulu term for placenta literally means "the nurse." But when

challenged, they were unable to explain how the placenta could feed the baby.

The result of these and similar discussions was to arouse a desire for further knowledge. When the functions of the placenta and umbilical cord were explained by use of posters and models, the interpretation was readily accepted. Over a period of time the concept that body tissues were nourished by the food via the bloodstream became generally agreed upon.

As the concepts about the digestion and absorption of food began to change, it was possible to direct discussion to the function of different types of food. The view generally held was that all food had only one function, to fill the stomach and relieve hunger. This view was challenged, however, by a number of people who maintained that certain foods gave strength and others were fattening—fat being a bodily attribute much valued in this community.

Gradually it became more generally accepted that different foods had different functions, particularly in regard to infant health and nutrition. In attempts to improve infant nutrition, and in view of the existing resources, it appeared that green vegetables, eggs, and milk were the foods on which greatest emphasis should be laid.

INTRODUCTION OF GREEN VEGETABLES

Even though green vegetables had rarely formed part of the diet, and then only in a few families, their introduction for a number of reasons appeared to be a promising first step. Wild greens gathered from the forests had always been eaten and were enjoyed, but due to soil erosion these were becoming increasingly difficult to find. There were no very strong feelings or set of beliefs about green vegetables. The climate was suitable for their cultivation and each home had adequate space for use as a garden. Thus no marked objections to green vegetables was anticipated, their similarity to wild greens could be usefully exploited, and resources, at present unused, could be brought into use to make the vegetables available. In view of these circumstances the procedures that appeared to be necessary consisted of attempts to:

1. Increase the motivation of the people to eat vegetables—These included emphasis on their nutritional value in terms of concepts previously discussed; relating their taste to that of familiar wild greens; and emphasis on the financial savings that could be made by replacing part of the purchased food by home-grown vegetables. These views were presented at the group discussions held in the homes of the people and reinforced at the clinical sessions, particularly the prenatal and mother and baby sessions.

2. Demonstrate the means by which these vegetables could be obtained by the families - A demonstration vegetable garden was developed at the Health Centre where members of the staff could themselves gain practical gardening experience and where the families could have the methods of growing vegetables demonstrated. Vegetables from this garden were made available to certain families on prescription. The more cooperative families were assisted in starting their own gardens and as the number increased a seed-buying cooperative was initiated. Over the years a small market was established where families could sell their surplus produce, and annual garden competitions and an annual agricultural show were organized. As members of the Health Centre staff, particularly the health assistants, became more proficient gardeners themselves they were able to give practical advice and assistance to any of their families who desired it.

3. Demonstrate the palatability of the vegetables and the best means of incorporating them into the diet without destroying their nutrient values—Cooking demonstrations were organized both at the discussion groups and for the waiting patients at the Health Centre sessions. As far as possible, traditional means of preparing food were adhered to. Over the years different recipes were introduced and women from the area who were successfully incorporating vegetables into their families' diets were encouraged to organize these demonstrations.

The response of the community to these programs was satisfactory. In 1941 the first garden survey revealed that only 3 per cent of the homes had any form of vegetable garden, and that a total of five varieties of vegetables were growing. By 1951, 80 percent of the homes had gardens and were growing more than 25 varieties of vegetables. This response was all the more gratifying as the size of the area served had been increased sixfold over the course of the 10-year period, so that some families had had only relatively recent contact with the program.

INCREASING EGG CONSUMPTION

The attempts to increase the consumption of eggs encountered somewhat greater resistance than did the vegetable program. Surveys revealed that over 95 percent of the families had poultry, and that eggs were relatively plentiful at certain seasons of the year. Eggs were infrequently eaten, however, but in contrast to vegetables, very definite views

were held about them. It was considered uneconomical to eat an egg that would later hatch and become a chicken; egg eating was regarded as a sign of greed; and finally eggs were thought by some people to make girls licentious.

After prolonged staff meetings the consensus of team opinion was that none of these beliefs had any deep emotional associations and that probably the concept of it being poor economy to eat eggs was the most important of the factors preventing further use being made of eggs. The program based on this analysis therefore was patterned on the same general lines as the vegetable program. Technics for improving the egg yield whereby eggs could both be eaten and leave sufficient for breeding were discussed; the nutritional value and palatability of eggs were stressed both in the community education programs and for each specific patient at the prenatal and mother and baby sessions; and various methods for incorporating eggs into the diet without any marked modification of prevailing recipes were demonstrated. The fact that certain adverse views were held about eggs was of importance, however. Even though these views were not strongly held they indicated that relatively greater effort would be required to motivate the community to use eggs.

As might have been anticipated, therefore, response to this program was slower than to the vegetable garden program. In the course of 12 years however the technics proved relatively successful, and the consumption of eggs, particularly in infant diets, steadily increased. Furthermore, toward the later years of the program, families began to put excess eggs on the local market and found purchasers among their neighbors, a phenomenon totally foreign to this community.

THE MILK PROGRAM

Increasing milk consumption proved to be a considerably more difficult and complex problem. Not only was the supply extremely limited, but frequently the available milk was not being consumed. In particular, women took no milk whatsoever, and this became a matter of some concern to the Health Centre in regard to expectant and lactating mothers. Investigation disclosed that milk drinking was associated with very deep-seated beliefs and customs. Only members of the kin group of the head of a household could use milk produced by that man's cattle. This restriction applied equally to men, women, and children, so that no family could supplement its milk supply from another family outside the kin group, but the situation was more complex in the case of women.

During her menses or when pregnant, a woman was thought to exert an evil influence on cattle and was not allowed to pass near the cattle enclosure or partake of any milk. This applied even in her own home. Since it was usually impossible for men to know when a woman was menstruating, it was customary to exclude milk from the diet of the majority of girls once they had passed puberty. When a woman married and went to live with her husband's family group she fell under a double restriction. Not only was she a woman, but she was now in the home of a different kin group. Consequently, of all people in the community, married women were most rigidly excluded from partaking of milk. Under two conditions only might she have milk. If her father presented her with her own cow at the time of her marriage she could use milk from that cow, or if her husband performed a special ceremony involving the slaying of a goat, she would be free to use any milk in his home. Because of the general poverty of the area neither of these two procedures was common although several decades ago they were not uncommon in certain families.

The reasons for these customs were lost in the mists of antiquity, not even the oldest people in this community being able to explain them. "This is our custom, and this is how it has always been" was the only explanation offered. In all probability the customs are closely related to native religious beliefs centered around ancestors. Even today when some 60 percent of the community are Christians, ancestors play a very important part, protecting a person from all manner of misfortune and ill health. The link between a man and his ancestors is his cattle, and ceremonies of propitiation involve the slaying of cattle. Accordingly, anything that might have an evil influence on the cattle would endanger the relationship between a man and his ancestors. In addition, a married woman continues to have her own set of ancestors derived from her family and may not interfere in any way with her husband's.

The degree to which concepts concerning milk were enmeshed in the over-all cultural pattern made it obvious that the general approach used for increasing vegetable and egg consumption would be of little avail. Considerable motivation for drinking milk already existed in the community, and attempts to increase this motivation would be unlikely to change existing practices. Attempts

were nevertheless made on several occasions to change these practices by stressing the importance of milk as an ideal food for lactating and expectant mothers, but were futile.

Analysis of the underlying beliefs indicated clearly that the barrier to greater milk consumption lay in the link between milk and the specific cattle from which it came. If milk from cows which did not belong to any member of the tribe or any other related groups could be introduced into the community, presumably the barrier would be overcome. The most practical method of accomplishing this was to make powdered milk available. Accordingly, supplies of powdered milk were obtained by the Health Centre and offered to families on prescription. No secret was made of the fact that this powder was a form of milk, but it was stressed that this milk did not originate from cows belonging to any of the Bantu people.

From the inception of this program it was clear that no stigma was attached to the use of the powdered milk. Even the most orthodox of mothers-in-law or husbands had no objection to this powder being used by the young women of their families. The only barrier that remained was that it was an unfamiliar food stuff with a strange taste. To overcome this, the familiar formula —increase the motivation, help make the product easily available and demonstrate methods for incorporation into the diet—proved all that was necessary. The application of these technics produced a demand for the powdered milk that, over the years, steadily outgrew the Health Centre supply. Families were then encouraged to budget their slender incomes to allow for its purchase from the local stores. By 1954 it became apparent that in some families the demand for milk had become too great to be met from either of these two sources. To the gratification of the Health Centre staff permission was given in a number of the more educated families for women to consume milk from the family cows without any marked reaction from the rest of the community.

Some of the results of the Health Centre's programs are of interest. In the course of 12 years the infant mortality rate dropped from 276 to 96, pellagra and kwashiorkor all but disappeared from this area and the average weight of the babies at one year had increased by two pounds. This trend, as far as was known, was not occurring elsewhere in the country, and in fact in neighboring areas where the Health Centre provided only medical care but no community health education program, no such changes were discernible.

SOME GENERAL PRINCIPLES

From this illustration it is now possible to derive some guiding principles indicating the significance of social and cultural factors to health programs in general.

The first is self-evident. Health workers should have an intimate detailed knowledge of the people's beliefs, attitudes, knowledge and behavior before attempting to introduce any innovation into an area. While this principle is frequently violated in practice it is certainly no new concept in public health. What is not so well recognized however is that the intimate knowledge of these factors is but the initial step in the evaluation of cultural factors.

The second principle, which is usually more difficult to apply, is that the psychologic and social functions of these practices, beliefs, and attitudes need to be evaluated. As stated by Benjamin Paul (1955), "It is relatively easy to perceive that others have different customs and beliefs, especially if they are 'odd' or 'curious'. It is generally more difficult to perceive the pattern or system into which these customs or beliefs fit." It is in this area of determining the pattern or system into which these customs or beliefs fit that social scientists can probably make their greatest contribution to health programs. This is the knowledge that will help determine why certain practices obtain, help predict how difficult it will be to change them, and give indications of the technics that can be expected to be most helpful.

A third principle that should be emphasized was unfortunately not well illustrated in the example but is of fundamental importance in the United States. It should be appreciated that while it is permissible for some purposes to consider an overall "American Culture," numerous distinct subcultures exist, sometimes even within a single county. These subcultural groups must be carefully defined, as programs based on premises, true for one group, will not necessarily be successful in a neighboring group. This also is an area in which we as health workers can receive invaluable assistance from social scientists.

There are a number of very interesting illustrations of the importance of these subcultural groupings in regard to food habits in the United States, of which only two will be mentioned. Margaret Cussler and Mary de Give (1952) differentiate five major subcultural groupings in the South: white owners, white sharecroppers, Negro owners, Negro sharecroppers and wage laborers. To a greater or lesser degree each of these groups

had different attitudes, beliefs, knowledge, and practices in regard to food, and for different reasons. John Bennet, Harvey Smith, and Herbert Passin (1942) have distinguished in the southern part of Illinois eight separate and distinct subcultural groups, each with its own set of attitudes toward food. Interestingly enough, none of these attitudes were based on a knowledge or interest in the nutritional value of the foods.

Many nutrition programs today are designed for application to the total area under the jurisdiction of the nutritionist —an area that is often very large. The underlying premise would be that this total area—region, state, district, or even county—has a completely homogenous culture. Greater concentration of a nutritionist's time in a carefully defined subcultural group with the development of programs specifically designed for that group would eventually be more efficient in effecting permanent changes in food habits.

There are a number of other social and cultural factors, some of which have been implied in the case study and which, though important, have not been emphasized. These include a determination of the leadership patterns within a community, a definition of the decision makers in a family or larger institution, and determination of the status of various groups within the community and the status of the health worker in comparison to these groups. The importance of these factors has been well recognized by health workers in planning and executing programs and requires little further emphasis here. In this connection it is of interest to note however that several investigators (Cussler and De Give 1952; Lewin 1952) report that in the majority of American homes all major food decisions are made by the housewife. Concentration of the nutritional education on the housewife alone therefore would probably be more beneficial than attempts to educate the total community in many subcultures in the United States.

A further area in which profitable collaboration between social science and health could occur is concerned with the prediction of the long-range effects of any program. Following a program two interrelated series of questions would be of interest to all health workers. Are there likely to be any unanticipated repercussions of the program and, secondly, how permanent can the changes introduced be expected to be?

The answers to these questions will be largely determined by the degree to which the innovations are absorbed into the cultural framework and the impact they have on other facets of the culture. For example, in our illustration the effect of the acceptance of new dietary patterns by the Zulu women might conceivably eventually threaten the status of males as decision makers in that society. Similarly permission given to women to take milk might eventually be one of the factors changing the intensity of religious beliefs.

By indicating the possible effects on other segments of the culture, and by determining the degree to which the concepts embodied by the new program have been absorbed into the existing cultural framework, social scientists would have a significant contribution to make in sensitizing health workers to the possible long-range effects of their programs.

The concluding consideration is the degree to which we as health workers are "culture bound." Unless we can avoid cultural bias or ethnocentricity in our dealing with people, much of the crucial data will not even be made available to us by members of the community. Even should we be fortunate enough to have access to some of these facts, our analysis of the situation will continually be distorted through imposition of our own culturally determined system of values onto the behavior of others. In this regard social scientists can perhaps help us to become somewhat less "culture bound," but fundamentally for many of us this requires a major reorientation of our own philosophy which only we ourselves can achieve.

27 Ritual Purity and Pollution in Relation to Domestic Sanitation

R. S. Khare

The subject of cultural beliefs in ritual cleanness and uncleanness is one that has long occupied the attention of anthropologists, at least since the time of Sir James Frazer, Edward Tylor, and Hutton Webster. The topic is also dear to the hearts of theologians and students of comparative religion. The relation between the concepts and practices of ritual purity and those of

Reprinted with abridgments from R. S. Khare, 1962, "Ritual Purity and Pollution in Relation to Domestic Sanitation," *The Eastern Anthropologist*, *15*:125–139, with permission of the author and the Ethnographic and Folk Culture Society (Lucknow, India) and *The Eastern Anthropologist*.

a pragmatically hygienic nature are not always clearly spelled out in ethnographies or in studies of religions. And theologians not infrequently attempt to "prove" that the original purpose of many ritual practices and taboos was actually health-directed, often thousands of years after such concepts first appeared. In many anthropological accounts, moreover, little effort may be made to link the two, in either an emic or etic sense. This study strives to do just that, drawing upon the data of the complex religious system of Hinduism in a Northern Indian village. This is, in R. S. Khare's phrase, "an anthropological inquiry on domestic sanitation."

We see that whereas the extermination of the sources of disease—germs and insect pests—may be uppermost on the agenda of public health experts, they are not part of the villager's notion of cleanliness (or of preventive medicine). Nevertheless, although in some instances practices of ritual purification are irrelevant to health, and some practices of cleanliness are, in some instances, irrelevant to ritual purification, there are many instances and areas where they overlap, and as Khare shows, promote and reinforce each other. Even practices that are presumably secular and with a specifically hygienic intent may in time become ritualized. Khare shows that pollution varies by gradations in different sectors of life and of the domestic household, and that the same act or object may be variously polluting depending upon circumstance, actors, and so on. Thus, the system differently affects members of the several castes, with a claim of greater purity as one ascends the caste ladder though not in a simple fashion. Finally, Khare discusses some of the changes taking place under "modernization" and presents a three-stage model of modifications occurring in practices and ideas relating to ritual purity and pollution, both in terms of modernization and of Sanskritization. Khare urges public health workers to attempt to keep this process in mind when instituting changes in health and sanitation policies and behaviors. As he points out in a concluding passage not included in this abridgment, both the public health system and the traditional one of ritual purity promote health and have several points of overlap and potential linkage or transference, and the public health worker can utilize this knowledge to increase communication, efficiency, and understanding between himself and the villager.

I

The present paper seeks to establish that a relationship exists between the ritual state of purity and pollution and the physical states of cleanliness and uncleanliness in domestic surroundings. This relationship, it is assumed, will help in clarifying the role played by *Chhoot-Pak* concepts in domestic sanitation, and may also be helpful in pointing out its importance in orienting Action Programmes on environmental sanitation in rural India. This paper attempts to point out the nature of ritual purity and pollution involved in the domestic activities and also underlines the importance of studying it with public health aims in view. The data for this study were collected during 1958–1960 in a peasant village, namely Gopalpur, in Uttar Pradesh.

II

... It is a multi-caste village with a population of 572 people. Among the "twice-born" there are only Brahmins and Kayasthas. The other prominent castes are: the Ahir, the Kurmi, the Pasi and the Barhai. The numerically dominant castes of the village, viz. the Kurmi, the Pasi and the Ahir, are sanskritizing as well as modernizing themselves. They are observing rules of ritual purity and impurity, on the one hand, and taking to modern practices, on the other.

III

In Gopalpur, as in the whole of Northern India, several words, such as, *Pavitra-Apavitra*, *Suddha-Asuddha*, and *Chhoot-Pak* are used to denote the ritual states of purity and impurity. Similarly, the words *Saaf-Ganda* are used to refer to physical states of cleanliness and dirtiness.

Ritual purity and physical cleanliness are two things: one refers to the ritual or religious state and the other to the physical state. The terms, *Chhoot-Pak*, are used in two senses: one, in which they refer to situational purity-impurity of an individual "at a given moment in relation to other individuals, as determined by events occurring in his family or in his own life", and the other in which they refer to the place of a caste group in the ritual hierarchy of caste system (Dumont and Pocock 1959). Human beings suffer ritual impurity through *natural happenings* (birth, death, menstruation, etc.), association with *material objects* (human skin, excreta, menstrual blood, spittle, reproductive fluids, etc.), and *human beings* (untouchables, Mahabrahmins, etc.). Human beings also suffer from ritual impurity owing to the internal or intrinsic nature of individual selves—the *Suddha-Asuddha Atman* (pious and impious soul).

Ritual impurity calls for remedial actions on two planes: of relations with materials or men, and of

relations with the soul. These concepts of ritual purity and impurity (*Chhoot-Pak*) give meaning to individual as well as group conduct. As exemplified below, in the case of homesteads in Gopalpur these concepts thoroughly control the pattern of behaviour of an individual within the domestic group. These elaborate ideas are ultimately based on social inoculation, especially in childhood.

Here we are primarily concerned with the relations that various beliefs and practices of ritual purity and impurity have with physical cleanliness. Though the activities concerned with ritual purity and impurity are oriented towards the patterning of human behaviour in relation to a non-empirical referrent—the welfare of the soul, and the religious piousness—they also have a physical dimension which consists in employing certain operations resulting in the removal of dirt and filth, disposal of refuse, use of air, water, and the sun as cleaning agents, and avoidance of contact with dirty objects. Although these activities result in both ritual purification as well as physical cleanliness, the villagers often give much less importance to the latter than to the former, so much so indeed that a casual observer may come to conclude that the people are surprisingly unconscious of the physical aspect of the situation; almost as if they did not understand it. On certain other occasions, for example, a sacrifice, the people resort to physical cleaning of person, place or thing as a preliminary measure for leading up to the requisite state of ritual purity. It is due to these implicit or explicit hygienic dimensions in the concepts of ritual purity and impurity that they are relevant to an anthropological inquiry on domestic sanitation.

Ritual impurity may be of several kinds: it may be external or internal, temporary or permanent, voluntary or involuntary. The internal impurity is more serious than the external: the intensity of impurity through impure food and water will be greater than a physical (external) contact with a lower caste man. Dumont and Pocock (1959) analyse the severity of pollution in the latter case as attributable to "the notion of appropriation of the object; when a man uses an object it becomes a part of him, participates in him" (....). "And the point is that appropriation precedes absorption". "In the kitchen, cooking may be taken to imply a complete appropriation of the food by the household" (p. 37). The temporary pollution, which is more significant for domestic sanitation, is "situational, it is a function of the relations between an individual and the commensal, endogamous and local groups of which he or she is a member", while the permanent pollution "is a function of the relations between commensal and endogamous groups, between men and phenomena of the natural world, and among these phenomena themselves" (Srinivas 1952). The voluntary pollution is acquired through wrong behaviour, such as touching of excreta, taking water from a low caste-man, etc., while the involuntary pollution is the result of natural crisis over which man has no apparent control; for example, defecation, menstruation, birth, death, etc.

In an Indian village, the values of ritual purity and impurity form an important part of the *Karma* (action) and the *Dharma* (religion) of an individual; they help in attaining the *summum bonum* of one's life. Pollution from various evitable and inevitable sources is expunged as readily as possible so that one may live in a state of ritual purity—a condition favourable for "upward spiral of soul". According to Srinivas (1952), ritual purity, normal ritual status, and ritual impurity form a hierarchy. The normal ritual status, "or neutral state" is enjoyed by the person most of the time. The deviations from it towards purity and impurity are mostly temporary.

For our purpose, the word "cleanliness" may be used in the sense of habitual actions designed to free a person, place or thing from unclean matter that soils, anything which is out of place, or worthless. This definition of cleanliness is based upon the *Oxford English Dictionary*, and it avoids various specialised notions of the public health expert in order to fit in with the villagers' notion of cleanliness. The killing of germs, insects, and pests, as they are vectors of all sorts of diseases, forms the major part of the expert's notion of cleanliness, but this has scarcely any place in the villager's notion.

The villagers use the word "*Ganda*" for: (a) the unclean state of persons, places, and things which need to be washed, cleaned, removed or replaced, and (b) the actions or habits which tend to create unclean state or disorderliness. The words "*Ganda*" and "*Gandgi*" as used by the villagers are therefore something more than "*dirty*" and "*dirtiness*" of the English language. Hands in the cooking activities become dirty (*Ganda*) when they are soiled by ashes, earth, flour, etc., and have to be washed to be clean again. Similarly, once the cooking and the eating in the kitchen is over, everything in the kitchen contributes to the "dirtiness" (*Gandgi*) of the kitchen and is required to be removed, washed and cleaned before the commencement of next cooking. Not only this, actions leading to disorderliness, such as throwing

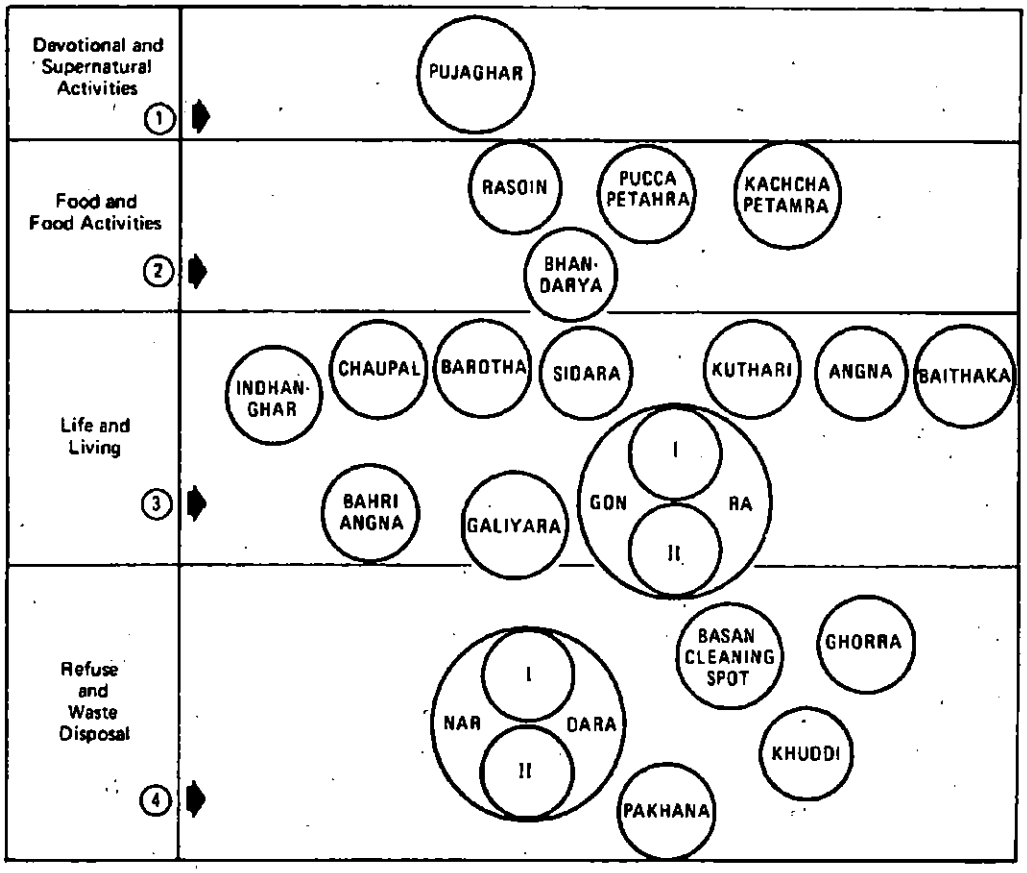


Figure 1.

of clothes on the floor, scattering of grain and scraps of vegetables in the courtyard or verandah make the place dirty. In a room if various articles, such as, clothes, utensils, provisions, baskets, etc., are out of their *proper* places (these are largely subject to a cultural definition) and are scattered in the room, they contribute to the "dirtiness" of the room. The habits or actions leading to such a state are called *gandi-adat* (unclean habits). Those habits, which in the villager's opinion help create dirty or unclean state, evoke disapproval. There can also be *ganda-kaam* ("dirty", obscene, indecent or unbecoming act) which may not have any relation with physical uncleanliness, as for example, the violation of a commensal or a sexual taboo.

In the domestic life of a villager, the difference between physical cleanliness and ritual purity is striking in some situations, while in certain others the two overlap and become coterminus. Sometimes the values of cleanliness and ritual purity are clearly identified without harming the role various practices of ritual purity play in promoting sanitation. The following situations may be recognized in this respect.

A. Dichotomous Situations

Within a homestead, certain places and things are dirty but not polluting, or certain things are dirty and polluting but the members of the household do not come into frequent contact with these, and pollution through them can therefore be easily avoided. The examples of the first category are cob-webs in the place of worship, vegetable peels, and scrap-heap within the kitchen, rodent excreta in flour, mud, and tiny worms in the drinking water, and flies. The examples of the second category are: drains, sewage, soakpits and ditches. These places are polluting as well as dirty. They can be cleaned but cannot be freed from the pollution which they permanently carry. A drain is a polluting object, howsoever clean it may be. Whenever contacted, it will make a person ritually impure.

The dichotomous situations can be of several types from the points of view of cleanliness: first, in which the purificatory steps also help promote cleanliness; second, in which the purificatory steps become irrelevant to it. In a third case, physical cleaning may be resorted to for its own sake. Let us take examples. The removal of scrap-heaps, vegetable peels and leavings of food from the kitchen, the disposal of ashes, and the plastering of hearth and floor with mud-paste are villager's essential purificatory steps which remove the ritual impurity of the kitchen and make it fit for the next cooking. These very steps promote the removal of dirt, ashes, filth and useless matter from the spot and accomplish cleanliness. But I wrote earlier that the notion of ritual purity is always a primary concern in the kitchen because the slightest ritual impurity in cooking or taking food immediately causes internal pollution--a state which should be scrupulously avoided. Permanent pollution can be caused in the household group due to such natural phenomena as, birth, death, and menstruation. In all these examples, the ritual impurity cannot be expunged at once. It takes its specified period of time before it can completely be expunged. But, it must, however, be noted that the severity of pollution decreases as the time that elapses after the polluting act. The impurity is also successively diminished (Dumont and Pocock 1959), as several intermediate purificatory steps are undertaken. For example, in birth-pollution, several *Nahans* (baths) successively diminish impurity of the mother and the child, and these allow them to return gradually to the normal state. These *Nahans* are essentially purificatory steps, but they also insure physical cleanliness. The room where the child is born (*Saur*) is thoroughly swept, washed, and plastered on the days of these *Nahans*; the bedsteads are washed with water; the beds are changed; and the clothes on the persons of the mother and the child are changed after they have undertaken the ritual baths. These baths are preceded by rubbing and massaging the bodies of the mother and the child with oil and/or *ubtan-paste* (a paste of flour in water and oil). The *saur* (nursery) is finally fumigated with double ends in view; that it cleans the atmosphere, and that it protects the mother (*Jachcha*) and the child (*Bachcha*) from evil spirits. Such *Nahans* are three in number and each of them marks a "stage" in the diminution of impurity.

Similarly, menstrual blood is not only polluting but also dirty and repelling. The purificatory rites on the third day, are preceded by several cleaning steps. The period of pollution extends as long as the blood does not stop flowing completely. However, in practice, after the first ritual bath, women begin doing limited domestic work, such as, peeling and scraping of vegetables, but not cooking them.

Not only in the removal of impurity, but also in raising the ritual purity of a person, place or thing, physical cleaning is resorted to. An apartment for normal living, is swept and washed afresh for making it fit for any ritual ceremony or for any religious festival. The *Satya-Narain-Katha* cannot be performed without cleaning an apartment which is considered fit for normal life. In any religious sacrifice (*Hawana*, or *Yajna*), such purificatory steps as the repeated plastering of floor with cowdung paste, repeated baths of the participating individuals, not only elevate them to the required level of ritual purity, but also promote cleanliness on the occasion.

In the above examples, the notions of ritual purity and impurity are dominant, and are the main concern on these occasions. The dichotomy between the ideas of physical cleanliness and ritual purity is marked. Yet, the purificatory steps involved on all these occasions promote sanitation or physical cleanliness. It can, therefore, be observed that the dichotomy need *not* necessarily preclude sanitation. In all the above cases, the practices of physical cleanliness ultimately help in expunging or diminishing pollution or raising the level of ritual purity. On such occasions the villager's view of cleanliness is "ritualised"; the steps of physical cleanliness find their full meaning and expression only in the context of ritual states. A place swept with a broom is clean but ritually impure, and hence as such it is unsuitable for worship, unless it is washed by water or cowdung paste. If one soils his feet with a polluting object like human excreta which is dirty as well as polluting, a wash with mud and soap may "clean" the feet but cannot "purify" the person, unless bath has been taken. In such examples, the practices of physical cleanliness are stepping stones, are a means to the end of acquiring ritual purity.

In the kitchen, even when cooking, the washing of hands and feet may be done with three purposes—for cleaning, or for ritual purity, or for both. One washes one's hands when they are soiled with, for example, flour. The aims of washing hands may be many: sometimes purely for cleanliness; and at other moments it may be necessary to wash them for ritual purity, for example, the hands require washing before serving the cooked vegetable kept "outside" the *Kachcha*-kitchen because the "contact" of a speck of wet

flour (*atta*) with vegetable will render the whole vegetable impure and will no longer be fit for consumption as Pucca food. In a third situation, while finishing the cooking activity, the hands are washed with water so as to remove the *atta* (flour) which is dirty as well as impure.

Under certain circumstances, the purificatory steps become irrelevant to physical cleanliness. Here the dichotomy becomes still more marked, and there is relatively more emphasis on ritual states. Water in a jug (*Lota*) may be clean but not pure on account of the impure state or nature of the receptacle. For a Brahmin, an earthen vessel is greatly polluting if it has been touched by a low caste individual. A brass or bronze utensil, which has been used for drinking and has not been rinsed, is impure and the water kept in it will also be consequently impure. The level of ritual purity of the container directly affects the ritual level of the thing contained in it. Similar is the case of *Dhoti* (loincloth), which is clean after it has been washed by *Dhobi*, but not pure (this example has been taken by Srinivas 1952, as well as Dumont and Pocock 1959, to elucidate the difference between ritual and physical states). Taking of *Panchgavya* is another example; it is extremely beneficial mixture for ritual purity (See, Dubois, 1906) and is a purificatory step of the highest order in Hindu religion, nevertheless it has a negative value with regard to physical cleanliness. It is, by physical standards, even a "dirty" preparation. Ceremonial *achman* (a purificatory step, usually before worship, in which a few drops of water are poured on the flattened palm and sipped three times), the sprinkling of *ganga-jal* (water of the Ganges) over a person, place or thing, and the differentiation between the *Kachcha* and *Pucca* foods, are some more examples which are symbolic of ritual states and may help raise the level of ritual purity. But they do not lead to physical cleanliness. The latter is rather put out of question or precluded in such cases.

In the third case, physical cleanliness is practised for its own sake in the normal life and even on those occasions in which the dichotomy is quite marked. Even the most backward people have notions of physical cleanliness and dirtiness, and they differentiate between what is clean and what is unclean. They have a set of practices designed to keep up the cleanliness. In normal living, cleanliness is valued, the cleaning practices, such as, washing of hands and feet, bathing when dirty, perspiring or feeling hot, sweeping of floors, and arranging articles in various apartments are followed by all the villagers according to their awareness and the emphasis laid by them on such activities. While worshipping a deity or cooking in the kitchen, cleaning steps are undertaken alongside the ritual practices of purity and impurity. It is a general habit to wash hands, whenever handling anything that soils. Those who work with mud in the fields, take bath to "clean" themselves when the work is over. In Gopalpur, as mostly in the whole of Northern India, a bath may serve a double purpose: first, that of removing dirt and second, that of either making the person ritually pure, or ending the impure state (Srinivas 1952:83; See also Dumont and Pocock 1959:17). In practice, a ritual bath is not characterised by any special feature; no *mantras* are chanted on the occasion.

The presence of rodent excreta in the flour, and mud and tiny water worms are dirty but not polluting. As far as they can, people remove the physical impurity from the water by sieving, decanting and picking, etc. Similar is the case with flies. These are dirty, but not ritually impure. If they settle on food, or on the person, it makes the object *dirty* but does not *pollute* it. If there is a fly in milk, or in a curry-bowl, most of the villagers throw the fly away as soon as they notice it and make use of the milk and curry. Such notions of cleanliness have become a part of the life of the villager and health educators would do well to attend them in right earnest.

B. Overlapping Situations

In those places where the values of ritual purity and physical cleanliness overlap, the ways of acquiring one also lead to the promotion of the other. The natural functions of the body are a great source of pollution; they are also dirty; for example, faeces, menstrual blood, spittle, hair, nail clippings, reproductive fluids, etc. These things are dirty as well as polluting, and therefore, those purificatory steps which are resorted to are also the steps for physical cleanliness. On all occasions, the contact with the above noted things is polluting, and in order to remove the pollution, one has either to take a bath or at least wash the affected part of the body. In such circumstances, the pollution as well as dirtiness are removed by one and the same method; the values of physical cleanliness and ritual purity, both are emphasised. In the above, such actions as: washings, ablutions, baths, removal of clothes, sweeping and plastering are the usual steps which purify as well as clean a person, place or thing. In domestic life, the occurrence of temporary and external pollution is inevitable as

well as more frequent (the casual permanent pollution due to birth, death and menstruation, has already been dealt with earlier). Since the household group is composed of those only members who belong to one endogamous and commensal group, the permanent ritual impurity from that source within the homestead is ruled out. Also in the house some places like drains, which carry permanent pollution, cannot be purified but only cleaned by such methods as flushing and scavenging.

IV

The degree of pollution is different in different apartments of the homestead. Thus the cleaning of the place of worship, the kitchen, the upper and lower portions of the courtyard (usually having drains, etc.), the drawing room, the compound and the cattleshed do not involve the equal degree of pollution.

This scale of ritual purity and pollution for various apartments of a homestead in North India offers sufficient support to the observation that physical cleanliness is practised mostly according to the main cultural theme of ritual purity and pollution. In the villager's opinion, ritual purity and physical cleanliness are two words necessarily implying each other, especially in domestic life.

Division of a homestead into various levels of ritual purity (see the Diagram) promotes as well as prevents sanitation. This is largely dependent upon: (a) the occasion, and (b) the strictness with which elaborate practices of ritual purity and pollution are observed in a household. In lower castes, for example, the Pasis (*Shakta* group) and the Chamars, it was observed that while their women were engaged in cooking food, they did not hesitate to touch broom, and, if necessary, even clean the courtyard or the verandah (but not the drains). This is never found with the women of the Kayasthas and the Brahmins. Once the women have taken bath, they do not engage in domestic cleaning; they usually do worship, or cooking or any other activity which cannot harm their normal ritual status. The cleaning of polluted places, if still undone, is postponed either to the evening or to the next morning.

Varying with situations, men, materials, places, and actions are all endowed with a certain kind and degree of ritual impurity in the Indian village. Accordingly, the associated cleaning activities are also polluting, and this pollution also requires to be expunged by following appropriate practices of regaining ritual purity.

... Individual as well as group differences in opinion and practice are always possible. For some Kayastha individuals, cleaning of cattle shed must be followed by a purificatory bath. For others, it is not so.

