



L'évaluation : pourquoi et comment ?

L'évaluation de l'action publique est une pratique qui s'est développée ces dernières années avec la volonté de quantifier les performances. Dans le domaine de la santé publique, l'évaluation est multiforme et a nécessité l'élaboration de stratégies et d'outils spécifiques.

L'évaluation en santé publique en France : diversification et consolidation

Chantal Cases

Directrice de l'Ined,

Isabelle Grémy

Directrice de

l'Observatoire

régional de la santé

d'Île-de-France,

Bernard Perret

Ingénieur général des

ponts, des eaux et

des forêts, Conseil

général des ponts et

chaussées

Membres du Haut

Conseil de la santé

publique

L'évaluation de l'action publique connaît depuis quelques années un développement important et multiforme dans les administrations publiques françaises. La santé publique ne fait pas exception à ce mouvement.

Une brève histoire de l'évaluation des actions publiques

L'évaluation a commencé à s'institutionnaliser aux États-Unis dès avant la Seconde Guerre mondiale avec la création, en 1921, du General Accounting Office, qui associe contrôle et analyse de l'efficacité. Elle s'y renforce et diversifie ses méthodes après la Seconde Guerre mondiale, avec le développement des politiques de *welfare*. Dans cette phase initiale, l'évaluation est conçue avant tout comme une mesure empirique des impacts de programmes simples, principalement dans les domaines de l'éducation et des politiques sociales. En Europe, son développement est plus tardif et date des années soixante. La Grande-Bretagne, les Pays-Bas et les pays scandinaves sont les premiers à suivre l'exemple américain. En France, la première tentative

visant à systématiser et institutionnaliser l'évaluation est la rationalisation des choix budgétaires (1970-1984), centrée sur l'aide à la programmation budgétaire et l'évaluation *ex ante*. C'est dans ce contexte qu'a été mise en place, en 1970, une première évaluation en matière de santé publique, celle du programme périnatalité, par le ministère de la Santé.

Au début des années quatre-vingt-dix, le gouvernement dirigé par Michel Rocard impulse le développement de l'évaluation *ex post* des politiques publiques. Sous l'influence du rapport Viveret¹ (1989), le développement de l'évaluation des politiques publiques est alors conçu comme un outil de modernisation du débat démocratique. Pour remplir cette fonction, l'évaluation doit respecter quatre principes de base : indépendance, rigueur, pluralisme et transparence². La Société française d'éva-

1. *L'évaluation des politiques publiques. Rapport au premier ministre.* Paris : La Documentation française, 1989.

2. Viveret P. *L'évaluation des politiques et des actions publiques. Propositions en vue de l'évaluation du revenu minimum d'insertion. Rapport au Premier ministre.* Paris : La Documentation française, 1989, coll. des rapports officiels : 193 p.

luation, créée en 1999, élargira ces principes à sept : pluralité, distanciation, compétence, respect des personnes, transparence, opportunité et responsabilité³. Un ensemble d'initiatives sont prises par le gouvernement Rocard pour inscrire l'évaluation dans les processus de décision publique. Un Conseil scientifique de l'évaluation est créé (1990) ; il deviendra ultérieurement le Conseil national de l'évaluation (1998), avant d'être supprimé en 2002. En 1999, une mission d'évaluation et de contrôle, appuyée sur l'expertise de la Cour des comptes, est créée au sein de l'Assemblée nationale pour veiller à l'efficacité de la dépense publique. Cette tentative d'institutionnalisation n'a pas eu l'impact escompté sur le processus d'élaboration des politiques, mais elle a contribué à la diffusion d'une culture de l'évaluation dans l'administration française. Parallèlement, à partir de 1993, l'évaluation a connu un fort développement au niveau régional, dans le cadre des contrats de plan entre l'État et les régions et, surtout, de la politique européenne d'aide aux régions défavorisées par les fonds structurels.

C'est à cette période qu'un dispositif d'évaluation est explicitement prévu pour la première fois en matière de santé publique, dans la loi Evin de lutte contre le tabagisme et l'alcoolisme (1991). En 1991 est créé le Haut Comité de santé publique, chargé d'orienter la politique de santé à travers, notamment, des analyses critiques du système de santé. La production du HCSP contribuera ainsi, au fil des années, à diffuser la culture et les méthodes de santé publique en France. D'autres institutions y contribuent : ainsi, en 1995, l'Andem publie un premier recueil de recommandations sur l'évaluation d'une action de santé publique⁴, dont la revue *adsp* rend d'ailleurs compte l'année suivante⁵.

Évaluation et culture du résultat

Au cours de la période récente, les initiatives des pouvoirs publics ont été recentrées sur l'amélioration du compte rendu budgétaire, les indicateurs de performance et l'utilisation de l'évaluation pour réduire les dépenses publiques, dans une optique souvent assez éloignée des principes fondateurs énoncés en 1990. La loi d'orientation sur les lois de finances (LOLF, 2001) met en avant une nouvelle culture d'évaluation de l'efficacité de la dépense à travers des procédures de *reporting* : aux grands objectifs de dépense publique sont associés des indicateurs chiffrés et des cibles.

