

Acteurs, objets et méthodes de l'évaluation

Il existe différents types d'évaluation en santé : plan, programmes et politiques, pratiques professionnelles, certification et performances des établissements. Qui sont les évaluateurs et quelles sont les méthodologies mises en œuvre ?

Acteurs et objets de l'évaluation en santé

Bernard Perret
Ingénieur général des ponts, des eaux et des forêts, membre permanent du Conseil général des ponts et chaussées, membre du Haut Conseil de la santé publique

L'évaluation de l'action publique présente de multiples visages. Entre les indicateurs de performance, la recherche évaluative, l'évaluation participative et l'audit, entre l'évaluation des politiques, programmes, projets et celle des pratiques professionnelles, etc., les opérations relevant du domaine de l'évaluation mobilisent différentes postures professionnelles et modèles méthodologiques, et font appel à des compétences techniques variées. Les activités d'évaluation se différencient par leurs objets et finalités, leurs méthodes et leur cadre institutionnel. Cette complexité du champ rend malaisée sa description et l'analyse des tendances qui s'y manifestent. En termes d'objets, une typologie grossière conduirait à distinguer les actions (lois, politiques, programmes, projets...), les structures (organismes, établissements, services...), les pratiques professionnelles, les techniques et les protocoles de soins. À chacune de ces classes d'objets correspond un *corpus* de méthodes ainsi que des finalités/fonctions particulières. Ces dernières s'inscrivent cependant toujours, comme l'indique la charte de la Société française de l'évaluation (SFE), « dans la triple logique du management public, de la démocratie et du débat scientifique ».

La seule définition officielle de l'évaluation figurait dans le décret actuellement abrogé du 18 novembre 1998 : « L'évaluation d'une politique publique a pour objet d'apprécier l'efficacité de cette politique en comparant ses résultats aux objectifs assignés et aux moyens mis

en œuvre. » Il ressort de cette définition que l'évaluation s'entend le plus souvent comme un examen rétrospectif des effets d'une action, étant précisé qu'elle s'intéresse aussi aux conditions de mise en œuvre de celle-ci, à la compréhension de ses mécanismes d'action, à sa pertinence et à sa cohérence.

Cette prééminence de l'évaluation *ex post* ne doit pas faire oublier qu'il existe une évaluation *ex ante* (ou étude d'impact) et une évaluation concomitante (ou « chemin faisant »). L'évaluation *ex ante*, plus directement liée à l'exigence de rationalisation des décisions publiques, devrait se développer au cours des prochaines années, suite notamment à la loi constitutionnelle du 23 juillet 2008 et la loi organique 2009-403 du 15 avril 2009 (article 8), qui imposent à partir du 1^{er} septembre 2009 une étude d'impact pour tout projet législatif. L'influence des procédures mises en place à Bruxelles devrait également pousser au développement des études d'impact, particulièrement dans l'analyse des projets à fort impact budgétaire. L'évaluation socio-économique *ex ante* des projets d'infrastructures de transport est une pratique courante depuis plusieurs décennies ; ses méthodes sont actuellement en cours de révision pour qu'elles prennent mieux en compte les enjeux du développement durable, et notamment la lutte contre le changement climatique.

La complexité des jeux d'acteurs autour de l'évaluation induit une grande diversité de configurations institutionnelles. Il n'est pas rare qu'une même action



donne lieu à plusieurs évaluations à l'initiative d'acteurs différents. Le risque de foisonnement qui en découle est d'autant plus grand que l'évaluation est perçue, non sans raison, comme un enjeu d'influence : être à l'initiative d'une évaluation donne généralement la possibilité d'en orienter le questionnement.

Pour se limiter au domaine de la santé, le paysage de l'évaluation est d'une grande complexité et l'on peut prévoir que la diversification des pratiques ne fera que s'accroître. En première analyse, deux grands blocs se dessinent : l'évaluation des pratiques et techniques de soins et l'évaluation des politiques et programmes de santé publique. Mais il est impossible de ne pas tenir compte de plusieurs champs connexes – recherche en santé publique, économie de la santé, statistiques, observation, contrôle et audit –, qui interfèrent à des degrés divers avec les activités d'évaluation. Il arrive fréquemment que des opérations lancées sous ces différents labels constituent de fait des évaluations. Le monde de la santé se caractérise enfin par une multiplicité d'acteurs aux statuts les plus divers : services administratifs, organismes de recherche, instances publiques autonomes. Outre la Haute Autorité de santé (HAS) et le Haut Conseil de la santé publique (HCSP), il convient de mentionner à ce titre le Conseil national de l'évaluation sociale et médico-sociale créé en 2002, dont le champ de compétence jouxte étroitement le domaine sanitaire proprement dit.

L'évaluation des pratiques et techniques de soin

Il s'agit de la partie la plus structurée de l'évaluation en santé. Avec plusieurs dizaines de salariés permanents, la HAS – qui a succédé en 2004 à l'Anaes – est l'organisme le plus important dans le champ de l'évaluation médicale. Les objets évalués par la HAS sont eux-mêmes divers : médicaments, technologies, dispositifs de soin, actes professionnels, actions ou programmes de prévention. L'évaluation n'est toutefois que l'une des deux grandes missions de la HAS, parallèlement à l'amélioration de la qualité et de la sécurité des soins (accréditation, certification, guide de bonnes pratiques, etc.). Au plan institutionnel, la HAS est dotée d'une forte autonomie : c'est une autorité administrative indépendante au sens plein du terme. Dans le domaine de l'évaluation des soins, la HAS est un acteur dominant. Toutefois, d'autres instances sont susceptibles d'intervenir sur ce même terrain. Ainsi, l'Office parlementaire des politiques de santé (aujourd'hui supprimé) a procédé à plusieurs évaluations de politiques de prévention et de dépistage, dont l'objet et le questionnement relèvent directement de la compétence de la HAS. Il existe par ailleurs une large zone de recouvrement entre l'évaluation des soins et la recherche clinique. La HAS s'appuie largement sur les recherches et expérimentations effectuées en milieu hospitalier, à travers notamment la pratique des conférences de consensus.

En termes de méthodes, l'évaluation des pratiques de soin repose principalement sur l'expérimentation avec

échantillon témoin et les revues (débouchant ou non sur des conférences de consensus¹). Elle est centrée sur la question de la mesure des bénéfices pour les malades, mais intègre également, de plus en plus, la question du coût (études coût-efficacité). Elle est fortement influencée par le concept anglo-saxon d'*evidence based medicine* (ou « médecine fondée sur des preuves »).

L'évaluation des politiques et programmes de santé publique

Le Haut Conseil de la santé publique est un organisme récent (2007), créé en application de la loi de santé publique de 2004 (LSP). Ses missions en matière d'évaluation, définies par la loi, comprennent plusieurs volets :

- l'évaluation des 100 objectifs de la loi de 2004,
- l'évaluation des programmes de santé publique mentionnés dans la loi,
- une contribution à l'évaluation globale de la loi elle-même.

Les travaux du HCSP relèvent pour une part d'une approche classique d'évaluation de politiques publiques. Les plans de santé publique (Plan cancer et Plan maladies rares, notamment²) ont été évalués sous l'égide de comités d'évaluation pluralistes et ont reposé sur la mise en place d'un programme d'étude et d'investigation visant à répondre à un questionnement élaboré par les comités. L'évaluation des objectifs de la LSP constitue en revanche un exercice plus original au regard des typologies usuelles de pratiques d'évaluation. Cette évaluation a débouché sur la rédaction de fiches commentant l'évolution des indicateurs associés aux objectifs (lorsque ces indicateurs existent et peuvent être renseignés), analysant sommairement les causes des évolutions constatées et faisant des propositions pour les objectifs de la future LSP qui devrait prendre la suite de la loi de 2004.

La contribution du HCSP à l'évaluation globale de la loi de santé publique se traduit par un ensemble d'observations rassemblées dans le rapport rédigé par le Haut Conseil, en s'appuyant notamment sur les deux exercices précédemment mentionnés (évaluation des plans et programmes de santé publique et évaluation des objectifs).