Broom is a polluting object and therefore touching of broom is also polluting. The degree of pollution that the broom carries is dependent upon the apartment which is cleansed with it. Thus, a broom used for clearing a courtyard, a verandah, and a room will be ritually more polluting than the one used for cleaning a kitchen or a place of worship. In terms of the diagram on the level of ritual purity in a homestead, the pollution carried by a broom increases from the 1st to the 4th level. Therefore, a broom used for the apartments of the 1st level is fit to be used for all others. In practice, as the broom gets worn out and old, it is employed for cleaning the apartments of 2nd or 3rd, or 4th level.

Only a new broom and unused broom (*bina barti jharoo*) is ritually fit for use at the 1st level. No broom which has even once been used at 2nd or 3rd, or 4th level, remains fit for the first level. In other words, 1st level cannot receive a broom used at 2nd, 3rd, or 4th level, while the 4th level can receive all brooms used in the homestead.

Touching of a broom is as much polluting as cleaning an apartment with it. Hence a touch with the broom is avoided especially when engaged in worshipping or cooking food, and if touched, appropriate purificatory steps should be undertaken, depending upon the ritual level at which the broom was used.

V

As cultures change, for whatever reason, people learn new things. The adults adopt new ways, "take them by and by under their moral umbrella by usage, and recommend them to the younger generations" (Dollard, 1945). Today in Gopalpur, the concepts of ritual purity and pollution are no exceptions to this remark. New learning situations are created in the village day by day. In Gopalpur, some people like Kayasthas are fast undergoing modernization and westernization. There are others like Kurmi, Ahir and Pasi who are undergoing sanskritization as well as modernization. Such a dynamics of the two processes of culture change is also operating on the concepts of ritual purity and pollution. Within one social

group, a caste or a sub-caste, the two standards, one sanskritic and the other modern are acquired together. Youths are claiming whatever is modern; while the old people, especially of lower castes, are increasingly observing details of ritual purity in the rules of commensalism, and in the activities of kitchen. Even within one household, the dual standards are applied today. A few educated Kayastha youths would not think it "clean" to squat in the kitchen-yard for taking food. Instead, they would like to sit on chairs and place the food on a table, while old people think it to be ritually polluting. But old people do not quarrel with youths on this matter, they are allowed to go their own way. Today, those people in Gopalpur who got one or the other apartment of their house sprayed with D.D.T. do not think that this spray is ritually polluting for those apartments. Still no body would like a place of worship, a kitchen, a pantry to be sprayed with D.D.T. because the ritual purity of these places will be affected. D.D.T. is a polluting substance for these places, a harmless and even useful substance for living verandahs, rooms, scullery, etc., and an "essential" thing, according to some Kayasthas and Kurmis, for the drains and ditches, refuse dumps which are the seats of flies and mosquitoes.

There is another example: thirty households of Gopalpur were sprayed with D.D.T. in the summer of 1959 because the youths of the house pressed their elders for it. Youths compete for such new "acquisitions". With the youths of the hamlet of Gopalpur, the spray of D.D.T. became an issue of social prestige—if Kurmi youths got it done, why should the Pasi youths lag behind. It is an important tendency where the old ways are less valued in the face of a *new* practice. The polluting dimensions of such things as D.D.T. gradually shrink and may even be lost with time.

The value of cleanliness is always culturally defined and changing with time; it may on occasions correspond to the modern notions of cleanliness owing to the vigorous emphasis on modernism. For example, among the Kayasthas of Gopalpur, the use of soap is generally not ritually polluting: it cleans the body as well as it does no harm to the normal ritual purity attained by a bath. For Kayasthas, the soap is no longer an item of polluting value. They use soap in bathing and then go for daily worship.

Formerly, (See also Dubois 1906) people did not eat tomato and carrot, because these had a "foreign" origin and were treated as non-vegetarian (ritually polluting) diet. Today, nobody avoids eating them. They are eaten by the people of upper and lower castes alike. Similarly, leather shoes were formerly thought of as polluting, so much so that they were always kept outside the main domestic building. Feet were washed regularly before one entered the inner courtyard. At present, except the kitchen and the place of worship, people enter all apartments with shoes on without any harm to the normal ritual status of various apartments. In the past, villagers avoided drinking water from water-taps and hand pumps in the cities because both of these employed washers made of leather. Today hand pumps are greatly in vogue in the villages and the villagers prefer them over wells. To satisfy themselves the more strict persons would first rub the cock with muds and wash it and then take the water from it for worshipping and drinking purposes. In Gopalpur, however, there is no water pump but the people of the village, whenever they visit the cities, have experienced such difficulties. Today, no villager objects to the dropping of potassium permanganate (red medicine, *lal dawa*) in a well which has dirty and smelling water.

In Gopalpur the Kayasthas and a section of the Kurmis and the Ahirs represent the trend of modernization, while the Pasis, a section of the Ahirs, and the Brahmins represent the trend of sanskritization. Brahmins of the village are still the fountainheads of sanskritic tradition for the low and the artisan caste people. In the areas of ritual purity and pollution the change is occurring in the following stages:

Stage I

This stage is characterized by dropping out of the details of practices connected with ritual purity and pollution. The *Chhoot-Pak* complex in the kitchen is becoming less rigid with use, still the majority of them do not disregard it completely. At present, the youth do not mind taking food (*Kachcha*) outside the kitchen-yard; however, they may not wash hands and feet prior to taking *Kachcha* food. The woman who is preparing food in the kitchen does not think herself as polluted by touching and feeding her child, yet she may avoid touching cots, beds, clothes, brooms, etc.

Stage II

This stage consists in practising the "atrophied" version of ritual purity *along with* occasional resort to the modern methods of cleanliness. We may take as examples: the use of soap; the wearing of leather-band wrist-watches and leather shoes by

the people. While even going for worship, the nephew of the village priest thinks that mere washing of hands and feet does away with the "pollution" acquired through wearing leather. Generally, the people use double standards for observing *Chhoot-Pak*: one for intra-domestic life and the other for extra-domestic one. *Chhoot-Pak* can be maintained only in the former case, while in the extra-domestic life one has to be less rigid, adaptive and tolerant. The example of this trend can be found among the Ahirs of Rampur who claim to sanskritize fastest. Normally, they observe the commensal taboos rigidly while within the village. But when they go to sell milk in the city of Lucknow, they have to eat, out of necessity, at various way-side shops and restaurants. Among the Ahir youths, the laxity of behaviour in observing the roles of pollution is widely manifest outside and is finding its way into domestic life. This is affecting the socialisation of children in the family. They are acquiring what the villagers themselves call "*newlight*" (*nai roshani*).

Stage III

It consists in dropping out the old practices of purity and pollution without systematic replacement by the new ones. This stage is not well represented in Gopalpur. The educated "elite" of the village dropped several practices of *Chhoot-Pak* as "superstition", but did not replace it with, for example, sanitary practices. A gap which has thus been caused by the process of culture change may be used with advantage by health educators. Such people (only Kayasthas) know a good deal about the *Chhoot-Pak*, as well as the principles of sanitation, but they have not been geared to action to reorganise their existing system of practices. They need to be won round to the hygienic ideas. Such a section of the population may be easily amenable to health education, if it is given properly. . . .

Les mécanismes de reproduction du tabagisme chez les préadolescents québécois : mémo août 1987

6. VERS UNE STRATEGIE GLOBALE D'INTERVENTION PREVENTIVE

Lorsque l'on examine les différents programmes anti-tabagiques mis au point au cours des vingt dernières années, on se rend compte que les concepteurs de ces programmes ont successivement et à tour de rôle réinventé la roue. Avant de présenter, dans une première section, les principaux axes autour desquels se sont précisément organisés ces programmes, il nous apparaît important d'introduire le sujet par quatre (4) remarques générales. Ces remarques valent également pour la seconde section de ce chapitre, qui sera dédiée à un examen sommaire de ce qui se fait présentement au Québec, dans le domaine de la prévention du tabagisme auprès des jeunes. Nous conclurons notre chapitre par une troisième section qui discute les cinq (5) principes fondamentaux, qui devraient à notre avis servir d'assises aux interventions préventives, que le Québec se doit d'entreprendre rapidement et systématiquement, auprès des enfants et adolescents.

La première remarque a trait au fait que l'efficacité des interventions anti-tabagiques a été très rarement mesurée de façon rigoureuse. Bernstein et McAlister ont considéré l'absence quasi-totale de devis évaluatif dans l'actualisation des programmes, comme l'un des problèmes cruciaux en matière de "smoking cessation programs" : "Specific design problems have included omission of relevant control groups, confounding of independent variables, lack of baseline smoking data, failure to equate groups on important environmental, subject, or task variables, experimenter bias, and use of unreliable outcome data sources" (1976:90). Nous pourrions ajouter, si nous étions le moins du monde cyniques, que les chercheurs ont parfois crié victoire, à la vue de leur taux de succès dans la semaine qui suivait l'intervention, tout en s'abstenant de nous dire ce qui en était un mois ou un an après la cessation de l'expérience.

- Notre seconde remarque introductoire veut rappeler que les interventions anti-tabagiques ont surtout été dirigées vers les adultes jusqu'au milieu des années 70. Lorsque les enfants et les adolescents ont davantage occupé les préoccupations des planificateurs et des chercheurs, comme cibles d'actions préventives, les interventions ont essentiellement cherché à rejoindre les jeunes dans leur milieu scolaire, en négligeant à peu près tous leurs autres milieux de vie.

Il nous semble important de noter, en troisième lieu, que les programmes anti-tabagiques souffrent d'un manque chronique de base conceptuelle, qui intègre dans un modèle unique les différentes dimensions, psychologique, sociale et culturelle, qui sont à l'oeuvre dans l'initiation au tabac chez les jeunes. Les interventions s'adressent le plus souvent à un nombre limité d'éléments, intervenant dans la chaîne de production des fumeurs: par exemple, certains suggèrent de travailler au niveau de la personnalité en diminuant l'anxiété, d'autres enseignent à résister à l'influence des parents et des amis, d'autres encore insistent sur les dangers associés au tabac et ainsi de suite. L'échec de nombreuses interventions est probablement imputable à l'absence de globalité dans l'approche préventive.

- Notre dernière remarque générale touche à l'absence de coordination dans les actions entreprises au niveau d'une même région. Différentes instances gouvernementales reconnaissent que ce domaine est de compétence mixte, mais elles s'articulent mal les unes aux autres. A leurs côtés et indépendamment d'elles, nous trouvons les organisations les plus diverses, qui apportent leur contribution, sans que celle-ci prenne vraiment place dans un ensemble cohérent.

6.1 Les programmes anti-tabagiques (1960-1980) auprès des jeunes

La plupart des programmes, réalisés avant 1977 et décrits dans la littérature scientifique, ont été analysés et critiqués par Thompson (1978) et Schwartz (1978; 1979). Nous ne croyons pas utile de reprendre leurs conclusions, puisque celles-ci sont disponibles. Nous nous limitons à signaler que les programmes d'intervention auprès des enfants et des adolescents ont pris quatre (4) formes principales. Les campagnes sporadiques dans les écoles à l'aide de posters, de films, de discussions de groupe, entre autres, ont été le plus fréquemment employées, bien que l'on sache depuis longtemps qu'elles n'ont aucun effet significatif sur les comportements tabagiques. Les programmes, dans lesquels des étudiants d'un niveau supérieur sont utilisés à titre d'éducateurs des cadets, ont également été employés dans de très nombreuses écoles, mais nous ne disposons malheureusement d'aucune évaluation d'impact. On a aussi comparé, à plusieurs endroits, la performance respective de divers média d'enseignement ou/et de divers enseignants: enseignants réguliers versus enseignants spécialisés en santé, enseignement en classe régulière versus discussion hors-classe, etc. Tous les auteurs notent, enfin, que certains thèmes ou messages sont récurrents dans toutes les interventions: la cigarette dangereuse pour votre santé, la mise en évidence des effets physiologiques immédiats et des effets à long terme du tabac sur la aanté, les efforts pour changer les images associées au fait de fumer et à celui de ne pas fumer.

Thompson (1978) a clairement dégagé les quatre (4) concepts théoriques qui parcourent l'ensemble des programmes que nous venons de décrire succinctement. Nous mettons en exergue ces quatre (4) concepts, à cause de leur importance:

Principe I: la pression sociale et les supports naturels, spécialement de la part des pairs, sont utilisés afin de susciter des changements de comportement;

Principe II: les enfants et les adolescents doivent participer activement au processus d'apprentissage de l'abstention de tabac;

Principe III: le tabagisme est considéré comme étant une réponse conditionnée, développée face à certains stimuli, et cette réponse ne peut être éliminée qu'à travers un nouveau conditionnement;

Principe IV: la personne humaine est un être rationnel et elle agit dans son propre intérêt.

Depuis cinq (5) ans environ, les équipes de Evans (1978) à Houston, de Hurd (1980) à Philadelphie, de Perry (1980) à Stanford, de McAlister (1980) en Californie et plusieurs autres en différents pays essaient d'intervenir auprès des jeunes, dans une perspective un peu plus globale. Ces programmes, qui se déroulent sur un espace de temps relativement long, conjuguent plusieurs formes d'intervention: inoculation psychologique pour résister aux pressions de la publicité et à l'influence des personnes significatives qui fument dans leur entourage, expérimentation des effets physiologiques immédiats du tabac, utilisation des aînés pour influencer les cadets, discussion de groupe dirigée par le professeur, etc. La plupart de ces programmes sont encore en cours et nous ne disposons que d'évaluations très partielles de leurs résultats à court terme.

Dans l'ensemble, ces études laissent penser que les comportements tabagiques seraient réduits par les programmes utilisés. Bien que présentant certaines tentatives de contrôle, par l'utilisation de groupes-témoins, il reste que ces études souffrent encore de biais importants, qui empêchent toute interprétation valide. Ainsi, l'étude de Perry (1980) souffre d'un biais de sélection des groupes, vis-à-vis une variable importante et confondante: l'attitude de l'administration

scolaire face au problème du tabagisme est choisie d'une façon qui favorise la vérification des hypothèses de recherche, en raison de l'utilisation de groupes de convenance. L'étude d'Evans (1978) contamine allègrement le processus de mesure de la variable dépendante (comportement tabagique), par l'utilisation de pressions normatives confondantes, dans l'espoir d'améliorer la véracité du rapport verbal. McAlister (1980) et Hurd (1980) emploient une telle variété de modalités d'intervention, tel que souligné plus haut, qu'il est presqu'impossible de connaître l'élément actif ou les éléments actifs, afin de pouvoir les reproduire. Hurd (1980) reconnaît d'ailleurs cette limite à la validité externe, suggérant que la multiplicité des interventions pourrait expliquer les effets observés. Ces programmes ne pourraient donc pas être généralisés au niveau national, de façon économiquement acceptable, dans un univers sous contraintes.

Nous sommes encore loin d'une véritable intégration du problème, que représente la consommation du tabac, à l'intérieur d'un cours structuré d'éducation à la santé. Manque encore, dans tous les programmes, la nécessaire profondeur historique, profondeur qui va évidemment contre le désir d'avoir des résultats rapides. Nous aurons l'occasion de revenir sur ces deux (2) points.

6.2 Les interventions anti-tabagiques au Québec

Nous n'avons pas l'intention de présenter toutes les activités de lutte contre le tabac, qui sont présentement menées auprès des jeunes du Québec. Notre présentation veut surtout indiquer les orientations qui sont à la base des programmes, organisées à partir de trois (3) catégories d'institutions: les Départements de Santé Communautaire, le Ministère de l'Education et certains organismes privés, comme l'Association Pulmonaire du Québec.

Les programmes de santé scolaire dans les D.S.C.

Nous commençons par rappeler que la "Santé des jeunes de 5 à 20 ans" est de la responsabilité des Départements de Santé Communautaire. Pour remplir le mandat qui leur a été confié par le Ministère des Affaires Sociales, les D.S.C. ont mis en place des programmes de santé scolaire qui se limitent le plus souvent aux établissements scolaires. Il y a actuellement un mouvement critique important, chez les coordonnateurs(trices) et intervenants engagés dans ces programmes, afin de dépasser le "Concept de santé scolaire"; mais, la route conduisant à la mise au point de programme de "santé des jeunes" risque d'être encore longue.

Afin de pouvoir recenser avec précision les multiples activités entreprises par les infirmier(e)s des différents programmes de santé scolaire, nous avons pris contact en mars 1980 avec les responsables des programmes dans les trente-deux (32) Départements de Santé Communautaire qui oeuvrent en territoire québécois. Plus précisément, nous leur avons demandé de nous fournir une description de toutes les stratégies préventives employées dans les écoles primaires et secondaires de leur région, en indiquant les objectifs poursuivis, la population visée, les types d'activités réalisées, les ressources engagées et, éventuellement, les résultats obtenus.

Vingt-six (26) Départements de Santé Communautaire sur trente-deux (32) nous ont fait parvenir un dossier plus ou moins volumineux. Ce taux très élevé de réponse mérite à lui seul d'être signalé. Il est sans doute indicateur du fait que les activités ne manquent pas dans le domaine des interventions anti-tabagiques. Nous avons été impressionnés par la masse d'efforts fournis par de nombreuses personnes isolées, qui ont produit plusieurs documents fort voisins les uns des autres, sans qu'il y ait de contacts organiques. Les activités concrètes qui ont été réalisées varient, évidemment, dans la forme, d'un D.S.C. à l'autre, mais les sources d'inspiration se ressemblent assez souvent. Voici, à titre d'exemple ,

un certain nombre d'activités anti-tabagiques entreprises dans les divers D.S.C.:

- réalisation par les étudiants de diaporamas, de posters et de matériel visuel, servant de support à la lutte anti-tabagique;
- "les 24 heures sans fumée";
- création à l'intérieur de l'école d'un conseil étudiants-professeurs anti-tabac;
- adoption d'un règlement interdisant le tabac dans l'école, après référendum auprès de toutes les personnes de l'école;
- projection de films ("L'homme cheminée", "Brûler le temps", "Papa, pourquoi fumes-tu?"), avec séances de discussions;
- campagne pour un "air meilleur";
- invitation de sportifs connus et non-fumeurs, durant les cours d'éducation physique;
- mise au point de plans de cours, pour les professeurs désirant parler du tabagisme;
- concours d'affiches anti-tabac, avec prix;
- .....

Chacun crée, en quelque sorte, son propre programme et met au point son propre matériel pédagogique. Une telle dispersion a l'avantage de maintenir l'esprit de créativité chez chacun, mais elle a le net

désavantage d'user bien des forces et de consumer du temps, qui pourrait être employé à meilleur escient. Il nous faut reconnaître, de plus, que ces activités ont rarement été évaluées, qu'elles sont le plus souvent épisodiques (pendant la Semaine anti-tabac généralement) et qu'elles sont fortement structurées sur les liens entre santé et tabac. De plus, le contenu de ces programmes ne semble pas efficace, tel que souligné plus tôt, à l'examen de la littérature.

Nous signalons enfin que plusieurs D.S.C. utilisent, de plus en plus, le document Monigo pour intervenir dans les écoles primaires. Nous reviendrons plus loin à ce document.

Le cours de Formation Personnelle et Sociale du M.E.Q.

Il n'existe actuellement aucun cours systématique d'éducation à la santé dans les écoles primaires et secondaires du Québec et les dispositions organiques, qui régissent les "services personnels" aux étudiants (Ministère de l'Education du Québec, 1980), délèguent explicitement le domaine de la santé au Ministère des Affaires Sociales.

Le Ministère de l'Education du Québec vient de mettre au point un document comprenant cinq (5) volets (Ministère de l'Education du Québec, 1981), dont la santé et la sexualité, document qui devait initialement servir de base au cours "Formation Personnelle et Sociale", à implanter dans toutes les écoles publiques du Québec à compter de septembre 1981. Le volet du cours dédié à l'éducation à la santé, tout comme les autres volets, d'ailleurs, a été conçu de façon à être enseigné durant les onze (11) années de formation de base, depuis la première année du primaire jusqu'à la dernière du secondaire. Nous ne croyons pas nécessaire d'examiner ici le contenu de ce programme d'éducation à la santé, avec lequel nous sommes largement d'accord. Cependant, nous nous devons de signaler le progrès que représente l'entrée d'un tel cours dans le

curriculum de formation de tous les jeunes québécois. Nous regrettons, toutefois, l'accueil négatif que la population a réservé à ce document, sous la pression d'un certain nombre de groupes de citoyens et en référence unique au volet traitant de l'initiation à la sexualité. Nous espérons que la levée de boucliers et le recul prudent du Ministère de l'Education ne représentent qu'une phase transitoire, dans la voie conduisant à l'adoption d'un véritable programme d'éducation à la santé dans les écoles du Québec.

L'Association Pulmonaire du Québec

Plusieurs organismes interviennent présentement dans le champ de la prévention anti-tabagique au Québec, mais aucun n'a investi autant dans ce domaine que l'Association Pulmonaire du Québec. Cet organisme a lancé, en 1979, un "Programme de sensibilisation et d'information sur le tabagisme à l'intention des élèves du primaire", qui fournit aux maîtres tout le matériel didactique dont ils ont besoin pour actualiser le programme. Ce programme, qui est mieux connu sous le nom de MONIGO (initiales de Monoxyde de carbone, de Nicotine et de Goudron), s'adresse aux enfants de 5 à 12 ans, qu'il rejoint dans les sept (7) niveaux scolaires, qui vont de la maternelle à la sixième primaire. Les activités sont répétitives, d'une année à l'autre, et elles vont dans le sens d'un approfondissement progressif.

Lorsque nous examinons en détail les activités (fondation d'un club de non-fumeurs, organisation de comités anti-tabac, macaron, défilé anti-cigarette, affiche, serment d'abstention, pétition, etc.), nous réalisons qu'elles sont d'un intérêt certain pour les jeunes, qu'elles rejoignent dans leurs préoccupations. Cependant, nous avons l'impression que ces activités ressemblent d'assez près aux programmes que nous avons critiqués dans la section précédente, parce qu'ils n'avaient jamais été systématiquement évalués ou qu'ils sont inefficaces. De plus, la structure évaluative proposée, dans les milieux scolaires, ne permet pas de mesurer l'impact réel de la campagne sur les comportements.

Il faudra donc sérieusement s'interroger sur la pertinence de diffuser le programme Monigo, dans le contexte actuel des connaissances dont nous disposons, quant à l'efficacité de ce genre d'intervention auprès des jeunes. Le caractère militant et relativement violent du programme suscite également une sérieuse réserve.

6.3 Cinq (5) principes à la base de programmes d'intervention futurs

Les cinq (5) principes, que nous désirons mettre de l'avant, délimitent un certain périmètre ouvert à toutes sortes d'interventions préventives, à la seule condition que l'intervention en question soit une des combinatoires possibles. Nous n'allons donc pas, dans cette dernière section, proposer des recettes qu'il suffirait ensuite de cuisiner, après s'être procuré les divers ingrédients. En réalité, il n'existe pas de "ready-made packages" et nous sommes convaincus qu'il est urgent de prendre un peu de distance, par rapport au phénomène dans lequel nous voulons intervenir, distance qui permettra de jeter un regard davantage globalisateur et probablement générateur de nouvelles pratiques.

difficulté et fragilité politique

Le premier principe a trait à la nécessité de mettre en place une intervention globale et à long terme. La polarisation sur le tabac en négligeant l'alcool, les drogues, la mauvaise alimentation et l'absence d'exercice physique isole un élément de son contexte; on est, de plus en plus, convaincu chez les spécialistes de la prévention qu'il faut d'abord et avant tout faire émerger un certain style de vie. Le tabagisme pourrait n'être qu'un épiphénomène, qui resurgira aussi longtemps que le sol qui le nourrit n'aura pas été changé. Le cours que veut lancer le Ministère de l'Education du Québec possède justement ce caractère de globalité.

La durée des interventions doit être fondamentalement repensée. Sur ce point, nous disposons de l'exemple suédois: en l'an 2 000, au

terme d'un effort continu qui aura duré vingt-cinq (25) ans, les planificateurs espèrent avoir créé une première génération de non-fumeurs. Tous les enfants nés depuis 1974 grandissent dans un environnement totalement privé de tabac et, pour ce faire, il est bien évident qu'il faut la contribution de tous les adultes et la mise en place d'un contrôle gouvernemental strict. Cette intervention de longue durée repose sur la conviction que le contexte est primordial dans l'apprentissage du tabagisme et qu'il faut nécessairement créer une anti-culture, dans laquelle le tabac n'a plus de place.

- Le second principe concerne la positivité du message à transmettre. Il faut abandonner les attaques du tabac qui cherchent à développer un climat de peur et qui prennent souvent des allures de croisade; s'il faut faire une prédication, ce ne peut être que celle qui annonce des valeurs de remplacement, valeurs qui peuvent séduire et entraîner des adeptes. Il nous apparaît urgent de faire la liaison avec tous les mouvements sociaux qui cherchent à réaliser chez-nous une plus grande qualité de vie. Ces mouvements, qui s'inspirent généralement de l'écologisme, seront probablement dans l'avenir les principaux alliés des éducateurs de santé.

- Le troisième principe est de nature organisationnelle: il faut de toute urgence créer une structure permanente qui prenne en charge tout le domaine de la prévention. Cet organisme devra posséder des objectifs précis de travail, des moyens financiers et en personnel, ainsi qu'une autorité lui permettant d'inscrire ses approches dans le réel des lois et des programmes d'intervention. La coordination, que nous avons appelée de tous nos voeux, ne sera possible qu'avec la création d'un office possédant le mandat que nous venons de présenter. Le Conseil des Affaires Sociales et de la Famille constitue une pierre d'attente, sur laquelle pourrait se construire l'office dont nous parlons. Pour qu'un tel office voit vraiment le jour, il faudra que la prévention soit enfin prise au sérieux et

que l'Etat accepte d'y consacrer la part du budget qui lui revient.

Notre quatrième principe s'abreuve à la philosophie pragmatique: il faut immédiatement abandonner toutes les interventions, quel que soit le lieu où elles sont exécutées, dont l'efficacité n'a pas été démontrée. Les interventions indifférentes quant aux résultats représentent des charges importantes pour l'Etat et nous nous devons de rationaliser les dépenses des ressources. Il faut donc élever le "doute méthodique" au niveau d'un principe et accepter de soumettre toute intervention de prévention à un processus vigoureux d'évaluation d'impact, dans une perspective de contrainte des coûts à long terme.

Notre dernier principe est également pragmatique: il faut maintenir tous les efforts positifs de dissuasion. Parmi ces efforts, notons la limitation des lieux publics accessibles aux fumeurs, l'imposition de taxes sur le tabac, le fait que les hommes politiques limitent l'usage de la cigarette en public, etc. Ces différentes mesures concourent certainement à la création d'un contexte, qui ne peut à la longue qu'être bénéfique pour la qualité de vie, sans imposer de contraintes budgétaires supplémentaires à l'Etat.

Ces principes contiennent en germe toute la politique préventive qu'il faut appliquer dans le cas du tabagisme. Pour transformer ces principes politiques en actions concrètes, il nous faut nous rappeler ce que nous avons dit sur l'âge d'apprentissage et sur les différentes filières qui conduisent à l'habitude du tabac. C'est auprès des jeunes qu'il faut intervenir et il faut le faire en intégrant tous les facteurs qui entrent dans la genèse du tabagisme.

LECTURE POUR JEUDI LE 16 JUIN 1983

Rationalisation de l'évaluation et du contrôle dans les organismes de santé

par
Raynald Pineault, M.D., Ph.D.
Professeur adjoint
Département d'Administration de la Santé
Université de Montréal

INTRODUCTION

Les théoriciens de l'entreprise entrevoient les organismes sanitaires comme un type particulier d'entreprise qu'ils appellent : "organismes professionnels" [1]. Ils en ont identifié plusieurs espèces selon les pouvoirs exercés par le groupe dominant des professionnels sur l'orientation globale de l'entreprise [2]. Le trait caractéristique des organismes professionnels réside dans la dualité de leur structure des pouvoirs : (1) une structure administrative qui se limite à la gestion des services complémentaires (finances, personnel, entretien); et (2) une structure professionnelle qui réglemente la production des services rendus par les professionnels [3].

C'est ainsi que les médecins qui constituent le groupe dominant des professionnels dans les organismes sanitaires jouissent d'une grande autonomie en ce qui a trait au contenu et aux modalités de leurs activités [4]. Une conséquence importante de cette structure "schizophrène" c'est que l'administration de ces organismes en contrôle les ressources mais n'a pas grands pouvoirs sur l'élément le plus important de la production, à savoir les soins médicaux. Les médecins organisent leur propre système de réglementation qui a tendance à être officieux, peu structuré et conforme aux principes et aux normes institutionnelles qui régissent l'exercice de la médecine [5].

Une loi fondamentale en médecine, c'est qu'il faut laisser une grande latitude aux praticiens à cause du caractère "urgent" de leur travail et du grand risque d'échec qu'il comporte [6]. Les médecins considèrent donc qu'ils sont seuls à pouvoir pleinement comprendre les aspects techniques reliés à la pratique de la médecine et, par conséquent, que leurs confrères doivent être seuls en droit de juger lorsqu'il y a faute médicale. De tels principes sont en accord avec un système de réglementation professionnelle qui est fondé sur l'évaluation du processus des soins médicaux plutôt que sur

l'évaluation du résultat des soins, sur des mécanismes d'évaluation non structurés plutôt que structurés et sur l'évaluation confraternelle plutôt que la vérification externe.

Ces dernières années, la domination des médecins sur le système de distribution des soins a été très vivement critiquée. De plus, l'importance de leur rôle en tant qu'incitateurs de la demande de soins a été plus largement reconnue. Ajoutées à la montée en spirale du coût des soins, ces raisons ont incité les gouvernements de plusieurs pays — au Canada, aux Etats-Unis et ailleurs — à prendre des mesures pour stimuler la mise au point de méthodes et systèmes d'évaluation des soins médicaux. Cependant, il n'est pas facile d'implanter de tels systèmes, tout particulièrement à cause de la résistance déployée par les professionnels dont les activités doivent être soumises à une telle évaluation. Par ailleurs, les systèmes qui sont finalement appliqués restent encore sous le contrôle des professionnels et sont donc basés sur des mécanismes d'évaluation confraternelle et de vérification interne. Il est intéressant de noter que l'évaluation a été surtout florissante dans le domaine de la santé publique où les activités sont peut-être moins complexes que dans les organismes de soins et où les programmes peuvent être délimités clairement. Il convient de signaler aussi que dans le domaine de la santé publique, la domination des professionnels ne joue pas un rôle aussi déterminant ou du moins aussi évident que dans les organismes sanitaires.

Ces quelques remarques en guise d'introduction donnent une description des principales caractéristiques et contraintes du système institutionnel où se déroule l'évaluation des soins médicaux. Elles serviront de toile de fond à notre discussion. Le but de cet exposé consiste à montrer que l'évaluation des soins médicaux dispensés sous diverses formes est techniquement faisable dans les organismes sanitaires mais que le contexte institutionnel qui vient d'être décrit doit être pris en considération lors de la mise au point d'un système d'évaluation.

Plus précisément, les objectifs du présent exposé sont les suivants : (1) présenter une typologie d'évaluation qui serve de cadre de référence pour la classification des diverses approches employées dans l'appréciation des actes médicaux; et (2) proposer un système d'évaluation des organismes sanitaires, fondé sur une considération systématique et globale des soins médicaux qui y sont dispensés et dont un des éléments principaux est l'évaluation des résultats des soins.

CADRE DE REFERENCE

Le cadre de référence que nous proposons n'est pas une création originale. Il est fondé sur l'abondante documentation qui traite de la qualité des soins et de l'évaluation. Il puise tout particulièrement dans l'oeuvre de Donabedian et quelques autres auteurs qui ont apporté des éclaircissements conceptuels dans ce domaine [7].

Le cadre de référence suggéré comporte trois paramètres qui correspondent aux principaux éléments de la distribution des soins : (1) le niveau d'agrégation du "malade" (ce mot servant à désigner l'entité dont le service ou l'organisme de santé assume la responsabilité); (2) le niveau d'agrégation du fournisseur des soins; et (3) l'objet ou le sujet de l'évaluation [8]. Ces trois aspects ou éléments de la distribution des soins sont illustrés dans le modèle à trois dimensions qui est présenté à la Figure 1.

Le niveau d'agrégation du client peut être le malade, les membres de sa famille, la clientèle du médecin ou de l'établissement, une population-cible (allant d'une clientèle à une population participant à un programme), et, finalement, l'ensemble de la population d'une localité. Ainsi, bien que trois niveaux seulement figurent dans le modèle pour le simplifier, il y a plus de trois niveaux d'agrégation possibles.

Le second aspect du cadre de référence est la progression de l'agrégation du fournisseur, allant d'un seul praticien à une série de professionnels plus ou moins liés (liens non structurés à très structurés), à un programme, à un organisme ou à un système de distribution des soins.

Le troisième aspect du cadre de référence se rapporte à l'objet ou au sujet de l'évaluation, soit à la composante des soins médicaux que l'on veut évaluer. Ces composantes identifiées par Donabedian comme étant la structure, le processus et le résultat correspondent aux ressources, aux activités et aux objectifs d'un programme ou d'un organisme de santé [9]. A chacune des composantes des soins médicaux ci-dessus mentionnées correspond un type particulier d'évaluation qui a également été identifié par structure, processus et résultat. Nous discuterons de ces trois types d'évaluation dans la partie suivante. Avant d'aborder cette discussion, nous aimerions souligner l'utilité du cadre de référence qui vient d'être décrit tant pour le planificateur que pour l'administrateur sanitaire. Ce cadre montre bien que les trois paramètres doivent d'abord être définis pour amorcer l'évaluation. Par ailleurs, on se rend compte qu'il y a plusieurs combinaisons possibles (plus exactement 27) tel qu'illustré par les cases, et que la délimitation de chaque paramètre réduit le nombre de ces combinaisons du tiers dans notre modèle. Bien entendu, ces chiffres sont tout à fait arbitraires et dépendent du nombre de niveaux considérés pour chacun des paramètres. En fait, le cadre de référence aide l'administrateur sanitaire à procéder à l'évaluation d'une façon ordonnée, méthodique et logique.

Finalement, il convient de souligner que les niveaux d'agrégation des clients et des fournisseurs jouent un rôle déterminant dans le choix de la méthode d'évaluation. Par exemple, c'est plutôt l'évaluation des résultats que celle du processus qui est l'approche généralement adoptée lorsque le niveau d'agrégation de la clientèle correspondait à l'ensemble d'une population ou à des populations-cibles. On choisit par contre l'évaluation du processus lorsqu'il est question des soins sous la responsabilité directe des fournisseurs. Si cela s'explique dans une certaine mesure par une série de facteurs techniques et institutionnels, il n'en doit pas nécessairement être ainsi. En effet, l'évaluation des résultats est techniquement faisable lorsqu'il s'agit d'organismes de santé comme nous le verrons plus loin. En fait, on espère que ce cadre de référence aidera l'administration à se rendre compte que ce ne sont pas seulement — ni nécessairement — des considérations d'ordre technique qui justifient les modes d'évaluation appliqués traditionnellement aux soins médicaux.

ELEMENTS D'UN SYSTEME D'EVALUATION

Compte tenu du cadre de référence que nous venons de décrire, nous allons maintenant aborder l'aspect que nous avons appelé l'objet ou le sujet de l'évaluation et dont nous avons identifié trois composantes, à savoir la structure, le processus et le résultat. Tel que mentionné précédemment, on classifie habituellement les méthodes d'évaluation qui se rapportent à ces composantes en trois types ou approches du même nom : soit évaluation de la structure, évaluation du processus et évaluation du résultat.

Le Tableau 1 indique pour chaque composante des soins médicaux (soit les ressources, les activités et les objectifs), le type particulier d'évaluation, les indicateurs de mesure et les méthodes de collecte des données qui sont employés. Notre discussion de chaque type d'évaluation se fondera sur ce tableau.

A. Evaluation de la structure

Ce type d'évaluation mesure la qualité et la quantité des ressources utilisées. Il vise à déterminer si le genre et le niveau des ressources disponibles sont adéquats. Quoiqu'elle englobe les aspects matériels des installations et de l'équipement, l'évaluation de la structure va beaucoup plus loin et porte aussi sur les caractéristiques de l'organisation administrative et sur la compétence des professionnels de la santé.

