

## NIVEAUX ET MÉTHODES DE L'ÉVALUATION EN SANTÉ PUBLIQUE

par Brigitte SANTOS-EGGIMANN\*

Les définitions de l'évaluation sont aussi nombreuses que celles de son objet, en l'occurrence la qualité des interventions sanitaires, et les attentes ou les intérêts des personnes ou structures concernées. Globalement, l'évaluation en santé publique peut être définie comme une collection d'activités multiples et diverses se réclamant de l'évaluation, parfois clairement consacrées à la santé publique (s'intéressant à l'impact des services de soins préventifs, palliatifs et curatifs sur la santé de la population), parfois relevant davantage de la médecine clinique, produites par des intérêts divergents et dont la réunion permet la progression de la connaissance sur l'efficacité tant des composants du système de santé que de leurs rapports mutuels.

### Les objectifs de l'évaluation en santé publique

L'évaluation en santé publique est une démarche dont le but essentiel est de porter sur la qualité des soins un jugement destiné à la prise de décision concernant leur organisation. *Par soins, il faut comprendre ici l'attention portée à l'état de santé se traduisant par le recours à un large spectre de moyens, allant de l'usage de technologies médicales à l'adoption de stratégies d'intervention; sont indifféremment comprises sous le terme de soins les interventions sanitaires de nature préventive, palliative ou curative produites par les services de santé.* L'objectif ultime de l'évaluation est de maximiser l'état de santé de la population par l'adoption des soins les plus performants.

Idéalement, l'évaluation permet l'intégration de la notion de performance lors des décisions d'allocation de ressources entre divers services de santé par les pouvoirs publics, dont la préoccupation est d'obtenir la plus grande amélioration de l'état de santé de la population pour un investissement financier donné. Son rôle est également de guider la prise de décision des producteurs de soins, en désignant non seulement les activités les plus efficaces, mais aussi les groupes de population les plus susceptibles de bénéficier de ces activités, ainsi que les circonstances les plus favorables à leur exercice.

\* Médecin-adjoint, chef de la Division Services de Santé, Institut universitaire de médecine sociale et préventive, rue du Bugnon 17, CH-1005 Lausanne

Les attentes vis-à-vis de l'évaluation varient ainsi selon l'interlocuteur, les financeurs s'intéressant avant tout à la démonstration de l'efficience (mesure de l'effet rapporté à l'investissement), les administrateurs de services de santé à celle de l'efficacité des structures qu'ils gèrent sur le terrain, les soignants à l'efficacité des soins qu'ils prodiguent et les chercheurs à leur apport au domaine de la qualité des soins. Elle est cependant encore rarement utilisée par les patients pour sélectionner les soins les mieux adaptés à leurs attentes.

### **La qualité des soins comme objet de l'évaluation en santé publique**

La qualité des soins est un concept d'excellence des soins médicaux qui peut être évalué selon plusieurs perspectives. Donabedian (1) fait la distinction entre les soins justes sur un plan technique («right care») et les bons soins («good care») du point de vue du patient; il s'agit là de deux conceptions qui peuvent s'opposer, car l'adéquation technique des soins n'est pas nécessairement liée à la satisfaction du patient, qui intègre dans sa perception de la qualité d'autres considérations, telles que le caractère désirable du traitement et la relation établie avec les soignants. Les soins techniquement justes peuvent encore être évalués sur le plan de leur rendement, lorsqu'une alternative existe entre des soins produisant en termes sanitaires un effet comparable mais à des coûts différents.

Ainsi, l'efficacité technique des soins, leur efficience, leur accessibilité, leur adéquation à la condition traitée ainsi que la satisfaction des patients sont des dimensions essentielles de la qualité des soins. L'importance relative que leur attribuent cliniciens, société et patients en tant qu'individus n'est cependant certainement pas équivalente. L'évaluation en santé publique porte sur l'ensemble de ces dimensions de la qualité des soins.