Les frontières entre les missions du HCSP et celle de la HAS sont globalement assez claires – d'un côté

1. Selon la présentation résumée des conférences de consensus par la HAS, « les recommandations sont rédigées en toute indépendance par un jury de non-experts du thème traité, dans le cadre d'un huis clos de 48 heures faisant immédiatement suite à une réunion publique. Les recommandations répondent à une liste de 4 à 6 questions prédéfinies posées au jury. Au cours de la réunion publique, les éléments de réponse à ces questions sont exposés par des experts du thème et débattus avec le jury, les experts et le public présent. La réalisation d'une conférence de consensus est particulièrement adaptée lorsqu'il existe une controverse professionnelle forte, justifiant une synthèse des données disponibles, une présentation des avis des experts du thème, un débat public, puis une prise de position de la part d'un jury indépendant. »

2. Le Plan national santé-environnement (PNSE) avait pour sa part fait l'objet d'une évaluation à mi-parcours par un comité ad hoc, avant la mise en place du HCSP.

les pratiques de soin, de l'autre la politique de santé publique – mais des zones de recouvrement existent. Des questions telles que le dépistage ou la vaccination sont susceptibles d'être abordés par les deux instances sous des angles en principe complémentaires. À titre d'exemple, le HCSP a été saisi d'une demande d'évaluation des mesures prises en matière de vaccination antituberculeuse, l'efficacité technique des vaccins ayant fait préalablement l'objet d'une expertise par la HAS. D'une manière générale, le travail du HCSP a un caractère plus « politique » que celui de la HAS. Les évaluations réalisées sous la responsabilité du Haut Conseil s'efforcent de prendre en compte plus largement les aspects sociaux (inégalités en matière de santé, impacts psycho-sociaux...), les processus administratifs de mise en œuvre des actions, et les enjeux stratégiques de la politique de santé publique.

Le rôle des services du ministère

La Direction générale de la santé et la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees) apportent diverses contributions à l'évaluation des politiques du ministère. La Drees publie chaque année depuis 2006 un *Rapport de suivi des objectifs de la loi de santé publique*, bilan chiffré de l'état de santé de la population élaboré à partir des indicateurs d'atteinte des objectifs de la loi. Ces rapports fournissent au HCSP des matériaux essentiels pour son travail d'évaluation et constituent par eux-mêmes une contribution à l'évaluation de la loi. La DGS possède en son sein un bureau « programmation-synthèse-évaluation ». Il a notamment pour mission de « développer les méthodes de programmation et d'évaluation ». Il apporte notamment un soutien méthodologique aux services déconcentrés du ministère dans l'évaluation des plans régionaux de santé publique.

L'évaluation comme prolongement du contrôle

L'Inspection générale des affaires sociales (Igas) participe à divers titres à l'évaluation des politiques de santé. Elle a contribué à plusieurs reprises aux évaluations que le gouvernement s'est engagé, par la loi, à adresser au Parlement. On peut citer l'exemple de la couverture maladie universelle (2001). L'Igas réalise en outre un certain nombre d'évaluations de sa propre initiative (ainsi, en 2002, l'évaluation de l'informatisation des professionnels de santé et de la gestion des praticiens hospitaliers). Pour réaliser ces évaluations, l'Igas s'appuie largement sur les travaux d'étude existant, complétés par le recueil des points de vue des acteurs institutionnels, administratifs et associatifs.

La connaissance du système de soin de l'Igas peut également être mise à contribution dans le cadre d'évaluations interministérielles, conjointement avec d'autres corps d'inspection (organisation des soins en prison, 2001), ou encore dans le cadre des évaluations du HCSP (Plan cancer, PNMR).

Le Parlement

Un Office parlementaire d'évaluation des politiques de santé (Opeps) avait été créé en 2002 par la loi de financement de la Sécurité sociale. Il a été supprimé par la loi constitutionnelle du 23 juillet 2008. Ses missions seront désormais exercées par les commissions des affaires sociales des deux assemblées et par le « Comité d'évaluation et de contrôle des politiques publiques », qui doit prochainement être créé à l'Assemblée nationale.

Entre 2003 et 2008, l'Opeps a remis une dizaine de rapports portant sur l'évaluation de pratiques de soin : dépistage des cancers du sein et de la prostate, prise en charge de la maladie d'Alzheimer et des accidents vasculaires cérébraux, utilisation des médicaments psychotropes... À l'avenir, le futur Comité d'évaluation et de contrôle des politiques publiques devrait permettre à l'Assemblée nationale de contribuer plus largement à l'évaluation des politiques de santé.

La Cour des comptes

À l'instar du Parlement – et à la demande expresse du président de la République –, la Cour des comptes envisage de se doter d'un dispositif spécifique d'évaluation des politiques publiques, en élargissant à cette occasion le spectre de ses compétences internes et en instituant de nouvelles formes de travail collectif. En rupture avec la posture traditionnelle des magistrats, la cour envisage d'associer à certaines évaluations les administrations en charge de la politique examinée et d'autres parties prenantes concernées. Jusqu'à la réforme en cours, les contributions de la cour à l'évaluation des politiques de santé sont restées limitées. À titre d'exemple, on peut mentionner le « rapport thématique » sur le Plan cancer (juin 2008), qui inclut notamment un audit du fonctionnement de l'Institut national du cancer (INCa). Dans son rapport, la cour précise elle-même que son examen a porté sur « l'efficacité et l'efficacité de la gestion du Plan par ces intervenants, au regard des objectifs initiaux », mais qu'il « n'a pas inclus une évaluation de l'efficacité du Plan en termes d'impact médical ».

L'évaluation des plans régionaux de santé publique (PRSP)

La loi de santé publique prévoit que, dans chaque région, « le plan régional de santé publique ainsi que les programmes définis par la région font l'objet d'une évaluation ». Pour procéder à ces évaluations, les régions ont pu s'appuyer sur un guide méthodologique diffusé par le ministère³. Ces évaluations ont été généralement conduites, avec l'aide d'un prestataire, par des comités d'évaluation mis en place sous l'égide des groupements régionaux de santé publique (GRSP). Un certain nombre de rapports intermédiaires d'évaluation (qui portent rappelons-le sur la période 2006-2009) sont d'ores et déjà publiés et consultables sur Internet. **F**

3. *Évaluation des PRSP, quelques repères pour les acteurs* (Direction générale de la santé, avril 2008).

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



La charte de la Société française de l'évaluation : des principes pour les participants à l'évaluation des politiques et des programmes publics

Jean-Claude Barbier
Directeur de recherche CNRS, Centre d'économie de la Sorbonne (CES), université Paris 1 Panthéon Sorbonne CNRS.

Le besoin de définir des critères de bonnes pratiques et de déontologie est particulièrement fort en matière d'évaluation des politiques et des programmes publics¹ (EPP, en abrégé dans le texte), compte tenu du caractère politique de cette activité, même si la tendance contemporaine des professions à se doter de règles normatives de conduite et de guides propres est générale. Il s'agit d'un mouvement qui relève de la standardisation, mais il est loin de s'y résumer. Dans pratiquement tous les pays où l'évaluation est pratiquée à grande échelle, les associations d'évaluateurs ont élaboré des systèmes de normes ou standards, relatifs, d'une part, à la déontologie (des évaluateurs, ou, plus largement, des commanditaires et des acteurs qui participent au pilotage de l'évaluation, comme les membres des instances d'évaluation) et, d'autre part, à la qualité des évaluations (qualité de ses produits, qualité de ses processus, qualité des systèmes d'évaluation). Le choix de ces normes, que les participants à l'évaluation se donnent le plus souvent volontairement, est essentiel pour la définition de ce qu'ils entendent par évaluation des politiques publiques.

La charte de la SFE comprend désormais sept principes, dont on dira brièvement quelle fut la genèse. On conclura par quelques considérations sur l'usage de la charte en France.