Pour évaluer la structure, il faut avoir des données sur :

1. la qualité des installations et de l'équipement;

2. les caractéristiques de l'organisation : taille, catégorie, agrément, affiliation, organisation des médecins, spécialisation;

3. l'organisation administrative et la structure hiérarchique;

4. l'organisation du personnel : effectifs, catégories et qualifications, politiques établies;

5. les aspects financiers et connexes;

6. la disponibilité de médecins;

7. la disponibilité d'installations spécialisées : banque de sang, laboratoires et services de diagnostic.

On peut obtenir ces données à l'aide d'un simple système de comptabilité des diverses ressources utilisées.

C'est ce type d'évaluation qui a toujours été pratiqué par les administrateurs de la santé. Bien entendu, c'est à la fois le plus facile et le moins coûteux puisque la majeure partie des données requises est déjà disponible. Toutefois, il faut reconnaître que l'utilisation de cette méthode pose en principe que des soins de meilleure qualité seront dispensés lorsque sont réunis un personnel plus qualifié, des installations plus perfectionnées et une organisation financière et administrative plus saine. Les moyens comme les permis d'exercice, les permis d'exploitation et l'agrément sont en grande partie fondés sur cette hypothèse. Or si elle est valable dans certains cas qui exigent des soins spécialisés, elle ne l'est pas nécessairement dans tous les cas. Par exemple, Rhee a montré que le degré de spécialisation des médecins est lié à la qualité des soins seulement lorsque les médecins exercent dans leur spécialité [10]. Ce qui est difficile à déterminer c'est le niveau des ressources qu'il faut atteindre pour produire des soins médicaux d'une qualité donnée. Afin d'y parvenir, on est obligé d'établir la relation entre la structure et les deux autres aspects de la qualité, soit le processus et le résultat.

B. Evaluation du processus

L'évaluation du processus se rapporte à la manière dont on a utilisé les ressources et aux méthodes employées pour réaliser certains objectifs. Elle vise à déterminer si les soins ont été dispensés d'une façon appropriée. Le nombre des visites était-il suffisant ? Le séjour hospitalier était-il trop prolongé ou trop bref ? Les méthodes de diagnostic et de traitement ont-elles été appliquées en nombre suffisant et d'une manière adéquate ? Telles sont les questions que l'on se pose lorsqu'on procède à ce type d'évaluation.

Les données nécessaires à l'évaluation du processus proviennent des rapports administratifs courants, de la revue des dossiers médicaux et de systèmes d'information spécialement conçus à cette fin. Nous allons discuter maintenant du système d'évaluation du processus plus en détail.

1. Système d'appréciation du processus

Un système d'appréciation du processus comporte au moins trois étapes : (1) la ratification de l'admission et du séjour, (2) la ventilation et l'analyse des statistiques, et (3) l'examen et l'analyse des dossiers [11].

a) *Ratification de l'admission et du séjour*

La ratification de l'admission et du séjour a été le mode d'évaluation du processus le plus couramment utilisé. C'est une composante importante du programme PSRO (*Professional Standards Review Organization*) récemment implanté aux Etats-Unis [12].

On peut employer deux méthodes pour passer en revue les admissions et les séjours : (1) la revue rétrospective qui a été celle généralement appliquée par les comités des normes d'admission et de séjour; et (2) la revue simultanée qui se fait pendant l'hospitalisation du malade et exige une nouvelle ratification du séjour. C'est là la méthode exigée par le programme PSRO. La revue simultanée, contrairement à la revue rétrospective plus courante, fournit un outil de gestion qui permet au comité des normes de séjour d'aider, par la voie hiérarchique appropriée, le médecin responsable d'un programme ainsi que les directeurs du programme ou de l'établissement à garder la durée des séjours dans les limites appropriées et à réduire l'hospitalisation dans les cas où l'admission est injustifiée.

b) *Ventilation et analyse des statistiques*

La deuxième étape du système d'évaluation du processus réside dans la ventilation et l'analyse des statistiques. L'avènement de l'ordinateur a permis de

traiter, de ventiler et de résumer de grandes quantités de données en vue de l'évaluation administrative. C'est probablement la CPHA (Commission on Professional and Hospital Activities) d'Ann Arbor, Michigan, et l'HMRI (Hospital Medical Records Institute), en Ontario, qui ont été les premiers à faire usage de cette technique pour faire des études de la qualité des soins à une vaste échelle. La ventilation des statistiques, ou analyse des profils comme on l'appelle, constitue une autre composante importante du programme PSRO.

Dans ces différents programmes, les résumés des données hospitalières fournissent des données sur l'âge, le sexe, les diagnostics, opérations et maladies de chaque malade. Les résumés renseignent aussi sur un grand nombre de techniques thérapeutiques et diagnostiques et donnent diverses normes d'utilisation. On peut ainsi obtenir une répartition des données par diagnostic; par âge et par sexe des malades; et par diverses caractéristiques des médecins et des hôpitaux. A partir de ces analyses, il est possible de comparer les activités des différents médecins d'un hôpital, ainsi que le rendement des hôpitaux, et d'identifier ensuite les écarts de l'un à l'autre.

Un écart par rapport à la moyenne statistique ne signifie pas automatiquement que les soins sont de qualité médiocre [13]. La moyenne non plus ne représente pas nécessairement des soins de haute qualité. Les écarts attirent simplement l'attention sur les points qui nécessitent un examen plus minutieux. Juger si ce que révèlent les statistiques est "bon" ou "mauvais", acceptable ou inacceptable, appartient en fin de compte à l'hôpital et à ses médecins.

c) *Examen et analyse des dossiers*

Il est donc évident que la ventilation et l'analyse des statistiques ne représentent qu'une étape du mécanisme d'évaluation. Il faut la compléménter par un examen approfondi des dossiers médicaux car certains procédés peuvent exiger une investigation et des mesures correctrices. L'examen des dossiers revêt généralement la même forme que l'appréciation des actes médicaux où les dossiers sont passés en revue et évalués en fonction de critères professionnels explicites ou implicites. Les normes sur lesquelles on se base peuvent aussi être tirées empiriquement de l'observation des techniques thérapeutiques appliquées dans la pratique. Comme l'illustre la Figure 2, on peut classer les mesures de contrôle d'après diverses caractéristiques [14]. Ces mesures peuvent être internes ou externes. Le contrôle interne est exécuté par des confrères attachés au même organisme ou établissement; le contrôle externe est assuré par un expert ou un groupe d'experts agissant au nom de l'établissement ou de l'organisme responsable d'un programme d'action sanitaire. Point n'est besoin de dire que la première méthode est généralement mieux acceptée par les professionnels. Cependant, pour des raisons d'objectivité, le contrôle externe est probablement plus indiqué. C'est d'ailleurs l'argument invoqué par les fondateurs du programme PSRO en optant pour le contrôle externe.

Il y a une autre manière de classer les méthodes de contrôle selon qu'elles comportent l'examen des dossiers ou qu'elles se limitent à l'examen des statistiques. On a déjà souligné la complémentarité de ces deux méthodes. C'est donc dire qu'un système d'appréciation du processus doit englober à la fois l'analyse statistique et l'examen des dossiers. L'analyse statistique donne rapidement une vue d'ensemble des profils de rendement et permet d'identifier les points faibles où s'impose une analyse plus approfondie à l'aide d'un examen détaillé des dossiers. Nous avons mentionné plus haut qu'un écart par rapport à la moyenne n'implique pas automatiquement des soins de mauvaise qualité et que la moyenne n'équivaut pas non plus à des soins de bonne qualité. Pour plus de précaution, il faut établir un système d'examen des dossiers qui soit indépendant de celui de l'analyse statistique [15]. Ce système parallèle permet d'examiner des échantillons aléatoires ou appropriés de tous les cas ou encore des échantillons appropriés des cas qui peuvent être particulièrement révélateurs soit du point de vue de la détection ou de la correction des erreurs, soit du point de vue de l'atteinte d'un objectif particulier. L'échantillonnage peut se faire au niveau des catégories suivantes : hospitalisation de longue durée, hospitalisations de courte durée, admissions d'urgence, diagnostics à partir des symptômes, traitements médicaux ou chirurgicaux particuliers, cas dont l'issue défavorable est peut-être évitable [16].

Finalement, on peut classer les activités relatives à l'examen des dossiers selon qu'elles sont professionnelles ou administratives. En réalité, ces deux extrêmes sont rares. Le plus souvent, l'administration veillera à ce que des mécanismes de revue appropriés soient établis et appliqués. Cela ne veut pas dire que le jugement final à propos de la qualité des soins n'est plus du ressort des professionnels. Néanmoins, c'est le directeur général qui est responsable du maintien d'un mécanisme de contrôle de la qualité dans son établissement.

2. Validité de la méthode

Il faut admettre au départ que l'évaluation du processus suppose un lien entre des résultats sanitaires favorables et des aspects particuliers du processus des soins. Depuis quelques années, cette supposition est mise en question. Des études empiriques se sont attachées à l'examen des relations entre les processus de soins médicaux et les résultats en termes de santé. Les constatations les plus frappantes sont probablement celles auxquelles ont abouti Mather et ses collaborateurs en Angleterre [17]. A l'aide de la méthode des essais cliniques aléatoires (RCT), ces auteurs ont comparé le traitement hospitalier (y compris le séjour dans une unité de soins spécialisés) et le traitement à domicile d'un certain nombre de cas d'infarctus du myocarde. Les résultats de l'étude montrent que les malades admis dans une unité de soins spécialisés ne sont pas plus avantagés que ceux qui sont soignés à domicile.

Plus récemment, dans une étude prospective contrôlée, Hutter et ses collaborateurs ont comparé des hospitalisations de deux et trois semaines pour 138 malades ayant souffert d'un infarctus du myocarde sans complications [18]. D'après les résultats de l'étude, il semble qu'un séjour plus prolongé (3 plutôt que 2 semaines) ne rehausse en rien les "résultats sanitaires" pour le malade.

Ces deux exemples ne sont pas cités ici pour tirer des conclusions quant aux bienfaits de l'hospitalisation dans les cas d'infarctus. Ils visent à illustrer les moyens

de valider les techniques de soins médicaux qui sont appliquées par les médecins.

En omettant de le faire et en se fondant uniquement sur les normes établies par la profession médicale, on peut en arriver à une surutilisation ou à une utilisation peu appropriée des ressources. Les résultats présentés par Brook et Appel à la suite d'une étude de cinq méthodes d'évaluation confraternelle de la qualité sont fort révélateurs à cet égard [19]. L'évaluation du processus à partir de critères explicites a fait ressortir le moins grand nombre de cas acceptables (seulement 1.4 pour cent). L'évaluation du processus à partir de critères implicites a fait ressortir 23.3 pour cent de cas acceptables. A l'autre extrême, lorsqu'on demandait aux médecins si l'on aurait pu obtenir de meilleurs résultats à l'aide d'un meilleur processus thérapeutique, ceux-ci ont déclaré que le résultat n'aurait pas changé pour 63.2 pour cent des cas, ce qui revient à dire que les soins étaient adéquats pour 63.2 pour cent des malades. En résumé, les méthodes d'évaluation du processus ont indiqué beaucoup moins de cas acceptables que les méthodes d'évaluation du résultat. Ainsi que le souligne Brook, il y a une certaine ironie dans l'absence de corrélation entre les indicateurs de résultat et les évaluations implicites ou explicites du processus [20]. De telles constatations portent à croire que les médecins ont tendance à juger de la qualité des soins selon des normes de sagesse professionnelle et de technique médicale et non pas seulement en termes des procédés et des méthodes qui sont appliqués et qui sont susceptibles de rehausser nettement les résultats obtenus au niveau de la santé.

Tout ceci est en accord avec les critiques les plus acerbes que l'on a formulées à l'égard de l'évaluation des soins médicaux par l'évaluation du processus. Les adversaires de cette méthode considèrent qu'elle favorise la domination des professionnels sur la distribution des soins médicaux. Deux auteurs s'expriment en ces termes :

> "A l'heure actuelle, la situation est des plus confuses car les médecins ramènent le niveau de santé aux soins médicaux lorsqu'il s'agit de demander des ressources et de la déférence mais se déclarent par contre impuissants lorsqu'on leur demande de rendre compte de leur rendement et soutiennent que la santé est liée à tellement plus de facteurs que les soins médicaux" [21].

C. Evaluation des résultats

1. Efficacité, utilité et rendement

L'évaluation des résultats se rapporte aux objectifs du programme ou de l'organisme, exprimés en termes d'un état de santé bien défini pour la population-cible ou la collectivité servie par l'organisme. Plus précisément, l'évaluation des résultats sanitaires peut viser trois aspects selon qu'on cherche à évaluer l'utilité (efficacy), l'efficacité (effectiveness) ou le rendement (efficiency) [22]. Il convient de définir ces termes parce qu'ils peuvent avoir des significations différentes selon l'orientation et la discipline de celui qui les emploie. Il est indiqué d'adopter les définitions proposées dans le Rapport du comité OMS d'experts des statistiques sanitaires car elles sont couramment acceptées en médecine et en santé publique [23].

> "Efficacité : effet de l'activité et résultat final ou avantages qui en découlent pour la population par rapport aux objectifs fixés" [24].
>
> "Utilité : avantage que présentent pour l'individu le service, le traitement, le médicament, la mesure de lutte ou de prévention recommandés ou adoptés" [25].
>
> "Rendement : effets ou résultats obtenus par rapport aux moyens employés sous forme d'argent, de ressources et de temps" [26].

En ce qui a trait à ces définitions, il importe de noter que l'utilité se rapporte essentiellement aux individus qui sont rejoints par le programme ou à qui l'organisme a dispensé des services pendant un certain temps. Il faut mentionner ici qu'une distinction est faite entre les individus servis et les individus qui ont reçu des services, les premiers englobant à la fois ceux qui en ont reçu et ceux qui n'en ont pas reçu. Lorsqu'on mesure l'utilité, le dénominateur s'exprime en termes des individus qui ont reçu des services, indépendamment de la population-cible ou de l'ensemble de la population servie par l'organisme. L'efficacité est aussi une mesure des résultats mais par rapport à la population-cible du programme ou à l'ensemble de la population qui a besoin d'être servie ou qui est servie par l'organisme. Le rendement est simplement le rapport entre les ressources employées et les mesures de l'utilité ou de l'efficacité.

La définition que nous venons de donner de l'efficacité est celle qui est couramment admise en médecine et en santé publique. Signalons que les théoriciens de l'organisation font souvent usage du terme dans un sens beaucoup plus large pour exprimer le degré de réussite dans la réalisation d'un objectif pré-établi, qu'il soit ou non en termes de santé et indépendamment de la population servie par l'organisme [27]. Par exemple, le degré d'efficacité d'un organisme de santé peut être déterminé en fonction de son accessibilité et de la continuité ou de la globalité des services dispensés. Ces aspects représentent les objectifs intermédiaires qui servent à atteindre l'objectif final d'un organisme sanitaire, à savoir rehausser le niveau de santé. De ce fait, on a toujours évalué le degré d'atteinte des objectifs intermédiaires par une évaluation du processus ou de la structure. Les termes efficacité et utilité seront donc appliqués uniquement à l'évaluation des résultats sanitaires.

2. Indicateurs des résultats

Les résultats sanitaires ou les résultats obtenus se rapportent à un aspect mesurable de l'état de santé qui peut être influencé par un élément particulier des soins médicaux. Les résultats obtenus peuvent donc s'exprimer par des unités de mesure comme la mortalité et la longévité, la mortalité périnatale, la létalité, les courbes de survie liées à des maladies chroniques (y compris le cancer, les affections aiguës, l'incapacité physique, etc.) [28]. Les données nécessaires à ce type d'évaluation sont essentiellement tirées de statistiques épidémiologiques, démographiques et sociologiques. D'une part, la recherche clinique peut contribuer à déter-

miner l'utilité de diverses méthodes et, d'autre part, la recherche opérationnelle et l'économique sont d'un précieux concours dans l'évaluation du rendement.

3. Résultats propres aux organismes de santé

Avant d'aborder la discussion des systèmes courants d'évaluation des résultats, il nous faut souligner les difficultés que soulève ce type d'évaluation dans les organismes sanitaires. En parcourant la documentation médicale relative à l'évaluation, on s'aperçoit qu'il y a une curieuse polarisation des opinions en ce qui concerne les mérites relatifs de l'évaluation du processus et des résultats. Cela est dû à plusieurs raisons [39]. Tout d'abord comme nous l'avons mentionné au début, le cadre institutionnel y est pour beaucoup. Helt et Pelikan ont clairement exprimé ce point de vue en disant que l'évaluation confraternelle des processus de soins médicaux est bien moins redoutable pour les professionnels qui continuent ainsi à dominer les organismes de santé que l'évaluation du résultat qui peut mettre en doute la validité de leurs méthodes [40].

Néanmoins, c'est aussi pour des raisons d'ordre technique que l'évaluation du processus a été plus couramment employée dans les organismes sanitaires. En effet, nous l'avons dit plus haut, l'évaluation des résultats a été généralement appliquée aux programmes de santé publique dont les objectifs d'amélioration de la santé sont relativement faciles à définir et se prêtent à la mesure. De plus, on peut identifier la population-cible de sorte que non seulement l'utilité mais l'efficacité du programme puissent être évaluées.

Dans le cas des organismes de santé qui poursuivent une gamme d'objectifs et sont engagés dans les activités complexes et étroitement liées l'évaluation des résultats présente quelques difficultés d'ordre conceptuel et certains problèmes de mesure [41]. Les principales difficultés résident dans : (1) l'identification de la population servie; (2) le manque de spécificité des indicateurs des résultats; (3) certains effets paradoxaux; et (4) la dimension temporelle des résultats.

a) *Population servie*

L'identification de la population servie par l'établissement pose un problème au départ. Alors que les programmes de santé publique visent tous les membres d'une population-cible, les organismes sanitaires servent une population dont seulement un petit nombre des membres bénéficiera de leurs services pendant une certaine période à la suite d'une maladie. Les paramètres de la population-cible des programmes de santé publique peuvent et doivent être définis avant que ces derniers soient mis en oeuvre. Au contraire, les organismes sanitaires sont des entités vivantes et dynamiques et la population qu'ils servent n'est pas statique de sorte qu'il faut la redéfinir constamment pour pouvoir expliquer les changements survenus au niveau de l'utilisation.

En dépit de ces difficultés, les planificateurs sanitaires ont mis au point des méthodes qui permettent de définir la population servie par un organisme de santé [42]. Fondamentalement, ces méthodes consistent à déterminer les pourcentages de gens dans des régions géographiques déterminées qui s'adressent à divers organismes sanitaires au cours d'une certaine période. Ces pourcentages sont appelés indices d'attraction (*relevance indices*). Ils peuvent être calculés pour toutes les zones géographiques dont l'organisme sanitaire tire sa clientèle. Connaissant les caractères démographiques des malades admis dans l'établissement et les caractéristiques de la population régionale, on peut présumer que la population servie par l'organisme dans chaque zone géographique correspond au pourcentage de gens de la région qui ont effectivement utilisé ses services. En d'autres termes, connaissant la proportion des gens de chaque région qui a fait usage des différents services de santé et connaissant leurs caractéristiques, on peut déduire les caractéristiques et la structure de l'ensemble de la population servie, c'est-à-dire non seulement des individus qui ont fait usage des services mais de ceux qui n'en ont pas fait usage.

On peut procéder à une identification plus poussée de la population servie par un organisme de santé en définissant des populations différentes pour chacun des programmes mis en oeuvre par cet organisme. Cela peut avoir une certaine importance dans le cas des organismes qui fournissent une vaste gamme de services, allant de services de soins généraux à des services ultra-spécialisés. De toute évidence, la population servie par des programmes spécialisés est beaucoup plus dispersée que la population servie par des programmes de soins de première ligne. Si l'on veut évaluer l'efficacité des divers programmes, il faut en établir correctement la population-cible.

b) *Manque de spécificité des indicateurs des résultats*

On peut conclure de ce qui précède qu'il est techniquement possible de définir la population servie par un organisme de santé pourvu que soient acceptées les hypothèses sur lesquelles se fonde la méthode que nous venons de présenter. Ceci est une condition primordiale à l'évaluation de l'efficacité d'un organisme sanitaire. Une autre condition réside dans la mise en évidence du degré d'amélioration de la santé qui est dû à l'intervention de l'organisme en question [33]. Personne ne conteste le fait que l'état de santé d'une collectivité résulte dans une très petite mesure des soins médicaux. Aussi semble-t-il peu utile d'analyser le niveau de santé d'une population pour déterminer les résultats sanitaires obtenus par l'action d'un organisme. Inversement, il semblerait injuste d'imputer à un organisme sanitaire l'entière responsabilité du manque d'amélioration de l'état de santé de la collectivité qu'il sert. Ainsi, ce que Donabedian appelle la "Capacité d'intégration" des indicateurs des résultats n'est autre que leur globalité ou manque de spécificité qui en restreint souvent l'utilité en matière d'évaluation [34]. Le seul moyen de surmonter convenablement ce problème serait peut-être des expériences contrôlées. Grâce à ce genre de recherche, on aurait au moins quelque indice sur les possibilités d'amélioration du niveau de la santé offertes par les techniques sanitaires actuelles.

c) *Effets paradoxaux*

Il reste encore une difficulté à surmonter qui est due aux "effets paradoxaux" du progrès technique [35]. Chose curieuse, un système de distribution des soins efficace semble être lié à un taux plus élevé et non pas moins élevé de maladie et de morbidité. Un tel para-

coxe peut s'expliquer par les effets de "découvertes pathologiques" (discovery) et ceux de la "longévité accrue" (postponement). Un meilleur système de distribution des soins peut en effet comporter des effets de découverte, en ce sens que des maladies qui n'étaient pas diagnostiquées ou déclarées jusqu'ici sont désormais diagnostiquées et calculées dans les statistiques. Les effets de la "longévité accrue" découlent de ce qu'en traitant des maladies auparavant non traitées on peut réduire les taux de mortalité tout en accroissant ceux d'invalidité, d'incapacité ou de maladie chronique à tous les âges [36].

d) *Dimension temporelle des résultats*

Un quatrième problème de mesure des résultats découle de ce que les effets positifs des soins dispensés sont en partie directs et à court terme et en partie indirects et à long terme. L'action des organismes sanitaires sur l'état de santé est la partie immédiate et en partie prospective. Lorsqu'on mesure l'état de santé à partir d'indicateurs de morbidité ou de mortalité, on obtient un bilan, une vue instantanée et statique de la santé d'une population. La contribution possible des organismes ou des programmes de santé à l'amélioration de l'état de santé futur n'est pas mesurée. Nous n'avons pas de méthode précise pour évaluer cette amélioration mais il convient d'en tenir compte surtout face aux critiques dont font l'objet l'action et la légitimité sociale des organismes de santé [37].

Les éléments constitutifs des résultats sanitaires ne peuvent donc pas être définis sans se reporter aux paramètres présentés dans notre cadre de référence. Un paramètre en particulier qu'il importe de spécifier c'est le niveau d'agrégation de la clientèle (population-cible, collectivité, etc.) car il représente le dénominateur qui sert à mesurer l'utilité et l'efficacité. Un autre point à retenir de ce qui précède, c'est que plus le niveau d'agrégation est élevé, et plus l'indicateur des changements du niveau de santé engendrés par les soins médicaux manquera de spécificité. En présentant ainsi les difficultés inhérentes à l'emploi des indicateurs des résultats dans l'évaluation de l'efficacité des organismes sanitaires, nous n'avons nullement l'intention de décourager ceux qui veulent entreprendre de telles évaluations car s'il ne faut pas ignorer ces difficultés, elles sont parfaitement surmontables ainsi que le prouvent les nombreux rapports sur le sujet. Nous examinerons maintenant certains des systèmes d'appréciation des résultats qui ont été proposés ou mis à l'essai et qu'on retrouve dans la documentation sur les soins médicaux.

4. Système d'appréciation des résultats

L'implantation d'un système d'évaluation de la qualité fondé sur l'appréciation des résultats peut se faire selon au moins trois méthodes interdépendantes : (1) création d'un système de contrôle des résultats; (2) analyse des résultats défavorables; et (3) études spéciales orientées vers la recherche [38]. Chacune de ces méthodes est discutée ci-dessous.

a) *Système de contrôle des résultats*

Pour établir un système de contrôle des résultats, il faut d'abord répondre aux trois questions suivantes : (1) cherche-t-on à mesurer l'utilité ou l'efficacité ? (2) quelles sont les normes qui serviront à l'appréciation des résultats mesurés ? (3) quels sont les résultats que l'on évaluera ?

(1) Utilité ou efficacité

La question de savoir si c'est l'utilité ou l'efficacité que l'on veut mesurer est extrêmement importante. En effet, si l'on se reporte aux définitions de ces termes que l'on a données plus haut, on se rend compte que s'il s'agit de l'efficacité, c'est l'ensemble de la population qui est le dénominateur et non pas seulement les individus servis comme dans le cas de l'utilité. Tel que mentionné précédemment, l'identification de l'ensemble de la population servie par un établissement est techniquement faisable mais exige une grande quantité de données sur l'utilisation des services et les caractères démographiques des populations appartenant aux zones géographiques d'où l'établissement tire sa clientèle. Ces données de base sont indispensables à la mesure de l'efficacité. C'est probablement pourquoi on n'a pas évalué les résultats atteints par un organisme sanitaire au niveau de l'ensemble de la population servie (efficacité). On a pratiqué ce genre d'évaluation surtout au niveau macroscopique dans le but de déterminer l'efficacité relative des systèmes de distribution des soins. L'étude de Shapiro et ses collaborateurs qui ont utilisé des taux de mortalité précis comme indicateurs de la qualité des soins dispensés aux membres du *Health Insurance Plan of Greater New York* (H.I.P.) comparativement à d'autres populations de la ville de New York en représente un excellent exemple [39].

On a donc mesuré les résultats de l'action des organismes sanitaires en fonction des individus servis. En d'autres termes, c'est l'utilité plutôt que l'efficacité de cette action qui a été évaluée. Une des unités de mesure de l'utilité a été les taux de létalité, le dénominateur étant le nombre de cas traités contre une affection déterminée, pendant une période de temps donnée. Les études de Lee, Morrisson, Morris et Lipworth en Angleterre il y a plusieurs années en sont un exemple [40]. Ces auteurs ont comparé l'action des hôpitaux d'enseignement et des hôpitaux non enseignants en mesurant des taux de létalité précis. Leur approche s'appuyait sur l'hypothèse implicite que l'action des hôpitaux d'enseignement pouvait servir de norme empirique à l'évaluation de celle des autres hôpitaux. Evidemment, des différences quant au degré de gravité des affections traitées peuvent fort bien servir à expliquer les différences des taux de létalité accusés par les divers établissements. Aussi faut-il pouvoir contrôler cet effet. La notion des "phases et degrés de morbidité" (*staging concept*) appliquée par Gonnella et ses collaborateurs à l'évaluation semble utile en ce sens car elle fournit un moyen de classer les problèmes de santé selon la phase de l'évolution de la gravité au premier contact du malade avec le système de distribution des soins [41].

Roemer et ses collaborateurs ont aussi mis au point une méthode de pondération statistique du taux de létalité relatif à toutes les affections traitées dans un organisme, qui se fonde sur la mesure de la gravité de l'ensemble des cas traités [42]. Les indicateurs de la gravité des cas sont : (1) la durée du séjour et (2) le

taux d'occupation. Mis en corrélation avec d'autres mesures de la gravité, ceux-ci semblent constituer des indicateurs valables du degré de gravité. Toutefois, la durée du séjour et le taux d'occupation peuvent être influencés par des facteurs reliés à la quantité des ressources disponibles et cela met quelque peu en doute la validité des comparaisons faites entre des hôpitaux qui se situent dans des régions où le niveau des ressources n'est pas comparable. A la longue, un indice de la qualité qui se base directement sur les taux spécifiques de létalité est probablement plus sûr [43].

En tant qu'unité de mesure de l'utilité de l'action des organismes sanitaires, les taux de létalité constituent donc un moyen d'évaluation simple et peu coûteux. Il est surprenant de constater que ce moyen n'ait pas été appliqué davantage à des fins d'évaluation. Il serait possible d'améliorer les résultats qu'il fournit en étendant la portée de ces taux au-delà du séjour hospitalier puisque des changements importants peuvent survenir dans l'évolution de certaines maladies après un certain temps et bien après la sortie de l'hôpital. Une autre amélioration possible consisterait à y inclure des indicateurs de morbidité, tels que : incapacité physique, présence ou persistance de symptômes, complications, etc. En englobant les indicateurs de mortalité et de morbidité et en s'étendant au-délà de la durée de l'hospitalisation, on obtient un moyen facile et pratique d'évaluer l'utilité des activités d'un hôpital.

(2) Normes d'évaluation des résultats

La première étape de la mise en place d'un système d'appréciation des résultats sanitaires consiste à déterminer si c'est l'utilité ou l'efficacité de l'action qu'on veut évaluer. La deuxième étape vise à préciser les normes par rapport auxquelles se fera cette évaluation. Le principe sur lequel s'appuie le système de contrôle des résultats obtenus c'est que, dans des conditions acceptables, une certaine proportion de résultats défavorables est escomptée. Si cette proportion est dépassée, il faut aussitôt procéder à une enquête pour déterminer les facteurs en cause. Il n'est pas facile de formuler des normes précises d'évaluation des résultats. On s'accorde généralement sur ce qui constitue un bon résultat final mais on ne sait pas très bien comment parvenir à de bons résultats dans des cas particuliers, lorsque différents niveaux de soins médicaux sont appliqués. Par conséquent, on n'a guère établi des normes à priori pour évaluer les résultats obtenus [44]. Alors que dans l'évaluation du processus, les normes utilisées sont bien souvent établies de façon normative, on a tendance à s'appuyer sur des normes empiriques tirées d'études spéciales ou de comparaisons avec des organismes et des méthodes similaires lorsqu'il s'agit d'évaluer les résultats. Une approche pratique pour établir des normes de façon normative nous est donnée par Williamson [45]. On commentera cette approche plus loin. Disons pour le moment que les normes qu'elle utilise sont une combinaison d'éléments normatifs et empiriques puisés des statistiques courantes, d'études de mortalité et d'évaluations confraternelles. Bien entendu, les professionnels de la santé hésitent à formuler des normes de résultats trop élevées et s'empressent de rappeler l'importance d'autres facteurs susceptibles de neutraliser l'effet bénéfique de l'intervention médicale et même de susciter des résultats défavorables [46]. Ce point est illustré par l'étude de Brooke citée plus haut [50]. En effet, tandis que les médecins jugeaient que les résultats obtenus n'auraient pas pu être améliorés dans 63.2 pour cent des cas, soit que la qualité des soins était adéquate dans cette proportion selon le critère d'évaluation établi, la qualité des soins était jugée acceptable dans seulement 1.4 pour cent des cas lorsque des critères explicites d'évaluation du processus étaient appliqués et dans 23.3 pour cent des cas lorsque ces critères étaient implicites.

Le caractère empirique des normes d'évaluation des résultats semble donc être une condition essentielle à leur validité. Il importe que les administrateurs et les planificateurs sanitaires en soient conscients car l'évaluation des résultats, contrairement à celle du processus, n'appartient pas aux seuls professionnels qui dispensent les soins. Tout administrateur sanitaire doit reconnaître ce fait et s'attendre à la résistance qu'opposeront les professionnels de la santé lorsqu'il tentera d'implanter un système de contrôle des résultats. En effet, les médecins pourront y voir de la méfiance à l'égard de leur compétence ou encore une menace à leur autonomie professionnelle.

(3) Choix des résultats

La troisième question à laquelle on doit répondre avant d'établir un système de contrôle des résultats se rapporte aux résultats qui seront évalués. Le choix de ces résultats dépend de bien des facteurs, notamment : les relations entre un résultat donné et les éléments du processus qui y conduisent; la disponibilité des données et la possibilité de mesurer correctement le résultat obtenu. En réalité, on peut choisir certaines situations ou certaines affections comme base d'évaluation des résultats et en déduire ensuite la qualité des soins. Une telle approche a été proposée par Kessner qui a désigné du nom de "traceurs" (tracers) les affections choisies à des fins d'évaluation [47].

La méthode mise au point par Kessner est simple. Les "traceurs" sont des problèmes sanitaires précis qui permettent aux évaluateurs d'identifier les qualités et les défauts d'un milieu particulier ou de l'ensemble d'un système de soins. Le concept des "traceurs" a été emprunté aux sciences physiques. Par exemple, les endocrinologistes emploient des traceurs radio-actifs pour étudier comment la glande thyroïde réagit à une substance comme l'iode. Ils observent comment elle absorbe une petite quantité d'iode radio-actif et supposent qu'elle absorbe de la même manière l'iode organique. Les "traceurs" choisis pour évaluer la qualité des soins sont aussi des problèmes sanitaires bien circonscrits et identifiables qui sont traités à l'intérieur du système de distribution de soins et qui mettent en lumière la façon dont des composantes du système fonctionnent non pas isolément mais systématiquement.

La méthode des "traceurs" est applicable à divers types d'organismes sanitaires ainsi qu'aux praticiens médicaux. L'hypothèse fondamentale demeure la même dans les deux cas. Par exemple, la manière dont un médecin ou une équipe de médecins traite une affection courante ou la manière dont un organisme identifie les grossesses à risques élevés sont des indices de la qualité des soins et de l'utilité des méthodes employées par ces médecins dans cet organisme [48].

Pour que le choix des "traceurs" soit rationnel et uniforme, Kessner a établi les six critères que nous énumérons ci-dessus par ordre d'importance [49] :

1. un "traceur" doit avoir un impact fonctionnel manifeste;

2. il doit être relativement bien défini et facile à diagnostiquer;

3. les taux de prévalence doivent être suffisamment élevés pour que l'on puisse recueillir assez de données à partir d'un échantillon restreint de la population;

4. l'évolution naturelle de l'affection doit être modifiée par l'intervention médicale;

5. les techniques applicables doivent être clairement déterminées au niveau d'au moins un des processus suivants : prévention, diagnostic, traitement ou réadaptation;

6. les effets de facteurs non médicaux sur les "traceurs" doivent être connus.

Ces critères montrent bien que la structure par âge et par sexe de la population-cible servie et son profil épidémiologique sont des données indispensables à la sélection de "traceurs" appropriés. Il est donc évident qu'une même série de "traceurs" ne sera pas applicable à toutes les populations.

Une seconde caractéristique de la méthode des "traceurs" c'est qu'elle peut être utilisée pour mesurer à la fois le processus et le résultat des soins. Ceci est un avantage important dans un système d'évaluation. Tel que souligné plus haut, les qualités et les défauts du processus sont difficiles à identifier quand on n'en connaît pas le résultat. En portant sur le processus et le résultat, la méthode des "traceurs" offre une approche globale pour l'évaluation de la qualité.

Le principal problème que pose l'application de cette méthode réside dans l'hypothèse fondamentale qui la sous-tend. Le comportement d'un médecin est-il uniforme ou change-t-il selon la catégorie de maladie qu'il traite ? Dans son étude de cinq méthodes d'évaluation confraternelle de la qualité des soins, Brooke a constaté que les critères implicites dont se servent les médecins varient avec le genre de cas qu'ils évaluent [50]. Il en était ainsi alors même que toutes les maladies dont il était question exigeaient une post-observation, et que tous les malades étaient traités dans le même établissement par le même groupe de médecins. Que le comportement des médecins varie selon le cas auquel ils font face est aussi une de nos constatations personnelles à la suite d'une étude au niveau d'un groupe d'internistes exerçant à l'intérieur du régime conventionnel Kaiser, à Portland, Oregon [51]. Cette variabilité laisse supposer qu'une méthode d'évaluation qui s'appuie sur un nombre restreint de "traceurs" comme indices de la qualité globale comporte des risques d'erreur, si l'on extrapole ces résultats à toutes les conditions médicales traitées.

Au niveau des organismes, le problème se pose en termes similaires. La qualité des soins dispensés dans un service hospitalier est-elle comparable à celle des soins dispensés dans un autre service de l'établissement ? Les réponses à ces questions ont une grande influence sur la manière dont on doit évaluer les soins. Si la pratique d'un médecin est uniforme ou si les pratiques au sein d'un établissement sont uniformes, l'évaluation de la qualité s'en trouve considérablement simplifiée. Il est alors possible de choisir quelques situations et quelques composantes des soins pour représenter le modèle général de pratique.

