



Société Française de l'Évaluation

LES SYNTHÈSES DE LA SFE

Promouvoir l'utilité et l'utilisation
de l'évaluation dans le secteur de
la santé

Par le groupe SFE « Evaluation en Santé »

Progressivement ces dernières années, l'exigence d'évaluation s'est intensifiée dans les différents secteurs de politiques publiques, tant pour rendre compte que pour rationaliser l'action publique, voire pour apporter la solution à de multiples problèmes. Le discours affiché en faveur de l'évaluation est parfois accompagné de clauses évaluatives et de procédures pour promouvoir et encadrer la démarche. Cependant, en dépit de cette rhétorique sur l'évaluation, on déplore une faible inscription de l'évaluation dans les mécanismes institutionnels, organisationnels et décisionnels. Pourquoi engage-t-on des travaux d'évaluation ? Quel est le devenir de ces travaux ? Parviennent-ils à alimenter la décision ? En d'autres termes, l'évaluation est-elle utile ?

En mars 2015, le groupe Santé organisait une journée nationale pour débattre de ces questions. Les trois sessions de cette journée ont successivement abordé : l'utilisation des travaux d'évaluation à partir d'évaluations conduites aux niveaux national (évaluation du plan VIH-IST Outremer/encadré 1), régional (évaluation de la politique régionale de santé des personnes en grande vulnérabilité/encadré 2) et local (évaluation des contrats locaux de santé sur les inégalités sociales et territoriales de santé/encadré 3) ; les avancées méthodologiques pour l'évaluation des interventions complexes, illustrées de quelques approches novatrices (encadré 4) ; la place de l'évaluation dans les institutions.

Ce document présente la synthèse des réflexions issues des débats de cette journée, notamment à partir des trois retours d'expériences.

Quelles sont les formes d'utilisation possible de l'évaluation ?

L'exploitation d'une évaluation à travers la prise en compte des conclusions et recommandations du rapport ne représente qu'un aspect de l'utilisation (révision du programme, de sa gouvernance, identification de nouvelles priorités, par exemple). D'autres formes d'utilisation sont identifiées. L'évaluation peut : éclairer le jugement en améliorant la compréhension de l'intervention (politique/programme), de ses effets et de ses marges d'amélioration ; générer des apprentissages du fait de la participation à l'évaluation et susciter une transformation de la conception et de la reformulation du problème ; soutenir, légitimer ou imposer des positions.

Quels sont les ingrédients nécessaires pour faire de l'évaluation, une démarche utile ?

L'analyse de la littérature nous enseigne que plusieurs facteurs interviennent en faveur (ou pas) de l'utilisation de l'évaluation. Certains tiennent à l'évaluation proprement dite, d'autres relèvent du contexte institutionnel et organisationnel dans lequel elle se déroule ou de l'engagement des acteurs qui y participent.

♦ Les facteurs liés à l'évaluation

La commande d'évaluation. La clarté et la précision de la commande sont un gage de bon déroulement de l'évaluation. S'il est nécessaire que les attentes soient clairement formulées, consensuelles entre les parties prenantes, l'utilisation des résultats des travaux d'évaluation doit être anticipée : révision de l'intervention dans le cas d'une évaluation à mi-parcours, décision de poursuite ou non renouvellement dans le cas d'une évaluation finale. Les évaluations du plan VIH-IST Outremer et des contrats locaux de santé en Ile de France en témoignent avec, dans le premier cas, des mesures relatives à la poursuite du plan et, dans le second, l'engagement d'un travail pour la restructuration des CLS. Ces deux exemples illustrent par ailleurs que l'importance accordée à l'évaluation est en général proportionnelle à celle accordée à l'intervention évaluée.

La conduite de l'évaluation. Le caractère collégial de la conduite de l'évaluation est d'autant plus nécessaire que l'intervention évaluée concerne un ensemble d'institutions, ce qui est généralement le cas, s'agissant des politiques de santé : accord sur les objectifs et questions de l'évaluation, pilotage et validation en commun des travaux. Dans les trois expériences, un dispositif d'évaluation en place intégrait les différentes parties prenantes et une concertation continue a été entretenue avec les partenaires.

La communication. Elément essentiel de l'appropriation des résultats, elle relève de deux mécanismes distincts : la mise en place d'un processus continu de partage d'information avec les acteurs, que ces derniers participent ou non à la démarche, et l'organisation de la diffusion des résultats du rapport, d'autant plus efficace qu'il crée un espace de débat avec les parties prenantes et instances de démocratie de concertation. Dans la région Pays de la Loire, des livrets d'évaluation sont régulièrement diffusés à l'intention d'un large public.

La méthode. Bien que relevant principalement de la responsabilité de l'évaluateur, certains aspects nécessitent des choix qui engagent l'institution commanditaire. Partager et débattre des choix méthodologiques accroît la crédibilité de l'évaluation, donne les clés d'une lecture critique des résultats et contribue à renforcer la culture d'évaluation des parties prenantes.

♦ Les facteurs contextuels

Le contexte dans lequel se déroule l'évaluation joue une influence considérable.

L'engagement des décideurs. Il est indispensable autant pour le lancement et le devenir d'une évaluation singulière que pour la reconnaissance de la fonction évaluative dans l'institution.

L'organisation de l'évaluation dans l'institution. L'existence d'un service et d'une équipe dédiée et formée à l'évaluation donne de la visibilité à la fonction. C'est un élément indispensable pour réaliser des évaluations mais aussi, créer des référentiels, les partager avec les autres services et ainsi, renforcer les pratiques et la culture d'évaluation au sein de l'institution. L'ARS de la région Pays de la Loire a mis en place une programmation des activités d'évaluation et une équipe dédiée qui réalise les travaux et développe la culture évaluative dans l'institution.