L'adoption de la charte

Dès sa création, la question des « standards de qualité » et de la « déontologie » s'est imposée à la SFE comme l'un des enjeux majeurs de la structuration d'un milieu professionnel de l'évaluation en France. Les statuts de la société prévoyaient expressément, dans leur article 3 (alinéa b), que la société « se fixe notamment pour objectifs de faire progresser les techniques et méthodes et de

1. Il n'est pas nécessaire de souligner le caractère éminemment polysémique du mot « évaluation ». On évalue constamment dès qu'on pense, et on classe, comme l'a souligné le regretté Georges Pérec. On peut évaluer des projets, des personnes, etc. La charte dont on parle ici a été adoptée, dans une première version, en octobre 2003, par la Société française de l'évaluation. L'auteur de ces lignes a été le fondateur du groupe « Standards de qualité et déontologie » de la société et son animateur dans ses débuts. À ce titre, il a participé, avec d'autres membres de la SFE et du groupe, aux étapes de l'élaboration de la charte, révisée en 2006. Les Cahiers de la SFE n° 1 (2005) retracent l'histoire de cette charte. Ils sont accessibles sur le site web de la société. On emploiera dans le texte « évaluation » ou « EPP » pour signifier l'évaluation des politiques et des programmes publics.

favoriser le respect des règles éthiques et procédurales propres à garantir la qualité des évaluations ainsi qu'un usage approprié des résultats ». Entre 1999 et l'adoption, dans son assemblée générale de Limoges, en 2003, les membres de la SFE ont été associés, par son conseil d'administration et son bureau, au débat sur le caractère opportun de l'adoption de « standards » ou de « principes éthiques ». Dans un premier temps, les adhérents se sont révélés pour une part rétifs, pour une autre part favorables, mais en majorité plutôt indifférents, avant de voter la charte à une large majorité. Le choix s'était porté sur l'affirmation d'un petit nombre de principes simples, plutôt que, comme l'ont fait beaucoup de sociétés étrangères, d'adopter une liste de prescriptions précises concernant l'utilisation, la conduite, les méthodes, etc., de l'évaluation². En 2006, la conscience de l'intérêt de la charte était nettement plus vive, quand, en assemblée générale, un débat a porté sur l'opportunité d'ajouter à la charte un commentaire affirmant que « les membres de la SFE déclarent adhérer aux principes de la Charte de l'évaluation, s'engagent à les appliquer en tenant compte des contextes particuliers rencontrés et à promouvoir la charte tant dans leur milieu professionnel qu'auprès des personnes et institutions concernées par l'évaluation et la délibération publique ». Ce texte adopté figure désormais dans le préambule de la charte.

Les principes

La première liste de principes en comprenait six (voir encadré page 31). Celui qui provoqua le plus de débats fut celui de « pluralité », tant il semblait à certains restreindre l'application du terme « évaluation » pour des opérations qui ne concernaient pas simplement la relation d'un commanditaire et d'un évaluateur.

En 2006, la formulation des principes fut légèrement remaniée et un septième principe, celui d'opportunité, fut adjoint, selon lequel :

« Une évaluation doit être décidée lorsqu'elle est susceptible de – et organisée afin de – produire des résultats à l'égard des finalités mentionnées au préambule de cette charte : compte rendu démocratique, efficacité de la dépense, apprentissage organisationnel, facilitation d'évaluations ultérieures ».

2. On se fera une idée d'une telle liste de prescriptions en consultant, sur le site de la SFE, le premier projet de texte discuté en 2001 (<http://www.sfe.asso.fr/fr/charte-sfe.html>).

La principale autre modification concernait le principe de pluralité, dont la rédaction est désormais la suivante : « L'évaluation s'inscrit dans la triple logique du management public, de la démocratie et du débat scientifique. Elle prend en compte de façon raisonnée les différents intérêts en présence et recueille la diversité des points de vue pertinents sur l'action évaluée, qu'ils émanent d'acteurs, d'experts, ou de toute autre personne concernée. Cette prise en compte de la pluralité des points de vue se traduit – chaque fois que possible – par l'association des différentes parties prenantes concernées par l'action publique ou par tout autre moyen approprié. »

La charte réaffirme ainsi une caractéristique de l'EPP inscrite dans un univers qui croise science, démocratie et gestion publique. La pluralité des points de vue qu'elle invite à rassembler ne transforme pas le pilotage de l'évaluation en une représentation politique des intérêts ; elle invite à y associer le plus grand nombre de ces points de vue parce qu'ils sont pertinents, pour des raisons cognitives : d'où le remplacement du mot « légitimes » par le mot « pertinents » dans le premier principe.

Logique de la charte et utilisation

Comme on pourra le constater en consultant les étapes de l'adoption de la charte que nous n'avons pas la place de détailler ici, la logique initiale – qui vaut encore aujourd'hui – est celle d'un document de référence que les commanditaires sont invités à mettre en annexe de leurs appels à propositions, ou les évaluateurs en référence de leurs propositions. La charte est aujourd'hui utilisée par les masters en évaluation, dont le nombre croît, dans les universités ; elle sert aussi de référence pour les travaux de méta-évaluation.

Les protagonistes de l'adoption et de la discussion de la charte de la SFE ont été étonnés, avec le recul d'une expérience de près de sept ans, donc relativement récente par comparaison avec les sociétés sœurs, de constater l'importance symbolique croissante de la charte. Ce document est un marqueur de l'identité de la société, mais il va désormais bien au-delà et fait l'objet d'une reconnaissance croissante, même si la SFE n'a pas la prétention de représenter tous les types d'acteurs qui s'occupent en France d'évaluation. **F**

Les principes

Principe de pluralité

L'évaluation prend en compte de façon équilibrée les différents points de vue légitimes qui ont été exprimés sur l'action évaluée. Cette prise en compte de la pluralité des points de vue se traduit – chaque fois que possible – par l'association des différentes parties prenantes au processus d'évaluation, ou par tout autre moyen approprié.

Principe de distanciation

L'évaluation est conduite de façon impartiale. Les personnes participant à l'évaluation à titre professionnel informent les autres partenaires de tout conflit d'intérêt éventuel. Le processus d'évaluation est conduit de façon autonome par rapport aux processus de gestion et de décision. Cette autonomie préserve la liberté de choix des décideurs publics.

Principe de compétence

Les personnes participant au processus d'évaluation à titre professionnel mettent en œuvre des compétences spécifiques en matière de conception et de conduite de l'évaluation,

de qualité de la commande, de méthodes de collecte de données et d'interprétation des résultats. Elles ont le souci d'améliorer et de mettre à jour leurs compétences. Elles s'appuient sur des méthodes reconnues, notamment celles en usage dans la communauté internationale de l'évaluation. Elles utilisent ces méthodes en explicitant leurs limites.

Principe de respect des personnes

Les participants au processus d'évaluation respectent les droits, l'intégrité et la sécurité de toutes les parties concernées. Ils s'interdisent de révéler l'origine nominative des informations ou opinions recueillies, sauf accord des personnes concernées.

Principe de transparence

La présentation des résultats d'une évaluation s'accompagne d'un exposé clair de son objet, de ses finalités, de ses destinataires, des questions posées, des méthodes employées et de leurs limites, ainsi que des arguments et critères qui

conduisent à ces résultats. La diffusion publique des résultats d'une évaluation est souhaitable. Les règles de diffusion des résultats sont établies dès le départ. L'intégrité des résultats doit être respectée, quels que soient les modalités ou les supports de diffusion retenus.

Principe de responsabilité

La répartition des rôles entre les différents acteurs de l'évaluation est établie dès le départ de façon à ce que toutes les fonctions de celle-ci soient bien prises en charge (définition du mandat, pilotage du processus, enquêtes et analyses, formulation du jugement et des recommandations, diffusion des résultats). Les personnes et institutions impliquées dans le processus d'évaluation mobilisent les moyens appropriés et fournissent les informations nécessaires à la conduite de l'évaluation. Elles sont conjointement responsables de la bonne application des principes énoncés dans cette charte. **F**



La place des données qualitatives dans les évaluations en santé publique

Valéry Ridde

Chercheur, Ph. D.,
Centre de recherche
du centre hospitalier
de l'université de
Montréal (CRCHUM),
Département de
médecine sociale
et préventive de
l'université de
Montréal, Institut
de recherche en
sciences de la santé/
CNRST du Burkina
Faso

Geneviève Mak

Doctorante en santé
publique, Centre
de recherche du
centre hospitalier
de l'université de
Montréal (CRCHUM),
Département de
médecine sociale
et préventive de
l'université de
Montréal

« **S**i vous ne pouvez pas le mesurer, cela n'existe pas », voici un adage qui revient régulièrement lorsque l'on souhaite entrer dans le débat de l'utilisation des méthodes quantitatives ou qualitatives dans le domaine de l'évaluation. Hier encore (pour plagier Aznavour), à la veille d'écrire ce texte, nous présentions à des médecins du Burkina Faso les résultats d'une évaluation de processus ayant eu recours à des données qualitatives. La synthèse des discours prononcés par les infirmiers que nous avons interrogés n'a pas été facilement perçue par cet auditoire car, nous a-t-on dit, « il faut des chiffres, nous dire quel est le pourcentage d'infirmiers qui ont dit cela ». Si pour les évaluateurs ce type de questionnement reste contemporain et rémanent à la fameuse « guerre des paradigmes », nous sommes au cœur d'un débat qui semble bien dépassé pour ceux qui suivent l'histoire des sciences. Car nos collègues médecins ne souffrent pas plus de quantophobie aiguë, cette fameuse maladie de la mesure décrite par De Gaulejac, que la plupart des personnes concernées par les évaluations. On ne peut donc faire l'économie d'une rapide mise en perspective pour éclairer les discours et pratiques actuelles de l'usage des données qualitatives dans les évaluations.