Lancée fin 2007, la « révision générale des politiques publiques » (RGPP) a été présentée comme une version française de la « revue des programmes » entreprise par le gouvernement canadien de 1995 à 1998, période au cours de laquelle les crédits des ministères ont été amputés de 5 % à 40 %. Mais cette comparaison est

trompeuse : la réforme canadienne – on aurait aussi pu s'inspirer des exemples suédois et néerlandais – reposait sur un large consensus social et une étroite collaboration entre le ministère des Finances et les ministères dépensiers. Le cadre de la RGPP s'inspire formellement de la démarche canadienne. L'ensemble des activités et des missions des ministères a été passé au crible sur la base d'un questionnement rigoureux : « Que faisons-nous ? », « Quels sont les besoins et les attentes collectives ? », « Faut-il continuer à faire de la sorte ? », « Qui doit le faire ? », « Qui doit payer ? », « Comment faire mieux et moins cher ? », « Quel scénario de transformation ? ». En dépit des intentions affichées, la RGPP n'a pas donné lieu à de véritables évaluations de politiques publiques et ne s'est guère appuyée sur les évaluations existantes. Compte tenu des délais et des objectifs qui leur étaient impartis, les équipes d'audit se sont rapidement focalisées sur les enjeux de rationalisation de l'organisation des services administratifs, sans se donner les moyens de questionner les objectifs et les priorités de l'action publique.

Ces développements s'inscrivent dans une volonté déjà ancienne d'introduire dans l'administration française les principes du *New Public Management* britannique : « agencisation », contractualisation entre l'État et les établissements publics, quantification des objectifs, indicateurs de performance, développement des obligations de compte rendu et de mesure du résultat à tous les niveaux, tentative de rémunération au mérite, etc. Dans le monde de la santé, les contrats d'objectifs et de moyens des agences de santé, ceux des établissements de soins participent de cette évolution générale. La santé publique a suivi ce mouvement de développement du *reporting*, notamment à travers la loi de santé publique de 2004 qui a entériné ce principe avec ses 100 objectifs annexés et ses grands programmes et plans de santé. La loi a fait davantage : elle s'est également dotée d'un dispositif d'évaluation périodique au niveau national, en constituant dans ce but et à des fins plus générales d'expertise un Haut Conseil de la santé publique qui a pris le relais du Haut Comité en élargissant ses missions. Elle a également prévu le suivi et l'évaluation des programmes régionaux dont elle a organisé la préparation et que les groupements régionaux de santé publique se sont pour la plupart appropriés même si cela constituait pour eux une démarche novatrice.

Au plan institutionnel, l'un des changements les plus significatifs est très récent, et élargit solennellement, et pour l'ensemble des politiques publiques, le *reporting* à l'évaluation : depuis le 1^{er} janvier 2009, l'évaluation des politiques publiques fait partie des missions fixées au Parlement par la Constitution⁶. Celui-ci sera désormais assisté par la Cour des comptes et va se doter d'une mission d'évaluation des politiques publiques.

3. <http://www.sfe.asso.fr/fr/charte-sfe.html>

4. Andem, « Évaluation d'une action de santé publique : recommandations », octobre 1995.

5. Blum-Boisgard C., Gaillot-Mangin J., Chabaud F., Matillon Y., « Évaluation en santé publique », *adsp* n° 17, décembre 1996, p. XVIII-XXII.

6. Par la loi constitutionnelle du 23 juillet 2008 de modernisation des institutions de la V^e République.



Une grande diversité de pratiques et d'institutions

Ce bref (et incomplet) historique fait apparaître une multiplicité de pratiques, d'acteurs et de contours de l'évaluation, assortis de méthodes variées. De plus, l'évaluation des actions publiques s'inscrit dans un ensemble toujours plus diversifié de pratiques de management, de compte rendu et de contrôle (certification, évaluation des protocoles de soins et de la performance des établissements...) entre lesquelles les frontières sont parfois floues. Les articles réunis dans ce dossier donnent un aperçu de cette diversité (cf. l'article « Acteurs et objets de l'évaluation en santé », p. 27). Sans prétendre à l'exhaustivité, la première partie présente les grands acteurs de l'évaluation en santé publique aujourd'hui, ainsi que certaines méthodes dont le développement permet d'enrichir les analyses. Dans une seconde partie, le propos sera illustré d'exemples récents d'évaluation de politiques nationales et régionales de santé, de leurs résultats, mais aussi des difficultés auxquelles elles se sont heurtées, notamment en matière de systèmes d'information et de déficience du suivi.

Sans surprise, on retrouve dans le champ de la santé les grandes oppositions qui structurent plus largement l'ensemble du domaine de l'évaluation de l'action publique. Pour simplifier, ces oppositions peuvent se lire selon deux axes principaux :

- un axe méthodologique, qui permet d'opposer des évaluations centrées sur la mesure d'un impact bien délimité (à l'image de l'évaluation des effets d'un médicament) et, d'autre part, des évaluations portant sur des actions plus complexes, accordant nécessairement davantage de place à la compréhension des processus sociaux, aux aspects descriptifs et aux méthodes qualitatives ;
- un axe relatif aux usages de l'évaluation, qui voit s'opposer des évaluations centrées sur le compte rendu, à visée de contrôle, de vérification et d'optimisation des moyens, et des évaluations dont les objectifs sont plus larges et variés : production de connaissances scientifiques, aide au pilotage des actions (grâce notamment à une meilleure appréhension des processus de mise en œuvre), interrogation sur la pertinence des objectifs et la cohérence des politiques.

Cependant, au-delà des inévitables controverses propres à tout domaine de connaissance (à l'exemple de celle portant, depuis des décennies, sur la place de l'expérimentation avec groupe de contrôle), il s'agit moins d'opposer différentes approches que de montrer leurs interpénétrations et leurs complémentarités. L'exemple du Haut Conseil de la santé publique illustre parfaitement ce point, puisque ses missions relèvent à la fois d'une approche centrée sur le compte rendu et la mesure (évaluation sur la base d'indicateurs de la réalisation des objectifs nationaux de santé publique annexée à la loi quinquennale débouchant sur la proposition d'objectifs quantifiés pour la prochaine loi) et d'une approche plus compréhensive, pluridisciplinaire et « pluraliste » (évaluation des plans de santé publique tels le Plan cancer et le Plan maladies rares).