Lorsque les pratiques ne sont pas uniformes ou homogènes, le choix de l'échantillon sur lequel on se base devient un élément capital de la méthode d'évaluation [52].

En dépit de ces problèmes, la méthode des "traceurs" reste un excellent moyen de sélectionner certaines affections à des fins d'évaluation et peut être profitablement utilisée par les administrateurs sanitaires.

b) *Analyse des résultats défavorables*

A ce stade, on a répondu aux trois questions requises et le système de contrôle des résultats peut être mis en place. Reste à savoir comment fonctionnera ce système. Passera-t-on en revue tous les cas ou seulement ceux dont l'issue est défavorable ? L'analyse des résultats défavorables est plus économique et constituera donc la seconde étape de l'évaluation des résultats.

La principale caractéristique de l'analyse des résultats est qu'elle englobe les trois approches de l'évaluation en rattachant le résultat au processus ou à la structure. Dans ce sens, c'est un complément de la méthode des "traceurs" que nous venons de décrire. L'analyse des résultats défavorables est à la base du système d'évaluation proposé par Williamson [53]. Ce système suppose l'établissement d'une liste des problèmes sanitaires selon l'ordre de priorité en vue de permettre une étude systématique. En fait, il porte sur quatre éléments : (1) les résultats diagnostiques (soit les données nécessaires au choix du traitement et au pronostic); (2) le processus diagnostique (soit les méthodes appliquées en vue de la pose du diagnostic); (3) les résultats thérapeutiques (soit l'état de santé du malade à un moment donné après le traitement); (4) le processus thérapeutique (soit la planification, l'administration et l'évaluation du traitement).

Les résultats diagnostiques et thérapeutiques sont d'abord évalués. Les processus connexes ne sont évalués que si les résultats n'atteignent pas les normes établies. Cela revient donc à une analyse des cas dont l'issue est défavorable. Lorsque les résultats relatifs à un cas se situent dans les limites acceptables, l'étude des processus diagnostiques et thérapeutiques qui s'y rapportent est suspendue en faveur d'une étude des résultats, conformément à l'ordre de priorité fixé.

Il importe de reconnaître que la portée de cette méthode est limitée si elle n'est pas appliquée conjointement avec une autre méthode de revue qui s'attachera, par exemple, à l'examen d'échantillons de tous les cas ou d'échantillons appropriés des cas qui sont particulièrement exposés à des résultats inférieurs aux normes. En effet, il est difficile d'identifier les qualités et les

défauts du processus lorsqu'on n'en connaît pas le résultat puisque c'est précisément ce résultat qui sert à justifier l'intervention médicale. Inversement, comme c'est le cas ici, le résultat à lui seul peut être trompeur si le malade subit des examens diagnostiques inutiles ou reçoit un traitement peu approprié qui n'engendre pas des résultats défavorables mais ne contribue pas non plus à une amélioration des résultats et peut même représenter un mauvais usage des ressources.

En résumé, la revue des résultats défavorables ne peut pas être pratiquée indépendamment d'un système global d'examen des dossiers qui assure non seulement l'examen des résultats médiocres mais celui des résultats acceptables qui découlent d'une mauvaise utilisation des ressources sanitaires.

c) *Etudes spéciales*

La troisième composante du système d'évaluation des résultats consiste en des études spéciales. On en distingue deux genres : les études fondées sur l'observation des faits et les études expérimentales [54].

Les études des faits se fondent sur des méthodes épidémiologiques bien établies pour examiner les rapports entre les résultats obtenus et les modes d'organisation des programmes existants. L'étude de la mortalité périnatale et des naissances prématurées chez les personnes inscrites au *Healt Insurance Plan of Greater New York*, comparativement à l'ensemble de la population de cette ville [39], est un exemple de ce genre de recherche tout comme celles des taux de létalité dans les hôpitaux d'enseignement et les hôpitaux d'enseignement que nous avons précédemment citées [40], [41]. Le trait caractéristique de ces études c'est l'absence de tout contrôle sur la répartition des sujets entre les programmes sanitaires qui sont étudiés. Quoiqu'il en soit, les études des faits sont souvent le seul moyen dont on dispose pour comparer le rendement de différents établissements et en tirer des normes d'évaluation.

Les études expérimentales consistent à introduire soit une "intervention" soit un changement dans le système de distribution des soins et à répartir les sujets au hasard entre deux ou plusieurs programmes ou établissements [55]. Contrairement aux études des faits, ici le chercheur exerce un contrôle sur les facteurs extérieurs qui peuvent modifier les résultats de l'expérience. On est ainsi plus sûr que les différences dans les résultats obtenus sont dues à des différences dans les programmes de soins plutôt que dans les caractéristiques des individus qui reçoivent des soins. Le protocole suivi correspond à l'essai clinique contrôlé (RCT) [77], [18].

Il convient de souligner que les résultats des études des faits et de la recherche expérimentale servent à établir les normes empiriques destinées à l'évaluation de la qualité. Par conséquent, même si l'élaboration de telles études n'est pas une responsabilité primordiale de l'administrateur sanitaire, il est indispensable qu'il se tienne au courant de leurs résultats. De plus, dans les organismes où la recherche représente une activité importante, l'administrateur veillera à encourager ce genre d'étude car elle peut fournir des données utiles pour le système d'évaluation.

CONCLUSION

La diversité des méthodes d'évaluation des soins médicaux que nous venons de présenter soulève plusieurs questions dont les administrateurs et planificateurs sanitaires doivent tenir compte lors de l'implantation d'un système d'évaluation de la qualité des soins dispensés. En conclusion, nous voudrions résumer la discussion qui précède en faisant ressortir les principaux éléments d'un système d'évaluation ainsi que les principes qui doivent lui servir de base dans les organismes sanitaires.

Tout système d'évaluation appliqué dans un organisme de santé doit comporter les éléments suivants :

1. Un sous-système d'évaluation de la structure, y compris les installations et le matériel, les caractéristiques de l'organisation administrative et la compétence des professionnels de la santé.

2. Un sous-système d'évaluation du processus englobant :

   2.1 la ratification de l'admission et du séjour s'appuyant de préférence sur un mécanisme de revue simultanée;

   2.2 la ventilation et l'analyse des statistiques complémentées par un examen des dossiers dans les cas qui s'écartent des statistiques;

   2.3 l'examen des dossiers relatifs aux cas qui sont particulièrement susceptibles de révéler des insuffisances et des erreurs.

3. Un sous-système de contrôle des résultats basé sur :

   3.1 l'identification de l'ensemble de la population servie ou des différentes populations servies par les différents programmes (dénominateur de l'efficacité);

   3.2 l'identification de la population qui a reçu des services de l'organisme ou des populations qui ont reçu des services des différents programmes (dénominateur de l'utilité);

   3.3 le choix de certaines affections (ou traceurs) à des fins d'évaluation;

   3.4 la revue des résultats défavorables.

L'application du système doit être guidée par les quatre principes suivants :

Un principe fondamental mentionné à plusieurs reprises au cours de la discussion réside dans les relations mutuelles qui existent entre la structure, le processus et le résultat. Avec une organisation adéquate (structure), on favorise la qualité du comportement professionnel (processus) et par conséquent, celle des résultats sanitaires obtenus (résultat); leur système d'évaluation devrait donc réunir des éléments des trois approches respectivement appliquées à l'évaluation de ces aspects.

Il faut reconnaître que les relations mutuelles précises entre ces aspects ne sont pas toujours bien comprises. Ainsi, les modes d'évaluation des soins médicaux sont-ils restés conservateurs et c'est l'évaluation de la structure ou du processus plutôt que des résultats qui a servi de base à la répartition des ressources et à la prise de décision. En d'autres termes, lorsque la relation entre structure et résultat ou entre processus et résultat n'est pas établie, on présume que les ressources ou les méthodes utilisées sont justifiées quoique leur relation avec des résultats particuliers ne puisse pas être mise en évidence. En résumé, c'est aux résultats que la preuve incombe. Nous avançons comme deuxième principe que le fardeau de la preuve repose sur la structure et le processus et non sur les résultats. En d'autres termes, l'administrateur doit se montrer très méfiant et très prudent lorsque l'addition de nouvelles ressources ou l'application de nouvelles méthodes ne sont pas liées à des objectifs déterminés d'utilité ou d'efficacité.

C'est de ce deuxième principe que découle le troisième qui veut que l'évaluation des soins médicaux fasse partie intégrante du système de gestion et guide les décisions relatives à la répartition ou à l'addition de diverses ressources. Les méthodes qui servent à l'évaluation sont généralement rétrospectives. Elles permettent d'identifier les insuffisances et d'éviter qu'elles ne se reproduisent grâce à l'implantation de mécanismes de contrôle efficaces. Cependant, pour que les décisions prises au cours d'un traitement soient appropriées, un mécanisme de revue simultanée est nécessaire. Jusqu'ici, aucune méthode n'a été établie à cette fin à part celles qui se basent sur la ratification de l'admission et du séjour. Cette ratification est au moins un point de départ mais on pourrait chercher à formuler de nouvelles méthodes.

Nous n'avons pas du tout abordé la question des coûts liés à l'implantation et à l'exploitation des différents systèmes d'évaluation décrits. Pour certains, ceci apparaîtra comme une omission sérieuse dans un domaine où la prise de décision se fonde trop souvent sur les ressources (inputs) et non sur le rendement (outputs). Toutefois, on peut dire sans crainte que ces coûts resteront marginaux si l'on tient compte du budget de plusieurs millions de dollars dont disposent la plupart des hôpitaux. Est-ce vraiment trop cher pour obtenir des renseignements précis sur ce qui est produit avec cet argent et sur la façon dont on réalise cette production ?

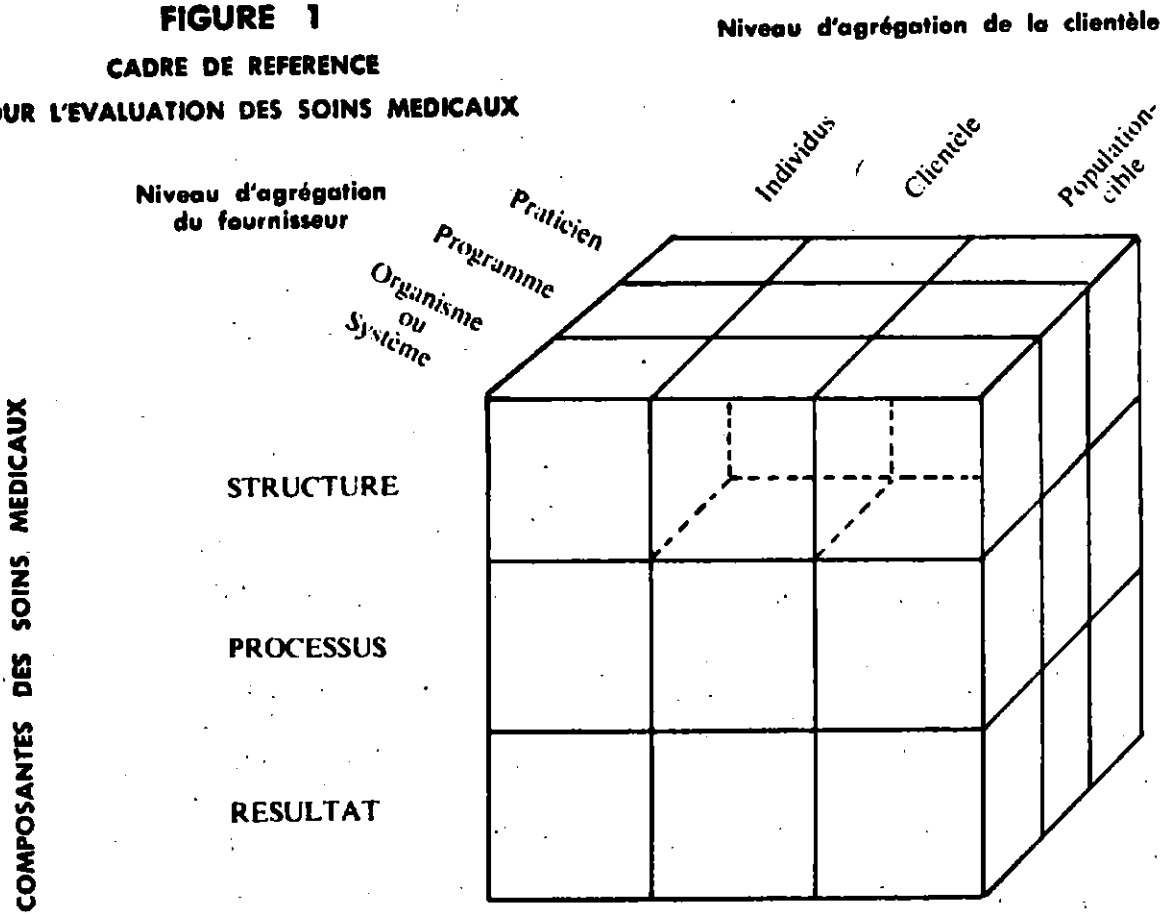
TABLEAU 1

Indicateurs servant à l'évaluation et méthodes de collecte des données selon les différents aspects évalués

COMPOSANTES DE L'ORGANISME OU DU PROGRAMME SANITAIRE	Ressources	Activités	Objectifs sanitaires		
EVALUATION GENERALE	Structure	Processus	Résultat		
EVALUATION SPECIFIQUE	Suffisance (adequacy)	Justesse (appropriateness)	Efficacité	Utilité	Rendement
INDICATEURS SERVANT A L'EVALUATION	Quantité et qualité des ressources humaines, matérielles et financières	Méthodes et techniques appliquées	Effets sanitaires au niveau de la population servie par le programme ou l'organisme	Effets sanitaires au niveau des individus rejoints par les programmes ou traités par les organismes	Rapport entre les résultats sanitaires et les ressources déployées
METHODE DE COLLECTE DES DONNEES	Comptabilité	Evaluation confraternelle	Epidémiologie Démographie Sociologie	Epidémiologie Recherche clinique	Recherche opérationnelle Economique

FIGURE 1

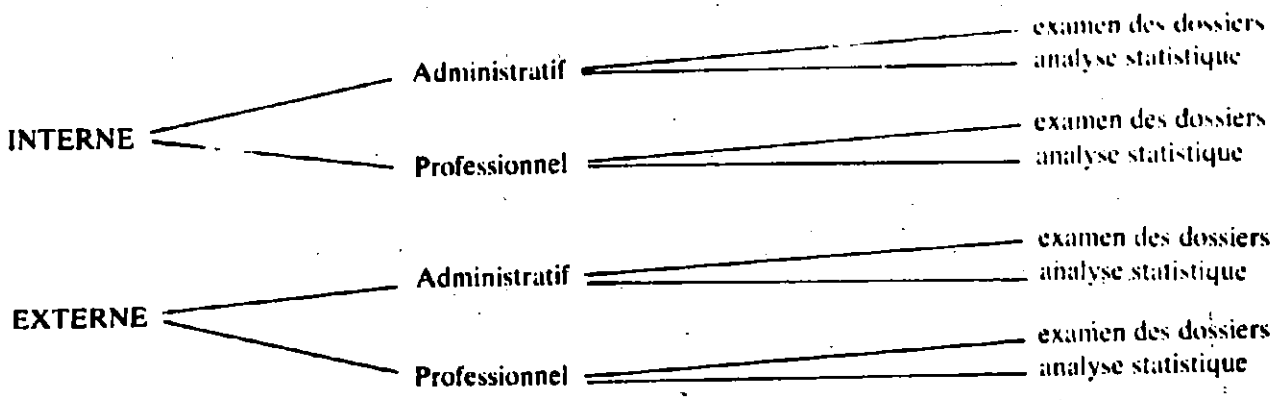
CADRE DE REFERENCE POUR L'EVALUATION DES SOINS MEDICAUX



Source: Adapté de Donabedian A. "A Perspective on Concepts of Health Care Quality". Conférence présentée à l'assemblée de l'Institute of Medecine, tenue à l'automne. Washington, D. C., 6 novembre 1974.

FIGURE 2

Classification des mesures de contrôle



Source: Donabedian A. "Promoting Quality through Evaluating the Process of Patient Care". *Medical Care*, 1968, 6, p. 190.

BIBLIOGRAPHIE

1— Etzioni, A. **Modern Organization**, Englewood Cliffs : Prentice-Hall, 1964, pp. 75-93.

Bucher, R. and Stelling J. "Characteristics of Professional Organizations". Journal of Health and Social Behavior, 1969, 10, 3-15.

2— Scott, R. W. "Reactions to Supervision in a Heteronomous Professional Organization". Administrative Science Quarterly, 1965, 10, 65-81.

3— Smith, H. L. "Two lines of Authority Are One too Many". Modern Hospital, 1955, 84, 59-64.

4.— Freidson, E. Professional Dominance, New York: Atherton Press, 1970.

5— Freidson, E. and Buford, R. "Processes of Control in a Company of Equals". Social Problems, 1963, 11, 119-131.

Goss, M. E. "Influence and Authority Among Physicians in an Outpatient Clinic". American Sociological Review, 1961, 26, 39-50.

6— Hughes, E. C. Men and Their Work. Glencoe: The Free Press, 1958.

7— Donabedian, A. **A Guide to Medical Care Administration : Medical Care Appraisal.** American Public Health Association, New York, 1969.

Donabedian, A. "A Perspective on Concepts of Health Care Quality". Paper Presented at the Fall Meeting of the Institute of Medecine, Washington, D.C., Nov. 6, 1974.

Donabedian, A., "Evaluating the Quality of Medical Care". Milbank Memorial Fund Quarterly, 1966, 44, 166-203.

Dobabedian, A. "Promoting Quality Through Evaluating the Process of Medical Care". Medical Care, 1968, 6, 181-202.

De Geyndt, W., "Five Approaches for Assessing the Quality of Care". Hospital Administration, 1970, 15, 21-42.

Rochon, J. "Measurement of Quality of Care : Concepts and Strategies for Evaluation". Annals of the Royal College of Physicians and Surgeons of Canada, 1974, 7, 301-311.

Sockett, D. and Raskin, M. S. (eds.) **Methods of Evaluating Health Services.** McMaster University, Hamilton, Ont., 1976, (3rd ed.).

Sheps, M. C. Approaches to the Quality of Hospital Care". Public Health Reports, 1955, 70, 877-886.

8— Ce cadre de référence est adapté de :
Donabedian, A. "A Perspective on Concepts of the Health Care Quality". Op. cit., pp. 1-3.

9— Donabedian, A., "Evaluating the Quality of Medical Care". Op. cit., 167-170.

10— Rhee, Sang-O. Relative Influence of Specialty Status and Organization of Office Care and Organization of Hospital Care on the Quality of Medical Care: "A Multivariate Analysis. Doctoral Dissertation, University of Michigan, 1975.

11— Donabedian, A. **A Guide to Medical Care Administration.** Op . cit., pp. 14-28.

12— Goran, M. J. et al. "The PSRO Hospital Review System". Medical Care, 1975, 13, (Supplement), 1-33.

13— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit., p. 24.

14— Donabedian, A. "Promoting Quality Through Evaluating the Process of Medical Care". Op. Cit., 189-191.

15— Donabedian, A. Ibid., p. 191.

16— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit., pp. 59-60.

17— Mather, H. G. et al., "Acute Myocardial Infarction: Home and Hospital Treatment". British Medical Journal. 1971, 3, 334-338.

18— Hutter, A. M. "Early Hospital Discharge After Myocardial Infarction". New England Journal of Medicine, 1973, 228, 1141-1144.

19— Brook, R. H. and Appel, F. A. "Quality of Care Assessment: Choosing a Methode for Peer Review". New England Journal of Medicine, 1973, 228, 1323-1329.

20— Ibid., p. 1328.

21— Helt, E. H. and Pelikan, J. A. "Quality: Medical Care's Answer to Madison Avenue". American Journal of Public Health, 1975, 65, 287.

22— Rochon, J. "Measurement of Quality of Care: Concepts and Strategies for Evaluation". Op. cit., 309.

23— Fourteenth Report of the WHO Expert Committee on Health Statistics. "Statistical Indicators for the Planning and Evaluation of Public Health Programs". World Health Organization. Technical Report Series No. 472, 1971, 10-14.

24— Ibid., p. 12.

25— Ibid., p. 11.

26— Ibid., p. 14.

27— Etzioni, A. **Modern Organizations, Op. cit., pp. 5-19.**

Perrow, C. "The Onalysis of Goals in Complex Organizations". American Sociological Review, 1961, 26, 854-866.

Donabedian, A. **Aspects of Medical Care Administration: Specifying Requirements for Health Care.** Cambridge, Mass.: Harvard University Press, 1973, pp. 31-57.

28— Shapiro S. "End Result Measurements of Quality of Medical Care". Milbank Memorial Fund Quarterly, 1967, 45, 7-30.

29— Donabedian, A. **A Guide to Medical Care Administration,** Op cit., pp. 3-4.

30— Helt, E. and Pelikan, J., Op. cit., p. 286.

31— Donabedian, A. "Models for Organizing the Delivery of Personal Health Services and Criteria for Evaluating Them". Milbank Memorial Fund Quarterly, 1972, 50, 103-154.

32— Morlin, J. R. **Comprehensive Health Planning : Analytic Concepts.** Blue Cross Association, 1974, pp. 85-103.

Griffith, J. R. **Quantitative Techniques in Hospital Planning and Control.** Lexington, Mass.: D. C. Health and Co., 1972.

33— Donabedian, A. "Models for Organizing the Delivery of Personal Health Services". Op. cit., 106-107.

34— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit., p. 35.

35— Donabedian, A. "Models for Organizing the Delivery of Personal Health Services". Op. cit., 107.

36— Ibid., p. 107.

37— Pineault, R. "Les établissements de santé dans le contexte de la santé communautaire". Administration Hospitalière et Sociale, 1976, 22, 11-13.

38— Donabedian, A. **A Guide to Medical Care Administration.** Op. cit., pp. 38-41.

39— Shapiro, C. et al. "Further Observations on Prematurity and Perinatal Mortality in a General Population and in the Population of a Prepaid Group Practice Medical Care Plan". Am. Journal of Public Health, 1960, 1304-1317.

Shapiro, S. et al. "Patterns of Medical Care Use by the Indigent Aged Under Two Systems of Medical Care". Am Journal of Public Health, 1967, 57, 784-790.

40— Lee, J. A. H. et al. "Fatality from Three Common Surgical Conditions in Teaching and Non-Teaching Hospitals". Lancet, 1957, 785-790.

Lipworth, L. et al. "Case-fatality in Teaching and Non-Teaching Hospitals 1956-1959". Medical Care, 1963, 1, 71-76.

Lee, J. A. H. et al. "Case-fatality in Teaching and Non-Teaching Hospitals". Lancet, 1960, 170-171.

41— Gonnella, J. S. and Goran, M. J. "Quality of Patient Care — A Measurement of Change : The Staging Concept". Medical Care, 1975, 13, 467-473.

Gonnella, J. S. et al. 'The Staging Concept — An approach to the Assessment of Outcome of Ambulatory Care". Medical Care, 1976, 14, 13-21.

42— Roemer, M. I. et al. "A Proposed Hospital Quality Index : Hospital Death Rates Adjusted for Case Severity". Health Services Research, 1968, 3, 96-118.

43— Goss, M. E. and Reed, J. I. "Evaluating the Quality of Hospital Care Through Severity-Adjusted Death Rates: Some Pitfalls". Medical Care, 1974, 12, 202-213.

44— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit., p. 36.

45— Williamson, J. W. "Evaluating Quality of Patient Care: A Strategy Relating Outcome and Process Assessment". Journal of the American Medical Association, 1971, 218, 564-569.

Williamson, J. W. "Outcomes of Health Care: Key to Health Improvement", in: **Methodology of Identifying, Measuring and Evaluating Outcomes of Health Service Programs, Systems and Subsystems.** Conference Series Outcomes Conference I-II, Dept. of U. E. W. Public Health Service, U. S. Government Printing Office, 1970, pp. 75-101.

46— Brook, R. H. and Appel, F. A. Op. cit., 1323-1329.

47— Kessner, D. M. et al. "Assessing Health Quality — The Case for Tracers" New England Journal of Medicine, 1973, 288, 189-194.

48— Ibid., p. 189.

49— Ibid., p. 190.

50— Brook, R. H. and Appel, F. A., Op. cit., p. 1326.

51— Pineault, R. "The Effect of Prepaid Group Practice on Physicians Utilization Behavior". Medical Care, 1976, 14, 121-136.

52— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit., pp. 53-56.

53— Williamson, J. W. "Evaluating Quality of Patient Care: A Strategy Relating Outcome and Process Assessment". Op. cit., 564.

54— Donabedian, A. **A Guide to Medical Care Administration,** Op. cit. pp. 40-41.

55— Lewis, C. E. et al. "Activities, Events and Outcomes in Ambulatory Patient Care". New England Journal of Medicine, 1969, 280, 645-649.

6

Modèles expérimentaux et quasi expérimentaux en recherche évaluative: une introduction

Thomas D. Cook, Fay Lomax Cook et Melvin M. Mark

Les demandes d'appréciation de l'impact de programmes sociaux actuels et nouveaux se multiplient constamment. Les spécialistes en recherche évaluative ont tenté de répondre à ces demandes, mais plusieurs programmes, y compris les programmes pilotes, n'ont pas encore été évalués par des procédés empiriques (Rutman et de Jong, 1976). En outre, des évaluations accusent tellement d'imperfections méthodologiques qu'il est difficile d'en conclure si les programmes *sont la cause* des résultats qu'on leur attribue (Bernstein et Freeman, 1975; McTavish et coll., 1975; Rutman et de Jong, 1976). Le présent chapitre vise principalement à proposer des méthodes susceptibles de réduire l'ambiguité des postulats de causalité en recherche évaluative.

Nous fixons notre attention sur le plan d'opération des évaluations sommatives, de préférence à celui des évaluations formatives. Les évaluations formatives sont effectuées pour procurer à un directeur ou à un gestionnaire de programme les moyens qui lui permettront d'améliorer le programme ou sa gestion. En général, ce type de recherche est entrepris par des membres du personnel de l'organisme qui accordent

Remarque des auteurs: Les auteurs expriment leur gratitude à Deborah Potts, Vita Carulli Rabinowitz et Myra Roche pour leurs observations sur une version antérieure du présent chapitre.

606

moins d'importance à l'aspect scientifique qu'au caractère d'urgence des retours d'information. Les évaluations sommatives, par contre, sont effectuées pour évaluer l'impact d'un programme, et elles sont généralement conçues et effectuées par des chercheurs étrangers à l'organisme évalué. Ordinairement, leur souci d'assurer la validité scientifique dépasse grandement celui que manifestent les personnes engagées dans des évaluations formatives.

La plupart des évaluations sommatives exigent que la collecte des données se fasse dans le cadre d'une expérience planifiée. Un tel cadre comporte: 1. l'intervention d'un événement quelconque que l'on peut qualifier de "traitement", et 2. une appréciation des effets de ce traitement au moyen d'un agencement systématique de groupes contrôle et(ou) d'un agencement systématique de mesures des groupes à l'étude. (Ces groupes sont généralement constitués de personnes, mais pas nécessairement. Les traitements peuvent aussi s'adresser à des quartiers, à des groupes de travail, à des écoles, etc.)

On peut distinguer deux types importants de recherche: la recherche expérimentale où les groupes expérimentaux sont assignés aléatoirement aux traitements, et la recherche quasi expérimentale où l'assignation aux traitements ne se fait pas d'une façon aléatoire parce que, le plus souvent, les individus ont choisi le traitement qui leur est assigné ou qu'ils sont assignés à un traitement par des administrateurs ou des professionnels convaincus que certains types d'individus doivent recevoir des traitements particuliers. Les recherches expérimentales produisent presque toujours des postulats de causalité plus fiables que les recherches quasi expérimentales (Campbell, 1969; Rivlin, 1971), mais de nombreux facteurs font qu'il est difficile de former ou de maintenir longtemps des groupes assignés aléatoirement. Le présent chapitre examine comment il faut s'y prendre pour prévoir et surmonter ces obstacles. Dans certains cas, cependant, il est tout simplement impossible de procéder à une recherche expérimentale, ou encore cette recherche échoue après avoir été mise en oeuvre. Par conséquent, la seconde préoccupation majeure de ce chapitre consiste à se pencher sur la conception et l'interprétation des recherches quasi expérimentales, mais nous présenterons toutefois, en premier lieu, quatre types de validité.

Quatre types de validité

La capacité de faire des inférences de causalité est capitale pour les évaluation sommatives, parce que la recherche vise à découvrir les effets que produit un programme. Il n'est pas question d'entreprendre à ce moment-ci une longue discussion philosophique sur l'inférence de causalité. Qu'il suffise de dire que pour établir des inférences de causalité il

faut démontrer: 1. que la cause est reliée à l'effet présumé; 2. que les modifications apportées à la cause précèdent celles qui surviennent dans les effets, et 3. qu'on ne peut trouver qu'une seule explication de la relation de cause à effet, à savoir que la cause telle que traitée a influé sur l'effet tel que mesuré.

Ces trois conditions étant acquises, il importe de faire des inférences de causalité pour établir si un traitement et ses effets éventuels sont *reliés* entre eux, sans tenir compte s'il s'agit d'une relation de causalité. On utilise la statistique inférentielle pour tirer des conclusions concernant une association simple. Il arrive malheureusement, toutefois, qu'on puisse l'utiliser d'une manière incorrecte, mais même lorsqu'on l'utilise correctement sa base de probabilité signifie que l'on peut tirer parfois des conclusions positivement fausses et négativement fausses (c.-à.-d. que des erreurs de type I et de type II se produiront). Nous emploierons l'expression *validité de la conclusion statistique* pour décrire la validité des conclusions concernant la relations statistique entre une cause présumée et un effet présumé.

Conclure en toute certitude qu'il existe une association statistique ne permet pas de conclure logiquement qu'il s'agit d'une association de causalité. L'inférence de causalité n'intervient que lorsqu'on a écarté toutes les explications concurrentes des raisons expliquant pourquoi le traitement tel que mis en oeuvre et les effets tels que mesurés sont reliés mutuellement. Par exemple, ils pourraient être reliés parce qu'une intervention survenue au même moment que le traitement a influé sur la mesure des résultats, ou parce que la mise en oeuvre d'un prétest a accru le rendement d'une mesure de post-test. Nous utiliserons l'expression *validité interne* pour désigner la validité des conclusions énonçant si l'on peut raisonnablement considérer que l'association statistique d'un traitement tel que mis en oeuvre et d'un effet tel que mesuré est une association de causalité.

Il arrive rarement que nous soyons intéressés à tirer des conclusions concernant un traitement tel que mis en oeuvre ou un effet tel que mesuré. Nous préférons nous en rapporter à un concept hypothétique général plutôt qu'à un mode particulier de traitement ou de mesure des résultats. Par exemple, nous voulons tirer des conclusions concernant une intervention personnelle à court terme par opposition au cas de "tel ou telle intervenant aidant une personne dans un climat de confiance qui leur permet d'aider le client à affronter des problèmes concrets comme administrer un budget, traiter avec des propriétaires, etc." Le problème consiste à savoir comment utiliser des techniques et des méthodes particulières pour tirer des conclusions plus générales, plus hypothétiques et plus pertinentes théoriquement. Nous entendrons par *validité de construct* la validité qui permet de formuler des conclusions concernant les concepts hypothétiques à partir d'opérations et de mesures. Aux fins de simplification, nous nous limiterons à la validité de construct des causes et

des effets, malgré le fait qu'il faille reconnaître que la question de définition est plus générale que cela.

La *validité externe* porte également sur la question de généralisabilité. Dans ce cas, cependant, il ne s'agit pas de généraliser, à partir de traitements et de mesures particuliers, à des concepts hypothétiques plus généraux. Nos efforts de généralisabilité visent plutôt des personnes, des lieux et des moments historiques. La question de validité s'énonce donc ainsi: jusqu'à quel point peut-on généraliser une relation de causalité à d'autres personnes, lieux ou périodes, ou à leur ensemble? Nous mentionnons la généralisation "à d'autres" et "à l'ensemble" parce que certaines questions de recherche exigent la généralisation à des groupes de personnes et à des époques particuliers (par ex. est-ce que l'intervention personnelle à court terme est efficace auprès d'adolescents dans une école secondaire?), alors que d'autres questions de recherche sont moins précises et exigent une généralisation à des groupes variés (par ex. avec quels groupes l'intervention personnelle à court terme est-elle efficace?). En outre, même si un chercheur peut façonner le modèle qui facilitera la généralisation des résultats de son étude à une groupe particulier, d'autres chercheurs ou parrains de recherche peuvent s'intéresser davantage à d'autres groupes et dire: En admettant que l'intervention personnelle à court terme est efficace auprès d'adolescents dans une école secondaire, et en admettant que c'est là l'objectif principal du chercheur initial, nous aimerions néanmoins savoir si cette méthode sera efficace auprès de la classe ouvrière défavorisée, auprès de névrosés de classe moyenne, etc.

On évalue les programmes pour savoir s'ils *produisent* certains effets. Par conséquent, la validité interne joue un rôle spécial dans les évaluations sommatives parce qu'elle s'intéresse à déterminer si les relations découvertes sont des relations de causalité. En accord avec cela, la discussion qui suit portera en grande part sur la validité interne. Il importe toutefois de ne pas confondre la primauté de la validité interne avec les coûts inadmissibles qui peuvent s'accumuler si l'on sacrifie d'autres types de validité pour accroître la validité interne. Pour citer un exemple extrême, une préférence non critique pour la validité interne pourrait inciter quiconque à n'entreprendre que des expériences aléatoires, se contentant dès lors de traitements à court terme dans des milieux qui portent au maximum le contrôle de l'expérimenteur. Le summum en ce domaine, c'est une expérience de cinquante minutes en laboratoire avec des élèves de 2e année du cours secondaire, situation expérimentale qui est nettement moins qu'optimale pour tirer des conclusions concernant les politiques de bien-être social. Les objectifs de recherche, le bon sens et les vérifications auprès des collègues stipulent jusqu'à quel point l'on peut compromettre la validité interne pour favoriser l'accroissement d'autres types de validité, quand et si un tel compromis s'avère nécessaire.

Accroître la validité

Dans ce contexte, nous allons expliquer brièvement comment il faut s'y prendre pour accroître chacun des quatre types de validité. Le lecteur doit se rendre compte jusqu'à quel point le plan de recherche est une question de choix, — les mesures prises pour accroître un type de validité peuvent souvent diminuer un autre type de validité. Vu que la suite de notre exposé traite des moyens d'accroître la validité interne, nous dresserons d'abord une liste des obstacles communs à la validité interne.

Accroître la validité statistique des conclusions

L'accroissement de la validité statistique des conclusions se produit en augmentant le pouvoir statistique des analyses et en utilisant la statistique inférentielle de manière à ne pas porter atteinte aux postulats importants ou à ne pas tirer profit du hasard. Les problèmes liés au pouvoir statistique sont minimisés:

1. en obtenant des échantillons de "grande" taille (la taille désirée dépend, bien entendu, de la taille anticipé de l'effet qui se rapporte à la variance anticipée);
2. en diminuant les sources accessoires d'erreurs (par ex. utiliser des groupes homogènes de répondants et uniformiser les agencements de mesure);
3. en rendant compte des sources accessoires de variance dans les analyses statistiques (par ex. par blocage ou par une analyse de covariance);
4. en augmentant la précision des mesures de résultats, et
5. en uniformisant l'application du traitement, de préférence à un degré élevé d'exposition au traitement.

On utilise validement la statistique: 1. lorsque les postulats importants servant de base à leur utilisation sont réalisés de façon évidente (par ex. le postulat de variances homogènes dans l'analyse de variance, ou le postulat d'erreurs non corrélatives dans l'analyse de séries chronologiques), et 2. lorsqu'on dispose de façon appropriée le nombre des comparaisons statistiques qui sont effectuées, étant donné qu'à moins de prendre des mesures de correction un nombre plus élevé de comparaisons que prévu sera statistiquement important par pur hasard au niveau choisi. Ces mesures de correction consiste à ajuster le niveau selon le nombre de comparaisons effectuées; à utiliser les tests de Newman-Keuls ou toute autre technique de comparaison multiple; ou fonder des inférences sur des essais de répliques multiples de la même relation pour l'ensemble des études ou, disons, en des droits distincts dans une seule étude[1].