Des expérimentations à l'utilisation des résultats des évaluations

Le domaine de l'évaluation a d'emblée été largement dominé par la vision du monde conçue par les sciences naturelles, les méthodes expérimentales et l'approche hypothético-déductive qui donnent une place prépon-

dérante aux mesures quantitatives et aux analyses statistiques. C'est ainsi qu'aux États-Unis d'Amérique de multiples évaluations se sont développées dans les années 1960 en usant d'une myriade d'expérimentations en comparant des groupes bénéficiant d'un programme à d'autres n'en bénéficiant pas, pour mieux évaluer l'efficacité. Certains évaluateurs pensent donc que seules ces expériences contrôlées sont en mesure de vérifier les effets et les impacts des programmes. Par la suite, et la taille de ce texte nous contraint à un réductionnisme historique, on s'est rendu compte de la difficulté (encadré « De l'efficacité des parachutes ») à organiser de telles évaluations et à expliquer comment ces effets ont été produits. Or les chiffres sont souvent moins utiles que les mots pour comprendre les processus. De surcroît, ces interrogations d'ordres méthodologique, éthique et politique ont émergé en parallèle à une nouvelle vague de pratiques évaluatives plus soucieuse de l'utilisation des résultats. C'est ce que l'on qualifie schématiquement de 4^e génération d'évaluation (après celles de la mesure, de la description et du jugement), une évaluation plus ancrée dans le constructivisme. Il s'agissait aussi, pour de nombreux évaluateurs, de trouver de nouvelles stratégies pour accroître le potentiel d'utilisation des résultats des évaluations face à la rédaction de nombreux rapports dont les recommandations n'étaient pas toujours utilisées. Ainsi, plusieurs nouvelles approches sont nées dans les années 1980 où les données qualitatives prirent plus de place afin de faire en sorte que les évaluations soient plus réactives, plus participatives et mieux centrées sur l'utilisation des recommandations.

De l'efficacité des parachutes

Dans un article célèbre du *British Medical Journal* paru en 2003, Smith et Pell présentent leur étude qui visait à vérifier si le parachute est un instrument efficace pour prévenir les traumatismes liés à la chute. Pour cela, ils ont entrepris une revue systématique des articles scientifiques présentant les résultats d'essais randomisés — la seule méthode valide pour juger de l'efficacité du parachute. Leur recherche a été infructueuse, aucun essai randomisé n'a été retrouvé. Pourtant, les études observationnelles montrent l'efficacité des parachutes. Les auteurs en concluent que les pourfendeurs des études observationnelles et ceux qui ne jurent que par l'évaluation expérimentale doivent redescendre sur terre.

De la complexité des interventions en santé publique

Le recours à des données qualitatives s'explique certainement en partie par cette impasse méthodologique des évaluations expérimentales pour expliquer comment produire des effets. Mais une autre clef d'explication doit être utilisée, celle de la complexité des interventions en santé publique. Ces interventions ont une histoire, elles sont enchâssées dans un contexte particulier qu'il faut obligatoirement prendre en compte. Elles sont des systèmes ouverts, dont l'évaluateur qui les observe fait lui-même partie. Compte tenu de cette complexité, le recours à des démarches évaluatives expérimentales paraît peu pertinent. Les données qualitatives sont donc ici d'un grand secours et pas uniquement pour comprendre le point de vue (*emic*) des acteurs, mais aussi pour rendre compte des effets des interventions. Car les données qualitatives peuvent aussi être utiles pour expliquer des liens de causalité, comme nous l'avons

Valéry Ridde dispose
d'une bourse de
recherche du Fonds pour
la recherche en santé
du Québec (FRSQ) et
l'analyse des écrits de
l'exemption en Afrique
s'inscrit dans une étude
financée par les Instituts
de recherche en santé
du Canada (IRSC).

Merci à
J.-P. Olivier de Sardan
pour ses commentaires
d'une version précédente
de ce (trop court) texte.

L'approche réaliste

S'inspirant des méthodes d'évaluation fondées sur la théorie et d'une épistémologie réaliste, Pawson et Tilley proposent de mettre en lumière le mécanisme (M) actionné par l'intervention pour produire ses effets (E) dans un contexte donné (C), ce qu'ils appellent la configuration « Contexte-Mécanisme-Effet » (CME). On ne répond plus seulement à la question : « L'intervention fonctionne-t-elle ? », mais : « L'intervention fonctionne-t-elle, comment, pour quels bénéficiaires, et dans quelles circonstances particulières ? »

fait par exemple au Bénin ou au Québec. Mais pour cela, la théorie de l'intervention doit avoir été bien décrite lors de la phase pré-évaluative (*evaluability assesment*), autrement dit la chaîne causale doit être explicitée. Les données quantitatives peuvent aussi manquer. Dans une étude entreprise en 2007 pour le compte de l'INPES dans trois régions françaises, nous souhaitions rendre compte des interventions qui favorisaient la réduction des inégalités sociales de santé. Or, dans toutes les interventions étudiées, aucune ne disposait de données quantitatives pour vérifier si elles avaient réellement réduit les inégalités sociales de santé, alors même qu'elles avaient été sélectionnées par des experts locaux sur la base de ce critère. Nous avons donc eu recours à des données qualitatives pour amorcer une explication en usant notamment de l'approche Réaliste (encadré « L'approche réaliste »).

De quelques incompréhensions à l'égard du qualitatif

Lors de l'allocution d'ouverture de la conférence américaine d'évaluation en 1991, son président se plaignait de la place de plus en plus prépondérante des évaluations qualitatives parce que, disait-il, « *elles sont attractives car elles sont si superficiellement faciles* ». Cette vision quelque peu éculée semble malheureusement encore perdurer. Nombre de personnes pensent que l'usage de données qualitatives est facile, et nombreux sont encore les évaluateurs qui, en effet, les utilisent sans grande rigueur. Or, avoir recours à des données qualitatives n'est pas une mince affaire, et celles et ceux qui, préalablement attirés par les analyses quantitatives, ont effectué des analyses qualitatives rigoureuses ont fini par avouer que les statistiques sont parfois plus faciles à mettre en œuvre et souvent plus rapides. L'incompréhension vient évidemment en partie de la perception erronée de certaines personnes quant aux approches qualitatives. Mais la responsabilité est aussi du côté des évaluateurs, qui, parfois, acceptent de faire des dizaines de groupes de discussion alors qu'ils n'auront que quelques heures pour en faire l'analyse et ne pourront jamais rendre compte de la complexité des propos ainsi recueillis. La quantophrénie gagne

aussi les « qualitatistes », qui, face aux demandes de pourcentages et d'échantillonnages représentatifs de la population à l'étude, tombent dans le panneau du nombre. Combien d'évaluations (rapides) sont aussi réalisées en allant poser quelques questions auprès de personnes clefs, prétextant ensuite avoir recours à une évaluation participative et qualitative ? Dans le domaine de la santé publique, les anthropologues ont l'habitude de ces demandes où le qualitatif vient souvent jouer un rôle d'alibi, « l'anthropologue de service » venant étudier les aspects culturels, sans pour autant, donner les moyens à cette étude, ni réellement prendre en compte les résultats mis au jour.