### **Activités d'évaluation**

Les activités d'évaluation sont souvent qualifiées de recherche. L'importance de la recherche dans les travaux d'évaluation est cependant très variable. Holzemer (2) met l'accent sur la nature des questions posées lors de l'évaluation, la recherche se préoccupant avant tout de la détection de relations causales, de façon aussi indépendante que possible des circonstances, et cherchant à conclure sur la vérité d'hypothèses sans viser en premier lieu à fonder une décision à prendre ponctuellement dans un contexte donné. A l'opposé, l'évaluation souhaitée par les gestionnaires doit permettre la prise de décision immédiate en déterminant la valeur d'une intervention spécifique menée dans des circonstances précises, et dépend moins strictement du respect des règles méthodologiques. Pratiquement, la démonstration utile aux gestionnaires prend cependant du poids lorsqu'elle inclut une dimension de recherche, et inversement, l'évidence scientifique d'une relation causale entre soins et santé est souvent constituée d'une accumulation d'évaluations convergentes dans leurs conclusions, qui souvent ne sont pas fondées sur des modèles expérimentaux.

La recherche, menée au moyen de méthodes scientifiques, occupe une place importante dans le domaine de l'évaluation. Elle intervient notamment dans la

procédure de vérification et de documentation de l'efficacité des soins («technology assessment»), permettant de détecter, qualifier et quantifier leurs effets. La mesure de l'impact économique des soins fait également partie de la recherche évaluative, mais il s'agit d'une préoccupation plus récente.

L'utilité de l'évaluation, au moins au niveau de l'efficacité, est peu contestée lorsqu'elle porte sur des *soins nouveaux*; ainsi, la mise sur le marché de produits médicamenteux passe par une procédure préalable d'évaluation systématique de leur efficacité sous forme d'essais cliniques. Une évaluation est également souvent demandée lors de l'installation ou l'extension de programmes de prévention. Elle est moins facilement reconnue pour d'autres innovations thérapeutiques (techniques opératoires, modes de prise en charge psychiatrique, par exemple), et souvent contestée lorsqu'elle porte sur des technologies anciennes largement utilisées, même si ces dernières n'ont jamais été formellement éprouvées dans leurs applications traditionnelles, pour de nouvelles indications, ou dans des conditions nouvelles d'application.

A un second niveau, la recherche évaluative porte sur les *pratiques professionnelles* («quality assessment») par l'étude du décalage entre la pratique réelle et les possibilités théoriques offertes par les technologies médicales (3). Cette discrédence est à l'origine de la définition de la qualité des soins donnée par Lohr et Brook (4) comme la minimisation de la différence entre l'efficacité théorique évaluée dans des conditions expérimentales ou optimales et l'efficacité mesurée sur le terrain.

Le troisième niveau de la recherche évaluative est l'*épidémiologie de la qualité*, comprenant l'étude descriptive des variations des pratiques professionnelles ainsi que l'analyse des facteurs associés. Ce troisième niveau fait largement appel aux statistiques sanitaires disponibles mais peut également faire l'objet d'évaluations très formelles lors d'interventions visant à modifier la pratique médicale.

L'*assurance de qualité* («quality assurance») est plus proche du domaine de la gestion des structures sanitaires, bien qu'elle s'appuie en partie sur des méthodes familières à la recherche évaluative, et se définit surtout comme une démarche formelle permettant de systématiquement vérifier, améliorer ou redresser en tous temps la qualité des soins fournis. Elle inclut un bilan des problèmes sanitaires, la sélection des soins à évaluer, la définition de priorités parmi les investigations retenues, l'évaluation elle-même, ainsi que les mécanismes d'information permettant de corriger les faiblesses observées.

Enfin, les récents développements de l'évaluation s'appuient sur l'intégration d'un modèle d'évaluation emprunté au monde industriel, décrits sous une variété de termes (*amélioration continue de la qualité, amélioration de la performance, «total quality management»*). Cette évaluation est souvent opposée à la pratique de l'assurance de qualité par son but de parvenir à une amélioration de la qualité des soins par une élévation du niveau général d'information plutôt que par la détection et la dénonciation de pratiques sous-optimales (5, 6). Elle se caractérise également par une approche plus globale de la notion de qualité, mettant l'accent sur la satisfaction du consommateur, et par une démarche d'évaluation plus systématique. Les deux approches sont cependant davantage en continuité qu'en opposition (7). Ainsi, l'adoption du modèle industriel a été précédée d'un élargissement de la définition de la qualité des soins, associé à

l'évolution d'une évaluation médicale essentiellement médico-technique vers l'intégration de la satisfaction du patient et la considération des dimensions économiques des soins.