6/0

Même si l'on a pu prendre toutes ces précautions, il demeure possible de conclure qu'un traitement est relié à un effet présumé alors qu'il n'en est pas ainsi (erreur de type I), ou qu'un traitement et un effet ne sont pas reliés alors qu'en réalité ils le sont (erreur de type II). Ces deux dernières menaces ne sont jamais écartées; voilà pourquoi les inférences concernant la corrélation sont en fin de compte probables mais non déterminantes. Néanmoins, ignorer les méthodes d'accroissement de la validité statistique des conclusions que nous avons très brièvement décrites, c'est accroître la probabilité de tirer de fausses conclusions au sujet de la corrélation.

Accroître la validité externe

La validité externe traite concerne la généralisabilité de la relation de causalité à d'autres groupes de personnes, à d'autres milieux ou à d'autres époques. La généralisabilité se limite aux catégories de personnes, de milieux ou d'époques représentées dans une évaluation *au moment où elle prend fin.* Ces catégories peuvent, ou non, représenter équitablement les groupes visés *initialement.*

On peut s'appuyer sur trois modèles principaux pour accroître la validité externe. L'*échantillonnage aléatoire* au sein d'une population choisie permet de généraliser les résultats à l'ensemble de cette population (c.-à.-d. qu'un échantillon national choisi au hasard permet de généraliser les résultats à l'ensemble des Canadiens). Bien qu'une sélection aléatoire soit désirable, sa valeur pour accroître la validité externe est limitée, parce que la sélection au hasard est rarement valable au titre de représentativité. Par exemple, on est chanceux si l'on peut découvrir un organisme social qui accepte de collaborer à une expérience, ou si l'on peut procéder à un échantillonnage aléatoire d'organismes au sein d'une population connue (par ex. celle d'une province) et, même, si l'on peut tirer un échantillon au hasard parmi les clients d'un organisme particulier. (Prenez note qu'il est question à ce moment-ci d'un échantillon tiré aléatoirement d'une population pour déterminer quelles sont les personnes qui feront l'objet de l'étude expérimentale, — et non de soumettre aléatoirement des groupes à des situations de traitement, — une question de validité interne. Nous expliquerons davantage cette question plus loin.)

Un modèle plus pratique d'accroître la validité externe consiste à choisir des groupes *hétérogènes* de personnes, de milieux ou d'époques. Par exemple, si l'on veut savoir dans quelle mesure la prestation de bons de logement permet de louer des maisons de meilleure qualité et mieux entretenues, on pourrait vouloir inclure dans l'échantillon des ménages dirigés par des femmes, des ménages dirigés par des vieillards, des familles dont le chef est sans emploi, etc. On tenterait alors de déterminer les groupes qui ont le mieux profité du traitement, indépendamment du fait que les familles ont été choisies au hasard pour être l'objet de l'expérience et, donc, indépendamment du fait qu'il existe un rapport *formel* entre l'échantillon et sa population référente. (Bien sûr, l'échantillonnage au hasard nous permet également de généraliser les résultats aux groupes qui font partie de l'échantillon.)

La *généralisation à des cas modèles* constitue un autre moyen pratique d'accroître la validité externe. Il s'agit d'expliquer comment un traitement serait vraisemblablement effectué s'il a perdu son caractère "expérimental" pour devenir une politique officielle et ensuite, d'identifier ou de créer à tout le moins un milieu de recherche où le traitement est effectué d'une façon qui correspond étroitement à la définition de milieu modèle. Par exemple, si l'on envisage d'effectuer un vaste programme de sessions d'orientation de groupes pour des étudiants défavorisés qui ne réussissent que peu au plan académique dans des écoles du centre-ville, nous chercherons à nous renseigner sur l'efficacité du programme d'orientation de groupes dans des banlieues. Par conséquent, ce modèle exige qu'on précise d'avance les groupes cibles de généralisabilité, et qu'on planifie ensuite le choix des personnes, des milieux et des périodes pour qu'il y ait une corrélation logique entre les groupes cibles et l'échantillon choisi.

Accroître la validité de construct

Accroître la validité de construct c'est pour ainsi dire adapter les opérations et les mesures aux concepts hypothétiques rigoureusement définis qu'elles ont pour fin de représenter. Etant donné que toute opérationalisation d'un construct est imparfaite, il est fort utile de mesurer ou d'agencer les concepts hypothétiques de diverses façons lorsqu'il est possible de le faire. Pour des raisons bien évidentes, il est généralement plus facile de mesurer diverses versions d'un construct de la variable dépendante que de manipuler plusieurs versions du même construct présumé de causalité. Par conséquent, nous considérerons à part l'une de l'autre la validité de construct de la cause et la validité de construct de l'effet, même si la même explication vaut pour les deux.

L'accroissement de la validité de construct de l'effet s'effectue habituellement en démontrant, d'abord, que les différentes mesures du même construct de causalité présumé sont en corrélation et, ensuite, que les mesures de construct ne sont pas en corrélation avec les mesures de constructs connexes mais différentes. Donc, la *logique* de l'analyse factorielle fournit probablement le meilleur modèle de validité de construct (parce que l'analyse factorielle vise à découvrir jusqu'à quel point plusieurs variables se raccordent comme élément d'un même phénomène ou facteur fondamental.) Deux évaluations de l'émission spectacle "Sesame Street" pour enfants donnent un exemple des questions de validité de construct soulevées dans un cas concret. Les évaluations initiales soutenaient que le simple fait de voir "Sesame Street" enseignait aux enfants comment lire (Ball et Bogatz, 1970; Bogatz et Ball, 1971). Une deuxième évaluation (Cook et coll., 1975) a toutefois démontré que

cette affirmation s'appuie sur des tests d'identification des lettres, mais que les *vrais* tests de lecture ne corroborent pas cette assertion.

Pour accroître la validité de construct d'un effet, nous examinons les relations de différentes mesures de construct de l'effet présumé entre elles et avec des mesures de construct semblables mais connexes. A cause de contraintes pratiques évidentes, nous ne considérons pas habituellement les agencements multiples d'un construct de causalité, ni les agencements de constructs semblables mais connexes. Par conséquent, la validité de construct de la cause est plus souvent accrue en démontrant qu'un traitement a modifié ce qu'il devait supposément modifier, mais rien d'autre. On appelle souvent cela la "prise" d'une variable autonome. Dans le cas de l'évaluation expérimentale de "Sesame Street", la validité de construct de la cause peut être mise en doute parce que des enfants avaient été visités chaque semaine ou chaque mois dans leur foyer par des chercheurs qui les encourageaient à visionner l'émission et qui leur ont offert des livres, des jouets et des ballons se rapportant à cette émission (Ball et Bogatz, 1970; Bogatz et Ball, 1971). Evidemment, ce traitement a été modifié davantage que le construct pertinent visé, — visionner "Sesame Street".

Il importe de souligner que les méthodes disponibles d'accroissement de la validité de construct ne se limitent pas aux seules périodes où l'évaluation est effectuée. La meilleure occasion d'accroître le lien entre un construct et son opérationalisation se présente plutôt si l'on effectue un test pilote avant l'évaluation.

Accroître la validité interne

La validité interne vise à découvrir si le traitement tel qu'utilisé a pu causer des changements dans l'effet tel que mesuré. Etant donné que le présent chapitre traite en très grande part de questions de validité interne, nous présenterons maintenant, en très peu de mots, quelques menaces courantes à la validité interne en nous référant à des remarques de Campbell et Stanley (1963). La liste n'est pas complète, mais on peut trouver une liste plus détaillée dans Cook et Campbell (1976).

Histoire. L'histoire constitue une menace quand l'effet étudié pourrait dépendre d'un événement, autre que le traitement, qui s'est produit entre le prétest et le post-test.

Maturation. La maturation constitue une menace quand l'effet étudié pourrait dépendre, non du traitement, mais de participants qui prennent de l'âge, deviennent plus sérieux, plus fatigués, etc., entre le prétest et le post-test.

Régression statistique. La régression statistique constitue une menace quand l'effet étudié pourrait être attribué aux participants assignés à des traitements en raison de mesures de prétest faillibles (ou d'autres mesures faillibles, quant à cela). Parce que les mesures ne sont pas fiables, ceux qui obtiennent des notes supérieures lors du prétest tendent à n'obtenir que de faibles notes lors du post-test, et ceux qui obtiennent de faibles notes lors du prétest tendent à en obtenir de meilleures lors du post-test. Une telle régression statistique déforme souvent l'effet réel du traitement.

Sélection. La sélection constitue une menace quand l'effet étudié peut être attribué à différents types de personnes dans divers divers groupes de traitement, plutôt qu'au traitement en tant que tel.

Interactions avec la sélection. Plusieurs menaces à la validité interne peuvent agir de concert avec la sélection pour falsifier les effets. Le cas probablement le plus courant est une interaction "sélection-maturation", où les groupes d'expérimentation se composent de divers types de personnes qui manifestent différents niveaux de maturité. Un autre cas pourrait être une interaction "sélection-histoire", laquelle se produit quand divers groupes d'expérimentaux visent des événements différents entre le prétest et le post-test; cette dissimilitude dans les événements historiques pourrait toucher par action différentielle les résultats du post-test.

Chacune de ces menaces particulières à la validité interne devient non plausible si l'on assigne au hasard les groupes expérimentaux aux traitements. Dans les études quasi expérimentales, vu qu'il n'y a pas d'assignation aléatoire, chacune de ces menaces doit être écartée, ce qui se fait normalement: 1. en incluant sagement les particularités du modèle que nous énumérerons plus loin; 2. en examinant les données supplémentaires qui pourraient influer sur la plausibilité de chaque menace, ou 3. en présumant, à cause de raisons théoriques ou du bon sens, qu'une menace particulière, — quoique possible, — n'est pas plausible en tant qu'explication de rechange dans l'évaluation particulière à l'étude.

Expériences aléatoires

L'assignation au hasard vise deux objectifs en recherche. Le premier objectif, — la représentativité, — a déjà été mentionnée comme étant un moyen d'accroître la validité externe. En tirant un échantillon aléatoire dans un univers bien délimité, nous pouvons généraliser nos résultats (dans les limites d'erreurs d'échantillonnage reconnues) de l'échantillon à l'ensemble de l'univers. Bien que cet univers soit normalement constitué de personnes, il peut aussi comprendre des milieux (par ex. organismes sociaux, écoles ou hôpitaux publics). Cependant, alors qu'un échantillonage aléatoire aux fins de représentativité constitue un bon moyen d'accroître la validité externe, cela s'avère souvent peu pratique.

Le deuxième objectif d'une assignation au hasard, — la comparabilité, — s'effectue en assignant aléatoirement des groupes expérimentaux aux traitements. L'assignation aléatoire aux fins de comparabilité est la carac-

téristique d'une expérience véritable. Cette méthode est importante, car elle garantit que les divers groupes traités ne diffèrent pas les uns des autres, *en moyenne*, au début d'une expérience. Une telle comparabilité écarte la plupart des menaces à la validité interne que nous avons mentionnées précédemment, parce qu'elles étaient fondées sur des traitements assignés à des groupes non comparables (par ex. sélection, sélection-maturation, sélection-histoire et régression statistique). L'assignation aléatoire aux groupes de traitement et la comparabilité du prétest qui s'ensuit accordées, nous savons que les différences de traitement postérieur observées entre différents groupes traités ne sont pas attribuables aux différences dans la composition des groupes. C'est dire que rares sont les autres interprétations des divergences du post-test qui sont viables; elles sont certainement moins nombreuses que lors des études quasi expérimentales où le chercheur doit énumérer toute les menaces pertinentes à la validité interne pour, ensuite, les écarter une à une.

Ces explications ne signifient pas que la planification d'une expérience véritable mènera nécessairement à une évaluation sommative réussie en tous points, car plusieurs problèmes assaillent l'assignation au hasard, en particulier: 1. les planificateurs et le personnel qui préparent et exécutent le programme peuvent s'opposer à une telle assignation comme moyen de répartir les traitements, en soutenant que d'autres critères d'assignation sont préférables, par ex. le besoin ou le mérite; 2. le processus d'assignation aléatoire peut ne pas être mis en oeuvre convenablement, ce qui produit des groupes expérimentaux non équivalents; 3. l'assignation au hasard peut échouer parce que différents types de personnes peuvent refuser de participer aux différents groupes de traitement ou décider de s'en retirer, et 4. l'assignation au hasard crée un climat d'injustice puisque certaines personnes bénéficient de meilleurs traitements que d'autres, ce qui peut entraîner des réactions que l'on pourrait faussement interpréter comme des effets du traitement. Nous allons donc analyser chacune de ces quatre difficultés, en précisant les moyens qui peuvent aider parfois à les surmonter.

Sources de résistance à l'assignation au hasard

Deux raisons générales expliquent pourquoi les responsables et le personnel de programme peuvent objecter de bons motifs contre l'utilisation de l'assignation au hasard aux fins de traitement: 1. la croyance que des critères d'assignation mieux appropriés (par ex. le besoin ou la compétence) sont disponibles et qu'il faudrait les utiliser, et 2. la croyance connexe que le traitement devrait être accessible à tout client éventuel, surtout s'il est perçu comme une promesse de situation meilleure.

La résistance fondée sur la croyance que le traitement ne devrait pas être soustrait aux contrôles est une chose assez courante. En effet, quand il a évalué un programme de counselling pour délinquants, Ward (1973) a dû faire face à une telle résistance de la part de ceux-là mêmes qui lui

avaient demandé d'entreprendre cette étude. Dans certains cas, on peut résoudre le problème en comparant un type de traitement (par ex. counselling auprès d'individus) avec un autre type de traitement (par ex. counselling auprès de groupes). Cela permet évidemment d'éviter le besoin d'un groupe sans traitement, et c'est utile quand on doit ou, à tout le moins, quand l'on accepte d'entreprendre une évaluation des avantages respectifs des deux situations. Parfois, les ressources disponibles ne sont pas suffisantes pour que tous ceux qui demandent ou ont besoin de ces ressources (ou traitements) puissent les obtenir. Dans ce cas, on peut soutenir que la loterie (ou la sélection au hasard) est le principe de distribution le plus équitable, spécialement lorsque l'impact positif des ressources n'est pas définitivement connu et que l'éventualité d'effets secondaires négatifs ne peut être écartée. *En outre*, on peut souvent dire aussi que des ressources, autres que celles procurées par le programme, peuvent plus vraisemblablement être disponibles à l'avenir s'il est possible de prouver sans ambiguité, — nous l'affirmons fortement, — s'obtient plus vraisemblablement quand les groupes ont été affectés aléatoirement aux traitements que lorsqu'ils l'ont été par d'autres critères de distribution.

L'argument le plus communément exprimé contre l'assignation aléatoire est que les ressources peu abondantes devraient être distribuées à ceux qui en ont besoin ou qui le méritent le plus. Ce point de vue postule que le mérite et le besoin sont validement mesurés. Dans la mesure où ils le sont, l'assignation au hasard n'est peut-être pas appropriée, mais dans la mesure où ils ne le sont pas l'assignation au hasard peut être une règle plus équitable de distribution que les autres solutions qui ne donnent qu'une apparence de distribution raisonnée. Même si l'on se fie grandement à l'exactitude des mesures de mérite ou de besoin, le fait de reconnaître son imperfection peut nous permettre d'effectuer une sélection au hasard à l'intérieur du groupe de sujets éligibles où l'incertitude est le plus répandue. Par exemple, une école régionale peut vouloir mettre sur pied et évaluer un programme visant à améliorer les défauts de lecture. On pourrait vraisemblablement s'opposer à ce que la sélection des sujets de l'expérience se fasse au hasard, sous prétexte que l'admission au programme doit se faire en fonction des besoins tels qu'indiqués par les faibles résultats obtenus lors des tests de lecture. On pourrait aisément accepter que les élèves qui ont obtenu les plus basses notes méritent de faire partie du groupe de traitement. Mais si, par exemple, cent enfants seulement peuvent prendre part au nouveau programme, qui est en mesure d'affirmer que ceux qui ont obtenu le 100e, le 101e ou le 99e rang se distinguent l'un de l'autre, ou le 110e et le 90e, ou encore le 125e et le 75e? En d'autres termes, on peut souvent cerner un secteur flou permettant que les élèves les mieux doués bénéficient du traitement, mais permettant également de choisir au hasard des élèves qui se rapprochent par leurs notes des "mieux doués" et de les inscrire au nouveau programme.

Considérer plusieurs sources de résistance à l'assignation au hasard, c'est supposer implicitement qu'un nouveau traitement sera avantageux. Ce n'est là qu'une supposition, toutefois, et faire remarquer qu'une telle croyance repose sur un manque de données fondamentales peut aider à diminuer la résistance jusqu'à un certain point. En outre, cela peut aider à diminuer la résistance jusqu'à un certain point. En outre, cela peut aider à démontrer que des évaluations rigoureuses de plusieurs programmes effectués naguère ont obtenu moins de succès que ne l'avaient prévu le grand public et, en particulier, les partisans du programme. La plupart des évaluations de l'efficacité des traitements thérapeutiques ou des tentatives de réhabilitation de prisonniers en sont des exemples éloquents. Une lecture rapide de ces évaluations pourrait abaisser la confiance naive de certaines personnes croyant qu'un programme sera avantageux, et peut les convaincre que l'évaluation rigoureuse de la valeur du programme vaut la peine d'être entreprise.

Dans les cas où l'on s'attend à rencontrer, au tout début, de la résistance à une assignation au hasard, l'acceptation de vraies expériences dépend de la possibilité de convaincre le personnel et les responsables de programme que l'évaluation de vérification de la réalité est effectuée avec le plus de validité au moyen de l'assignation au hasard, et que cette assignation au hasard peut se faire en pratique sans trop d'inconvénients pour les parties en cause. Il est donc avantageux que le chercheur ait préparé à l'avance: 1. des explications très simples sur les procédés d'assignation au hasard; 2. une liste d'évaluations dans un secteur particulier important où une expérience d'assignation au hasard a obtenu un plein succès (voir Boruch, 1974), et 3. des énoncés démontrant que l'assignation au hasard a été appuyée par des personnes estimables (par ex. Campbell, 1969; Rivlin, 1971) qui ont fait ressortir les biais que l'on retrouve généralement dans des modèles expérimentaux qui seraient probablement utilisés si l'assignation au hasard n'avait pas été effectuée.

Enfin, la résistance à l'assignation au hasard peut être un symptôme d'une résistance générale à l'évaluation. Entretenir de bonnes relations avec les personnes évaluées, faire appel à leur désir d'améliorer leur programme dans les cas où il fait preuve d'inefficacité, et prouver que les résultats de l'évaluation ne nuiront à personne en ce qui concerne la sécurité d'emploi, voilà des moyens susceptibles de mettre fin à toute résistance.

Assignation au hasard défectueuse

Problèmes de procédure dans l'assignation au hasard. L'assignation au hasard est une procédure garantissant que chaque groupe expérimental a autant de chance que les autres d'être choisi pour chaque traitement. Dans certains cas, l'assignation au hasard échoue à cause de vices de forme dans le mode d'assignation que l'on aurait pu corriger. Par exemple, lors du tirage au sort pour le service militaire en 1969 aux

Etats-Unis, les hommes dont la date de naissance se situait en fin d'année avaient plus de chance d'être enrôlés au bas de la liste les rendant admissibles au service militaire. Ce biais s'est produit parce que les mois de l'année ont été mis en ordre, de janvier a décembre, dans une urne d'où on les a tirés par la suite. L'urne n'a pas été secouées adéquatement, de sorte que les dates déposées les dernières avaient tendance à être tirées les premières, accordant ainsi aux personnes nées plus tard dans l'année de plus grandes chances d'être choisies (Feinberg, 1971; Notz et coll., 1971). On peut résoudre facilement ce problème en consultant des ouvrages expliquant comment l'on peut effectuer correctement une assignation au hasard. Par exemple, on trouvera dans Riecken et Boruch (1974) un exposé sur les procédures d'assignation au hasard.

L'assignation au hasard échoue parfois lorsque le personnel chargé de la mise en oeuvre de la procédure n'accomplit pas entièrement sa tâche. Souvent, la responsabilité de répartir les candidats a l'intérieur d'un programme est confiée au personnel permanent de l'organisme ou à des aides temporaires, et ces personnes peuvent être d'avis qu'il faut utiliser des critères autres que le hasard, par ex. le besoin ou le mérite, pour décider qui devraient participer au programme, ou elles peuvent ne pas être entièrement gagnées au principe d'un quelconque système d'assignation. L'assignation au hasard peut donc être complètement ou partiellement "faussée" pour qu'elle corresponde à l'opinion personnelle du responsable sur le mode de sélection à adopter, ou simplement pour satisfaire des candidats "récalcitrants" qui expriment ouvertement leur désir d'obtenir un traitement particulier. Autant que possible, on doit prévenir un tel sabotage bien intentionné en établissant des systèmes d'assignation au hasard qu'il sera difficile de contourner. Une simple procédure d'assignation au hasard, qui comporte deux étapes, aide souvent à le faire: la première étape exige que la personne chargée de sélectionner des candidats obtienne tous les renseignements nécessaires; la deuxième étape exige que l'assignation subséquente soit faite par une personne ignorant l'identité et les traits caractéristiques du candidat. Que l'on utilise ce système ou un autre pour réduire les erreurs dans l'assignation au hasard, On ne devrait jamais présumer qu'un processus bien planifié sera mise à exécution tel que prévu initialement.

Taille de l'échantillon et variabilité de l'échantillonnage. On décide parfois d'accorder un traitement à des groupes plus vastes qu'une seule personne. Par exemple, on pourrait soumettre tous les bénéficiaires de l'assistance publique de certaines villes à un traitement portant sur le mode de distribution de l'aide financière, alors que tous les bénéficiaires d'autres villes seraient choisis comme groupe contrôle sans traitement. De telles décisions sont prises souvent pour que les répondants soient moins au courant de l'existence d'autres situations expérimentales, car une telle prise de conscience peut entraîner de sérieux problèmes dont nous parlerons plus loin. Vu que la sélection de plus grands groupes

614

donne ordinairement comme résultat un moindre nombre de groupes à assigner à chaque traitement, il est moins probable que l'assignation au hasard de quelques groupes plus vastes donne comme résultat des groupes comparables.

La solution de ce problème consiste à jumeler les groupes en fonction de variables qui, — nous le croyons fortement, en nous appuyant sur la théorie ou sur le bon sens, — sont en corrélation avec la variable dépendante principale. Ensuite, nous choisissons au hasard, parmi les groupes jumelés, ceux qui seront soumis au traitement. Une deuxième solution, bien qu'elle semble être une approche plutôt banale, consiste à utiliser un plus grand nombre de petits groupes observés, c.-à.-d. qu'au lieu de soumettre des villes entières à un mode particulier d'assistance publique, on pourrait le faire au niveau des bureaux locaux de bien-être social, de sorte que plusieurs bureaux dans chaque ville seraient soumis à des traitements différents. Une troisième solution serait d'utiliser un modèle quelconque qu'on pourrait désigner comme "réplique de rechange", ce qui nécessite que l'on pourrait désigner comme "réplique de rechange", ce qui nécessite que l'on utilise d'abord un système de distribution dans un certain nombre de villes et un autre système de distribution en d'autres villes, et qu'ensuite les villes du premier groupe passent au système de distribution du deuxième groupe, et vice versa.

Comme alternative, on pourrait s'efforcer de mettre en oeuvre un plan de séries chronologiques, en utilisant les archives pour rassembler les mesures de prétraitement répétées de la variable dépendante, ou encore pour rassembler les mesures répétées durant la période d'exposition au traitement. Cette dernière partie des séries sera particulièrement utile si nous nous attendons à ce que le traitement produise à la longue des effets cumulatifs. Ces effets cumulatifs évalués en fonction de l'interaction statistique du traitement et de la durée du test. En d'autres termes, nous nous attendrons à ce que la différence entre les groupes augmente régulièrement après le traitement, et beaucoup plus qu'elle ne l'aurait fait avant le traitement.

Pertes de participants reliées au traitement

Refus reliés au traitement. Dans certains cas, les groupes observés seront assignés aléatoirement à des expériences, mais des individus préféreront ne pas participer au traitement qu'on leur avait destiné. Par exemple, Meyer et Borgatta (1959) mentionnent une étude qui visait à déterminer l'efficacité d'une méthode particulière de réhabilitation de personnes hospitalisées pour troubles psychiatriques. Le traitement a été effectué dans le cadre d'un atelier où les malades devaient séjourner pendant toute une année. Naturellement, plusieurs personnes ont préféré ne pas accepter de participer au traitement indiqué, et donc un tiers seulement des personnes choisies pour l'expérience ont vraiment participé au programme. Une telle décision personnelle se présente plus probablement dans deux cas: lorsque les candidats éventuels perçoivent qu'il leur en coûtera en efforts ou autrement pour recevoir le traitement indiqué, ou lorsque les traitements diffèrent intrinsèquement en fait d'agrément, comme dans le cas d'expériences en matière de revenu garanti qui offrent des garanties, disons, de 12 000$ au lieu de 3 600$. Les refus de participation au traitement indiqué contribuent à constituer des groupes qui, en moyenne, se composent de types différents de personnes, ce qui cause des problèmes de sélection et permet difficilement d'interpréter la relation entre le traitement et toute variable dépendante.

Le biais relié au traitement qu'on retrouve dans les groupes qui participent au traitement peut provenir de processus autres que l'autosélection. Par exemple, il peut être également le résultat de pressions qui font que certains groupes ou individus (par ex. des amis du directeur du projet) sont trop fortement représentés dans une situation particulière de traitement.

Abondons reliés au traitement. Même si l'assignation aléatoire aux traitements peut se faire avec succès, les participants peuvent encore se retirer pour des motifs différents d'un programme avant que l'évaluation soit terminée. C'est fondamentalement un problème de recherche longitudinale où les participants perçoivent que les différences de coût ou de gains sont reliées à différents traitements. Par exemple, lors de l'expérience effectuée au New Jersey à propos de l'impôt négatif sur le revenu, on a constaté que les cas d'abandon de l'expérience, après une période de trois ans, ont été de 25.3 p.c. dans le groupe contrôle à 6.5 p.c. dans le groupe des mieux nantis. Un abandon différentiel de ce type ne nous laisse que les groupes qui diffèrent en bien des façons importantes, ce qui menace sérieusement la validité interne.

Il existe trois méthodes principales de vérification de l'abandon différentiel. La première, et la plus simple, méthode consiste a vérifier si le pourcentage d'abandons varie d'un groupe de traitement a l'autre. La deuxième méthode consiste a clarifier les motifs d'abandon dans les différents groupes. Cela est utile parce qu'il devient possible d'obtenir les mêmes taux d'abandon, même si différents *types* de personnes peuvent avoir abandonné chaque groupe de traitement. Par exemple, imaginons un programme conçu pour améliorer réellement le sort des défavorisés. Les membres d'un groupe contrôle sans traitement pourraient abandonner parce qu'ils ne pouvaient payer leur loyer et qu'ils ont dû déménager ailleurs, alors que les membres d'un groupe observé pourraient abandonner parce qu'ils ont bénéficié du traitement et qu'ils ont pu obtenir un meilleur emploi ou un meilleur logement. Malheureusement, cette seconde méthode d'inspection des cas d'abandon différentiel exige des mesures directes du motif de l'abandon ou, lorsque ces mesures ne sont pas disponibles, elle exige l'expression d'une opinion éclairée.

La troisième méthode d'analyse des causes d'abandon différentiel exige que l'on rassemble des données antérieures au traitement. En gé-

néral, il s'agit de données de prétest, mais parfois des covariantes ou des variables de blocage sont également disponibles. Ces données sont ensuite examinées pour vérifier si les moyens de prétraitement sont différents dans l'ensemble des groupes de traitement *après que les personnes qui n'ont pas fourni de données de post-test ont été éliminées.* De telles différences laissent entendre que l'abandon différentiel a eu lieu et procurent une première, bien qu'imparfaite, estimation de l'ampleur de tout biais.

Il ne faut pas oublier que, si l'assignation au hasard a été faite au début, l'abandon d'une expérience signifie que le traitement a *causé* les différences dans la décision de poursuivre l'expérience. C'est souvent un effet interprétable et important. Par exemple, si l'on crée un programme pour encourager la stabilité d'emploi, l'abandon de l'expérience est l'ultime variable dépendante d'intérêt. Dans d'autres études, cependant, l'abandon n'est pas un résultat particulièrement intéressant, et il constitue généralement un problème parce qu'il rend les groupes de traitement non comparables, ce qui permet difficilement de tirer des conclusions de causalité précises relativement à tout résultat autre que l'abandon.

Solutions possibles. Il existe plusieurs moyens de traiter le problème que causent les refus et les abandons reliés au traitement. Plusieurs sources de pertes reliées au traitement peuvent être identifiées avant l'évaluation, et des mesures de correction peuvent être prises à cette étape pour surmonter les difficultés anticipées. On pourrait ainsi prévoir les pressions faites pour que des individus bénéficient d'un certain traitement. Si tel est le cas, l'assignation sélective peut être contrôlée et les personnes qui en bénéficient pourraient être éliminées de l'étude tout en bénéficiant du traitement, ce qui redéfinit le groupe expérimental et limite l'assignation aléatoire aux personnes membres du groupe redéfini.

Comme alternative, nous pourrions encore davantage redéfinir le groupe en n'assignant aux traitements que les personnes qui acceptent de participer à l'expérience, *quel que soit le traitement auquel elles pourraient être assignées au hasard.* La question de savoir quand procéder à une assignation au hasard est un problème complexe, et les problèmes de validité interne, de validité externe et d'éthique entrent en jeu. Logiquement, nous pouvons assigner aléatoirement des personnes aux traitements à trois moments précis: 1. dès que la réserve de groupes disponibles et éligibles a été identifiée, que les participants acceptent, ou non, d'être évalués ou "traités"; 2. quand les groupes ont accepté de participer à toutes les exigences de mesure de l'étude, même s'ils peuvent cesser de recevoir le traitement, et 3. quand les groupes ont accepté à l'avance de participé à l'évaluation et de faire partie du groupe de traitement qu'on pourrait leur assigner. L'assignation au hasard, au premier moment décrit, comporte le plus grand risque quant à la validité interne, étant donné que les participants ne se sont aucunement engagés et qu'on n'a donc pu identifier les abandons possibles reliés au traitement. Cependant, l'assignation à ce moment-là peut présenter le moindre risque de limiter la validité externe, vu que tous les groupes disponibles sont inclus dans l'étude et que leur ensemble n'a pas été redéfini à part de la population visée initialement par l'évaluation. Toutefois, nous devons nous rappeler que la généralisabilité se limite à la catégorie de personnes représentées dans l'évaluation, une fois celle-ci terminée, de manière que la validité externe peut être limitée par la perte de participants dans les études qui effectuent l'assignation au hasard à ce moment précis.

L'assignation au hasard, au troisième moment indiqué précédemment, devrait réduire les pertes de participants reliées au traitement plus que celles des autres moments et satisfaire aux règles d'éthique par le fait que les participants sont pleinement renseignés sur l'évaluation, tant en ce qui concerne les demandes de traitement que celles de mesure. Cependant, le fait d'avoir bien renseigné les participants peut causer d'autres problèmes. Par exemple, les groupes contrôle peuvent causer d'autres problèmes. Par exemple, les groupes contrôle peuvent réagir aux injustices provenant du fait qu'ils sont moins favorisés que d'autres groupes, et ces réactions peuvent induire le chercheur à interpréter faussement les effets du traitement. (Il sera question plus loin des réactions à ces injustices.) La deuxième solution, qui se situe dans un juste milieu entre les réactions extrêmes, est souvent préférable. Dans bien des cas, les individus continueront naturellement de participer au mécanisme de l'évaluation, comme, par exemple, dans les milieux d'éducation, ou ils consentiront à se plier aux exigences du chercheur en retour de petits avantages financiers. Toutefois, l'objet propre d'une évaluation particulière suggère souvent que l'on utilise l'une ou l'autre alternative, et chaque évaluation doit être ouverte aux transactions inhérentes à son choix du moment de l'assignation au hasard.

Une discussion antérieure à l'expérience peut également aider à identifier certaines sources du coût différentiel pour les participants, que l'on peut solutionner facilement. Par exemple, Hudson (1976) voulait évaluer les effets d'un dispositif d'enseignements autolétique qui se trouvait à plusieurs milles de l'endroit où résidaient les répondants. Pour atténuer l'abandon possible de membres du groupe de traitement, lesquels devaient se rendre jusqu'au dispositif (alors que les membres du groupe contrôle n'avaient pas à le faire), il leur a simplement fourni les moyens de transport jusqu'au dispositif. Un autre moyen connexe, qui permet de minimiser le problème de pertes de participants reliées au traitement, consiste à offrir des avantages financiers supplémentaires qui incitent à participer aux groupes de traitement les moins intéressants. Ce moyen semble avoir réduit le taux d'abandons au sein du groupe contrôle dans l'expérience effectuée au New Jersey a propos de l'impôt négatif sur le revenu, lequel taux était initialement de beaucoup supérieur à celui des groupes expérimentaux. A d'autres moments, on peut simplement offrir un choix de traitements aux participants, plutôt que d'assigner des personnes à un groupe contrôle sans traitement. En autant que les partici-

pants soumis à des traitements différents perçoivent que les traitements sont également attrayants, les refus et les abandons différentiels devraient être moins nombreux. Il est souvent possible d'élaborer des plans qui utilisent des traitements différents. Par exemple, Atkinson (1968) a étudié les effets d'un cours programmé par ordinateur en assignant au hasard des enfants à un cours de mathématiques ou d'anglais programmé par ordinateur. Chaque groupe a été évalué par rapport aux deux cours, de sorte que les notes d'anglais des élèves qui avaient suivi le cours de mathématiques servent de cadre de référence sans traitement pour le cours d'anglais, et que les notes de mathématiques des élèves qui avaient suivi le cours d'anglais servent de contrôle pour le cours de mathématiques. Comme alternative, on peut imaginer des cas où les personnes sont assignées à un type de psychothérapie de préférence à un autre type, chacun bénéficiant d'un mode de traitement.

Les plans offrant différents traitements à tous les participants soulèvent un problème. C'est que, si l'on peut évaluer l'efficacité d'un traitement par rapport à un autre, il est impossible d'évaluer aussi facilement leur efficacité absolue. C'est particulièrement problématique lorsque les traitements sont en fait également efficaces (ou également nocifs) car, sans un groupe contrôle qui n'est pas soumis à un traitement, l'effet de chaque traitement passerait inaperçu. Voilà pourquoi il importe tout particulièrement d'essayer d'établir quelle serait la base de référence sans traitement quand on utilise un plan grâce auquel tous les groupes reçoivent un traitement. Des méthodes autres qu'un groupe contrôle sans traitement sont disponibles pour réaliser cet objectif. Ce qui est le plus logique, même si ce n'est pas nécessairement le plus avantageux, c'est d'utiliser une longue série chronologique de traitements préliminaires, plutôt qu'une seule mesure de traitement. Il est également possible de recourir à l'une des nombreuses stratégies que nous décrirons plus loin quand il sera question des plans quasi expérimentaux.

On peut parfois réduire les pertes de participants reliées au traitement en utilisant avec discernement l'autorité dont disposent la plupart des organismes auprès de leurs membres. Par exemple, les écoles peuvent s'assurer que leurs élèves participent à certaines activités, dans le cadre d'un taux normal d'absentéisme, d'abandon et de déménagement; les forces armées peuvent s'assurer que leurs membres reçoivent certains traitements dans les limites reconnues de défections et de congédiements différentiels. Mais même si le recours au soutien organisationnel peut être d'une aide inestimable pour obtenir la collaboration et pour réduire au minimum les refus, sa nature *coercitive* est parfois douteuse. En premier lieu, la validité de construct est moindre parce que la contrainte se mêle au traitement. En deuxième lieu, il existe des questions d'éthique évidentes que très souvent l'on ne peut résoudre facilement, dont la plus importante est celle-ci; même lorsque certains organismes ont, de par leur nature, autorité sur des individus, quand, — si jamais, — est-il souhaitable qu'ils se servent de cette autorité à des fins d'expériences, particulièrement si l'utilisation de ce pouvoir accroît le ressentiment ou la passivité.