De la rigueur du qualitatif

Or analyser des données qualitatives demande de la rigueur, du temps, de la patience, de la persévérance et de la transparence. Bien que toutes les procédures ne soient pas toujours faciles à expliciter, et qu'un pacte qualitatif, pour paraphraser les termes d'Olivier de Sardan, doit être passé entre l'évaluateur et son client, il existe aujourd'hui une multitude de réflexions sur les critères de scientificité des démarches qualitatives (encadré « Quelques critères de qualité »). Des manuels entiers présentent les différentes manières de choisir des échantillons dans le domaine qualitatif et proposent des démarches pour analyser de manière rigoureuse les données qualitatives. Yin a notamment rendu populaire la stratégie méthodologique des études de cas où les données qualitatives sont largement, mais pas uniquement, utilisées pour évaluer un programme. Les matrices et autres instruments d'organisation des données qualitatives ont été décrites il y a bien longtemps par Huberman et Miles [17]. Il existe sur le marché libre ou commercial un nombre important de logiciels d'organisation (l'analyse reste du ressort de l'évaluateur !) des données qualitatives. En Amérique du

Quelques critères de qualité

- Triangulation
- Réflexivité
- Validation par les enquêtés
- Transparence des méthodes de collecte et d'analyse
- Prendre en compte les cas déviants/témoins
- Imprégnation/observation en contexte naturel
- Itération
- Saturation
- Plausibilité
- Concordance entre observations empiriques et interprétation

Sources : Laperrière, Olivier de Sardan, Pope & May, Huberman & Miles



Nord, les évaluateurs ont la possibilité de trouver des cours spécifiques d'analyse qualitative dans les universités, ce qui manque certainement encore en France.

Le qualitatif et les méthodes mixtes

Nous pourrions demeurer pessimistes car, par exemple, alors que nous avons cru que le débat était clos sur la suprématie des approches expérimentales en évaluation, le ministère fédéral de l'Éducation aux États-Unis propose un retour vers le futur en affirmant en 2003 que « *les devis expérimentaux sont les meilleurs pour déterminer l'efficacité des projets* ». Cela a suscité d'énormes débats outre-Atlantique. Mais au plan international, nous pourrions retrouver le sourire. Le réseau des réseaux sur l'évaluation d'impacts (Nonie) a affirmé en 2007 que « *la norme (gold standard) dans les évaluations d'impacts doit être la pertinence du choix de la méthode* », appuyant ainsi une déclaration de la Société européenne d'évaluation. Autrement dit, que nous ayons recours à des données quantitatives ou qualitatives, l'important est qu'elles soient utiles pour répondre aux questions posées par l'évaluation et ses commanditaires. Nous sommes là au cœur du débat contemporain de l'utilisation des méthodes mixtes dans les évaluations. Dans ces approches, les données qualitatives peuvent ainsi disposer de plus ou moins de prépondérance sur les données quantitatives (tableau 1). Elles peuvent également être employées avant, après ou en même temps que les données quantitatives.

Pour illustrer la place du qualitatif dans les recherches évaluatives, nous avons analysé les écrits scientifiques qui concernent les interventions visant à exempter du paiement des soins les patients en Afrique.

Dans les dix-neuf études scientifiques publiées à ce jour, on remarque que la place des données qualitatives reste minime. Dans douze études, seules les données quantitatives ont été employées et, dans les

sept autres, seulement une a donné une place prépondérante aux données qualitatives. Deux études ont eu recours exclusivement à des données qualitatives. Les composantes qualitatives ont permis, d'une part, de confirmer les résultats des données quantitatives (par ex : augmentation de l'utilisation des services, impact ambigu sur la qualité des soins, modeste protection financière), et, d'autre part, d'étudier les autres barrières à l'accès aux soins et les facteurs influençant la qualité des services. La qualité des composantes qualitatives de ces études est satisfaisante au regard des critères de validité interne énoncés par Laperrière : prise en compte de la subjectivité humaine, observation en contexte naturel et concordance entre observations empiriques et leur interprétation (voir encadré « Quelques critères de qualité », p. 33). Pour chacun de ces critères, différentes stratégies sont proposées par Laperrière. Ainsi, nous avons analysé les sept études en fonction de la stratégie décrite par les auteurs dans leur article pour observer ces critères. Il s'agit donc d'un classement très conservateur puisque les auteurs ont pu utiliser plus de stratégies qu'ils n'en présentent dans leur article. Cependant, nous constatons que deux études disposent des trois critères, deux autres de deux critères et enfin deux d'un seul. Une étude ne dispose d'aucun renseignement permettant de porter un jugement. Mais il faut noter que la méthodologie pour l'analyse des données qualitatives est peu explicitée dans tous les cas.

Conclusion

Les données qualitatives ont évidemment leur place dans le domaine de l'évaluation, mais ni plus ni moins que les données quantitatives. Cependant, elles ne doivent pas être cantonnées à un seul type d'évaluation (e.g. processus, besoin) et doivent être analysées avec rigueur et transparence pour accroître la crédibilité des conclusions et le potentiel d'utilisation des résultats. **F**

tableau 1

Classification et thématique des études sur l'évaluation des exemptions en Afrique

Idéal type QUANT (n = 12)	Méthodes mixtes (MM)			Idéal type QUAL (n = 2)
	QUANT mixte QUANT > QUAL (n = 2)	Idéal type MM QUANT = QUAL (n = 2)	QUAL mixte QUANT < QUAL (n = 1)	
<ul style="list-style-type: none"> ● Utilisation des services ● Qualité des services ● Protection financière ● État de santé ● Stigmatisation du VIH 	<ul style="list-style-type: none"> ● Qualité des services ● Utilisation des services ● Perception du coût du traitement 	<ul style="list-style-type: none"> ● Qualité des services ● Stratégies pour faire face à la maladie 	<ul style="list-style-type: none"> ● Utilisation des services ● Charge de travail du personnel 	<ul style="list-style-type: none"> ● Utilisation des services ● Protection financière ● Viabilité à long terme ● Qualité des services

L'évaluation en santé publique à la Haute Autorité de santé

Au terme d'un processus entamé au début des années 1990 et dont les lois des 9 et 13 août 2004 ont constitué une étape majeure, plusieurs institutions se sont vu reconnaître des compétences dans le champ de l'évaluation en santé publique. C'est le cas de la Haute Autorité de santé (HAS), créée par la loi du 13 août 2004, dont l'une des missions est d'élaborer des recommandations en santé publique afin d'aider les décideurs publics dans leurs choix en matière d'actions et de programmes de santé publique.

Missions de la HAS en matière d'évaluation en santé publique : rappels

Selon la loi n° 2004-810 du 13 août 2004 portant réforme de l'assurance-maladie¹, « au titre de sa mission d'évaluation de la qualité de la prise en charge sanitaire de la population », la HAS est chargée, en vertu du 3^e alinéa de l'article L. 161-40 du Code de la Sécurité sociale (CSS), « d'évaluer la qualité et l'efficacité des actions ou programmes de prévention, notamment d'éducation pour la santé, de diagnostic ou de soins ». Le législateur a précisé les missions de la HAS par la loi n° 2006-1640 du 21 décembre 2006 qui a ajouté à ce même article L. 161-40 du CSS un alinéa 2^o bis lui donnant compétence pour « rendre un avis sur la liste des consultations médicales périodiques de prévention et des examens de dépistage mis en œuvre dans le cadre des programmes de santé visés à l'article L. 1411-6 du Code de la santé publique ».

L'article L. 1411-6 du Code de la santé publique (CSP) résultant de la même loi prévoit ainsi : « Sans préjudice des compétences des départements prévues à l'article L. 2111-2, des programmes de santé destinés à éviter l'apparition, le développement ou l'aggravation de maladies ou incapacités sont déterminés par arrêté du ministre chargé de la Santé et de la Sécurité sociale et, en tant que de besoin, des ministres intéressés.

Dans le cadre de ce programme sont prévus des consultations médicales périodiques de prévention et des examens de dépistage, dont la liste est fixée, après avis de la Haute Autorité de santé, par arrêté du ministre chargé de la Santé, ainsi que des actions d'information et d'éducation pour la santé. »

La loi du 13 août 2004 précise par ailleurs que « dans l'exercice de ses missions, la Haute Autorité tient compte des objectifs pluriannuels de la politique de santé publique [...] ».