En tant qu'elle participe d'une « culture du résultat », l'évaluation s'inscrit dans une démarche de rationalisation de l'action publique. Cependant, la rationalité de l'action (qu'elle soit individuelle ou collective) suppose toujours la donnée d'un cadre normatif et cognitif stabilisé (des objectifs clairement fixés, des valeurs et des représentations de la réalité suffisamment stabilisées et partagées). Or, comme l'ont bien montré plusieurs sociologues contemporains (Ulrich Beck, Antony Giddens...), la société « post-moderne » se caractérise par l'instabilité croissante des cadres de l'action humaine. Sous l'effet de différents facteurs (pluralisme culturel, complexité de la réalité sociale, importance croissante des risques et des incertitudes...), l'individu contemporain est contraint de réévaluer en permanence les paramètres et les buts de ses actions. Cette conceptualisation, d'abord conçue pour comprendre l'action des individus, s'applique parfaitement aux actions et politiques publiques. De fait, l'évaluation participe presque toujours d'un processus réflexif de réinterprétation et de réappropriation du sens de l'action. En effet, même lorsqu'elle obéit à des impératifs légitimes de compte rendu et de rationalisation, l'évaluation évite rarement de s'interroger sur le bien-fondé et le sens des objectifs politiques. Ce faisant, à travers notamment la discussion collective d'un questionnement évaluatif, elle permet à un cercle élargi de « parties prenantes » (experts, professionnels de santé, associations d'usagers et de malades, élus locaux...) de participer à la reformulation des référentiels de la politique de santé. Sans renoncer aux exigences du *reporting*, l'évaluation constitue donc l'une des expressions les plus caractéristiques du développement de la réflexivité sociale.

Un défi pour l'avenir : mettre l'évaluation au service d'une utilisation plus rationnelle des ressources publiques

L'un des objectifs de ce dossier est de réfléchir à ce que devrait être un « bon usage » de l'évaluation en santé publique, et aux conditions de son développement. L'une des conclusions qui s'en dégage est que l'évaluation doit répondre à une pluralité de besoins également légitimes. Comme on vient de le voir, elle ne s'adresse pas uniquement aux besoins des décideurs centraux : c'est aussi un outil de connaissance et d'apprentissage collectif appropriable par l'ensemble des acteurs du système de santé. Reconnaître cela n'enlève cependant rien à l'exigence de fonder la décision politique sur une base plus rationnelle. De ce point de vue, le faible usage des méthodes de l'évaluation économique n'est pas sans poser problème. L'accroissement prévisible de la contrainte financière, l'accroissement des exigences de qualité et de sécurité, le progrès technique, le vieillissement de la population et, plus généralement, l'explosion des demandes de toutes natures adressées au système de santé sont autant de facteurs qui rendent plus que jamais nécessaires l'instauration de processus décisionnels à la fois rationnels

et transparents. L'évaluation pourrait y contribuer, en s'appuyant notamment sur les outils décrits dans l'article de Marie-Odile Carrère, page 57. Encore faudrait-il vaincre toutes les méfiances et incompréhensions qui s'attachent à la mise en œuvre des méthodes d'aide au choix rationnel. On peut penser que le développement d'une culture de l'évaluation donnant droit de cité aux différents aspects de la démarche évaluative créera progressivement les conditions d'une appropriation

des outils économiques. Il deviendra alors possible de déployer toutes les potentialités de l'évaluation, y compris celle de fonder sur de meilleures bases l'affectation de ressources publiques limitées aux différents objectifs de santé. Le respect des exigences indissociables de rigueur et de transparence, de pluralisme et de pluridisciplinarité permettra ainsi à l'évaluation de remplir son objectif ultime : au-delà du jugement sur le passé, orienter l'évolution des politiques. **F**

L'évaluation de la performance au service d'un pilotage décentralisé du système de santé

Les outils nécessaires pour piloter les transformations requises des systèmes de santé, en France, au Québec, comme dans tous les pays développés, font l'objet de discussions et d'efforts considérables. Partout, les effets conjugués du développement des connaissances et des techniques, du vieillissement de la population et de la dégradation de l'environnement créent de nouvelles demandes pour les systèmes de santé. Les pressions qui s'exercent sur les finances publiques obligent également les États à contrôler avec vigueur les dépenses de santé. La crise qui résulte de la rencontre de ces deux mouvements oblige à proposer des réformes majeures des systèmes de santé. Ces réformes visent toutes à améliorer l'intégration des soins. Elles proposent de nouvelles instances se situant entre, d'une part, le niveau micro où les activités cliniques des professionnels se déroulent sans véritable coordination dans des organisations peu articulées les unes avec les autres et, d'autre part, le niveau macro où se prennent les décisions stratégiques (l'État, le ministère de la Santé). Ces instances visent, sur un territoire donné, à coordonner les ressources pour permettre à toute la population d'avoir accès de façon efficiente à des services de qualité ; exemples : au Québec, la loi 25 qui a créé les centres de santé et de services sociaux et modifié les responsabilités des agences régionales de santé et de services sociaux. En France, la loi « Hôpital, patients, santé, territoires », qui crée les agences régionales de santé (ARS) « pour renforcer résolument le pilotage territorial de notre système de santé » et exercer, en reprenant les termes du Rapport Ritter « un pilotage territorial unifié et responsabilisé tant en matière de politique de santé que de maîtrise des dépenses ».