En dépit des meilleurs efforts du chercheur, des participants peuvent se retirer, pour différentes raisons, de différentes situations d'expérimentation. Il nous faut alors recourir à des stratégies d'analyse à posteriori des données; l'une d'elles consiste à effectuer l'analyse statistique en utilisant tous les individus qui avaient été assignés à un groupe particulier de traitement même si, en fait, quelques-uns d'entre eux n'ont pas participé au traitement. Par conséquent, si l'on se rapporte au cas (cité précédemment) d'un traitement de réhabilitation de malades hospitalisés pour troubles psychiatriques dans le cadre d'un atelier où ils devaient séjourner pendant une année entière, la stratégie serait de regrouper ensemble tous ceux qui avaient été choisis pour faire partie du groupe de traitement, qu'ils participent à l'atelier ou qu'ils refusent de le faire, et à analyser ensuite les données du prétest et du post-test qu'ils ont fournies comme étant l'ensemble des données du groupe. Ce mode d'analyse des données n'est possible que dans les cas où les gens se retirent d'eux-mêmes du traitement *sans quitter pour autant le cadre d'évaluation* et fournissent ainsi des renseignements postérieurs au test. Un tel cas se produit, par exemple, lors d'évaluations d'innovations académiques au niveau collégial. Des étudiants pourraient refuser de participer à un cours expérimental de comptabilité, par exemple, tout en continuant de suivre le cours régulier de comptabilité qui les mène à un diplôme de fin d'études.

Variations dans la mise en oeuvre du traitement

Nous avons soulevé auparavant le problème des participants qui ne reçoivent pas le traitement planifié auquel ils étaient assignés. Un problème connexe se pose quand les participants au programme ont la capacité de déterminer jusqu'à quel point ils désirent recevoir un traitement. Par exemple, lors d'un test portant sur un programme de counselling individualisé pour des personnes qui ne bénéficient que récemment des allocations sociales, Mullen et ses collègues (1972) ont constaté que le nombre de rencontres thérapeutiques individualisées varie de 1 à 129, la médiane se situant à 29, et que le nombre d'appels téléphoniques et de lettres échangées varie de 0 à 81, la médiane se situant à 9.5. Donc, l'exposition au traitement varie de la seule rencontre personnelle initiale requise avec chacun à un nombre de beaucoup supérieur. Cet écart n'est pas nécessairement dû à la seule bonne volonté des répondants. Il peut aussi s'expliquer par le fait que les personnes chargées de la mise en oeuvre du traitement, c.-à.-d. travailleurs en service social personnel, enseignants, conseillers, etc., se distinguent, en général, quant aux périodes de temps qu'ils peuvent consacrer à chaque répondant ou quant au nombre de cas à traiter. En outre, les personnes qui exécutent un traitement, par ex.

une méthode d'assistance individualisée, peuvent se distinguer quant à la qualité du traitement qu'ils offrent au moment de leur rencontre avec le répondant.

Standardiser le service de traitements, bien former le personnel du service de traitements, encourager les participants à se prévaloir sérieusement des services offerts et contrôler sans relâche ces activités, voilà des moyens aptes à réduire grandement le problème de l'hétérogénéité dans la mise en oeuvre des traitements. Néanmoins, si une telle hétérogénéité se manifeste en dépit de toutes ces précautions, la meilleure stratégie consiste à effectuer une analyse semblable à la méthode classique décrite dans le paragraphe précédent, c.-à.-d. que l'on devrait simplement traiter tous ceux qui sont affectés au groupe auquel ils avaient été initialement assignés. Mais cette analyse devrait être accompagnée d'une analyse quasi expérimentale qui consiste à classer les individus selon la durée ou l'intensité de leur exposition au traitement. Plusieurs analyses quasi expérimentales de ce genre ne se prêtent pas facilement à l'interprétation, mais habituellement on peut mieux les interpréter si le type de résultats obtenus est semblable au type de résultats de l'analyse expérimentale qui conserve l'assignation au hasard planifiée. Bien entendu, "semblable" ne veut pas dire "identique"; il faut user de bon sens pour établir à quel degré la similitude justifie la conclusion que les effets d'une analyse correpondent à ceux d'une autre analyse.

Réactions aux inégalités accentuées que cause l'assignation au hasard

La mise en oeuvre d'une assignation au hasard crée des inégalités entre les groupes, particulièrement quand l'on compare un groupe contrôle sans traitement avec un groupe soumis à un traitement. La prise de conscience de ces inégalités peut susciter des réactions que l'on peut interpréter par mégarde comme effets du traitement. Quatre types de réactions sont décrits dans la présente section, mais vu qu'ils bénéficient d'un ensemble commun de solutions, nous n'analyserons ces réactions qu'à la fin de cette section. On peut considérer les quatre réactions comme des menaces à la validité interne plutôt qu'à la validité de construct parce que leurs effets ne sont pas dus au traitement. Plus exactement, elles sont causées par une réaction atypique quelconque de la part de membres des groupes contrôle.

Décider d'inclure l'analyse de toutes les personnes qui se sont retirées peut paraître mauvais, parce que cela laisse place à un biais conservateur qui laisse entendre que les programmes sont moins efficaces qu'ils ne le sont en réalité. Comment se peut-il, dira-t-on, qu'un programme puisse être sérieusement évalué, quand l'analyse comprend des personnes qui,

dans le groupe évalué, n'ont pas ou ont bien peu bénéficié du traitement? Par contre, on peut soutenir que refuser d'inclure dans l'analyse les personnes qui n'on pas reçu de traitement c'est introduire des biais de sélection qui pourraient fausser les résultats. Il s'agit donc de se demander: Que préfère-t-on? Un test conservateur au point de vue statistique et sans biais, ou un test plus rigoureux au point de vue statistique mais avec biais? On peut, bien entendu, faire ces deux analyses, l'une incluant les personnes qui se sont soustraites au traitement et l'autre, les excluant. Si les deux analyses produisent des résultats comparables, il est facile d'établir un lien de causalité. Dans le cas contraire, il faut être très prudent quand l'on tire les conclusions, particulièrement celles de l'analyse où des personnes ont été exclues vu que des biais de sélection y ont probablement été introduits.

Une autre approche a posteriori consiste à examiner les données pour savoir si l'assignation au hasard est demeurée intacte dans certains sous-groupes, mais non en d'autres. Par exemple, dans les cas où les évaluations sont effectuées en plusieurs endroits, l'assignation au hasard peut n'avoir pas obtenu les résultats attendus en certains endroits, mais les avoir obtenus en d'autres, ou elle peut ne pas les avoir obtenus à l'égard de certains groupes quant à l'âge, au sexe ou à l'ethnie, mais les avoir obtenus à l'égard d'autres groupes. Après avoir découvert les sous-groupes où l'assignation au hasard est demeurée intacte, les données qu'on en recueille peuvent être analysées comme provenant d'une expérience d'assignation au hasard, avec tous les avantages qui en découlent pour la validité interne. Un examen a posteriori des données a été effectué à cette fin par Cook et ses collègues (1975) dans leur nouvelle analyse des évaluations de "Sesame Street" qu'avaient faites Ball et Bogatz (1970) et Bogatz et Ball (1971). Cook et ses collègues ont constaté que les groupes constitués au hasard étaient demeurés intacts en certains endroits d'expérimentation, mais non en d'autres, et ils ont mis l'emphase au cours de leur analyse des données sur l'expérience véritable qui avait été complétée en certains endroits.

Cette méthode d'étude de sous-groupes exige que l'on se concentre sur une partie seulement de l'ensemble des données, ce qui peut diminuer la puissance statistique. Elle devrait donc être utilisée parfois de concert avec les analyses quasi expérimentales utilisant l'ensemble des données mais qui n'ont pas le même impact que les approches d'assignation au hasard pour en induire la cause. L'utilisation des deux méthodes, — l'approche expérimentale avec un sous-groupe de répondants et l'approche quasi expérimentale pour l'ensemble des données, — renseigne plus clairement sur les liens de causalité quand les analyses tirent des conclusions identiques.

Diffusion ou imitation du traitement. La diffusion ou l'imitation du traitement se produit quand des membres du groupe contrôle sans traitement apprennent qu'il existe d'autres types de traitement et qu'ils sont

exposées à des événements réservés aux seuls groupes expérimentaux, ce qui réduit le contraste voulu entre les expériences et les contrôles et, par conséquent, embrouille les résultats réels du traitement. Par exemple, un travailleur social, dont le collègue applique une méthode de counselling expérimentale, peut être informé de cette méthode et l'utiliser auprès de son groupe contrôle. Ou encore, le groupe contrôle intentionnellement étranger à l'émission "Sesame Street" pourrait de sa propre initiative se mettre à regarder l'émission "Sesame Street" étant donné sa grande popularité (Ball et Bogatz, 1970). En outre, des considérations d'ordre judirique, d'éthique et de relations publiques peuvent inciter des membres du groupe contrôle à recevoir le traitement. Par exemple, lors d'une enquête sur l'efficacité des rencontres entre clients et avocats avant le procès (Rosenberg, 1964), un groupe choisi au hasard devait obligatoirement participer à de telles rencontres, alors qu'on avait simplement demandé au groupe contrôle de s'en abstenir. On ne pouvait leur refuser l'accès au traitement pour des raisons d'éthique et juridiques et, donc, plus de la moitié des personnes assignées au groupe contrôle ont décidé de participer à ces rencontres.

Rivalité compensatoire. Les membres d'un groupe contrôle sans traitement apprennent parfois qu'ils ne reçoivent pas de traitement, contrairement à d'autres groupes: cela peut parfois les pousser à modifier leur comportement pour que leur groupe ne soit pas surpassé au plan de la performance. Ceci peut se produire si les groupes existants, par ex. des groupes de travail, sont assignés à des situations contrôlées alors que des groupes voisins reçoivent les traitements. Il est également probable que les membres du groupe contrôle risquent d'être perdants si le traitement s'avère efficace. Saretsky (1972) cite le cas de l'OEO Performance Contracting Experiment pour mettre ce dernier point en lumière. Il démontre que la sécurité d'emploi des enseignants aurait été menacée si les résultats de l'expérience avaient démontré que les enseignants contractuels donnaient un meilleur enseignement que les enseignants réguliers. La meilleure façon d'écarter cette menace serait que les enseignants réguliers enseigne exceptionnellement mieux pendant l'année d'expérimentation de manière à réduire la possibilité que les enseignants contractuels obtiennent de meilleurs résultats lors du post-test.

Démoralisation rancunière des participants qui reçoivent les traitements les moins attrayants. Le fait de savoir que d'autres personnes font partie d'un groupe de traitement mieux favorisé incitera parfois ces personnes à éprouver du ressentiment et à se décourager, plutôt que de se comporter de manière à les surpasser au plan de la performance. Prenons l'exemple d'un projet d'évaluation de l'effet d'activités récréatives sur le comportement d'adolescents dépressifs. Les jeunes assignés au groupe contrôle pourraient observer le traitement que les autres reçoivent, se sentir relativement défavorisés et, même, devenir encore plus dépressifs. Si tel est le cas, toute différence constatée entre les deux

groupes après l'expérience ne serait aucunement due au traitement, mais au ressentiment ou à la démoralisation des membres du groupe contrôle. La démoralisation rancunière peut donc produire de faux résultats.

Egalisation compensatoire des traitements. La perception qu'ont les groupes expérimentaux d'être moins bien traités que d'autres groupes peut inciter un administrateur, peut-être sous la pression des membres du groupe le moins favorisé en ce qui concerne l'allocation de traitement, à offrir d'autres services semblables à sa disposition de manière à mettre sur un même pied les groupes de traitement. Par exemple, si une expérience était effectuée dans un organisme à vocation sociale en vue d'évaluer les effets d'une réduction du nombre de cas individualisés, l'administrateur peut prendre l'initiative de réduire lui-même la charge de travail des membres du groupe contrôle afin d'éviter une disparité trop prononcée du nombre de cas dont s'occupent différentes personnes. Ce type de réaction administrative se rencontre plus probablement quand l'administrateur n'est pas trop engagé dans l'évaluation et qu'il prévoit que les membres du groupe contrôle critiqueront ouvertement les disparités. L'égalisation compensatoire, comme la rivalité compensatoire et la diffusion ou l'imitation de traitement, tend à traiter sur une même pied d'égalité les groupes contrôle.

Solutions. Les quatre menaces à la validité interne dont il est présentement question ne ressemblent pas à celles dont nous avons parlé précédemment parce qu'elles ne sont pas écartées au moyen de l'assignation au hasard. Au lieu de cela, elles peuvent être écartées séparément, en fonction des données de base dont dispose le chercheur. L'information la plus pertinente est la certitude que les groupes qui font partie de l'expérience ont été isolés avec succès les uns des autres, ou qu'ils ne se sont pas informés mutuellement au sujet de l'expérience. Quand nous séparons les groupes expérimentaux, les individus ne peuvent réagir contre les disparités étant donné qu'ils n'en sont pas conscients. Ordinairement, la séparation des groupes expérimentaux implique l'utilisation de groupes plus vastes et moins nombreux (par ex. des écoles plutôt que des individus) qui ne communiquent pas régulièrement entre eux.

Les données de prétest, particulièrement les données de séries chronologiques sur les groupes contrôle, constituent une aide additionnelle pour déterminer si l'un des quatre types de réaction s'est produit. Si l'imitation du traitement, la rivalité compensatoire ou l'égalisation compensatoire se produit, le groupe contrôle doit alors s'éloigner carrément de sa ligne de base sans justification de traitement. S'il le fait, ces menaces deviennent d'autant plus plausibles. S'il ne le fait pas, il est improbable que quelque chose d'exceptionnel se produise dans le groupe. Il vaut la peine de rappeler que les données de séries chronologiques sont valables même si les séries sont abrégées, car même alors elles fournissent une estimation raisonnable de ce que serait une ligne de base sans justification. Ainsi, dans l'exemple de l'OEO Performance Contracting

Experiment cité précédemment par Saretsky, nous constatons que les étudiants du groupe contrôle ont obtenu plus de points lors des tests standardisés de réussite, au cours de l'année d'expérimentation, qu'ils n'en avaient obtenus au cours de chacune des deux années précédentes. Ce résultat renforce la prétention que la rivalité compensatoire peut avoir joué un rôle dans cette évaluation.

Le chercheur peut également obtenir des données complémentaires (préférablement des données de séries chronologiques) de groupes semblables aux groupes contrôle qui n'ont pas participé à l'expérience. Par exemple, si dans l'OEO Performance Contracting Experiment nous devions obtenir des données sur des classes d'enseignants ressemblant aux classes contrôle, et si nous devions constater que ces classes ont obtenu pendant l'année de l'expérience le même pourcentage de succès qu'au cours des années précédentes, ces données seraient entièrement compatibles avec l'interprétation voulant que la rivalité compensatoire a joué son rôle entre les classes contrôle et les classes d'enseignants.

Enfin, une série de post-traitements peut être également utile pour établir la plausibilité de ces menaces, spécialement lorsque des répondants sont retirés du traitement. Ceci met probablement, — mais non immanquablement, — fin à la disparité trop prononcée entre les groupes, et les facteurs qui contribuent à modifier la performance du groupe contrôle devraient disparaître. Dès lors, la performance du groupe contrôle devrait revenir à son niveau initial, de sorte que l'on puisse observer si la performance se stabilise, ou non, au niveau obtenu pendant la période où le traitement était mis en oeuvre.

Situer les difficultés de réalisation de l'assignation au hasard dans une perspective réaliste

La discussion précédente pourrait inciter le lecteur à se poser la question suivante: Si la mise en oeuvre d'une assignation au hasard crée tant de problèmes dans les évaluations, pourquoi n'a-t-on pas au départ retenu un modèle opératoire quasi expérimental? On pourrait lui répondre: 1. Il est souvent possible de surmonter ces difficultés, à condition de les avoir identifiées à l'avance et d'avoir envisagé des solutions pertinentes; 2. La plupart des problèmes que causent les groupes expérimentaux peuvent également se poser avec les groupes quasi expérimentaux. Par exemple, les membres de groupes contrôle soumis à un traitement quasi expérimental peuvent également découvrir qu'il existe de meilleurs traitements, et donc éprouver du ressentiment ou se laisser démoraliser; les départs peuvent se produire aussi bien dans les groupes quasi expérimentaux que dans les groupes expérimentaux; 3. La discussion précédente a mis en lumière les menaces intentionnelles à la validité qui peuvent donner par inadvertance l'impression que le biais qui se produit lorsqu'une expérience d'assignation au hasard échoue est plus grave que le biais qui se produit lors de quasi expériences de rechange faisables. Nous

sommes fermement convaincus, au contraire, que le biais découlant d'une expérience véritable qui a échoué est presque toujours moindre que le biais découlant de biais systématiques liés à l'auto-sélection et au recrutement qui caractérisent le processus d'assignation au traitement dans la plupart des quasi-expériences planifiées ou ex post facto.

Modèles quasi expérimentaux

Il est impossible d'utiliser des plans expérimentaux dans les cas suivants: lorsque la résistance à la procédure d'affectation aléatoire est trop forte; lorsqu'une innovation doit être appliquée à l'ensemble, ou lorsqu'il s'agit d'examiner l'impact d'un changement alors qu'il s'est déjà produit. Dans de telles situations, il faut recourir à un plan quasi expérimental. Dans ce cas, l'affectation des individus au traitement ne se fait pas au hasard parce qu'ordinairement ils ont choisi eux-mêmes le traitement qu'ils reçoivent ou qu'ils ont été affectés à un traitement par des reponsables ou des professionnels persuadés que certains types de personnes doivent recevoir des traitements particuliers. Le problème principal que posent les programmes quasi expérimentaux est que cette affectation non aléatoire permet difficilement de tirer des inférences de causalité. Il existe toutefois des particularités du mode opératoire qui permettent aux chercheurs d'éliminer certaines autres interprétations de toute relation obtenue entre les variables indépendantes et les variables dépendantes. Ces particularités accroissent notre capacité de tirer des inférences de causalité moins ambiguës.

Nous ne décrirons pas les modèles opératoires quasi expérimentaux particuliers, car on l'a déjà fait dans des ouvrages antérieurs (Campbell et Stanley, 1963, 1966; Cook et Campbell, 1976). Au contraire, nous nous attarderons à décrire certaines particularités qui peuvent renforcer l'inférence de causalité dans un modèle opératoire quasi expérimental. Cette position s'accorde avec la position des auteurs susmentionnés selon laquelle l'énumération des modèles n'est pas immuable et les particularités de quelques-uns de leurs modèles pourraient être avantageusement combinées avec celles d'autres modèles pour répondre aux besoins particuliers d'une recherche particulière. En fait, les particularités dont il sera question plus loin ne s'appliquent pas aux seuls modèles opératoires quasi expérimentaux, car elles peuvent et elles devraient être utilisées également pour mieux clarifier certaines interprétations faites dans les recherches expérimentales.

Nous mettons donc l'accent sur les éléments du plan opératoire qui peuvent renforcer les inférences de causalité dans une évaluation. Mais malgré cela, nous décrirons certains modèles opératoires quasi expéri-

620

mentaux fréquemment utilisés pour montrer jusqu'à quel point l'utilité d'un élément particulier du modèle opératoire dépend d'autres éléments particuliers du modèle.

Prétest

Le prétest est une mesure effectuée avant le traitement auprès de répondants, sur une échelle identique ou équivalente à celle qui sera utilisée lors du post-test. Le prétest est important pour la plupart des recherches quasi expérimentales, et son utilité découle du fait qu'il consiste à apprécier les différences initiales entre des groupes non équivalents.

Groupe contrôle sans traitement

Le groupe contrôle sans traitement est composé d'unités expérimentales (généralement des individus) qui ne reçoivent pas le traitement. Leur post-test est comparé avec celui du groupe expérimental afin d'établir si le traitement a eu un impact. Par conséquent, les groupes contrôle sans traitement servent de ligne de départ. Mais vu que l'on utilise des groupes non équivalents dans les modèles quasi expérimentaux, il s'ensuit qu'il est particulièrement inapproprié de se contenter des seules mesures de performance d'après test comme lignes de départ. La raison en est simple: si la performance d'après test diffère entre les groupes expérimentaux et les groupes contrôle, il ne sera pas facile d'en déduire que cela est dû au traitement ou aux compositions différentes des groupes qui pourraient expliquer les différents résultats de post-test obtenus.

Pour expliquer l'utilité du prétest et du groupe contrôle, particulièrement lorsqu'ils sont combinés, prenons l'exemple d'un nouveau programme d'activités récréatives visant à réduire le nombre des bagarres dans un foyer pour jeunes délinquants. Si le programme est mis en oeuvre et que la seule mesure d'évaluation du nombre de bagarres est prise après le test, on n'obtiendra sûrement pas une idée exacte de l'effet du programme. Le mieux que le chercheur puisse faire serait de procéder à une estimation subjective du nombre de bagarres avant l'évaluation, lequel correspondrait, exactement ou non, au nombre réel de bagarres.

S'il effectue en plus un prétest, le chercheur disposera d'un ensemble de données permettant d'estimer les effets du programme (c.-à.-d. la différence entre la performance au moment du prétest et au moment du post-test). Mais plusieurs menaces sérieuses à la validité interne compliquent l'utilisation d'un seul groupe mesuré à deux périodes différentes. Supposons, par exemple, que l'on constate une réduction statistiquement fiable du nombre de bagarres entre le prétest et le post-test. En déduire que le programme est la cause de cette réduction des bagarres peut se révéler inapproprié, car la réduction peut être due à certains événements autres que le traitement intervenu entre le prétest et le post-test. Ces autres interprétations de la "réalité" peuvent comprendre: la surveillance des jeunes délinquants a été plus rigoureuse entre le prétest et le post-test; les cliques rivales au sein de l'institution ont accepté une trêve; les conditions de vie du centre ont été améliorées pour des motifs étrangers au traitement. Avant que les chercheurs puissent raisonnablement conclure que le programme d'activités récréatives a réduit le nombre de bagarres, il faut écarter ces causes et également toute autre cause plausible.

Le type d'événement décrit ci-dessus n'est pas l'unique menace à la validité interne quand un seul groupe est évalué au moment du prétest et au moment du post-test. La régression statistique pourrait être une menace plausible si le programme d'activités récréatives a été mis en oeuvre pour mettre fin à un taux exceptionnellement élevé de bagarres. Ce taux élevé pourrait être dû, au moins en partie, à des facteurs aléatoires, y compris des erreurs de mesure. Dans la mesure où cela est vrai, on est en droit de s'attendre à ce que le taux de bagarres soit moindre que son niveau anormalement élevé, sans ou avec traitement.

La plausibilité de la réalité des faits et de la régression statistique, ainsi que l'intervention possible d'autres menaces à la validité interne (par ex. la qualité de l'instrument, ou les effets dûs non au traitement mais à un changement dans l'instrument de mesure entre le prétest et le post-test) compliquent généralement l'interprétation de l'évaluation n'utilisant qu'une mesure de prétest et de post-test d'un seul groupe.

Considérons maintenant une évaluation qui mesure les données des post-test du groupe expérimental et du groupe contrôle, sans tenir compte des prétests. Dans l'exemple cité plus haut, le taux de bagarres est mesuré après que le traitement a été mis en oeuvre simultanément dans le centre expérimental des jeunes délinquants et dans une institution semblable qui n'a pas reçu le même traitement. La menace de biais de sélection est tellement présente dans ce plan qu'il est impossible d'attribuer au traitement des différences entre le nombre de bagarres dans les deux institutions. Les différences peuvent être dues plutôt aux compositions différentes des deux groupes. Selon notre exemple, plus de délinquants agressifs ont pu être placés dans une institution que dans l'autre, ou il peut y avoir plus de bandes de jeunes délinquants dans l'une que dans l'autre.

La combinaison des deux plans que nous venons de décrire, — comprenant les mesures de prétest et de post-test d'un traitement et un groupe contrôle, — réduit les menaces dues à la sélection. En outre, dans ce plan (que nous appellerons modèle prétest et post-test du groupe contrôle) les données historiques, la vérification et quelques autres menaces à la validité interne sont éliminées. Bien qu'amélioré, le modèle demeure déficient dans la plupart des situations où on l'utilise, et on ne peut recommander sont utilisation que dans les cas où on peut disposer d'un meilleur modèle. Cependant, parce qu'il demeure un modèle fréquemment utilisé en recherche évaluative, les problèmes qui y sont associés méritent d'être considérés avec circonspection.

La sélection-maturation est la menace la plus probable à la validité interne dans un modèle opératoire de prétest et de post-test avec groupe contrôle. La sélection-maturation, comme il a été mentionné précédemment, est une menace quand l'effet observé peut être attribué à différents types de personnes dans des groupes qui sont à des niveaux différents de maturité (c.-à.-d. devenant plus solides, plus modérés, plus expérimentés, plus frustrés, etc.). Les taux de croissance différentielle de ce genre sont particulièrement courants dans les milieux scolaires où les élèves provenant des milieux socio-économiques les plus favorisés obtiennent généralement des résultats supérieurs, lors des tests de connaissances, à ceux des élèves de milieux défavorisés dans une période de temps déterminée. Dans plusieurs programmes de formation compensatoire, par exemple, le groupe de traitement se compose d'élèves défavorisés sur le plan de la scolarité, alors que le groupe contrôle est formé d'élèves dont la performance scolaire est de beaucoup supérieure et dont la réussite atteint des niveaux beaucoup plus élevés que ceux de la plupart des groupes expérimentaux plus défavorisés. Compte tenu de cette situation, il est difficile de confirmer les effets du traitement, car tout gain obtenu par le groupe expérimental à performance moindre devrait surpasser les gains supérieurs naturellement acquis par le groupe contrôle.

Parfois, en vue de contrôler le facteur maturité dans la sélection des sujets, le chercheur peut procéder à un jumelage des sujets (en effectuant une analyse ne portant que sur les élèves du groupe de traitement et du groupe contrôle qui ont obtenu des résultats similaires lors du prétest) ou recourir à diverses techniques de régression (telles que l'analyse de covariance, la régression multiple ou la corrélation partielle). Chacune de ces approches est inappropriée pour la raison suivante: le prétest et les corrélations contiennent des erreurs en ce qui concerne les modalités de mesure, et ce manque de fiabilité entraîne un rajustement insuffisant des différences entre les groupes. Il s'ensuit que l'on peut aboutir à des pseudo-résultats. La question du rajustement insuffisant concernant le jumelage et les techniques de régression est amplement traitée ailleurs (Lord, 1960, 1967, 1969; Cronbach et Furby, 1970; Campbell et Erlebacher, 1970; Campbell et Boruch, 1975) et ne sera pas considérée davantage dans ce chapitre.

Des problèmes de test et de mesure frappent aussi le modèle opératoire qui utilise des groupes non équivalents avec mesures de prétest et de post-test. Par exemple, la somme des changements qui peuvent se produire dans un groupe peut être réduite parce que les points obtenus sont initialement très élevés (ce qu'on appelle "effet de plafond") ou très bas (ce qu'on appelle "effet de plancher"). Il est très probable que ce problème produira des différences faussées dans la somme des changements obtenus lorsque la différence initiale entre les groupes est grande. Il importe donc que les groupes expérimentaux et les instruments de mesure soient choisis de manière à minimiser les effets de plafond et les

effets de plancher. Un second type de biais de vérification et de mesure se produit quand les moyennes des groupes diffèrent avant test et que la fiabilité des tests s'accroît du moment du prétest à celui du post-test. Cela arrive souvent, en particulier dans les recherches portant sur des adolescents, ce qui crée une différence faussée d'orientations prévisibles. La cause en est qu'une fiabilité plus élevée permet de mieux distinguer les groupes non équivalents, de sorte que si le post-test est plus fiable, les différences entre les groupes que le post-test a indiquées peuvent être plus prononcées que celles du prétest. (Voir Campbell et Boruch, 1975, pour une explication plus détaillée de ces biais.)

Le facteur régression peut être parfois une troisième menace plausible au plan prétest-post-test avec groupe contrôle. Reprenons l'exemple du programme d'activités récréatives visant à réduire le nombre de bagarres dans une institution pour jeunes délinquants, et supposons que la raison d'être du programme est que le nombre de bagarres avait été exceptionnellement élevé l'année précédente. Si nous pouvons démontrer que le nombre de bagarres a beaucoup plus diminué pendant la mise en oeuvre du programme dans l'institution expérimentale que dans une institution contrôle, ce résultat pourra fortement s'expliquer par le facteur régression. Cela s'explique, car la moyenne des bagarres du groupe expérimental était exceptionnellement élevée au moment du prétest à cause peut-être de facteurs anormaux qui ont produit une erreur de mesure, mais elle revient ordinairement au niveau stable de performance du prétest, différent du niveau atypique qui a poussé les responsables à créer un programme de redressement. La régression se fait donc probablement sentir lorsqu'un traitement est mis en oeuvre pour mettre fin à une performance exceptionnellement mauvaise, et il est moins probable que ce facteur constitue une menace quand les différences de prétest entre les groupes non équivalents se stabilisent avec le temps. Dans ce dernier cas; les résultats obtenus ont plus de chance de refléter des résultats élevés réels plutôt que des résultats élevés erronés.

Nous avons examiné trois types de menaces à la possibilité d'interpréter un modèle quasi contrôle au moment du prétest et du post-test. C'est un modèle quasi expérimental qui est souvent interprétable, mais il s'expose néanmoins à d'importantes menaces à la validité interne et on doit l'utiliser avec beaucoup de prudence.

Cohortes considérées comme contrôles

Le va-et-vient des clients dans certaines organisations et la promotion du personnel en d'autres donnent une idée de ce que sont les groupes contrôle non équivalents, que nous appelons "cohortes". Les cohortes sont des groupes de personnes qui sont promus au sein d'une organisation ou d'une institution et qu'on retrouve à divers échelons dans l'organisation. Les cohortes sont utiles comme contrôles parce qu'elles tendent à se ressembler les unes les autres dans l'expérience organisationnelle et

quant à leurs antécédents, ce qui minimise, — mais n'élimine pas nécessairement, — les problèmes de sélection. (L'exemple parfait d'une cohorte est peut-être celui de jumeaux identiques élevés dans le même foyer, car ils connaissent des expériences comparables sur le plan génétique, social et du déroulement de leur vie.) Prenons l'exemple d'une école comme milieu où il est possible d'analyser une cohorte. Chaque année un nouveau groupe d'élèves entrent en 3e année; on peut donc consulter les dossiers pour comparer comment les classes de 3e année se sont comportées avant et après le traitement. La profession de travailleur social offre un autre exemple: le progrès régulier des clients, grâce au programme de counselling, fournit les éléments nécessaires pour évaluer les effets de ce programme au moyen d'une analyse de cohortes.

La plus grande menace associée à l'analyse de cohortes est "le déroulement de la vie" ou l'histoire, car chaque individu ou groupe doit être mesuré à différents moments ou à différents âges ou échelons dans l'organisation. Par exemple, si une classe de 3e année constitue le groupe expérimental recevant le traitement, et si la classe de 3e année qui l'a précédée et qui a eu le même enseignant constitue le groupe contrôle, il est évident qu'un groupe a pu connaître pendant sa 3e année des expériences différentes de celles de l'autre groupe pendant sa 3e année. L'une des meilleures façons de vérifier les données historiques consiste à diviser les groupes de cohortes selon la durée de leur exposition au traitement. Cela sert de contrôle pour l'histoire parce que l'effet d'un traitement peut être plus puissant pour les groupes expérimentaux les plus longtemps exposés au traitement, bien qu'il n'existe aucune raison pour laquelle l'histoire devrait faire surgir cette différence. C'est facile de diviser selon la durée d'exposition au traitement quand il s'agit de groupes expérimentaux, mais comment le faire pour les groupes contrôle d'avant traitement. Parfois, la réponse est manifeste, par ex. quand les cohortes se composent d'enfants d'une même famille, car on peut classifier les enfants, par exemple, d'après le nombre d'heures passées à regarder l'émission "Sesame Street", puis obtenir des données de l'aîné de chaque famille qui avait le même âge que les enfants du groupe expérimental avant la mise en onde de l'émission "Sesame Street". Le groupe contrôle est alors subdivisé selon que, par la suite, les aînés se sont mis à regarder l'émission. (Pour connaître cette stratégie plus en détail, voir Cook et coll., 1975.)

Non-équivalence des variables dépendantes

Il est parfois possible de préciser, de pair avec les variables dépendantes, qui nous intéressent le plus, d'autres variables dépendantes qui ne devraient pas être influencées par le traitement, mais qui le seraient si d'autres facteurs plausibles (par ex., histoire ou maturation) avaient influencé les principales variables dépendantes. Par exemple, un organisme de service social peut établir un programme pour accroître les données

d'information que chaque travailleur social porte au dossier de chaque client. Pour éliminer l'explication selon laquelle les résultats obtenus pourraient être dûs, non pas au programme, mais à une surveillance accrue, nous pouvons examiner les variables dépendantes qui ne devraient pas être influencées par le fait de rassembler plus de renseignements sur les clients, mais qui pourraient l'être par la surveillance, par ex. le nombre de nouveaux cas étudiées par le travailleur social.

L'utilité de variables dépendantes non équivalentes est plausible si l'on choisit une variable dépendante qui n'est pas influencée par le traitement, mais qui peut l'être par d'autres facteurs de causalité plus plausibles. Ce choix exige donc une théorie initiale rigoureuse. En outre, il exige que l'on accepte l'hypothèse nulle voulant qu'il ne s'est produit aucun changement dans la variable dépendante non équivalente et que, pour des raisons d'ordre logique, il n'est jamais facile d'accepter l'hypothèse nulle. Enfin, l'utilisation de variables dépendantes non équivalentes présuppose que l'on ne généralise pas à la variable contrôle secondaire le changement que l'on espère observer dans la variable principale. Si l'on généralise, — rappelons que le plan exige des dépendantes variables *semblables*, — chaque mesure de changement sera influencée et l'on ne pourra pas établir une inférence causale même si un changement de causalité a pu se produire. Ces problèmes laissent entendre que les variables dépendantes non équivalentes sont le mieux utilisées, non comme éléments fondamentaux d'une évaluation, mais comme complément à d'autres éléments, par ex. prétests, groupe contrôle sans traitement, ou groupes cohortes.

Elimination du traitement

Il est impossible parfois d'obtenir un groupe contrôle non équivalent, et le plan doit être réduit à un seul groupe de traitement. Dans ces cas, il demeure encore possible de vérifier la corrélation du traitement et de l'effet présumé en comparant les moments où le traitement est présent avec les moments où il est absent. Par exemple, envisager un système de récompenses visant à prolonger la période de temps que les élèves problèmes passent à leur pupitre pour étudier, puis mettre le programme en vigueur, pour constater que la période de temps que les élèves consacrent sérieusement à l'étude s'accroit entre le prétest et le post-test. Eliminer ensuite le traitement et vérifier s'il en découle une réduction des heures consacrées à l'étude pourraient parfois éliminer les menaces à la validité interne dues à l'histoire, à la mesure et à l'instrumentation. L'influence du facteur histoire est devenue peu plausible parce que les événements historiques qui peuvent produire des effets contradictoires coïncident rarement avec les périodes où le traitement est introduit, puis retiré. La maturation est souvent écartée parce que les formes les plus plausibles de maturation *linéraire* sont exclues par l'accroissement initial, suivi d'une réduction, bien que des formes de maturation cyclique ne le soient

pas. Enfin, le facteur "mesure" est souvent écarté parce qu'il n'est nullement plausible d'admettre que le fait de se soumettre à un test accroît les résultats obtenus à un moment donné et contribue à les réduire après que le traitement a pris fin.

Plusieurs questions se posent lorsque l'on utilise ce genre de stratégie: 1. un problème de validité de construct surgit si l'élimination du traitement cause de la frustration, car tout changement apporté dans un sens opposé au changement initial prétest-post-test peut être attribuable à cette frustration; 2. si l'on ne mesure qu'à trois temps (c.-à.-d. *prétest* — traitement — *post-test*), une moyenne anormalement élevée enregistrée au premier post-test va indiquer l'évolution des résultats qui seront interprétés comme effets du traitement. Il importe donc d'exercer un contrôle pour faire disparaître cette possibilité. Le meilleur contrôle consiste à multiplier les post-tests avant d'éliminer le traitement. S'il est impossible de le faire, il faudrait alors vérifier si la moyenne anormalement élevée enregistrée au premier post-test peut être que l'effet du hasard, ce qui implique que l'on devra procéder à un rajustement du taux d'erreurs (Ryan, 1959), et 3. il devient évident qu'on ne peut éliminer le traitement que dans les cas où l'on prévoit qu'un traitement n'aura que des effets transitoires, et que ces effets disparaîtront aussitôt que le traitement sera éliminé. Tout effet durable mènerait à des conclusions fausses sur l'impact réel du traitement.