Cette mission d'évaluation de la qualité de la prise en charge sanitaire de la population par le système de santé reprend celle qui avait été confiée à l'Agence nationale d'accréditation et d'évaluation en santé (Anaes) par la loi du 4 mars 2002. En effet, selon les termes de l'article L. 1414-3-1 du CSP inséré par la loi du 4 mars 2002 et explicitant l'article L. 791-1 du CSP, « l'Agence nationale [pouvait] également être chargée de l'évaluation d'actions et de programmes de santé publique ».

La HAS et l'évaluation des actions et programmes de santé publique

Dans le cadre des priorités de santé fixées par la loi de santé publique et à la demande des autorités publiques, la HAS exerce sa mission d'évaluation des programmes de santé publique dans une perspective d'aide à la décision publique. Elle peut également s'auto-saisir.

Historiquement, l'Anaes a rempli une partie de cette mission, notamment dans le cadre de l'évaluation a priori des programmes de dépistage, réalisée initialement par le Service d'évaluation des technologies. Ponctuellement, la réalisation d'auditions publiques coordonnées par le Service des recommandations professionnelles dans le domaine d'actions de prévention (traitements hormonaux substitutifs de la ménopause, vaccination contre l'hépatite B) participait également de cette mission.

Champs et objets de l'évaluation

Dans le cadre de sa mission d'évaluation en santé publique, la HAS intervient le plus souvent en amont de la mise en place d'interventions de niveaux de complexité variés (actions ponctuelles, programmes voire politiques de santé publique²) afin d'en déterminer la pertinence à partir d'une analyse systématique de l'évidence scientifique. L'évaluation de la HAS peut également porter sur l'intérêt d'une évolution des modalités d'un programme de santé et de son impact sur le système de santé.

Jusqu'à présent, l'investissement de la HAS dans ce domaine a concerné à titre principal le champ des programmes de dépistage (dans la suite des travaux d'évaluation des technologies effectués par l'Anaes). Douze recommandations en santé publique dans ce domaine ont

2. Selon l'Andem, une action de santé publique est définie comme une liste d'opérations à entreprendre, de caractère collectif, en vue de l'amélioration de la santé d'une population. Un programme de santé correspond à un ensemble coordonné d'actions. Une politique de santé peut être définie comme l'ensemble des orientations et des décisions prises par les pouvoirs d'un pays pour le maintien et/ou l'amélioration de l'état de santé de la population.

Olivier Scemama

Adjoint au chef du Service évaluation économique et santé publique

Catherine Rumeau-Pichon

Adjointe au directeur de l'Évaluation médicale, économique et de santé publique, Haute Autorité de santé

1. Modifiée par la loi n° 2006-1640 du 21 décembre 2006 (articles L. 161-37 et L. 161-40 du Code de la Sécurité sociale).



été publiées entre janvier 2005 et novembre 2009 ; huit autres sont en cours d'élaboration. Les problématiques de santé abordées sont très diverses : les pathologies cancéreuses arrivent au 1^{er} rang (colon-rectum, sein, prostate, peau, col de l'utérus), puis les pathologies infectieuses sexuellement transmissibles (VIH, syphilis, gonocoque), les pathologies maternelles ou périnatales, les maladies génétiques (trisomie 21, mucoviscidose). Par ailleurs, dans la quasi-totalité des cas, il s'est agi d'évaluations *a priori*, qui s'inscrivent cependant dans des contextes différents :

- avant la mise en place d'un dépistage dont il s'agit d'évaluer l'opportunité. La démarche méthodologique consiste à analyser cette stratégie de dépistage à l'aune d'un certain nombre de critères d'évaluation, définis par l'Organisation mondiale de la santé (Wilson et Jungner) et actualisés dans un guide méthodologique de l'Anaes en 2004 [2] ;

- dans un contexte d'évolution épidémiologique et/ou technologique, ou à la lumière d'une publication scientifique nouvelle susceptible de modifier un dépistage en cours. L'évaluation porte alors sur la confirmation de la pertinence et les modifications des conditions de réalisation du dépistage.

Le champ de la prévention primaire a été abordé plus récemment (ex. efficacité et efficience des stratégies thérapeutiques d'aide au sevrage tabagique, évaluation des stratégies de prévention de la carie dentaire en cours).

Enfin, les travaux d'évaluation peuvent porter également sur des modes d'organisation dont l'impact sur le système de santé doit être apprécié (réseaux de santé, télémédecine, recours à l'hôpital, coopérations interprofessionnelles, etc.).

Méthodes et processus d'élaboration

L'évaluation des actions et programmes de santé publique apprécie le rapport bénéfices/risques des différentes interventions possibles à l'échelle de la population et intègre de façon systématique une évaluation économique et organisationnelle. Elle aboutit à l'élaboration de recommandations en santé publique.

Pour réaliser ces évaluations, la HAS recourt principalement à des méthodes reposant sur l'analyse critique et systématique de la littérature scientifique. Outre la revue de littérature, des travaux de modélisation médico-économiques spécifiques sont réalisés en tant que de besoin. La HAS mobilise aussi d'autres sources d'information, comme le recueil de l'avis d'experts, réunis dans un groupe de travail et un groupe de lecture, tous deux pluridisciplinaires (médecine, économie, sociologie, etc.), pluriprofessionnels, avec la participation de représentants d'associations d'usagers ou de patients.

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

Quelques exemples d'évaluations en santé publique à la HAS

Évaluations publiées et disponibles sur le site de la HAS :

- Dépistage et diagnostic précoce du glaucome : problématique et perspectives en France (novembre 2006)
- Évaluation des stratégies de dépistage de la trisomie 21 (juin 2007)
- Place des tests immunologiques de recherche de sang occulte dans les selles (IFOBT) dans le programme de dépistage organisé du cancer colorectal en France (décembre 2008)
- Surveillance sérologique et prévention de la toxoplasmose et de la rubéole au cours de la grossesse (décembre 2009)

Évaluations en cours :

- Évaluation des stratégies de prévention de la carie dentaire
- Évaluation *a priori* de l'extension du dépistage néonatal d'une ou plusieurs erreurs innées du métabolisme par la technique de spectrométrie de masse en population générale en France
- Stratégies de dépistage du cancer du col de l'utérus en France

Au sein de la HAS, cette mission d'évaluation dans le domaine du dépistage, de la prévention et de l'organisation du système de santé est assurée par le Service évaluation économique et santé publique (SEESP). Les travaux du service sont présentés à la Commission d'évaluation économique et de santé publique (CEESP), puis validés par le Collège.

Conclusion : une nécessaire complémentarité

Les travaux d'évaluation en santé publique menés par la HAS s'inscrivent dans le cadre des priorités de santé définies par la loi de santé publique. Le rôle de la HAS en matière d'aide à la décision publique dans ce domaine ne peut se concevoir qu'en relation avec les autres institutions impliquées dans le champ de l'évaluation des actions et programmes de santé, et de la détermination des priorités de santé publique.

L'évaluation de l'impact attendu des actions de santé publique devrait constituer le principal élément d'aide à la priorisation des objectifs de santé publique, dans un souci d'optimisation des ressources disponibles. **F**

L'évaluation des politiques de santé et la Cour des comptes

La Cour des comptes tient depuis 2008 de la Constitution une mission d'évaluation des politiques publiques. Deux enquêtes sur la mise en œuvre de plans de santé publique montrent la complémentarité entre ses contrôles de performance et une évaluation. La réforme des juridictions financières qui est en préparation devra être accompagnée des moyens nécessaires à la conduite d'évaluations.

La Cour des comptes pratique de longue date le contrôle de performance. Il est proche de l'évaluation *a posteriori* des politiques publiques au sens du décret du 22 janvier 1990 (« *Évaluer une politique, c'est rechercher si les moyens juridiques, administratifs ou financiers mis en œuvre permettent de produire les effets attendus de cette politique et d'atteindre les objectifs qui lui sont fixés* »). Ses rapports évoquent souvent l'évaluation en santé publique, mais c'est principalement au regard de préoccupations quantitatives, et plus rarement pour souligner l'absence d'évaluation qualitative¹.