L'implantation de ces instances régionales consacre l'existence d'un pilotage décentralisé du système de santé. Elle implique une redistribution importante des responsabilités entre les différents niveaux du système de santé et elle oblige à concevoir de nouveaux outils pour permettre aux décideurs d'exercer leurs responsa-

bilités et, en particulier, de nouveaux systèmes d'évaluation de la performance des instances régionales, des établissements et des projets cliniques pour pouvoir donner un sens aux réformes en cours, pour orienter les changements et pour gérer de façon responsable.

À partir des études que nous menons au Québec, nous montrons qu'il est utile de concevoir les systèmes régionaux de santé et les organisations qui les composent comme des systèmes organisés d'action complexes qui s'emboîtent et s'articulent les uns avec les autres, et qu'il est possible d'évaluer de façon globale et intégrée la performance de ces systèmes complexes de façon à produire des informations utiles pour les différents preneurs de décisions, aux différents niveaux où ils exercent leurs fonctions. L'idée maîtresse qui a guidé nos travaux réside dans le fait que l'utilité d'une évaluation est d'autant plus grande qu'il existe une forte adéquation entre la complexité de l'intervention à évaluer et le dispositif d'évaluation choisi.

L'article est divisé en quatre parties : 1) l'objet de l'évaluation ; 2) la conceptualisation de la performance ; 3) le but de l'évaluation et le public visé ; et 4) la faisabilité opérationnelle de la démarche. Nous concluons par quelques considérations sur l'utilité de la mesure de la performance des systèmes et des organisations de santé dans l'implantation des réformes.

L'objet de l'évaluation

Représentation systémique du système et des organisations de soins

Le premier choix à faire dans l'élaboration des systèmes d'évaluation de la performance concerne l'objet de l'évaluation. S'agissant de systèmes régionaux de santé, il faut commencer par savoir, et préciser, si on veut mesurer la performance du système de santé au sens large ou la performance du système de soins.

Les systèmes de santé concernent l'état de santé et le bien-être des individus et des collectivités, tandis que l'objectif du système de soins est de « réduire la

**André-Pierre
Contandriopoulos
François
Champagne
Claude Sicotte**

Professeurs titulaires,
Département
d'administration de
la santé, Faculté de
médecine, université
de Montréal et
chercheurs à l'Institut
de recherche en
santé publique
de l'université de
Montréal (Irspum)

**Geneviève
Sainte-Marie**

Coordonatrice de
projet de recherche à
l'Irspum



prévalence, la durée, l'intensité ou les conséquences des problèmes sociaux ou de santé spécifiques des individus ».

Le système de soins ne peut pas avoir d'influence directe, *a fortiori* exclusive, sur l'ensemble des conditions sociales, économiques, culturelles, démographiques qui affectent la santé et le bien-être des populations. Il ne dispose pas des leviers nécessaires pour agir sur l'ensemble de ces conditions, mais il peut intervenir sur les problèmes physiques, mentaux, comportementaux et sociaux des individus.

L'évaluation des systèmes de santé revient en fait à évaluer la capacité de la société à permettre à ses membres de vivre bien et longtemps. Les frontières de ces systèmes sont très difficiles à tracer, elles tendent à se confondre avec celles de la société. En revanche, il est possible de tracer les frontières du système de soins, de mettre en place des dispositifs d'évaluation de sa performance et d'aider les décideurs dans leurs responsabilités.

Ces dispositifs d'évaluation doivent pouvoir rendre compte de la complexité des systèmes et des organisations à évaluer. Il faut donc commencer par appréhender la dynamique du système et des organisations de soins, en l'analysant en tant que *système organisé d'action* (figure 1).

Comme tout système organisé d'action, le système de soins est situé dans un contexte concret (par exemple la société québécoise, le Canada), à un moment donné. Sa structure repose sur l'interaction d'une structure physique particulière (bâtiments, architecture, plateaux techniques, budgets), d'une structure organisationnelle (gouverne) et d'une structure symbolique spécifique (représentations, valeurs, normes collectives). Elle délimite un

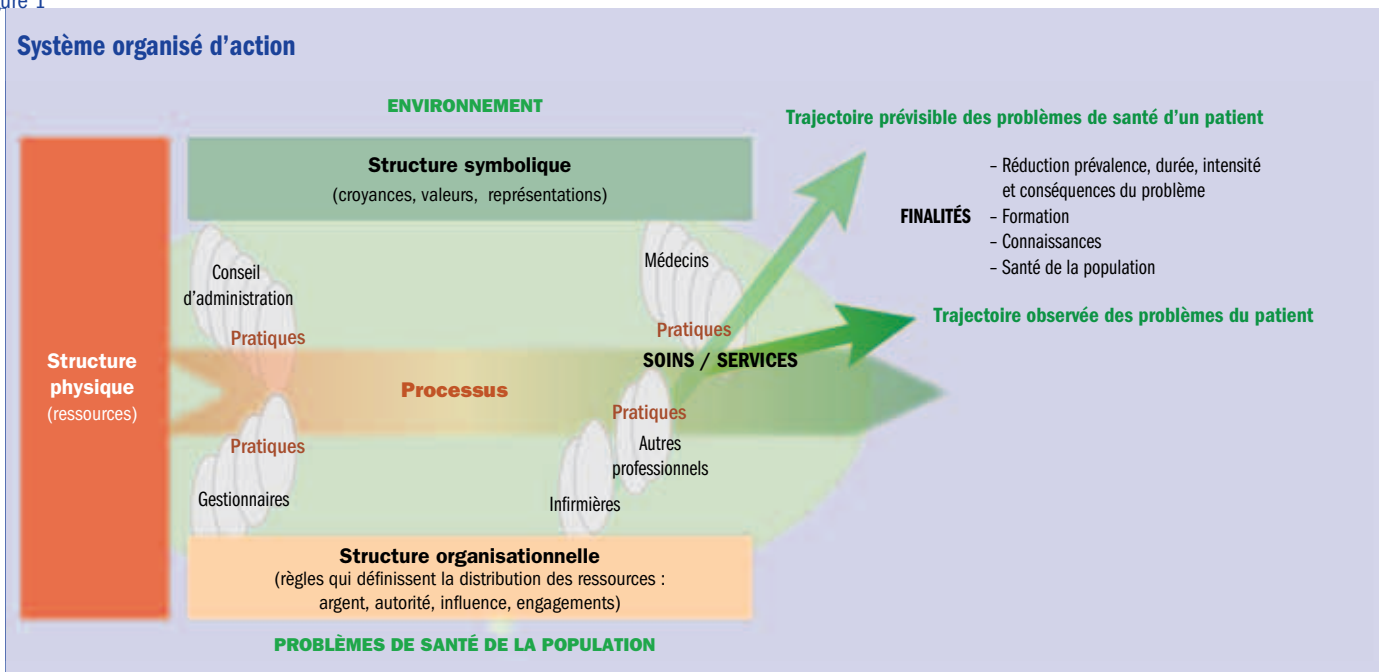
espace social dans lequel des acteurs (professionnels, gestionnaires, planificateurs, etc.) interagissent dans un jeu permanent de concurrence et de coopération, orienté par la ou les finalités du système, afin d'obtenir ou de contrôler les ressources pour améliorer leur position dans le système. Le système de soins en tant que système organisé d'action est lui-même constitué de plusieurs sous-systèmes organisés d'action (les systèmes régionaux de soins, les organisations de santé) interdépendants et ayant chacun un certain degré d'autonomie.