Répéter le traitement

Parfois il est possible non seulement d'introduire le traitement et de le retirer, mais aussi de le réintroduire une deuxième et, même, une troisième fois. En certains cas, la réintroduction se fait spontanément mais, en d'autres occasions, cette répétition est le résultat de décisions planifiées, car l'élimination du traitement laisse soupçonner que le traitement était initialement efficace. Réintroduire le traitement et retirer le traitement, de toute façon, c'est du pareil au même en ce qui concerne les avantages et les problèmes, c.-à.-d. que la plupart des menaces à la validité interne sont catégoriquement écartées, mais il demeure néanmoins possible que des effets apparents du traitement puissent être attribués au fait que la maturation cyclique ou l'élimination du traitement ait donné lieu à un désenchantement qui disparaît dès que le traitement est réintroduit. En outre, ceux qui participent à l'expérience peuvent s'apercevoir du changement survenu dans leur traitement et devenir méfiants. Le plan n'est valable que lorsque l'effet du traitement est tellement transitoire qu'il disparaît dès que le traitement est retiré, et qu'il peut réapparaître dès que le traitement est réintroduit.

Séries chronologiques interrompues

Retirer le traitement et le réintroduire laissent entendre que, dans certaines conditions, un groupe expérimental peut se contrôler lui-même.

La collecte de mesures multiples aux périodes de prétest et de post-test est utile à cet égard. Par exemple, Baldus (1973) voulait étudier les effets des lois de certains Etats américains exigeant que le gouvernement soit remboursé quand un bénéficiaire du Programme de sécurité de la vieillesse meurt et laisse des biens. Baldus voulait savoir si des personnes âgées ne réclameraient pas l'aide financière à laquelle elles ont droit si cette aide devait compromettre l'héritage qu'elles veulent léguer à leurs héritiers, particulièrement les habitations qu'elles désirent leur laisser. Le chercheur aurait pu simplement examiner le nombre de cas enregistrés auprès du Programme de sécurité de la vieillesse avant et après la promulgation des nouvelles lois de l'Etat, mais il ne l'a pas fait parce que ce procédé ne permet pas de contrôler toutes les menaces plausibles à la validité interne. Toute restriction observée dans le nombre de personnes âgées réclamant la pension de sécurité de la vieillesse pourrait, par exemple, être un effet de la régression statistique, comme ce serait le cas si l'assemblée législative de l'Etat avait voté ces lois au moment où le (ou à cause du) nombre anormalement élevé des demandes. On pourrait examiner la plausibilité de cette interprétation en établissant le nombre de demandes, pendant une longue période de temps, avant et après la promulgation de la loi. Si la régression à la moyenne est observée, la mesure d'avant l'introduction de la loi serait anormalement élevée pour la série chronologique ou la tendance dans le temps. C'est dire qu'une simple série chronologique peut écarter la régression (et, également, la mesure de maturation et d'instrumentation).

Par contre, le facteur histoire est le plus faible élément dans un plan de séries chronologiques. Dans le cas cité par Baldus, le facteur histoire serait tout événement responsable d'une réduction des cas et qui serait survenu au moment où la nouvelle loi sur le remboursement a été votée et mise en vigueur. Pour écarter ce facteur histoire comme menace, nous devons nous reporter aux séries chronologiques complexes qui combinent la simple série chronologique interrompue avec les autres facteurs étudiés antérieurement. Cette combinaison produit, par exemple, une série chronologique ininterrompue avec mesures dépendantes non équivalentes, ou une série chronologique avec groupe contrôle non équivalent.

Modèles complexes de séries chronologiques interrompues

Combinaison de séries chronologiques avec des mesures dépendantes non équivalentes. Supposons que Baldus croit qu'un changement dans le nombre de demandes d'aide est dû à la réduction générale des fonds de bien-être social, et non à la loi particulière dont il est question plus haut. En d'autres termes, les nouvelles lois peuvent avoir reflété les efforts déployés par l'Etat pour réduire les frais de bien-être social, et la réduction du nombre de cas peut découler du fait que les ressources de bien-être social diminuent et que les personnes âgées n'ont reçu aucune aide,

plutôt qu'au fait que ces personnes n'ont pas présenté de demandes d'aide. Si la réduction de l'aide aux personnes âgées est due au refus de consacrer des fonds publics aux bénéficiaires de l'aide sociale, on s'attendrait à ce que les montants consacrés aux autres catégories d'assistés sociaux diminuent au même moment ou les persones âgées connaissent une telle situation. Il y aurait moyen de vérifier cet énoncé en procédant à des mesures répétées auprès des autres groupes de bénéficiaires afin de voir si la même diminution a été observée au moment de la promulgation de la loi.

Combinaison de séries chronologiques avec un groupe contrôle non équivalent. La collecte de données sur des groupes comparables autant que possible aux séries chronologiques expérimentales permet de vérifier l'influence du facteur histoire. On suppose que les groupes contrôle non équivalents sont influencés par les mêmes facteurs historiques que les groupes expérimentaux. Il est clair que le fait d'ajouter la série chronologique du groupe contrôle non équivalent ne rend pas compte de tous les effets des facteurs historiques, mais simplement de ceux que les groupes ont en commun. Tout événement particulier qui touche un groupe mais non l'autre, à cause de différences entre les groupes, ne peut être contrôle dans le plan élargi.

L'addition d'en série chronologique contrôle n'écarte pas automatiquement toutes les menaces à la validité interne. Par exemple, une analyse chronologique des effets de la législation de l'avortement sur le nombre de naissances peut être sujette à des erreurs liées à la diffusion du traitement, particulièrement si l'on utilise les données de l'Etat et celles des Etats limitrophes comme contrôle. C'est parce que beaucoup de femmes dans les Etats contrôle peuvent franchir la frontière et subir un avortement ailleurs, que les différences prévues entre les groupes non équivalents et les groupes contrôle sont sensiblement réduites. Néanmoins, la série chronologique interrompue assortie d'une série de groupes contrôle non équivalents demeure un modèle très efficace.

Remarque finale

La mise en application fructueuse d'un modèle quasi expérimental authentique ou efficace est une tâche ardue en recherche évaluative. L'utilisation d'un tel modèle doit être appréciée en fonction d'autres modèles, lesquels sont des techniques de corrélation passive (comme une analyse causale) et des méthodes qualitatives (comme l'observation participante). Ces approches qualitatives et quantitatives passives ont des particularités qui peuvent bien compléter un modèle quasi expérimental authentique ou efficace. Cependant, on reconnaît communément que ces particularités ne sont pas suffisantes pour en déduire des inférences de causalité, alors qu'une inférence de causalité concernant l'impact du programme demeure le but principal d'une évaluation sommative. Parce qu'elles permettent des inférences de causalité plus fiables, on recommande donc l'utilisation d'un modèle quasi expérimental authentique ou efficace. Dans le présent chapitre, nous avons présenté certaines particularités qui devraient faciliter la mise en oeuvre de tels modèles et leur interprétation.

La prévoyance est un élément essentiel de l'exécution fructueuse d'une évaluation. Les problèmes que l'on peut rencontrer doivent être identifiés, et des mesures préventives doivent être prises pour qu'ils ne se produisent pas. En outre, lorsqu'on se propose d'enterprendre une évaluation, il faut y inclure des méthodes de découverte rapide de ces problèmes et d'autres problèmes imprévus. Cette prévoyance nécessaire s'acquiert le mieux en lisant de nombreux ouvrages portant sur des évaluations et au cours de longues discussions avec des collègues sur les problèmes qui peuvent surgir lors d'un travail d'évaluation.

Bien que l'exécution parfaitement réussie d'une évaluation ne soit pas facile, cela demeure une tâche stimulante, enrichissante et essentielle pour connaître l'impact d'une programme.

Remarques

1. Nous recommandons au lecteur de consulter les ouvrages de Myers (1972), Riecken et Boruch (1974) et Winer (1971) pour se mieux renseigner sur les questions soulevées dans cette partie du chapitre.

2. On discute sérieusement la question de l'analyse statistique appropriée du modèle prétest et post-test du groupe contrôle, que l'on désigne également comme modèle non équivalent (prétest — post-test) du groupe contrôle. Pour une étude des diverses méthodes d'analyse proposées, voir Cook et Reichardt (sous presse).

Références

Atkinson, R.C., computerized instruction and the learning process, *American Psychologist*, 1968, no. 23, pages 225 à 239.

Baldus, D.C., Welfare as a loan: The recovery of public assistance in the United States, *Stanford University Law Review*, 1973, no. 25, pages 123 à 250.

Ball, S. et G.A. Bogatz, *The First Year of "Sesame Street": An Evaluation*, Princeton: Educational Testing Service, 1970.

Bernstein, I. et H.E. Freeman, *Academic and Entrepreneurial Research*, New York: Russell Sage Foundation, 1975.

Bloom, B.J., J.H. Hastings et G.F. Madaus, *Handbook on Formative and Summative Evaluation of Student Learning*, New York: McGraw-Hill, 1971.

Bogatz, B.A. et S. Ball, *The Second Year of "Sesame Street": A Continuing Evaluation*, Princeton: Educational Testing Service, 1971, deux volumes.

Boruch, R.F., Bibliographie: Illustration d'expériences d'assignation au hasard pour planifier et évaluer des programmes sociaux, *Evaluation*, 1974, no. 2, pages 83 à 87.

Campbell, D.T., Reforms as experiments, *American Psychologist*, 1969, no. 14, pages 409 à 429.

Campbell, D.T. et R.F. Boruch, Illustration d'une assignation aléatoire au traitement en examinant des options: six façons dont les évaluations quasi expérimentales en éducation compensatoire ont tendance à sous-estimer les effets, A. Lumsdaine et C.A. Bennett (éd.), *Evaluation of Experience: Some Critical Issues in Assessing Social Programs*, New York: Academic Press, 1975.

Campbell, D.T. et A.E. Erlebacher, Comment les éléments de régression dans les évaluations quasi expérimentales peuvent à tort donner l'idée que l'éducation compensatoire est néfaste, J. Hellmuth (éd.), *Compensatory Education: A National Debate*, vol. 3, *Disadvantaged Child*, New York: Brunner/Mazel, 1970.

Campbell, D.T. et J.C. Stanley, Experimental and quasi-experimental designs for research on teachings, N.L. Gage (éd.), *Handbook of Research on Teaching*, Chicago: Rand McNally, 1963. (Publié également sous le titre *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally, 1966).

Cook, T.D. et D.T. Campbell, The design and conduct of quasi-experiments and true experiments in field settings, M.D. Dunnette (éd.), *Handbook in Industrial and Organizational Psychology*, Chicago: Rand McNally, 1976.

Cook, T.D., H. Appleton, R. Connor, A. Shaffer, G. Tamkin et S.J. Weber, *"Sesame Street" Revisited: A Case Study in Evaluation Research*, New York: Russell Sage Foundation, 1975.

Cook, T.D. et C.S. Reichardt, Statistical analysis of data from the non-equivalent control groups design: A guide to some current literature, *Evaluation*, sous presse.

Cronbach, L.J. et L. Furby, How we should measure "change" — or should we?, *Psychological Bulletin*, 1970, no. 74, pages 68 à 80.

Fienberg, S.E., Randomization and social affairs: The 1970 draft lottery, *Science*, 1971, pages 225 à 261.

Glass, G.V., V.L. Willson et J.M. Gottman, *Design and Analysis of Time-Series Experiments*, Boulder: Colorado Associated University Press, 1975.

Hudson, W.W., *Project Breakthrough: A Responsive Environment Field Experiment with Preschool Children from Public Assistance Families*, Chicago: Cook County Department of Public Aid, 1969.

Lord, F.M., Large-scale covariance analysis when the control variable is fallible, *Journal of the American Statistical Association*, 1960, no. 55, pages 307 à 321.

Lord, F.M., A paradox in the interpretation of group comparisons, *Psychological Bulletin*, 1967, no. 68, pages 304 et 305.

Lord, F.M., Statistical adjustments when comparing preexisting groups, *Psychological Bulletin*, 1969, no. 72, pages 336 et 337.

McTavish, D.G., E.E. Brent, J.D. Cleary et D.R. Knudsen, *The Systematic Assessment and Prediction of Research Methodology*, Minneapolis, Minn.: Minnesota Systems Research, Inc., septembre 1975, deux vol.

Meyer, H.J. et E.F. Borgatta *An experiment in Mental Patient Rehabilitation*, New York: Russell Sage Foundation, 1959.

Mullen E.J., R.M. Chazin et D.M. Feldstein, Services for the newly dependent: An assessment, *Social Service Review*, 1972, no. 46, pages 309 à 322.

Myers, J.L., *Fundamentals of Experimental Design*, Boston: Allyn et Bacon, 1972.

Notz, W.W., B.M. Staw et T.D. Cook, Attitude toward troop withdrawal from Indochina as a function of draft number: Dissonance of self-interest? *Journal of Personality and Social Psychology*, 1971, no. 20, pages 118 à 126.

Riecken, H.W. et R.F. Boruch (éd.), *Social Experimentation: A Method for Planning and Evaluating Social Intervention*, New York: Academic Press, 1974.

Rivlin, A.M., *Systematic Thinking for Social Action*, Washington, D.C.: Brookings Institution, 1971.

Rosenberg, M., *The Pretrial Conference and Effective Justice*, New York: Columbia University Press, 1964.

Rutman, L. et D. deJong, *Federal Level Evaluation*, Ottawa, Ontario: Centre for Social Welfare Studies, Carleton University, 1976.

Ryan, T.A., Multiple comparisons in psychological research, *Psychological Bulletin*, 1959, no. 56, pages 26 à 41.

Saretsky, G., The OEO P.C. experiment and the John Henry effect, *Phi Delta Kappan*, 1972, no. 53, pages 579 à 581.

Scriven, M., *Value Claims in the Social Sciences*, Lafayette, Ind.: Social Science Education Consortium, 1966, Publication No. 123.

Sween, J.A. The experimental regression design: An inquiry into the feasibility of nonrandom treatment allocation. Thèse de doctorat non publiée, Northwestern University, 1971.

Thistlethwaite, D.L. et D.T. Campbell, Regression-discontinuity analysis: An alternative to the ex post facto experiment, *Journal of Educational Psychology*, 1960, no. 51, pages 309 à 317.

Ward, D.H., Evaluation research for corrections, S.E. Ollin (éd.), *Prisoners in America*, Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Willson, V.L. Concomitant variation in the interrupted time-series experiment. Thèse de doctorat non publiée, University of Colorado, 1973.

Winer, B.J., Statistical Principles in Experimental Design, New York: McGraw-Hill, 1971.

8

La réalisation de l'évaluation dans la pratique: critères et limites

André Beaudoin

Tout au long de ce volume, l'évaluation a été présentée comme un processus formel de recherche et les méthodes d'évaluation sont empruntées au baggage des connaissances concernant les méthodes de la recherche existant dans les sciences de l'hommme. Cook, Lomax et Mark (cf. Ch VI) ont permis, en particulier, de comprendre comment les méthodes expérimentales et quasi-expérimentales satisfont le mieux les critères scientifiques de la validité (interne et externe), de la fidélité ou de la fiabilité et de l'objectivité. Malgré l'importance de ces critères scientifiques, il faut néanmoins comprendre que l'évaluation et la recherche scientifique ne poursuivent par tout à fait les mêmes fins: selon Nadeau "le chercheur vise des conclusions alors que l'évaluateur alimente des décisions" (Nadeau, 1981, p. 48). Autrement dit, l'évaluation, même lorsqu'elle utilise un modèle formel de recherche vise d'abord à fournir l'information nécessaire à l'utilisateur ou au preneur de décisions, alors que la recherche scientifique se concentre sur la conceptualisation et la compréhension d'un phénomène donné. Il ne s'agit pas là d'une conception opposant l'évaluation et la recherche mais plutôt d'une façon de voir l'évaluation comme étroitement liée à la prise de décision et à l'action. L'évaluation inclut cette dimension additionnelle d'être, comme le font ressortir clairement Stufflebeam et ses collaborateurs, "le processus par lequel on délimite, obtient et fournit des informations utiles permettant de juger des décisions possibles" (Stufflebeam, 1980, p. 48).

627

C'est pourquoi l'évaluation doit en plus des critères scientifiques satisfaire un certain nombre de critères pratiques si l'on veut qu'elle atteigne ses objectifs propres et que sa réalisation soit efficace. Le présent chapitre présente en premier lieu les critères pratiques de l'évaluation. En second lieu, on y trouve une analyse des problèmes qui proviennent de la place trop considérable accordée dans le processus d'évaluation aux seuls objectifs exprimés dans le programme et définis en terme d'efficacité et d'efficience alors qu'il n'existe pas en fait de méthodologie de la définition des objectifs. En troisième lieu, il s'adresse à la question de l'interférence entre les programmes et l'évaluation: interférence de la situation administrative et socio-politique du programme, ou à l'inverse interférence de l'évaluation sur l'opérationalisation et le fonctionnement du programme. Finalement, il présente les obstacles méthodologiques et la place qu'ils occupent dans le processus d'évaluation.

Critères de la réalisation de l'évaluation dans la pratique

En plus de critères scientifiques déjà énumérés, l'évaluation doit aussi correspondre, dans sa réalisation, à un certain nombre de critères pratiques qui servent de guide dans la recherche d'une information vraiment utile pour la prise de décision, que cette prise de décision soit celle de politiciens, d'administrateurs ou d'autres intervenants (praticiens sociaux, définisseurs de politique sur le plan de l'éducation, des affaires sociales, etc.) Ces critères impliquent une interaction réelle, qu'elle soit exprimée ou non, qu'elle soit consciente ou inconsciente, qu'elle soit formelle ou informelle, entre l'évaluateur qui agit comme un fournisseur de résultats et le décideur qui agit comme récepteur et utilisateur potentiel des résultats. Plusieurs auteurs ont d'une manière ou de l'autre explicité ce point de vue (Stufflebeam, 1967; Weiss, 1972; Edwards, Guttenbag et Snapper, 1975; Stufflebeam, 1980; Nadeau, 1981). On peut en fait identifier trois grands types de critères pratiques: l'isomorphisme, la crédibilité et le réalisme.

Isomorphisme

La première question qui se pose dans la réalisation pratique de l'évaluation est la suivante: jusqu'à quel point l'information que recueille l'évaluateur et qu'il se prépare à fournir au décideur est-elle du même type et de la même forme que celle nécessaire au récepteur ou à l'utilisateur compte tenue la réalité de ce dernier? L'isomorphisme de l'évaluation repose en fait sur trois caractéristiques essentielles: la pertinence de l'information, sa signification ou son importance et sa portée.

La pertinence réfère aux liens entre la nature des données recueillies par l'évaluation et les buts visés par le décideur ou l'utilisateur. Autrement dit, l'information est pertinente lorsque les données recueillies en cours d'évaluation seront en liaison étroite avec les buts poursuivis par l'utilisateur. Très souvent ce critère sera difficilement satisfait parce que l'évaluateur n'ira chercher que les données pour lesquelles il existe des instruments sans se préoccuper si ces données sont vraiment liées aux buts poursuivis par l'utilisateur. Bien souvent, rien ne risque d'être moins approprié à un niveau de décision d'ensemble que les seules données opérationnelles d'un niveau inférieur. Par exemple, il est sûr que les seules données provenant du niveau de satisfaction des étudiants ne sauraient être suffisantes pour établir si un programme atteint ses objectifs: ce serait aussi une erreur de croire que l'atteinte des objectifs de cours pris un à un est suffisant pour exprimer l'atteinte des objectifs d'un programme universitaire. En fait, il est souvent quasi-impossible d'en arriver à établir ces liens parce que tout simplement ils n'ont pas été clairement formulés au point de départ.

Il est évident qu'au niveau de l'établissement de la pertinence, l'utilisation de méthodes de recherche en évaluation peut être particulièrement appropriée. Elle est l'occasion pour l'évaluateur de se centrer sur un nombre limité de variables de sorte qu'il pourra plus facilement, en examinant l'information à recueillir dans l'évaluation, établir ses liens avec les buts qu'elle est destinée à atteindre.

Dans la plupart des situations d'évaluation, on peut songer à faire la cueillette d'une quantité considérable de données qui sont liées à un but ou un secteur de décision. L'information recueillie risque toutefois de ne pas être appropriée si on ne se préoccupe pas de sa signification ou de son importance. Elle risque d'être répétitive, redondante, triviale, ou sans lien évident avec ce qui est recherché par l'utilisateur. De la même manière; l'habitude fréquente de limiter d'étude des variables à ce qui est mesurable a aussi souvent les conséquences négatives: on risque d'ignorer des dimensions essentielles d'une réalité parce que tout simplement on ne dispose pas des moyens de les mesurer. En pratique, il s'agit pour l'évaluateur de sélectionner, parmi les données disponibles, celles dont les liaisons sont établies de manière évidente et claire avec la prise de décision. Cela nécessite qu'il porte un jugement de valeur sur les données à recueillir à la lumière des éléments susceptibles d'être plus significatifs pour l'utilisateur.

La troisième caractéristique du critère pratique de l'isomorphisme de l'évaluation est celle de la portée de l'information. Autrement dit, jusqu'à quel point l'information recueillie a-t-elle l'amplitude suffisante pour être vraiment utile au décideur ou à l'utilisateur? Ce n'est pas tout, en effet, d'avoir une information dont on a la certitude qu'elle est vraie, mais encore faut-il qu'elle soit complète, qu'elle soit suffisamment représentative du phénomène étudié et qu'elle permette de saisir toute la réalité. Ici, il faut souligner que la méthodologie de recherche peut même être nuisible puisqu'il n'existe, à notre connaissance, à peu près aucun

628

modèle de recherche qui soit réalisable en accordant toute la considération voulue à plus d'une poignée de variables. Par exemple, prenons un praticien social dans un service de consultation en planification des naissances; dans la consultation qu'il veut donner aux utilisatrices des différentes méthodes de contraception en vue de les aider à choisir celle qui est la plus efficace, il essaiera de fournir une information suffisante sur chacune des méthodes disponibles et sur son efficacité. Il ne se contentera pas de fournir l'information sur une seule méthode, car il ne pourrait vraiment pas dans de telles conditions permettre un choix véritable. Pour l'évaluateur, il y a là un dilemme car il ne s'agit par pour lui d'obtenir un éventail complet de renseignements sur une question. Ce serait risquer d'obtenir des renseignements qu n'ont pas de signification, sont redondants et trop nombreux. Il est essentiel que l'évaluateur en soit conscient et en tienne compte dans la dynamique de la réalisation de son évaluation.

En somme, ce premier critère pratique qu'est l'isomorphisme implique la réalisation simultanée des caractéristiques de la pertinence, de l'importance ou de la signification et de la portée. Il s'agit de produire une information qui "rencontre les buts de l'évaluation (la pertinence) sans trop entrer dans les détails (l'importance) ou qui soit trop étroite (la portée)" (Stufflebeam, 1980, p. 35).

Crédibilité

La crédibilité implique que les utilisateurs puissent accorder foi à l'information recueillie et utilisée par l'évaluateur. Assez souvent, les utilisateurs doivent tout simplement faire confiance à l'évaluateur parce qu'ils ne sont pas en mesure de déterminer les conditions de la validité, de la fidélité ou de l'objectivité de l'information.

L'établissement de cette crédibilité se fait en pratique au niveau de la qualité de la relation évaluateur-utilisateurs. Dans la mesure où très souvent il y a interaction entre les valeurs de l'évaluateur et celles des utilisateurs, on peut être en présence d'une menace potentielle à la crédibilité. L'emploi d'une méthodologie d'évaluation trop restreinte ou trop rigide risque d'engendrer chez les gestionnaires et les praticiens, en particulier, un manque de confiance dans l'évaluation parce qu'on aura tendance à douter de son utilité. Il est aussi assez fréquent que les données d'évaluation recueillies ne soient pas acceptées comme crédibles lorsqu'elles proviennent d'un autre palier ou niveau de la structure administrative, bureaucratique ou même politique. C'est la raison pour laquelle il faut s'assurer que l'évaluation soit réalisée le plus possible à des endroits et par des agents dans des conditions d'indépendance maximale, tout en étant en étroite liaison avec les utilisateurs de l'information. Il s'agit là d'une condition de concomitance importante pour la crédibilité. Par exemple, nous avons été à même de constater que l'introduction dans certains organismes gouvernementaux de divisions administratives ayant

des fonctions d'évaluation liées à des fonctions de planification est souvent source de distorsion sur le plan de la crédibilité. Rapidement, ces divisions en viennent à proposer des orientations qui apparaissent aux yeux des autres comme la promotion de leurs propres orientations et comme support à leurs propres choix de valeurs. Elles en viennent ainsi à réaliser leurs études selon leur propre philosophie de l'organisation de sorte que leur crédibilité est remise en cause par une partie des utilisateurs. A partir de ce moment, elles deviennent objet de résistance et identifiées à un groupe d'intérêt particulier ce qui les empêchera de produire des travaux crédibles.

Toute comme la validité interne, la crédibilité n'est pas généralisable mais reliée à la situation particulière sous étude. La crédibilité est plus facile à obtenir dans certaines circonstances que dans d'autres, comme nous avons été à même de nous en rendre compte dans deux études différentes. Dans une première étude, il s'agissait d'évaluer le fonctionnement et l'impact des services de consultation en planification des naissances sur les bénéficiaires dans trente (30) Centres locaux de services communautaires du Québec. L'évaluation avait été demandée et était faite grâce à un financement sous forme de subvention du Ministère des Affaires Sociales du Québec, dans le cadre d'une entente avec le gouvernement fédéral. L'équipe en charge de l'évaluation était rattachée au Laboratoire de recherche de l'Ecole de service social de l'Université Laval. Elle était donc externe au gouvernement et financée de manière assez indépendante puisque son financement était sous forme de subvention plutôt que contractuelle. Comme il s'agissait de services à leur phase d'implantation, le projet visait en même temps que l'évaluation à fournir un support au développement des intervenants (médecins, infirmières, travailleurs sociaux, sexologues, etc.). Cela a aidé dès le départ à impliquer ces intervenants dans l'étude évaluative et à obtenir leur contribution au moment de la réalisation de l'étude, en particulier au moment de la phase de cueillette des données. En fait, un seul des C.L.S.C. n'a pas participé à toutes les phases de l'étude auxquelles sa participation était prévue (cf. Varin, 1982; Bouchard, 1982).

Si dans l'étude précédente les conditions de réalisation ont augmenté la crédibilité, d'autres circonstances peuvent exercer leur influence de manière opposée. C'est ainsi que une étude destinée à évaluer un projet de développement rural intégré en Haute-Volta, la crédibilité a été beaucoup plus difficilie à établir (Beaudoin et Gagnon, 1978). L'équipe d'évaluation tout en étant située à l'extérieur de la compagnie chargée de la réalisation du projet a pu difficilement maintenir des conditions facilitant sa crédibilité. D'abord, même si le financement de l'évaluation était assuré par l'A.C.D.I., le budget affecté à l'évaluation était administré par la compagnie elle-même, créant un premier lien de dépendance sur l'utilisation des fonds. L'équipe en plus d'être composé d'un universitaire associé au projet à cause de ses connaissances des méthodes d'évaluation,

comprenait un autre chercheur qui avait été intéressé au projet comme ex-membre du conseil d'administration de l'association volontaire ayant passé le contrat avec la compagnie chargée d'exécuter le projet sur le terrain. Malgré sa connaissance directe du projet et le fait qu'il tentait au maximum de recueillir l'information complète et objective, ce deuxième membre était vu comme représentant de l'association volontaire. Le troisième membre de l'équipe était une sociologue qui avait été amenée au sein de l'équipe à cause de ses connaissances et son expérience sur le terrain dans cette région. Elle avait été sélectionée sur référence du président de la compagnie.

Comme l'évaluation s'est réalisée à la dernière étape de l'exécution du projet, il s'est agi, en premier lieu, après avoir formulé l'esquisse d'évaluation à partir des objectifs déjà énoncés, de procéder à la cueillette à partir des informations disponibles sur le projet au siège social de la compagnie (rapports d'étape, rapports d'activités, statistiques, etc.) avant d'aller sur le terrain. Malgré la collaboration accordée à l'équipe par la compagnie, il est évident que certaines informations ont été plus faciles à obtenir que d'autres. Pour le séjour d'observation sur le terrain (deuxième phase de cueillette des données), il avait été convenu que deux Voltaïques se joindraient à l'équipe d'évaluation. Toutefois, pour obtenir ces personnes et pour l'organisation du séjour, il a fallu nécessairement compter sur le p.d.g. de la compagnie qui a lui-même fait les démarches auprès des autorités voltaïques concernés. Une fois sur le terrain, il a été possible de se rendre compte que la crédibilité devenait difficile à établir pour l'obtention de l'information. L'équipe comprenait deux membres étrangers au pays et inconnus dans la région. Quant aux deux Voltaïques, ils étaient vus par les autochtones comme des contrôleurs gouvernementaux. Sur le terrain, on s'est rendu compte qu'il y avait aussi des conflits de valeurs et de perception entre plusieurs coopérateurs canadiens et la sociologue, à cause d'expériences antérieures. En plus, l'équipe était identifiée comme une extension de la compagnie. Il va sans dire que tous ces facteurs ont été des obstacles sérieux pour l'obtention de l'information nécessaire.

Réalisme

La réalisation de l'évaluation dans la pratique est vouée à l'échec si le processus d'évaluation n'est pas maintenu à l'intérieur de limites sensées quant à ses exigences de temps et de ressources financières et humaines. L'évaluation doit d'abord être définie de telle sorte qu'il soit possible de fournir l'information en temps opportun. Les évaluateurs, à cause de leur formation exclusivement scientifique ou à cause de leur insistance unique sur l'utilisation d'un modèle de recherche trop élaboré, sont très souvent peu enclins à rendre compte de leurs résultats tant que leur travail n'est pas entièrement complet et fait avec les nuances les plus infimes. Une telle façon de faire peut à la limite se retourner contre la possibilité d'utilisation des résultats. L'information, peu importe son degré de perfection, risque de devenir inutile pour le décideur ou l'utilisateur si elle est fournie trop tard. Souvent le décideur ou l'utilisateur peut avoir pris ses décisions ou avoir posé ses actes à partir d'autres données, sans incorporer celles de l'évaluation. De la sorte, des actions sont réorientées avant même que les résultats de l'évaluation ne soient connus parce que ces derniers ne sont tout simplement pas venus à temps. Comment être utiles en pratique si nos modèles et méthodes de recherche sont si immuables et dépendants d'une séquence si longue d'application qu'ils viennent trop tard pour ceux qui en ont besoin? Il faut dans la pratique faire preuve de créativité et réinventer l'utilisation des méthodes de recherche pour réussir à satisfaire les exigences du réalisme.

La prise en considération de ressources financières et le personnel est aussi une condition essentielle au réalisme de l'évaluation. L'emploi de moyens trop considérables en fait souvent déborder les coûts, en empêche la réalisation complète et diminuent la confiance qu'on peut lui accorder.

Etre réaliste dans la pratique de l'évaluation signifie être efficace dans la mise en application des moyens nécessaires à la réalisation des critères scientifiques de validité, de fidélité ou fiabilité et d'objectivité. Il s'agit aussi de garantir le respect des critères d'isomorphisme et de crédibilité en tenant compte des conditions de temps, d'argent et de personnel en liaison avec le décideur ou l'utilisateur.

Le processus d'évaluation: problème de la formulation des objectifs

Si les critères précédents permettent de situer la démarche pratique de la réalisation de l'évaluation dans sa perspective d'ensemble, le processus même de réalisation de l'évaluation n'est pas exempt de problèmes. Parmi ces problèmes, il y a la place accordée à la formulation des objectifs. Une des règles méthodologiques qui est la plupart du temps énoncée à la base de l'évaluation de programme, stipule que la mesure des résultats du programme se fait à partir des objectifs exprimés de sorte que la première étape de réalisation d'une recherche d'évaluation consiste précisément à déterminer les objectifs exprimés dans un programme.

Rutman fait ressortir cet aspect:

> "La majorité des définitions qui décrivent la recherche d'évaluation mentionnent que les programmes sont évalués en fonction des objectifs exprimés. Cette position implique qu'il y a une valeur sous-jacente ratta-

chée au but poursuivi et que la recherche consiste à identifier et à déterminer dans quelle mesure le programme atteint ce but" (1977, p. 17)*

C'est l'importance considérable accordée à la définition des objectifs qui a été remise en question par plusieurs auteurs. Dans une conférence qu'elle prononçait au Colloque international sur l'administration publique, à Québec en mai 1979, Elinor Ostrom qualifiait même cette règle de "méthodologie en débandade" (Ostrom, 1979).

En pratique, cela signifie que l'insistance quasi-rituelle mise sur l'identification des objectifs et de leurs composantes dans la réalisation de l'évaluation est remise en question. Y a-t-il une obligation "sine qua non" dans la pratique de l'évaluation de passer par cette étape? S'agit-il d'une étape-clé essentielle? Sur quoi se base-t-on pour déterminer les objectifs? Comment s'assure-t-on que les objectifs sont bien définis et que ceux qui ont été définis sont les "bons"?

L'importance trop considérable accordée aux seuls objectifs formels

L'importance que prend la formulation des objectifs comme première étape du processus d'évaluation est déterminante de toutes les autres étapes de sa réalisation. Cette première étape concerne l'incorporation des énoncés formels d'objectifs du programme comme point de départ de l'évaluation. On risque cependant d'être tout à fait incomplet si cela se fait de manière exclusive, sans tenir compte d'autres dimensions importantes.

Il y a d'abord risque de distorsion si on ne tient pas compte des différences entre les objectifs formels pour un programme et leur réalisation dans son exécution. L'exécution d'un programme ne se fait pas toujours, dans la pratique, dans la stricte actualisation des objectifs exprimés, même si ceux-ci le sont de manière opérationnelle. Les circonstances, les personnes ou les situations visées par le programme, le personnel, le plus ou moins grand manque d'adéquation entre la situation identifiée et la réalité existante, etc, peuvent affecter et même modifier assez considérablement des objectifs formels exprimés. Il y ainsi une distance entre les objectifs formulés et leur exécution opératoire. Les mêmes objectifs formulés pour un programme au niveau central d'une organisation, même acceptés par d'autres paliers ou composantes, prendront souvent une coloration différente dans leur formulation et leur exécution aux différents paliers de l'organisation. Dans la pratique de l'évaluation, s'en tenir aux énoncés exprimés à un seul niveau sans tenir compte de leur transposition dans l'exécution, c'est accorder une importance trop exclusive et déterminante à ce seul aspect et prendre un trop grand risque d'ignorer des éléments essentiels dans le processus d'évaluation.

*Texte reproduit dans ce volume, Chapitre II

Rutman attire aussi l'attention sur deux autres aspects importants: celui "de certains objectifs latents (ceux qui ne sont pas formellement exprimés) et des résultats non anticipés" (1977, 17)*. Il n'est pas nécessaire d'insister plus longtemps pour dire qu'il existe très souvent certaines intentions non exprimées dans les objectifs d'un programme. Elles n'en demeurent pas moins présentes tant dans le programme dans son ensemble qu'au niveau des différents paliers de l'exécution opératoire. Dans la pratique de l'évaluation, il s'agit de découvrir au mieux ces intentions et d'en tenir compte. Il en est de même pour les résultats non anticipés, qui sont des conséquences non prévues du programme. Même si elles sont beaucoup plus difficiles à inclure dans un processus formel d'évaluation, on peut le faire grâce à l'utilisation de méthodes plus qualitatives d'évaluation, comme l'indique Lecomte dans un chapitre ultérieur de ce volume.

A partir de considérations précédentes, il faut même se demander s'il ne faut pas accorder une importance plus forte dans la pratique de l'évaluation aux points de vue exprimés par les premiers analystes de systèmes en recherche opérationnelle (Hitch, 1960, 1; Wildavsky, 1966, 299) à l'effet de prendre pour acquis qu'il n'y a pas d'objectifs connus à l'avance ou qu'ils changent très rapidement. Il ne s'agit pas par là de nier l'importance de la formulation d'objectifs dans le processus d'évaluation, mais d'éviter de réifier les objectifs formels exprimés comme l'unique point de départ de l'évaluation à partir duquel toutes les autres étapes sont dépendantes.