### **Méthodes d'évaluation**

Pour Donabedian (8), l'évaluation de la qualité des soins peut porter sur l'adéquation des structures disponibles pour produire les soins (personnel, locaux, instrumentation), sur les activités de soins, ou directement sur les résultats obtenus. L'évaluation des seules structures comme mesure indirecte de la qualité des soins repose cependant sur l'hypothèse de leur influence sur les résultats; cette approche est limitée par notre connaissance insuffisante de la relation entre structures et performances. Lorsque l'évaluation est orientée sur les résultats obtenus, qui comprennent également l'impact des soins sur la fonction sociale et psychologique ainsi que les attitudes du patient (satisfaction), ses acquisitions en matière de connaissances relatives à la santé et ses modifications de comportement, la difficulté est d'exclure la responsabilité d'autres facteurs que les soins pouvant expliquer l'évolution de l'état de santé. Donabedian définit cependant avant tout l'évaluation comme un jugement concernant les activités de soins (dans leurs dimensions techniques et de relation interpersonnelle) fondé sur la mesure de leur contribution au résultat recherché.

#### *a) Efficacité des technologies médicales*

Le moyen privilégié de l'évaluation des technologies médicales est l'*essai clinique* visant la comparaison de l'état de santé de plusieurs groupes placés sous des régimes d'intervention différents. Lorsque les sujets inclus dans l'essai sont attribués au hasard aux différents groupes de traitement ou de prise en charge (allocation aléatoire ou «randomization» des sujets), les investigateurs s'assurent des meilleures chances (à défaut de garantie absolue) d'observer des groupes en tous points comparables, exception faite de l'intervention appliquée. La validité de l'essai clinique est encore améliorée lorsque les sujets, les intervenants, voire les chercheurs analysant les résultats ignorent quelle intervention est appliquée dans chaque cas («single, double, triple blind clinical trial»). Lorsqu'il est mené en recourant aux techniques d'allocation aléatoire des sujets et d'ignorance de l'intervention appliquée, l'essai clinique offre les meilleures garanties de validité et l'évolution différentielle de l'état de santé observée entre les groupes résulte le plus vraisemblablement des soins reçus et non de la sélection des cas dont le pronostic est le plus favorable vers l'une ou l'autre des interventions disponibles ou de la préférence des sujets pour certains soins. La participation de plusieurs centres thérapeutiques à de tels essais permet de vérifier l'universalité des résultats observés.

A l'évidence, l'exploration de la relation causale entre soins et état de santé ne peut pas faire appel aux méthodes expérimentales pour toutes les technologies médicales. Ainsi, l'allocation aléatoire des patients peut être rejetée pour des considérations éthiques lorsque un traitement est largement préjugé favo-

nable par les professionnels ou par les patients; en outre, l'ignorance du traitement appliqué ne peut certainement pas être supposée lorsque l'évaluation porte sur des modalités différentes de prise en charge thérapeutique ou sur des interventions chirurgicales. Ces considérations ont mené au développement de *méthodes quasi-expérimentales* (9), jouant essentiellement sur la séquence des interventions et des observations pour tenter de maîtriser les biais résultant de la comparaison de groupes de sujets qui ne sont pas équivalents.

#### b) *Evaluation économique*

L'évaluation économique porte sur le rendement des services de santé et vise la sélection des projets sanitaires les plus performants en termes de rapport entre les résultats qu'ils produisent et l'investissement financier qu'ils nécessitent. Sa pratique s'est étendue au cours des années 1960, notamment avec le développement des grands programmes sociaux aux Etats-Unis.

Les méthodes d'évaluation économique ont pour dénominateur commun l'estimation des coûts (directs, indirects, d'opportunité) exprimés en unité monétaire, et diffèrent essentiellement par leur mesure des résultats. L'*analyse coût-avantage* («cost-benefit») convertit les résultats obtenus en termes monétaires, ce qui permet d'effectuer des choix entre des programmes visant des buts différents. L'expression monétaire permet également d'actualiser les résultats, ou de les ajuster à leur valeur présente, lorsqu'ils seront atteints après plusieurs années de latence. L'*analyse coût-efficacité* («cost-effectiveness») en revanche ne permet de comparer que des programmes dont le but est semblable, mais quantifie les résultats en termes de santé (par exemple, nombre de décès infantiles évités ou années de vie gagnées pour une unité monétaire investie). Enfin, l'*analyse coût-utilité* («cost-utility») permet d'intégrer à la mesure des résultats obtenus les jugements de valeur portés sur des états de santé spécifiés.