La figure 1 propose une représentation générique d'un système organisé d'action, elle permet de voir comment interagissent les éléments qui le caractérisent : la structure, les acteurs et leurs pratiques, les processus, et les finalités.

La structure est constituée de trois composantes et de leurs interrelations : une structure symbolique (ensemble des croyances, des représentations et des valeurs), une structure organisationnelle (ensemble des lois, règlements, conventions et règles de gestion) et une structure physique (volume et structuration des différentes ressources mobilisées). Ces trois composantes définissent un espace à l'intérieur duquel des acteurs interdépendants pratiquent un jeu permanent de coopération et de compétition.

Les acteurs peuvent être des individus ou des groupes organisés d'agents (organisations, groupes de pression, syndicats, etc.). Ils sont interdépendants et interagissent dans un jeu permanent de coopération et de compétition pour améliorer leur position, avoir ou contrôler les ressources critiques du système d'action. Leurs pratiques sont simultanément influencées par les structures du

figure 1



système et constitutives de ces dernières, Elles organisent les processus par lesquels les ressources du système sont mobilisées et employées pour produire les biens et services requis pour atteindre la ou les finalités recherchées.

Le réseau des processus de production constitue en quelque sorte la colonne vertébrale d'un système organisé d'action. Chaque processus est un ensemble cohérent d'activités qui produisent, à partir des ressources du système, des biens ou des services. Ceux-ci, en répondant aux besoins d'un client interne ou externe, contribuent à la réalisation de la finalité du système [4].

De façon générique, la finalité d'un système consiste à transformer la trajectoire prévisible d'évolution d'un ou de plusieurs phénomènes en agissant, au cours du temps, sur un certain nombre de leurs déterminants (objectifs spécifiques de l'intervention) pour corriger une situation problématique. Dans le cas du système de soins, la finalité consiste à réduire l'incidence, la durée, l'intensité et les conséquences des problèmes de santé et des problèmes sociaux de la population, à répondre aux attentes de la population et à favoriser l'équité.

Pour résumer, le système de soins et les organisations de santé dont il s'agit d'apprécier la performance peuvent être conçus comme des systèmes organisés d'action situés dans un contexte concret (la société québécoise, la France), à un moment donné.

La conceptualisation de la performance

Dans le domaine de la santé, de nombreux concepts – *efficacité, efficacité, rendement, productivité, qualité, accès, équité, sécurité, performance* – sont très souvent utilisés pour qualifier la performance sans que les nuances entre ces notions – pourtant distinctes – soient clairement établies et sans qu'on puisse être en mesure de les intégrer dans un même concept.

Nous faisons ici le choix de considérer la *performance* comme un concept générique englobant l'ensemble de ces notions qui en constituent des sous-dimensions. En nous appuyant sur les travaux de T. Parsons [25-27], nous définissons la performance comme la capacité pour un système d'accomplir au cours du temps quatre fonctions interdépendantes : 1) s'adapter à son environnement (acquérir des ressources, répondre aux besoins), 2) atteindre ses buts (réduire l'incidence, la durée, la gravité, les conséquences des problèmes de santé), 3) produire avec productivité des biens et des services de qualité et 4) maintenir et développer des valeurs et un climat favorable à la réussite des trois autres fonctions.

«L'appréciation de la performance repose ainsi, non seulement sur la mesure d'indicateurs de réussite dans chacune des quatre fonctions de l'organisation, mais aussi sur le caractère dynamique de la tension qui existe entre elles, c'est-à-dire sur la capacité de la gouvernance à orchestrer les échanges et les négociations nécessaires pour maintenir un équilibre entre les quatre fonctions par les différents acteurs.

«La performance d'une organisation ne peut se concevoir qu'en relation avec la performance de chacune de ses composantes et avec celle du réseau constitué par les autres organisations avec lesquelles elle interagit. C'est seulement dans la mesure où la performance d'une organisation contribue à la performance du réseau qu'elle est véritablement performante» [32, 33].»