L'efficacité et l'efficience comme critères de mesure de l'atteinte des objectifs

En liaison étroite avec le problème d'une attention trop forte et exclusive attachée aux seuls objectifs formels exprimés, il y a le problème de l'exagération faite dans la mesure les résultats, aux seuls aspects de contenu concernant l'efficacité et l'efficience des objectifs identifiés. Pour l'évaluation, il est souvent tentant et plus facile de ne partir que de ce qui est défini comme l'efficacité du fonctionnement administratif du programme. A titre d'illustration, Ostrom (1979, 1626 à 1630) prend l'exemple d'un projet de la mesure de la performance de la police aux Etats-Unis. Le projet partait d'une définition de la performance de la police qui incluait un ensemble d'éléments destinés à montrer jusqu'à quel point les différents systèmes policiers correspondent aux exigences et attentes définies à leur égard. De telles attentes sont exprimés de manière plus ou moins explicite à partir des stipulations de la loi, des règlements de la police et des façons de fonctionner d'un service de police. Mais malgré une définition aussi large et aussi englobante de la perfor-

*Texte reproduit dans ce volume, chapitre II.

631

mance, lorsqu'on en est arrivé à définir le contenu servant à mesurer cette performance dans la pratique, on a restreint ce contenu presqu'exclusivement à des mesures d'efficacité et d'efficience des policiers, définis à partir de critères individuels comme le degré auquel les policiers atteignent ou réalisent les différents objectifs qui leur sont assignés, le rendement qu'ils fournissent dans l'accomplissement de ces activités et de ces tâches et leur niveau d'intégrité personnelle.

Comme le souligne Ostrom (1979, p. 1629), on accorde dans un exemple comme le précédent trop d'importance lorsqu'il s'agit de mesurer la performance, aux seuls éléments servant à d'écrire le fonctionnement d'un département de police. Cela conduit à ignorer d'autres facteurs tout aussi significatifs comme la conformité des opérations de police avec les spécifications de la loi, avec les politiques gouvernementales, avec les attentes du public et avec les valeurs fondamentales d'une société.

Le danger illustré par l'exemple précédent est très présent dans la plupart des situations d'évaluation. L'évaluateur doit faire un examen sérieux pour identifier les objectifs réels d'un programme et les distinguer des seuls objectifs de la gestion de ce programme. En particulier en matière d'évaluation de programme, ce serait risquer des erreurs sérieuses de restreindre le contenu des résultats anticipés à la seule mesure des objectifs de gestion d'un programme, sans tenir compte des objectifs plus essentiels qui, somme toute, sont reliés aux effets du programme à l'extérieur de son propre système de fonctionnement. L'évaluation n'est pas seulement une aide aux décisions de gestion, mais une aide aux décisions relatives au contenu de programme et à son orientation fondamentale.

L'absence d'une méthodologie de la définition des objectifs

Malgré l'accord généralisé à l'effet que l'évaluation doit se faire à partir des objectifs du programme, il n'existe pas de méthode bien déterminée et précise pour ce faire. Il n'y a pas de critère explicite pour savoir si on a réussi à définir les "bons objectifs". L'étude des différentes façons de définir les objectifs est assez révélatrice à cet égard.

A partir de ce qui a été dit dans les deux sections précédentes, il est clair que les seuls objectifs exprimés ou encore les seuls critères d'efficacité et d'efficience ne peuvent servir comme base à une méthodologie de la définition des objectifs. Il en est de même des énoncés de politique générale, qui sont souvent définis en termes si vagues qu'ils peuvent prendre toutes sortes de directions sans compter que bien souvent, ils ne contiennent pas d'objectifs à proprement parler.

Une autre façon a été suggérée pour définir les objectifs d'un programme: elle consiste à le faire en se tournant vers les membres de l'organisation ou ceux qui sont susceptibles d'être impliqués dans le programme. A qui faut-il s'adresser lorsqu'on décide d'utiliser une telle approche? Au directeur général? aux cadres supérieurs? aux responsables du programme? aux intervenants? Très rapidement on se rend compte que la décision de s'adresser à des membres ou des groupes de membres particuliers de l'organisation plutôt qu'à d'autres devient un choix qui peut faire varier beaucoup la définition des objectifs. En fait, il y a souvent des différences assez considérables dans la façon de concevoir les objectifs du programme non seulement selon les personnes ou groupes situés à différents paliers de l'organisation, mais même entre ceux qui sont situés au même niveau de l'organisation. Il est donc difficile d'affirmer comme règle pratique, qu'utiliser la définition des objectifs donnés par les membres de l'organisation est un moyen nécessairement adéquat pour s'assurer qu'on en arrivera à déterminer les "bons" objectifs du programme à évaluer.

On a aussi suggéré, comme point de départ à la formulation d'objectifs, l'utilisation des énoncés faits par des experts chargés d'étudier, d'analyser ou de proposer des orientations sur une question quelconque. Par exemple, pour évaluer le programme d'assurance-maladie, les programmes de santé et les programmes de services sociaux au Québec, on pourrait se référer aux objectifs formulés dans les rapports de la Commission d'Enquête sur la santé et le bien-être social où on retrouve de tels objectifs énoncés (vg: *L'assurance-maladie*, 1967, Vol. 1, p. 35-49; *La santé*, 1970, Vol. IV, Tome II, p. 11-24; *Le développement*, 1971, Vol. III, Tome I, p. 224-230; *Les services sociaux*, Vol. VI, 1972, Tome I, p. 156-170). Toutefois, très rapidement on se rend compte que la plupart du temps les objectifs formulés par de tels groupes d'experts ou dans les rapports de commissions d'enquête peuvent être difficilement transposables parce qu'ils manquent de l'une ou l'autre des caractéristiques nécessaires à leur évaluabilité:

1) Ils ne sont pas énoncés de manière assez précise dans leur définition pour qu'on puisse les transposer en éléments mesurables au minimum selon une échelle nominale, c'est-à-dire une échelle qui permet au moins d'identifier les différentes catégories d'objets mesurés. Par exemple, un des objectifs de l'assurance-maladie au Québec a été formulé comme suit: "rendre les soins accessibles à l'ensemble de la population de façon qu'elle puisse en faire une consommation optimale" (C.E.S.B.S., 1967, p. 35). Lorsqu'on essaie de transposer un tel objectif sur le plan de l'évaluation, on se rend compte qu'un tel objectif est difficilement mesurable. Que faut-il entendre par "rendre accessibles"? Il ne peut pas s'agir uniquement d'une question de coût pour le consommateur, comme cela a été souvent mentionné, mais cela réfère tout autant à la disponibilité des soins et des médecins sur une base régionale ou locale. Tant et aussi longtemps que la définition d'un objectif peut inclure d'autres éléments, c'est jusqu'à un certain point un phénomène différent qu'on étudie;

2) Les objectifs formulés dans de tels rapports le sont souvent en termes trop absolus; ils ne tiennent pas compte du caractère relatif de la réalité du milieu et des situations de cette réalité du milieu et des situations de cette réalité. Il est clair que l'objectif énoncé précédemment de "rendre accessibles les soins à l'ensemble de la population" ne signifie pas la même chose selon qu'il s'applique au milieu urbain ou rural.

3) Il faut aussi se demander si les éléments mentionnés dans la définition réflètent tous les résultats attendus de l'action d'un programme et non pas seulement une partie d'entre eux. Ce serait une erreur de parler uniquement d'accessibilité des soins quand il faut en même égaliser l'accès à des soins de meilleure qualité.

En somme, la nature des énoncés qu'on retrouve dans les rapports de commission d'enquête ou dans les rapports de groupes d'experts, ne peuvent que très rarement être directement transposés pour fins d'évaluation. On ne peut pas entreprendre leur évaluation à l'aide des méthodes de recherche avant qu'ils soient reformulés en termes mesurables dans le cadre de programmes plus précis et avant qu'ils ne soient définis comme réalisables en tenant compte des contingences de la situation du milieu particulier où ils se font.

Une dernière façon proposée pour définir les objectifs est de séparer chacun d'entre eux en une série de composantes qui réflètent soit des parties de résultats visés, soit l'actualisation de conditions associées à l'existence de résultats, ou la réalisation de certaines tâches susceptibles de produire les résultats anticipés. Une telle approche prend pour acquis qu'il existe des méthodes bien établies pour ce faire. En fait, c'est assez rarement le cas parce que tout simplement l'état des connaissances, surtout du côté des sciences du comportement, n'est pas suffisamment avancé pour permettre la plupart du temps d'accepter comme étant valides et fiables les éléments incorporés dans la définition d'un objectif.

Les trois problèmes discutés précédement — celui de l'importance trop considérable accordée à l'étape de la formulation des objectifs, celui de la référence unique à l'efficacité et l'efficience dans le contenu des résultats mesurés et celui de l'absence d'une méthodologie acceptée pour la définition des objectifs — sont autant d'obstacles à affronter dans la pratique de l'évaluation. Ils font, jusqu'à un certain point, mieux comprendre que l'étape de la définition des objectifs du programme a beaucoup de limites, même si elle n'en continue pas moins, à notre avis, d'être importante dans la pratique de l'évaluation.

Interférence évaluation-programme

Dans la réalisation concrète de l'évaluation, il peut y avoir des interférences soit pour empêcher, soit pour changer, soit pour amplifier le déroulement de l'évaluation ou modifier la nature de l'information recueillie. Les principales sources d'interférence peuvent provenir du programme lui-même. Les conditions dans lesquelles se trouvent le programme nuisent à la réalisation de l'évaluation: on est en présence d'obstacles administratifs ou socio-politiques. Il peut aussi se produire des interférences à partir du processus d'évaluation lui-même dans sa jonction avec le programme: il s'agit à ce moment-là des interventions de l'évaluation sur le programme.

Obstacles administratifs et socio-politiques à l'évaluation

Pour réaliser une étude d'évaluation d'impact, il faut un programme clairement articulé (Rutman, 1977, 19,20).* Dans le cas contraire, il est difficile d'entreprendre une étude d'évaluation d'impact: il faut songer plutôt à des études d'évaluation formative destinées à rendre les programmes évaluables (Rutman, 1977, 69-71).**

Mais au delà de cet obstacle de base à la réalisation de l'évaluation, il y a d'autres obstacles administratifs qui ont aussi leur importance. Un premier de ces obstacles est lié à la structure organisationnelle: c'est l'absence d'un point d'entrée précis pour l'évaluation dans la structure organisationnelle. En conséquence, l'évaluation devient une tâche surajoutée aux tâches de fonctionnement normal du programme, d'où une première source d'interférence en ce qui concerne l'évaluation.

Un deuxième obstacle lié au premier, est celui de la place qu'occupe l'évaluateur à l'intérieur de la structure organisationnelle. Comme l'ont fait ressortir Attkisson et Hargreaves (1979) les évaluateurs peuvent occuper une place plus ou moins importante dans l'organisation, certains exerçant uniquement un rôle de compilation de données statistiques alors que d'autres sont plus près des points où les décisions se prennent. Si dans la structure administrative, l'évaluateur n'occupe pas une place qui lui permet de fournir de l'information utile au processus de prise de décision, il ne peut pratiquement pas être efficace (cf. Schulberg et Jerrell, 1979, p. 7-20).

Un troisième obstacle administratif tient à la multiplicité des sources de contrôles "déjà existantes dans l'organisation". Comme le souligne Dubois "beaucoup de gestionnaires de programmes ont le sentiment qu'il est d'autant moins nécessaire d'entreprendre des études systématiques d'évaluation, qu'il existe, de par le fonctionnement des démocraties, un grand nombre de mécanismes de 'sanction naturelle' depuis le vote individuel de citoyen jusqu'à la possibilité de recours auprès des tribunaux administratifs" (Dubois, 1979, p. 1607). Même si Dubois parle pour la France, la transposition à notre situation est possible. Les nom-

*Reproduit dans ce volume, chapitre 4
**Reproduit dans ce volume, chapitre 4

693

breuses formes de contrôles utilisées incitent les gestionnaires à résister aux différentes entreprises d'évaluation, et à les considérer comme plus ou moins utiles d'autant plus que le contrôle leur en échappe.

Aux obstacles précédents, il faut rappeler les six sources de conflits entre l'évaluateur et le gestionnaire identifiés par Carol M. Weiss (1972, pp. 98-104):

- **Les différences de personnalité**: l'évaluateur est défini plutôt comme un individu plus indépendant, plus détaché alors que le gestionnaire et le personnel du programme sont considérés comme plus impliqués plus engagés dans le programme. Comme nous le verrons plus loin ces différences de personnalité se sont sans doute quelque peu estompées au cours des années.

- **Les différences de rôles**: le gestionnaire croit en ce que le programme accomplit alors que l'évaluateur remet en cause le programme. Le gestionnaire a la responsabilité de conduire le programme pour qu'il atteigne ses objectifs; l'évaluateur pose des questions aux résultats atteints par rapport aux objectifs visés.

- **Confusion dans la définition des responsabilités**: il y a souvent ambiguités entre l'évaluateur, le gestionnaire et les autres membres du personnel concernant l'autorité et le contrôle sur le fonctionnement du programme.

- **Conflit d'objectifs, de valeurs, d'intérêts et de schème de référence**: le gestionnaire essaie d'utiliser au maximum les ressources dont il dispose pour fournir des services au plus grand nombre de clients possible. C'est pourquoi toute tentative de diversion vers des tâches autres que les services — comme l'évaluation — n'est pas toujours bien acceptée.

- **Caractéristiques institutionnelles**: une structure formelle, hiérarchique, centralisée rend la communication plus difficile pour l'évaluation — n'est pas toujours bien acceptée.

- **Pression des méthodes et techniques d'évaluation**: les méthodes et techniques de recherche utilisées en évaluation créent une pression sur le fonctionnement du programme et de l'organisation. Nous reviendrons plus spécifiquement sur cet aspect.

En plus des obstacles administratifs et souvent intimement reliés à ces derniers, les obstacles socio-politiques constituent probablement les forces d'interférence les plus fortes pour empêcher la réalisation de l'évaluation. L'un des principaux obstacles socio-politique trouve sa source dans la contradiction entre la position de pouvoir liée à l'exercice de responsabilité de programme qui appartient au gestionnaire et le danger

perçu de perte ou d'affaiblissement de son pouvoir qu'il peut envisager suite à l'évaluation, surtout si les résultats en sont défavorables. A cause de cela, le gestionnaire préférera procéder lui-même à l'évaluation et la contrôler le plus possible, surtout s'il doit s'en remettre à une instance extérieure. De cette façon, l'évaluation utilisant des méthodes de la recherche scientifique devient pour lui la plus dangereuse parce qu'elle est beaucoup plus difficile à contrôler.

Le deuxième obstacle tient aussi à la différence entre l'idéologie propre à la méthode scientifique et l'idéologie politique. La méthodologie de la recherche évaluative s'articule autour d'un mode d'approche rationnel: des objectifs sont déterminés, un programme constitué d'une série de moyens est mis en place pour atteindre ces objectifs et des données sont recueillies pour vérifier les résultats obtenus par le programme. Dans l'idéologie politique, le processus est différent: les objectifs, les moyens et les résultats du programme sont considérés ensemble et il s'agit de faciliter l'émergence d'un consensus au moment où la décision doit être prise. En ce sens, une décision prise sur un programme implique que ce dernier est considéré comme "bon". Toute évaluation ultérieure risque de remettre en cause l'appréciation déjà portée qui a été le fondement de la décision. Dans notre système techno-bureaucratique à caractère étatique, cet obstacle socio-politique se renforcit du fait que les grands responsables de programmes font partie de cette structure. Ils auront tendance à mettre en place leurs propres mécanismes d'évaluation et à les confier à des individus ou des groupes ayant la même formation et partageant la même façon de voir les choses qu'eux. Ils n'auront que peu d'intérêt à leur proposer un mode de référence extérieur, à supposer qu'ils en soient conscients. (Dubois, 1979, p. 1609-1610).

L'intervention de l'évaluation dans le fonctionnement du programme

En plus des obstacles administratifs et socio-politiques précédents, une autre source d'interférence susceptible d'influencer la réalisation pratique de l'évaluation est l'intervention que constitue l'évaluation dans le fonctionnement du programme. En effet, l'évaluation constitue toujours plus ou moins un élément ajouté à un programme qui, à des degrés divers, intervient avec le fonctionnement même de ce programme. Il est entendu que si l'évaluation procède à partir de données déjà existantes, l'intervention avec le fonctionnement du programme est beaucoup plus faible et les risques d'interférence se trouvent réduits au minimum. Si, au contraire, l'évaluation suppose un système particulier de cueillette de données, si elle se fait à différentes étapes du déroulement de l'action, le risque d'interférence est beaucoup plus présent. C'est aussi le cas lorsque l'évaluateur découvre que le programme qu'on lui demande d'évaluer ne satisfait pas complètement les différentes conditions d'évaluabilité. L'évaluateur peut choisir, soit de se retirer, soit d'entreprendre une démarche de recherche formative, ou d'accepter le

634

compromis de réaliser dans les meilleures conditions possibles l'évaluation demandée en tâchant d'améliorer au maximum l'évaluabilité. Sauf s'il décide de se retirer — et on pourrait même prétendre le contraire — il y aura presque toujours intervention de l'évaluation dans le programme.

En fait, l'évaluation peut devenir une forme d'intervention dans le programme de deux manières différentes: la première concerne l'influence de l'évaluateur sur la prise de décision concernant la définition du contenu du programme, la seconde est l'intervention du processus d'évaluation lui-même sur le déroulement du programme. Nous laissons de côté une troisième forme d'intervention: celle de l'utilisation des résultats. Lecomte et Rutman traitement de ce sujet plus loin dans ce volume.

En ce qui concerne d'abord l'influence de l'évaluateur sur le contenu du programme, dans plusieurs situations d'évaluation, l'évaluateur est étroitement associé à différentes étapes de la programmation. Il est impliqué, d'une manière ou de l'autre, dans les différents aspects de la définition du contenu de ce programme. Il y a d'ailleurs, un lien très étroit entre cette démarche de définition opérationnelle du programme et la définition de l'objet même d'une étude évaluative. Dans de telles circonstances, l'évaluateur peut à la limite influencer la définition du contenu du programme soit en privilégiant certains aspects du programme à cause de ses propres valeurs, soit en n'accordant de l'importance qu'à ce qui est mesurable dans le programme.

Il y a aussi eu des situations où on a utilisé une analyse faite par un évaluateur pour la définition des objectifs d'un programme. C'est le cas d'une étude faite par Edwards (1979, p. 13-54). L'étude portait sur sept plans de déségrégation raciale des écoles à Los Angeles soumis pour acceptation au conseil scolaire de cette ville. Ces plans avaient été soumis par des groupes externes. L'évaluateur qui avait été mandaté pour évaluer la faisabilité de ces sept plans a utilisé une grille multimensionnelle pour étudier les différents objectifs proposés dans les différents plans. Cette grille a permis d'établir comment les différents groupes d'intérêts concernés (parents, membres du conseil scolaire, professeurs, etc.) réagissaient concernant l'influence des objectifs proposés sur l'apprentissage et la qualité de l'éducation, leur acceptation par la communauté, leurs implications pour le personnel, etc. De cette manière, il a été possible d'établir quelle était la possibilité de réussir. Les résultats de cette étude n'ont pas permis d'accepter l'un ou l'autre des sept plans soumis comme plus acceptable, mais a tout simplement eu comme conséquence qu'un huitième plan a été élaboré et accepté par le conseil. En fait, ce plan, selon Edwards (1979, p. 48) a été un compromis incluant des parties fragmentaires des plans dont les objectifs avaient été considérés et analysés antérieurement. Dans un exemple comme le précédent, l'évaluateur a eu une influence déterminante sur la formulation d'un nouveau programme

à travers l'analyse qu'il a faite des seuls objectifs de différents plans de la déségrégation soumis au Conseil.

Il est évident que les situations où l'influence de l'évaluateur est aussi considérable ne sont pas très fréquentes. Mais que doit faire l'évaluateur lorsqu'il est en face d'un programme dont les objectifs ne sont pas clairement énoncés ou dont la réalisation ne se fait pas conformément aux objectifs officiels déclarés? Est-il approprié d'intervenir dans le programme pour ensuite être en mesure de réaliser l'évaluation qu'on a le mandat de faire? Les dilemmes de l'évaluateur sont considérables comme l'a souligné Perloff (1979, p. 109-112). S'il intervient, il risque d'influencer la validité externe, c'est-à-dire la possibilité d'étendre quelque peu ses généralisations à d'autres programmes du même genre, susceptibles d'être dans des conditions similaires. S'il n'intervient pas, il n'accomplira peut-être pas ce qu'on attend de lui, même si les résultats qu'il fournira constitueront une information rétroactive qui peut être utile sur la réalisation du programme. Quelle que soit la voie choisie, il y a des avantages et des inconvénients.

Le processus d'évaluation intervient aussi dans la réalisation d'un programme. L'évaluateur, en définissant son objet d'évaluation, en explicitant la nature de sa démarche, en appliquant ses instruments de cueillette des données, est au coeur même de ce qui s'accomplit dans un programme et ce serait faire l'autruche que de ne pas admettre l'influence qu'il a sur le programme lui-même dans bien des situations. De quelle manière cette influence de l'évaluation sur le programme doit-elle être traitée? Doit-on l'ignorer? Doit-on la réduire au maximum? Doit-on l'utiliser dans le programme? Il s'agit là de questions qui ont, somme toute, été l'objet de peu d'attention mais qui n'en deviennent pas moins cruciales pour la pratique de l'évaluation.

En ce qui concerne les aspects favorables à l'intervention de l'évaluateur, comme l'indique Perloff (1979, p. 113), l'influence de l'évaluateur sur le programme est tout à fait acceptable dans le processus de l'évaluation formative. Le rationnel à la base de cette forme d'intervention que constitue l'évaluation est que l'information est recueillie précisément pour aider à une meilleure articulation du fonctionnement du programme. Il y a aussi ceux qui prétendent que, d'un point de vue humanitaire, "ce serait un gaspillage d'énergie et une menace au bien-être d'une population-cible ou de l'autre si les évaluateurs devenaient antiseptiques ou 'sans coeur', au nom de 'l'objectivité' et de la 'pure science', et abandonneraient leur préoccupation du bien-être de cette population en déférence à leur acceptation aveugle du faux bien du "scientisme . . . " (Perloff, 1979, p. 113). En somme, pour ceux qui sont favorables à accepter et même à encourager l'influence de l'évaluation sur le programme, il y a d'abord la préoccupation des effets du programme sur la population visée par ce programme.

635

Il y a, par ailleurs, tout autant de raisons pour lesquelles il faut minimiser le plus possible l'influence du processus d'évaluation sur le programme. La menace à la validité externe de l'évaluation doit aussi être soulignée ici. Jusqu'à quel point serait-il possible de généraliser les résultats d'une évaluation, qu'ils soient positifs ou négatifs, si le fonctionnement opérationnel du programme a été influencé par des évaluateurs qui normalement n'auraient tout simplement pas été présents dans l'exécution du programme et qui n'y seront pas après l'évaluation, si cette dernière n'est pas continue. L'utilisation de l'expérimentation (projet-pilote, projet expérimental, etc.) doit être aussi vue dans cette perspective. Le fait d'évaluer une programme qui se réalise de manière contrôlée, expérimentale, dans des conditions exceptionnelles, est aussi une sorte d'intervention de l'évaluation qu'il faut considérer avec beaucoup de réserve, car on peut se demander si un tel programme pourra être réalisé, l'évaluateur ne déboucherait-il pas sur une erreur de type I très dispendieuse, s'il s'y implique même pour en favoriser une meilleure définition? Une dernière objection qu'il faut apporter à l'intervention de l'évaluateur dans le programme concerne la remise en question bien souvent de la compétence de l'évaluateur pour faire des recommandations et des suggestions qui ont une portée sur la réalisation d'un programme. A moins qu'il ne soit à la fine pointe des connaissances et de la pratique sur la gestion, le risque est fort pour qu'il fasse fausse route.

Obstacles méthodologiques

Les limites mentionnées précédemment ont définitivement une influence sur l'application des méthodes de la recherche scientifique à l'évaluation des programmes. La plupart des volumes traitant des méthodes d'évaluation en font état. Hudson a présenté les problèmes liés à la mesure dans un autre chapitre alors que Cook, Lomax et Mark ont aussi traité de méthodologie en parlant du modèle expérimental et quasi-expérimental d'évaluation. C'est pourquoi nous nous contenterons de faire en bref survol des limites de l'application des méthodes scientifiques à l'évaluation.

Pour l'essentiel, ces limites naissent des contradictions opposant les conditions de mise en oeuvre des programmes et les exigences de la méthode scientifique. Ainsi, évaluer l'impact d'un programme revient à établir un lien de causalité entre les effets attendus et le contenu du programme. Dans la pratique de l'évaluation, cela signifie que l'attribution des effets au programme doit être la moins ambiguë possible. Cela revient à dire que l'évaluateur doit choisir la méthode de recherche qui lui permette le mieux d'assurer que toutes les autres formes d'explication que celle du programme pourront être éliminées. Très rapidement, la

méthode qui se présente à l'évaluateur est la méthode expérimentale. De par sa structure, elle apparaît offrir les meilleures garanties. Parce qu'elle incorpore à l'étude un groupe de contrôle et qu'elle affecte selon les règles du hasard des individus au groupe bénéficiant du programme, elle permet de prendre en considération de nombreux facteurs de "contamination" possible des résultats.

Mais cette procédure se heurte à des obstacles inévitables dans la pratique, comme l'ont expliqué précédemment Cook, Lomax et Mark: résistance à la distribution au hasard, problèmes liés à la sélection au hasard et à la grandeur de l'échantillon, problèmes d'attrition en cours de réalisation de programmes, variation dans l'exécution du traitement, interaction entre groupe expérimental et groupe de contrôle, rivalité compensatoire, etc.

En plus de ces difficultés internes de la méthode expérimentale, deux autres difficultés doivent être soulignées. La première concerne la validité externe du modèle expérimental. On reproche aux évaluations faisant appel à la méthode expérimentale leur manque de généralisation quant aux conclusions obtenues à cause du nombre trop restreint de sujets, des difficultés de représentativité et de la nature trop "obtrusive" des procédures d'évaluation. Si la plupart de ces difficultés sont surmontables sur un plan méthodologique, elles se traduisent par une augmentation du coût de l'évaluation qui peut être hors de proportion.

Il y a une deuxième difficulté additionnelle. Peu de programmes publics se prêtent à l'utilisation de la méthode expérimentale, parce qu'il est difficile de constituer un groupe de contrôle. L'égalité des citoyens devant la loi fait que les programmes qui sont issus de législation et de réglementation doivent être accessibles également à l'ensemble des citoyens. Ainsi, dans l'étude déjà mentionnée sur la consultation en planification des naissances au Québec (Varin, 1982), il était impossible de priver d'une partie de la clientèle s'adressant au Centres locaux de services communautaires pour constituer un groupe de contrôle. On aurait alors contrevenu à la loi sur les services de santé et les services sociaux du Québec concernant l'égalité d'accès. De la même manière, l'affectation aléatoire est considérée comme immorale et injuste et on lui substitue souvent d'autres critères. Par exemple, qui serait d'accord avec le refus aléatoire des étudiants à un programme d'études universitaires? Qui serait d'accord avec la répartition aléatoire de bourses d'études doctorales? L'utilisation de la méthode expérimentale pure est souvent inapplicable à moins de mettre en place des solutions compensatoires qui souvent alourdissent sa réalisation en enlevant sa pureté au modèle.

C'est pourquoi, on a souvent recours à d'autres formes d'évaluation comme des méthodes quasi-expérimentales ou même les méthodes non expérimentales. Dès que de telles méthodes sont utilisées, on se heurte à des difficultés additionnelles qui rendent encore plus difficiles d'attribuer les résultats mesurés au programme lui-même. En effet, il devient sou-

vent difficile d'isoler les effets du programme d'autres effets comme ceux provenant de la maturation, des instruments utilisés, de la régression statistique, etc.

Toutefois, tout en étant conscient de ces difficultés méthodologiques, il n'en reste pas moins que, dans la pratique, l'évaluateur doit souvent avoir recours à des méthodes moins parfaites s'il veut demeurer réaliste dans ces entreprises d'évaluation en tenant compte de critères pratiques que nous avons identifiés au début de ce chapitre.

Conclusion

La réalisation de l'évaluation, tout en respectant au maximum des critères scientifiques de validité, de fiabilité et d'objectivité, n'est efficace que si elle réussit à alimenter les décisions à prendre dans l'action. C'est pour cela qu'elle doit satisfaire au critère de l'isomorphisme, c'est-à-dire être en correspondance avec les besoins de l'action par sa pertinence, son importance et sa portée, à celui de la crédibilité qui implique qu'on accordera foi aux conclusions de l'évaluation et à celui du réalisme qui signifie que l'évaluation ne prendra pas un envol qui la conduira à l'extérieur des conditions de réalité qui lui sont imposées. Dans cette perspective, un certain nombre de limites et d'obstacles qu'on rencontre dans le processus formel d'évaluation ont été présentés et discutés: limites reliées à la formulation des objectifs; l'interférence entre le processus d'évaluation sur le programme; et finalement, les obstacles liés au type de méthode de recherche.

Il est évident que ces obstacles ou limites ne sont pas présentés pour décourager l'évaluation mais plutôt pour la situer dans une perspective de réalité. Prendre conscience de limites de l'évaluation empêche d'en faire une panacée. Nous n'aurions pas atteint notre but si cela en décourageait la réalisation, car mieux vaut évaluer avec des méthodes ayant des limites, que poser des jugements de valeur et prendre des décisions au petit bonheur en se fiant plus à l'instinct qu'aux faits établis. Une telle prise de conscience permet d'attirer l'attention sur les développements encore à accomplir en évaluation et elle ouvre la porte à d'autres approches comme celle de l'utilisation des méthodes qualitatives dont il est question dans le chapitre précédent.

Références

ATTKISSON, C.C. et HARGREAVES (1979) "A conceptional Model for program evaluation in health organizations" in H.C. Schulberd et F. Baker (eds.) *Program Evaluation in the Health Fields* (Vol. II) New York

BEAUDOIN, André "L'évaluation des services sociaux: défi de la prochaine décennie" *La famille*, Montréal, Fédération des services sociaux à la famille, Vol. 7, no. 81-82 (déc. 1971) p. 12-43

BEAUDOIN, André (1979) "Introduction: l'évaluation de programmes et le futur de l'administration publique dans Colloque International sur l'administration publique: perspectives d'avenir "*Actes du Colloque/ Conference Proceedings*, Vol. VII, *L'évaluation des programmes gouvernementaux*, Québec, ENAP.

BEAUDOIN, André et GAGNON, André, *Evaluation du projet de développement rural dans l'O.R.D. du Centre-Est en Haute Volta* — Rapport soumis à Opération Haute-Volta Inc. (Québec, 1978)

BOUCHARD, Odette (1982) *La planification des naissances dans les C.L.S.C. du Québec — Vol. II — Les services de consultation, organisation, fonctionnement et populations des services* (Ste-Foy, Laboratoire de recherche, Ecole de service social, Université Laval)

COMMISSION D'ENQUÊTE SUR LA SANTÉ ET LE BIEN-ÊTRE SOCIAL (C.E.S.B.S.) —

- (1967) Vol. I — *L'assurance-maladie*
- (1970) Vol. IV — *La santé* — Tome II — Titre deuxième, Le régime de la santé
- (1971) Vol. III — *Le développement* — Tome I
- (1972) Vol. VI — *Les services sociaux* — Tome I (Québec, Gouvernement du Québec

DATTA, LOIS ELLIN and ROBERT PERLOFF, *Improving Evaluations* (Beverly Hills, Cal. Sage, 1979)

DUBOIS, Bernard (1979), Les obstacles à l'évaluation des programmes d'action sociale et culturelle dans Colloque international sur l'administration publique: perspectives d'avenir — *Actes du Colloque/ Conference Proceeding* — Vol. VII, *L'évaluation des programmes gouvernementaux*, pp. 1595-1618

EDWARDS, W. GUTTENTAG, M. et SNAPPER, K. (1975) "A decision — Theoretic approach to evaluation research" in E.L. Struening and M. Guttentag (eds.) *Handbook of Evaluation Research*, Vol. I, Beverly Hills: Sage Publications, p. 139-181

EDWARDS, Ward (1979) "Multiattribute Utility Measurement — Evaluating Desigragation Plans in a High Political Context" in Robert Perloff *Evaluator Interventions — Pros and Cons*, (Beverly Hills, Sige) p. 13-54

HITCH, Charles J. (1960) "On the Choice of Objectives" in *Systems Studies*' Santa Monica, Cal.: The Rand Corporation, I, 555

NADEAU, March-André (1981) *L'évaluation des programmes d'études. Théorie et pratique.* Québec, Les Presses de l'Université Laval

OSTROM, Elinor (1979) "Performance Measurement in Practice: A Methodology Gone Amuck dans Colloque international sur l'administration publique: Perspectives d'avenir — *Actes du Colloque/ Conference Proceeding*, Vol. VII — L'évaluation des programmes gouvernementaux (Québec, E.N.A.P.) p. 1619-1652

PERLOFF, Robert (1979) *Evaluator interventions Pros and Cons*, (Beverly Hills, Col., Sage)

RUTMAN, Leonard (edit.) (1977) *Evaluation Research Methods: A Basic Guide* (Beverly Hills, Sage)

SCHULBERG, Herbert C. et JERRELL J., BANNETE M., *The Evaluator and Management* (Beverly Hills, Sage publications, 1979)

632

STUFFLEBEAM, Daniel L. (1967) "The Use and Abuse of Evaluation in title III" *Theory into Practice*, Vol. 6 (June) pp. 126-133

STUFFLEBEAM, Daniel L. et als. (1980) *L'évaluation en éducation et la prise de décision* (Traduit par Jean Dumas/Victoriaville, Que., Les éditions NHP)

VARIN' Christiane (1982) *La planification des naissances dans les C.L.S.C. du Québec — Vol. I — Description des bénéficiaires et impact de la consultation* (Ste-Foy, Laboratoire de recherche, Ecole de service social, Université Laval)

WEISS, Carol H. (1972) *Evaluation Research* Englewood Cliffs, N.J.: Prentice Hall

WILDAVSKY, Aaron (1966) "The Political Economy of Efficiency: Cost-benefit Analysis, Systems Analysis, and Program Budgeting" *Public Administration Review*, Vol. 26 (December) 292-310

9

La problématique de l'utilisation de la recherche évaluative

Leonard Rutman et Roland Lecomte

Depuis le milieu des années 1960, et plus particulièrement dans la dernière décennie, on observe une controverse face à l'utilité et à l'impact de la recherche évaluative. Plusieurs explications sont offertes pour justifier ce manque d'utilisation: des faiblesses méthodologiques[1], une mauvaise présentation et dissémination des résultats[2], la naïveté politique des évaluateurs[3], etc. Ce manque d'attention à l'utilisation a même fait récemment le sujet d'un symposium lors de la conférence mondiale du "Réseau d'évaluation"[4]. Toutes sortes d'études sont présentement en cours afin de développer une problématique de l'utilisation[5]. Il ressort de ces principales études qu'il existe une très grande ambiguïté concernant la définition même de l'utilisation de l'évaluation et les facteurs influençant cette utilisation. D'une part, certains auteurs laissent entendre qu'il y a utilisation chaque fois qu'un individu utilise un élément quelconque de processus ou des résultats de l'évaluation, et à n'importe quelle fin; d'autre part, certains auteurs estiment qu'il y a utilisation lorsqu'une décision particulière est prise par un utilisateur par suite des données d'un rapport d'évaluation fondé uniquement sur les constatations de ce rapport[6]. Nous tenterons de développer une problématique qui tient compte de ces deux extrêmes, tout en mettant davantage l'accent sur les liens entre les différentes dimensions de l'utilisation et les types d'approches évaluatives.

L'intérêt accru à l'égard de l'évaluation des programmes tient aux diverses fins pour lesquelles les études de ce genre sont censées être utili-

A 11,684 #9718

Gauthier, Yvon
Jenicek, Milos
AUTEUR Desrosiers, Georges et al.
«Approches et méthodes en santé communautaire» : cours intensif et

A 11,684