#### c) *Evaluation de la pratique médicale*

L'évaluation de la pratique médicale permet de vérifier la conservation de l'efficacité des soins sur le terrain, leur accessibilité et leur adaptation aux besoins de la population.

L'application la plus évidente de l'épidémiologie à l'évaluation en santé publique est l'étude des variations temporelles et géographiques des consommations de soins à un niveau de population. Des variations considérables des taux d'hospitalisation (10), des durées de séjour hospitalier (11) ou de l'utilisation d'interventions chirurgicales ont ainsi été mises en évidence, qui ne semblent pas expliquables par des différences de morbidité dans les populations comparées (12-15).

Parmi les hypothèses explicatives envisagées dans un contexte d'évaluation figurent un effet de l'offre en services de soins (16) et l'utilisation inappropriée des services de soins (17). Les variations de la pratique médicale font ainsi l'objet d'études confrontant la conformité des soins aux recommandations

explicites pour la pratique médicale («clinical guidelines») issues de revues de la littérature et de la consultation d'experts (18, 19).

L'évaluation des activités de soins au niveau des hôpitaux ou d'autres institutions fait souvent appel à la méthode de l'*audit médical* (20) ou revue d'une sélection de cas, soit rétrospectivement sur dossier, soit prospectivement. De telles revues peuvent être menées par la confrontation entre les soins effectués et des critères explicites, pré-établis, tels que ceux qui sont publiés sous forme de recommandations pour la pratique médicale. L'avantage d'une telle méthode est sa reproductibilité; elle a cependant l'inconvénient de ne pas permettre d'envisager les circonstances particulières qui justifient parfois une option thérapeutique. A l'opposé, les cas peuvent être revus sur la base de critères implicites par un groupe d'experts. La concordance entre experts est cependant généralement faible. Pratiquement, les deux approches sont parfois conjuguées, par la soumission à un groupe d'experts des cas de pratique sous-optimale détectés au moyen de critères explicites. Une difficulté particulière de l'*audit médical* est la sélection des dossiers devant faire l'objet d'une revue. Cette sélection peut être réalisée par le choix d'une pathologie, d'un acte thérapeutique, ou de résultats défavorables tels que le décès ou le transfert vers une unité de soins intensifs.

La *méthodologie des traceurs* représente une alternative à ce mode de sélection des dossiers. Les traceurs sont des affections sélectionnées pour leur fréquence et parce que leurs traitements respectifs permettent l'étude de divers aspects des activités de soins produites par une institution ou un système de soins (21). L'évaluation se fonde sur l'hypothèse que la qualité des soins mesurée pour ces affections reflète celle des soins offerts en général par l'institution.

L'*évaluation des résultats*, préconisée depuis le début du siècle par Codman (22), a mis longtemps à s'imposer. Outre l'intégration de cette dimension aux méthodes de l'*audit* ou des traceurs, la comparaison des taux de mortalité hospitalière après contrôle des facteurs de risque présentés par les patients est une illustration de cette démarche (23). Pratiquement, l'ensemble des méthodes épidémiologiques peuvent être adoptées à des fins d'évaluation, en particulier lorsque l'accent est mis sur le résultat. Ainsi, des groupes de patients survivants et décédés peuvent être comparés en termes de soins lors d'études cas-témoins, notamment pour investiguer des affections rarement mortelles; le stade de la maladie lorsque le diagnostic est posé est également un résultat d'intérêt si l'évaluation porte sur des soins préventifs.

## Conclusion

Les méthodes d'évaluation de la pratique médicale reflètent avant tout la diversité des objets de l'évaluation et des points de vue adoptés. Elles font appel à plusieurs disciplines telles que l'épidémiologie, les sciences sociales ou l'économie. L'évaluation en santé publique n'est que le reflet de l'ensemble des efforts investis permettant d'orienter le système de santé vers ce que la société juge souhaitable.