Le but de l'évaluation et le public visé

L'engouement actuel pour la mesure de la performance des systèmes et services de santé découle de deux exigences de plus en plus fortes. D'une part, la population, et par conséquent les appareils politiques et technocratiques, s'inquiètent de plus en plus de l'état de leur système de santé. Des réformes et des restructurations majeures ont eu lieu à peu près partout et l'impact de ces transformations reste à démontrer. Une première raison d'implanter des systèmes d'évaluation de la performance est donc de répondre à ces demandes croissantes d'imputabilité et de reddition de comptes.

D'autre part, on observe à peu près partout d'importantes variations dans les coûts et dans la qualité des services dispensés. Ces constats exercent de fortes pressions pour améliorer la performance du système de soins. La deuxième raison pour laquelle on veut implanter des systèmes d'évaluation de la performance est d'aider les décideurs à améliorer de façon continue les systèmes et les organisations.

La figure 2 illustre l'idée que le besoin d'une évaluation sommative pour rendre des comptes est d'autant plus grand qu'on se situe à un niveau macroscopique (la population, le gouvernement, à la fois l'appareil technocratique et l'appareil politique, dans un but d'imputabilité) et que, inversement, le besoin d'une évaluation permettant d'améliorer de façon continue le fonctionnement du système est une nécessité pour les acteurs qui se situent au niveau microscopique (les gestionnaires du système et de ses organisations, dans un but premier d'amélioration, mais aussi d'imputabilité à l'appareil technocratique ; les professionnels et les gestionnaires des opérations dans un but d'amélioration continue des processus).

La faisabilité opérationnelle

La mise en œuvre du modèle d'évaluation globale et intégrée du système de soins (modèle EGIPSS) se fait en parcourant les quatre étapes décrites sur la figure 3.

Étape 1 : Sélection des indicateurs

Le choix des indicateurs permettant de documenter de façon équilibrée et valide chacune des dimensions et sous-dimensions du modèle est critique. Elle devra être faite avec la double exigence de parcimonie et d'exhaustivité, et dépend de la disponibilité et de la qualité des données.

La parcimonie est certainement une vertu importante dans l'élaboration d'un système d'évaluation. Les indicateurs doivent être choisis avec minutie pour à la fois

Les références entre crochets renvoient à la Bibliographie générale p. 63.



refléter de façon suffisamment globale et exhaustive les différentes dimensions de la performance tout en évitant de noyer l'utilisateur dans une mer d'informations qui ne lui permettra pas de porter un jugement éclairé. La contrainte de parcimonie variera aussi selon le but poursuivi (et le public cible). Un système mis en place à des fins d'imputabilité contiendra un nombre moindre d'indicateurs qu'un système visant l'amélioration continue de la performance [34]. De même, un système visant l'amélioration comprendra moins d'indicateurs lorsqu'il vise et est utilisé par des gestionnaires œuvrant à un niveau stratégique que lorsqu'il est utilisé par des professionnels ou des gestionnaires très proches des opérations.

Le choix des indicateurs devra aussi être réaliste, le dilemme ici étant de limiter le fardeau de collecte de données tout en optimisant la qualité des indicateurs. Il y aura donc d'abord un choix à faire entre l'utilisation des données existantes, provenant de systèmes d'information administratifs, financiers ou cliniques, et la collecte de données originales. Il est important de noter que la collecte de données originales ne représente pas obligatoirement un fardeau important. Comme le suggère la figure 4, la stabilité d'un système d'évaluation de la performance repose sur un mélange judicieux de données existantes et de données originales. L'absence d'un de ces piliers risque de rendre le système boiteux.

Les données, secondaires ou originales, devront être fiables et valides.

Étape 2 : Collecte des données

Étape souvent longue et fastidieuse qui exige qu'une attention très grande soit portée à la validation des processus de saisie et de sélection des données. Les banques de données administratives et les systèmes d'information mobilisés sont souvent nombreux et n'ont généralement pas été constitués pour l'évaluation de la performance.

Étape 3 : Analyse des données

Le tableau 1 fournit une carte de route pour analyser les données de façon à ce qu'elles puissent servir à porter un jugement sur la performance du système évalué.

Ce jugement doit porter, d'une part, sur la capacité du système à accomplir chacune de ses fonctions et de leurs sous-dimensions et, d'autre part, sur les relations d'équilibre entre les sous-dimensions et les dimensions du modèle.

Étape 4 : Élaboration du jugement

Le jugement que les décideurs peuvent élaborer grâce au modèle EGIPSS repose sur le postulat qu'aucun indicateur, à lui tout seul, n'est suffisant pour révéler les problèmes ou la qualité des processus des systèmes ou des organisations de santé dans leur complexité et leur dynamique. Pour porter un jugement sur la performance d'une organisation, on a besoin de l'apprécier par des configurations d'indicateurs qui éclairent simultanément les quatre fonctions du modèle. Plus la conceptualisation des processus sous-jacents est valide et complète, plus le jugement a de chance d'être valide. Pour aider les décideurs à accroître la validité de leur jugement, il est utile de formaliser les modèles logiques permettant de mettre en relation les différentes sous-dimensions du modèle. Quatre types de relations peuvent exister entre les sous-dimensions :