Distance et empathie

Une nouvelle étape est annoncée : suite à la promulgation en 2008 de l'article 47-2 de la Constitution² ; l'évaluation elle-même devient le quatrième métier des magistrats de la rue Cambon³, après le métier historique de *jugement des comptes*, progressivement accompagné du *contrôle de la gestion et de la performance*, puis de la *certification des comptes* de l'État et de la Sécurité sociale. Ils n'ont jusqu'à maintenant pratiqué l'évaluation que de manière empirique et ponctuelle, alors que la définition de l'évaluation a beaucoup plus évolué depuis 1990 que les méthodes de contrôle de la performance par la cour.

1. Exemple récent : le *Rapport annuel sur la Sécurité sociale* soulignait en septembre 2009, à propos de la T2A, « *l'absence d'une évaluation indépendante et experte des effets de la réforme à laquelle il semble absolument nécessaire de remédier* » (page 211). Puis, page 250, « *il n'y a pas eu d'évaluation médico-économique des examens de santé* » (alors que la cour l'avait préconisée dix ans plus tôt).

2. « *La Cour des comptes assiste le Parlement dans le contrôle de l'action du gouvernement. Elle assiste le Parlement et le gouvernement dans le contrôle de l'exécution des lois de finances et de l'application des lois de financement de la sécurité sociale ainsi que dans l'évaluation des politiques publiques. Par ses rapports publics, elle contribue à l'information des citoyens [...]* ».

3. Danièle Lamarque, conseillère maître, a analysé l'évaluation pratiquée par les juridictions financières au regard de la charte de la Société française de l'évaluation (pluralité, distanciation, compétence, respect des personnes, transparence, opportunité, responsabilité...), dans : « *L'évaluation de politiques publiques, le quatrième métier de la Cour des comptes : le rapport sur la formation professionnelle tout au long de la vie* », Paris, *Gestion finances publiques*, n° 7, juillet 2009, p. 562-566.

La posture du magistrat financier demeure ainsi la distance, certes palliée par le dialogue qu'appelle la procédure contradictoire, alors que l'évaluation est réputée comme fondée sur l'empathie, ou à tout le moins un accord sur l'existence d'un problème⁴.

Deux exemples récents de travaux menés par la 6^e chambre seront évoqués ici, s'agissant du Programme national nutrition santé 2001-2005 (PNNS)⁵ et du Plan cancer 2003-2007 [7]⁶.

Le PNNS

La Cour avait de longue date recommandé à l'État d'assortir les dépenses de santé d'une évaluation. En procédant au contrôle de performance⁷ du PNNS 2002-2005, programme budgétairement modeste, elle s'y est essayée en 2006. Avec des moyens très limités, la tentative demeurerait implicite : trois mois d'un magistrat et six semaines au total de deux médecins, l'une généraliste, consultante rompue aux travaux de la cour, et l'autre, diététicienne. Aux côtés de l'examen traditionnel de la gestion des moyens, y compris avec un examen critique de la quantification des produits et des résultats affichés, ces expertes ont surtout évalué si les acteurs du PNNS avaient développé des indicateurs réalistes et des évaluations en termes d'efficience et d'efficacité, sans procéder directement à une évaluation.

Il était, comme quasiment toujours en ces domaines, impossible d'apprécier l'impact sanitaire du programme dans l'année suivant son achèvement. L'éventuel impact du PNNS I (et celui des constats et recommandations de la Cour) pourra être mieux apprécié lors d'une évaluation de plein exercice du PNNS II, 2006-2010.

Le Plan cancer [7]

Cette seconde enquête a comporté le traditionnel contrôle organique (l'Institut national du cancer, auquel un chapitre du rapport public annuel 2008 a ensuite été consacré),

4. Cf. les travaux de la Société française d'évaluation, et de Sylvie Trosa, rapporteur à la Cour des comptes.

5. Non publiée, hormis sur un point : La coordination interministérielle du PNNS, Cour des comptes, *Rapport public annuel*, février 2009, tome II, p. 55-59. La cour y a noté, entre autres, qu'elle avait sans succès « *souligné l'utilité d'une évaluation périodique de la diffusion et de l'impact des travaux émanant du Service d'information du gouvernement : la DGS [direction générale de la Santé] n'avait pas été destinataire d'un « baromètre » de l'opinion concernant le PNNS, alors qu'il fournissait des indicateurs utiles pour affiner le pilotage de ce dernier* », signe des obstacles que rencontrent même les évaluations les plus simples.

6. Le mot « évaluation » figure dans près d'une page sur trois, mais dans des acceptions multiples.

7. Assorti d'un contrôle de régularité des dépenses, sans donc s'éloigner de la vocation première de la cour.

Alain Gillette

Conseiller maître à la Cour des comptes
Doyen de la 6^e chambre



et le contrôle de performance du Plan cancer, avec le rapport public thématique précité. Le contrôle de performance s'est efforcé de reconstruire aussi complètement que possible la chaîne des objectifs, moyens et produits de cette politique publique. Le contrôle a aussi porté sur l'efficacité et l'efficacités de la gestion du plan au regard des objectifs initiaux. Faute de pouvoir mettre en œuvre des méthodes qualitatives, il n'a pas inclus une évaluation de l'efficacité du plan en termes d'impact médical et scientifique, que la cour n'était pas habilitée à réaliser. Cette dernière mission a d'ailleurs été confiée peu après par l'État au Haut Conseil de la santé publique, avec lequel la cour s'est coordonnée.

Le rapport final du HCSP s'est appuyé à maintes reprises sur l'analyse traditionnelle des défaillances et des incohérences effectuée par la cour. Elle a montré les insuffisances des outils statistiques et des données, mais sans avoir les moyens ni les données permettant d'évaluer les résultats et d'établir des analyses socio-économiques⁸ ou des calculs financiers quant à l'impact macro-économique des mesures du plan. Cette dernière

8. L'impasse a également été faite sur une approche du type « théorie d'action », qui rechercherait les relations de cause à effet entre les intentions implicites des auteurs du plan, les mesures choisies sur leur proposition et l'impact réel de ces dernières. Les magistrats de la cour ont l'habitude de rechercher des causalités, mais ils se limitent jusqu'à maintenant à ce qui est explicitement documenté dans les dossiers à l'appui de leurs rapports.

L'évaluation du Plan cancer par le HCSP

Comme indiqué dans ce numéro, le HCSP a réalisé deux rapports d'évaluation sur le même plan cancer. Le premier a été consacré principalement à la prévention et au dépistage (116 pages)¹, et le second, à une évaluation complète (495 pages)². La seule indication des moyens mobilisés montre la différence de nature des travaux : les deux auteurs du premier rapport ont bénéficié du concours d'une dizaine de spécialistes, et d'une relecture par des membres du Haut Conseil. Les auteurs du second ont souligné que « *L'évaluation du Plan cancer représente, au vu de l'ampleur des interventions, une métaévaluation. Elle a nécessité la conduite d'importants travaux d'évaluation par des contributeurs et évaluateurs issus du HCSP ou externes à celui-ci, répartis en dix axes d'évaluation. Une collaboration avec l'IgAs a pu être mise en place pour la prise en charge de deux axes de l'évaluation. Une vingtaine d'experts externes et de consultants ont été mobilisés* » (page 26). Une société de consultants a contribué à l'élaboration de la méthode. Un comité d'évaluation, un comité de projet, et dix équipes d'évaluation de 2 à 5 experts ont été constitués. Les deux rapports témoignent de la richesse des analyses ainsi obtenues, et aussi du fait qu'elles se sont heurtées aux mêmes insuffisances de données que la cour avait relevées. Le HCSP n'a pas eu les moyens d'évaluer le volet « recherche » du Plan cancer. **F**

1. *Rapport d'étape sur l'évaluation du Plan cancer, prévention et dépistage*, 6 juin 2008. L'Office parlementaire d'évaluation des politiques de santé avait aussi abordé ce dernier thème dans : « Les résultats du dépistage du cancer du sein », Assemblée nationale, n° 1678, 17 juin 2004, 83 pages, dont la cour a noté que plusieurs recommandations sont restées sans suites.
2. *L'évaluation du Plan cancer*, 30 janvier 2009.

approche serait logique dans le cadre d'évaluations confiées à la cour.