### Références bibliographiques

1. DONABEDIAN A.: The definition of quality and approaches to its assessment. Ann Arbor, Health Administration Press, 1980.
2. HOLZEMER W.L.: Research and evaluation: an overview. *Quality Review Bulletin*, 5, 31-34, 1980.
3. AMOURETTI M.; BERAUD C.: L'évaluation des activités de soins. *Sociologie Santé*, 3, 56-67, 1990.
4. LOHR K., BROOK R.: Quality assurance in medicine: experience in the private sector. Santa Monica, Rand Corporation, 1984.
5. ROPER W.L., WINKENWERDER W., HACKBARTH G.M., KRAKAUER H.: Effectiveness in health care. An initiative to evaluate and improve medical practice. *N. Eng. J. Med.*, 319, 1197-1202, 1988.
6. HEALTH SERVICES RESEARCH GROUP: Quality of care: 1. What is quality and how can it be measured. *Can. Med. Assoc. J.*, 146, 2153-2158, 1992.
7. DONABEDIAN A.: Continuity and change in the quest for quality. *Clinical performance and quality health care*, 1, 9-16, 1993.
8. DONABEDIAN A.: The quality of care: How can it be assessed? *JAMA*, 260, 1743-1748, 1988.
9. CAMPBELL D.T., STANLEY J.C.: Experimental and quasi-experimental designs for research. Boston, Houghton Mifflin Company, 1963.
10. PERRIN J.M., HOMER C.J., BERWICK D.M., WOOLF A.D., FREEMAN J.L., WENNBURG J.E.: Variations in rates of hospitalization of children in three urban communities. *N. Eng. J. Med.*, 320, 1183-1187, 1989.
11. MAYER W., CLINTON J., NEWHALL D.: A first report of the Department of Defense External Civilian Peer Review of Medical Care. *JAMA*, 11, 2690-2693, 1988.
12. WENNBURG J., GITTELSON A.M.: Small area variations in health care delivery. *Science*, 182, 1102-1108, 1973.
13. ROOS N.P., ROOS L.L.: Surgical rates variations: Do they reflect the health or socio-economic characteristics of the population? *Medical Care*, XX, 945-958, 1982.
14. BARNES B.A., O'BRIEN E., COMSTOCK C., D'ARPA D.G., DONAHUE C.L.: Report on variation in rates of utilization of surgical services in the Commonwealth of Massachusetts. *JAMA*, 254, 371-375, 1985.
15. LU-YAO G.L., McLERRAN D., WASSON J., WENNBURG J.E.: An assessment of radical prostatectomy. Time trends, geographic variation, and outcomes. *JAMA*, 269, 2633-2636, 1993.
16. SANTOS-EGGIMANN B., PACCAUD F., GUTZWILLER F.: Coronary arteriography rates in Switzerland - How do they vary? *Soc. Sci. Med.* 28, 115-120, 1989.
17. CHASSIN M., KOSEKOFF J., WINSLOW C.M., KAHN K.L., MERRICK M.J., KEESEY J., et al.: Does inappropriate use explain variations in the use of health care services? *JAMA*, 258, 2533-2537, 1987.
18. WINSLOW C.M., KOSEKOFF J.B., CHASSIN M., KANOUSE D.E., BROOK R.H.: The appropriateness of performing coronary artery bypass surgery. *JAMA*, 260, 505-509, 1988.
19. BROOK R.H., EDWARD PARK R., CHASSIN M.R., SOLOMON D.H., KEESEY J., KOSEKOFF J.: Predicting the appropriate use of carotid endarterectomy, upper gastrointestinal endoscopy, and coronary angiography. *N. Eng. J. Med.*, 323, 1173-1177, 1990.

20. LEMBCKE P.A.: Evolution of the medical audit. *JAMA*, 8, 111-118, 1967.
21. KESSNER D.M., KALK C.E., SINGER J.: Assessing health quality - the case for tracers. *New. Eng. J. Med.*, 288, 189-194, 1973.
22. NEUHAUSER D.: Ernest Amory Codman, M.D., and end results of medical care. *Int J of Technology Assessment in Health Care*, 6, 307-325, 1990.
23. DUBOIS R.W., BROOK R.H., ROGERS W.H.: Adjusted hospital death rates: a potential screen for quality of medical care. *Am J Public Health*, 77, 1162-1167, 1987.