- 1) Relation de causalité : la performance d'une sous-dimension est directement influencée par la performance d'une autre sous-dimension. Exemple : la relation qui existe entre l'Acquisition des ressources et l'Utilisation des services par la population ;
- 2) Relation d'arbitrage : dans ce type de relation, il faut maintenir un équilibre entre deux sous-dimensions – favoriser l'une d'entre elles pourrait se faire au détriment de l'autre. Exemple : accorder trop d'importance à la Productivité des soins et services pourrait hypothéquer la Qualité non technique ;
- 3) Relation contingente : la relation entre deux

figure 2

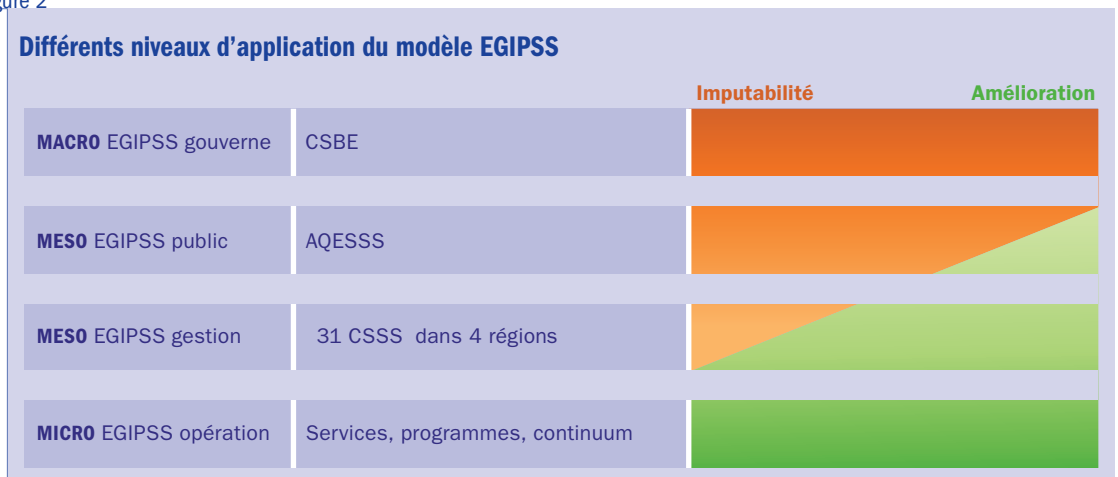


figure 3

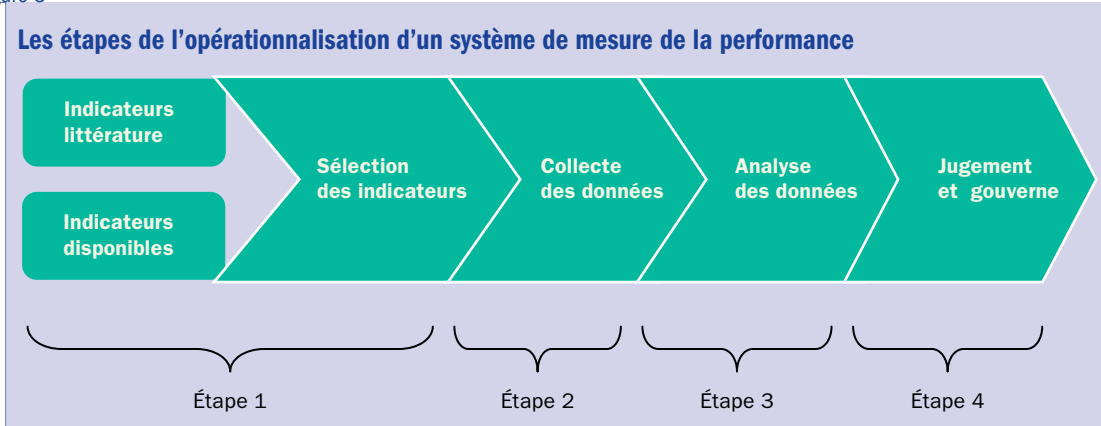


figure 4

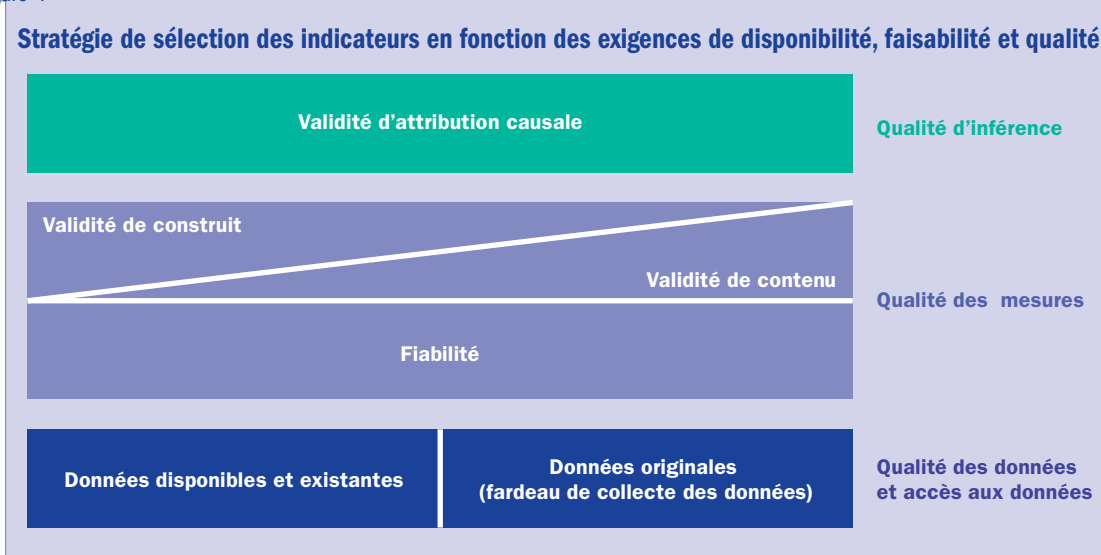


tableau 1

Stratégies d'analyse

Type d'analyse	Niveau d'analyse		
	Indicateurs	Sous-dimensions	Dimensions
Normative	<ul style="list-style-type: none"> Comparaison avec une norme et calcul d'indices relatifs (IR) (outils 1 et 2) Identification des leaders pour quelques indicateurs prioritaires (outil 3) 	<ul style="list-style-type: none"> Distribution par niveau catégoriel de performance (outil 4) Scores agrégés (outil 5) 	<ul style="list-style-type: none"> Distribution par niveau catégoriel de performance (outil 4)
Relationnelle	<ul style="list-style-type: none"> Cartes logiques (outil 6) 	<ul style="list-style-type: none"> Tableaux croisés des scores agrégés (outil 7) Analyses de regroupements (outil 8) 	



sous-dimensions dépend d'une troisième ; 4) Relation parabolique : en U ou en U inversé.