La cour a aussi en quelque sorte procédé à une évaluation de l'évaluation : absence de mise en place de celle-ci tant par la mission interministérielle qui en avait été chargée que, jusqu'à la parution dudit rapport public, par l'Institut national du cancer ou avant lui le groupement d'intérêt public (GIP) chargé de sa préfiguration. Cette absence d'évaluation par les organes qui en avaient été chargés ne résultait pas seulement de dysfonctionnements mais aussi de la grande difficulté à rassembler les connaissances appropriées⁹. La plupart des indicateurs prévus par le plan n'étaient à l'époque qu'incomplètement renseignés - alors que la cour avait déjà, mais en vain, souligné en 2000 puis en 2003 la difficulté d'évaluer, faute de données, les résultats de la prise en charge des cancers [8]. L'INCa s'est mobilisé pour répondre aux questions du rapporteur, et a ensuite mieux organisé la mission qui lui est assignée en ce domaine.

Toutefois, en cancérologie comme dans d'autres domaines, un délai de plusieurs années, voire de plusieurs décennies, est un préalable à toute évaluation de ce type. La cour ne pouvait donc que marginalement rendre compte de l'impact du plan en termes d'efficacité, et de causes avérées de cette efficacité. Les indicateurs intermédiaires, notamment sur les changements de comportement des professionnels et des usagers (vis-à-vis de l'alcool, du tabac, du dépistage...), ont été cités, mais eux-mêmes ne sont connus qu'après un certain délai.

Restructuration

Ces deux exemples montrent que la cour va devoir significativement faire évoluer ses moyens et ses méthodes pour assurer la mission qui lui est désormais assignée en matière d'évaluation par la Constitution. Aussi une évolution structurelle est-elle en préparation¹⁰. L'un des considérants est que la répartition des compétences entre la cour et les chambres régionales des comptes, qui travaillent efficacement ensemble notamment en matière hospitalière, ne permettrait pas d'organiser efficacement et de façon homogène sur le terrain l'évaluation des politiques publiques. C'est l'un des motifs du projet de loi transmis fin 2009 au Parlement en vue d'une fusion des juridictions financières en une seule entité, dotée d'un réseau restreint, la cour. Philippe Séguin y était très attaché.

9. L'INCa notait en 2005 que les études scientifiques en matière de soins cancérologiques ne faisaient « l'objet ni d'orientations coordonnées, ni de mise en commun, ce qui interdit de fait une connaissance véritable de la qualité des soins actuellement dispensés aux patients », cf. Cour des comptes, *La mise en œuvre du Plan cancer*, pages 36, 41-44 et 54.

10. Cette réforme est élaborée au moment où l'avènement des agences régionales de santé restructure l'action publique en ce domaine, ce qui pourrait à terme améliorer la production des données nécessaires à toute évaluation.

Suite à une décision rendue par le Conseil constitutionnel le 25 juin 2009, il est également prévu que la loi détermine les organes du Parlement compétents pour demander l'assistance de la cour en matière d'évaluation des politiques publiques. Elle habiliterait le gouvernement à fixer par ordonnance les procédures de la cour en ce domaine (les modalités traditionnelles de délibéré, notamment, ne correspondent pas aux canons des groupes pluridisciplinaires d'évaluation).

Il lui faudra ensuite trouver les moyens appropriés. L'exemple cité plus haut du HCSP en souligne l'ampleur. Une évaluation à le mérite d'apporter des ressources externes et temporaires, mais leur coordination et l'appropriation de leurs apports par les magistrats exigeront des moyens, ne serait-ce que pour coordonner des équipes d'experts et en valider les apports. Le projet parallèle d'une réduction des effectifs des juridictions financières pourrait à cet égard constituer un frein ; toutefois, le

recrutement de magistrats en service extraordinaire pour une durée limitée et d'assistants, venus de tous horizons, est également prévu ; il serait de nature à répondre aux exigences de pluralité et de compétence qui sont celles de l'évaluation.

À l'instar d'autres institutions supérieures de contrôle, la cour ne pouvait sans doute pas rester à l'écart du développement de l'évaluation des politiques publiques. Néanmoins, tant l'adjonction du rôle d'évaluateur à sa culture ancestrale de juge des comptes que la reconnaissance sur le terrain de cette extension de sa légitimité ne seront sans doute pas instantanées. Bref, comme il a été écrit à propos d'évaluations de politiques régionales de santé par d'autres acteurs, « l'organisation des relations entre les différents groupes et acteurs, la répartition des rôles et des responsabilités sont encore à clarifier » [18]. **F**

Extrait du discours prononcé par Philippe Séguin, président de la Cour des comptes, lors de la journée de clôture des travaux du Haut Conseil de la santé publique le 18 décembre 2010

« **J**e suis ici en effet pour évoquer, devant vous, les perspectives de collaboration qui nous sont ouvertes par la reconnaissance de l'importance de l'acte d'évaluation, inscrite, vous le savez, depuis juillet 2008 dans la Constitution.

Cette consécration constitutionnelle de l'évaluation des politiques publiques est, à n'en pas douter, une décision majeure. Elle résulte de la prise de conscience de l'exigence accrue d'efficacité et d'efficience des politiques publiques dans une économie mondialisée. Elle traduit également l'attente de nos concitoyens de toujours plus de transparence et d'une meilleure information sur les résultats réels des actions publiques engagées. Une information délivrée par des organismes qui présentent toutes les garanties d'indépendance...

La Cour a donc acquis une expertise reconnue dans l'analyse de grandes politiques de santé publique, et en particulier, mais c'est évidemment, dans notre « cœur de métier », sur tout ce qui porte sur l'organisation administrative et le financement de ces politiques.

Nous bénéficions pour ce faire d'un atout majeur : un champ de compétences très étendu, qui nous permet de suivre l'utilisation de l'ensemble des deniers publics, qu'il s'agisse des dépenses budgétaires de l'État, de celles des collectivités territoriales ou de celles de la sécurité sociale. Nous avons donc une vision globale du finance-

ment des politiques, notamment de santé publique, globalité sur laquelle vous avez pu, de votre côté, parfois atchopper, comme dans le cas de l'évaluation du Plan national sur les maladies rares, dans lequel vous n'avez pu examiner que les seules dépenses inscrites au budget de l'État...

Je voudrais plus précisément prendre quelques exemples rapides, afin de vous montrer notre intérêt pour l'évaluation des politiques de santé publique, mais surtout le caractère complémentaire, selon moi, de nos travaux et de nos analyses.

En 2006, nous avons, dans le cadre d'une enquête sur la politique de périnatalité dans notre pays, examiné les plans successifs qui ont été lancés par le Gouvernement pour améliorer la prise en charge sanitaire des nouveau-nés et des parturientes... Nous avons montré que la mise en œuvre de ces orientations s'est révélée insuffisante et qu'elle a été particulièrement lente dans le secteur public, du fait de fortes rigidités. Nous avons également souligné que les performances sanitaires globales de la France se sont améliorées, mais que de très fortes inégalités perdurent entre les régions, mais aussi entre les catégories sociales face aux risques liés, pour les futures mères, à la grossesse. Mais au-delà de la constatation basée sur des données disparates, les enquêtes de la Cour s'arrêtent aux frontières de l'évaluation de l'im-

pact sanitaire des politiques publiques, ou de la détermination des bonnes pratiques médicales. Et c'est ici que les analyses des professionnels de santé et des spécialistes en santé publique peuvent utilement prendre le relais, et d'ailleurs nous n'hésitons pas à faire reposer nos constatations et nos recommandations sur leurs conclusions... Reste que dans ces matières, votre intervention est irremplaçable car vous avez su mobiliser des compétences médicales et en santé publique reconnues.

Il nous faut désormais aller plus loin dans la convergence de nos travaux, car le Parlement ne se contentera pas de la juxtaposition des travaux du Haut Conseil et de la Cour. Nos concitoyens, quant à eux, comment peuvent-ils comprendre qu'une synthèse des démarches ne soit pas réalisée par nos soins sur un même sujet ? Il nous sera demandé à l'avenir, à vous comme à nous, d'intégrer l'ensemble des approches, des disciplines, des problématiques dans une démarche d'évaluation globale et unique.

Cela signifie qu'il nous faut sans doute bâtir une expertise collective, au sein de laquelle chaque instance aura sa place, chaque discipline aura sa contribution à apporter et disposera de la même légitimité. Cela signifie qu'il nous faut rapidement apprendre à travailler en commun, sur les méthodes, et demain, sur des sujets d'intérêt partagé. **F**