Discussion

Dans cet article, nous avons proposé que le contenu des systèmes d'évaluation de la performance des systèmes de santé reflète et soit cohérent avec un ensemble de choix portant d'abord sur l'objet de l'évaluation, la nature de la performance, le but de l'évaluation et le public visé. Le contenu devra aussi être adapté aux valeurs, intérêts, stratégies et priorités, et être réaliste par rapport aux contraintes du terrain.

Mais l'effort qui devra être mis pour développer et adapter de tels systèmes aux particularités locales en vaut-il la peine ? Que sait-on de l'utilité des systèmes d'évaluation de la performance ?

L'évidence sur l'utilité des systèmes visant l'imputabilité est limitée. La population, un des utilisateurs potentiels de systèmes d'imputabilité, ne semble pas les trouver utiles [14]. Bien que fort préoccupés par les questions de santé, les gens trouvent difficile de se prononcer sur l'information qu'ils souhaiteraient obtenir, particulièrement s'ils sont interrogés dans une perspective systémique plutôt qu'individuelle. De plus, ils trouvent l'information difficile à interpréter et n'ont

pas confiance dans les organisations pour fournir de l'information crédible sur leur performance, ni dans les gouvernements, ni dans les universités associées aux centres hospitaliers universitaires (CHU).

Pour les gouvernements, les systèmes d'imputabilité semblent au contraire être d'une grande utilité pour rendre des comptes à la population, transmettre des messages sur les valeurs et les priorités, et gérer le système en comparant et en motivant. L'utilisation par des gestionnaires gouvernementaux de systèmes d'imputabilité peut cependant amener des comportements indésirables de la part des gestionnaires des organisations dont la performance est évaluée.

Que ce soit dans un but d'imputabilité ou dans une perspective d'amélioration, l'élaboration de systèmes d'évaluation de la performance qui respecte les exigences d'architecture évoquées ici devrait permettre d'optimiser l'utilité de ces systèmes pour la prise de décision. L'élaboration de tels systèmes se fait souvent dans des contextes sous pression. Il semble cependant important d'éviter la simplification conceptuelle qui pourrait être tentante pour faire vite et simple. La finesse requise dans l'évaluation de la performance des systèmes et organisations de santé est nécessaire pour rendre compte de leur complexité et pour agir en toute connaissance de cause. **F**

Bien choisir et dimensionner l'évaluation : l'approche de la Banque mondiale

Les pratiques de suivi (« monitoring ») et d'évaluation (« evaluation ») des activités donnent aux politiques publiques, financeurs, managers et responsables de la société civile les meilleurs moyens de tirer parti de l'expérience, améliorer les services, planifier et allouer les ressources en rendant compte des résultats de l'action aux différentes parties prenantes.

C'est l'objet du court ouvrage publié par la Banque mondiale (deuxième édition 2004) intitulé : *Monitoring Evaluation: Some Tools, Methods & Approaches* consultable et téléchargeable sur le site de la Banque mondiale : [http://lnweb90.worldbank.org/OED/oeddoclib.nsf/DocUNIDViewForJavaSearch/A5EFB85D776B67D285256B1E0079C9A3/\\$file/MandE_tools_methods_approaches.pdf](http://lnweb90.worldbank.org/OED/oeddoclib.nsf/DocUNIDViewForJavaSearch/A5EFB85D776B67D285256B1E0079C9A3/$file/MandE_tools_methods_approaches.pdf)

Aujourd'hui, où l'accent est mis fortement sur l'objectivation des résultats mais aussi sur la nécessité

d'un calibrage réaliste et raisonnable du coût de l'évaluation par rapport au coût intrinsèque de l'action, ce document synthétique et pragmatique permet de guider le choix des méthodes à privilégier selon l'objectif et le point de vue d'où elle est conduite.

Il répertorie les principales pratiques possibles en ce domaine : outils, méthodes dont certains sont complémentaires, d'autres alternatifs et le choix relève de façon conjuguée de l'objectif de l'évaluation, du point de vue de l'évaluateur, du délai exigé de disponibilité des résultats de cette évaluation, et enfin du niveau des moyens mobilisables pour la réaliser.

Il donne sur une fiche consacrée à chaque méthode un descriptif synthétique et systématique des avantages et désavantages, coûts, temps, compétences requises et des références bibliographiques (et internet) permettant de compléter son information.

Sans prétendre à l'exhaustivité, ce guide est précieux par sa concision et son approche didactique tout autant pour le maître d'ouvrage — *a priori*, lors de l'élaboration du cahier des charges d'un appel à projets et, *a posteriori*, pour disposer d'une aide à la décision ou tirer enseignement d'une expérience — que pour le maître d'œuvre, pour répondre à un appel d'offres, assurer le suivi de son action, repérer les actions correctrices nécessaires et enfin « rendre compte » au commanditaire. **F**

Dominique Deugnier
Médecin de santé publique