Ces points évoqués font écho aux sept principes de la charte de la Société française de l'évaluation : pluralité, opportunité, responsabilité, distanciation, compétence, transparence, respect des personnes.

Trois leviers pour améliorer l'utilité de l'évaluation

L'activité évaluative est largement dépendante de (1) la motivation des acteurs à s'y impliquer, (2) des compétences disponibles, et (3) de sa reconnaissance au sein de l'institution.

Le contexte a fortement évolué ces dernières années et l'évaluation est désormais une **activité partagée** entre les acteurs institutionnels et les organismes externes (bureaux d'études ou laboratoires universitaires). Cependant, pour améliorer l'utilité de l'évaluation, il est nécessaire de **renforcer les capacités** en matière d'évaluation dans le secteur de la santé, avec un **ancrage plus solide de l'évaluation dans les institutions**, un soutien et un accompagnement pédagogique adapté et en favorisant la création de **communautés de pratiques**. Enfin, un **travail collaboratif** impliquant largement les parties prenantes, est souhaité pour que l'évaluation réalise pleinement ses fonctions.

Encadré 1. L'évaluation du plan VIH-IST Outremer (2012)

Thierry Troussier (DGS), Carine Chaix-Couturier (EQR Conseil), Hervé Creusvaux (DGOM), Bernard Faliu (DGS)

Contexte. Le plan VIH-IST est le premier plan spécifiquement adapté au contexte Outremer. Les directions des deux ministères (Outremer et Santé) en charge de ce plan ont décidé conjointement d'une évaluation intermédiaire pour procéder à des réajustements de la déclinaison de ce plan.

Méthode. La commande était précise, sur la base d'un cahier des charges co-rédigé par les deux directions. Cinq monographies ont été réalisées (une par région), relues et amendée par l'ARS concernée. Des entretiens ont été conduits auprès d'une trentaine d'acteurs institutionnels au niveau national et une centaine au niveau local. Il n'y a pas eu de rencontres avec les bénéficiaires finaux.

Processus. Le comité de pilotage était pluri institutionnel (plusieurs ministères, INPES, ARS, COREVIH). Les deux directions ont assuré le suivi de cette évaluation avec le cabinet EQR. Il y a eu une concertation étroite durant tout le processus et une restitution des résultats sur l'ensemble des DOM dans trois lieux. Le rapport a fait l'objet d'une communication immédiate (4 pages) et a été mis en ligne 18 mois plus tard. L'évaluation a par ailleurs donné lieu à des échanges de bonnes pratiques entre les régions.

Résultats. Il y a eu un bilan des réalisations, lesquelles étaient inégales selon les régions et plusieurs préconisations.

Utilisation. L'évaluation a permis une meilleure coordination des actions, une meilleure collaboration avec les institutions, la valorisation et la diffusion d'expériences originales, le financement de projets ainsi légitimés. Elle a servi de base pour le futur plan régional. Le Haut conseil de santé publique a bénéficié de ces apports pour l'évaluation globale du plan et la formulation de recommandations sur les stratégies en santé sexuelle. L'intérêt d'une évaluation intermédiaire résidait dans sa capacité à adapter le plan aux spécificités des départements Outremer et à préparer l'évaluation finale.

Questions en débat. Quel est le niveau le plus pertinent (national versus régional) pour l'évaluation d'un plan ? Quel suivi mettre en place à distance de l'évaluation ? Quelle aide à l'évaluation finale ?

Encadré 2. L'évaluation de la politique régionale de santé des populations en grande vulnérabilité par l'ARS Pays de Loire (2013)

Jean-Pierre Boulé, Xavier Brun, Catherine Ogé, ARS Pays de la Loire

Contexte. Il s'agit de la première évaluation intermédiaire du Projet régional de santé. Les évaluations intermédiaires portent sur les populations prioritaires de la politique régionale. La direction générale de l'ARS a mis en place une organisation et une équipe dédiées pour mener ces évaluations.

Méthode. L'évaluation vise à apporter un éclairage sur la politique en cours de déploiement, proposer des recommandations d'ajustement et alimenter le débat public. Elle repose sur une approche participative permettant de croiser une pluralité de points de vue et d'investigations (entretiens individuels, entretiens de groupe, méthode Delphi, enquêtes par questionnaire etc.).

Processus. Un directeur métier de l'ARS préside un comité de pilotage articulé avec les instances de démocratie sanitaire (CRSA) et associant des usagers. Il contribue à la définition des critères d'évaluation (référentiel), à la formulation des jugements et à la production de recommandations. Les recommandations sont élaborées à partir du croisement des travaux du comité de pilotage, des points de vue d'experts et du recueil de la parole des usagers.

Résultats. Le rapport final de l'évaluation comporte à la fois des recommandations stratégiques et opérationnelles. La CRSA émet un avis sur l'évaluation, lequel est repris dans le corps du rapport.

Utilisation. . Les résultats sont communiqués aux acteurs concernés, l'évaluation faisant aussi l'objet d'un document de synthèse (livret). Les feuilles de route annuelles de l'ARS reprennent, suivant les décisions du comité de direction de l'ARS, les éléments de l'évaluation figurant dans une note stratégique élaborée par l'équipe d'évaluation.

Questions en débat. Quelles sont les limites du rôle d'une équipe d'évaluation interne à l'institution quant à la hiérarchie des recommandations, à la communication sur l'évaluation et au suivi des recommandations ?

Encadré 3. L'évaluation de l'impact des contrats locaux de santé sur les inégalités sociales et territoriales de santé en Ile-de-France (2013-2014)

Christelle Berthon, Luc Ginot, ARS Ile de France

Contexte. Fortement engagée dans le déploiement des contrats locaux de santé (60 CLS signés), l'ARS de la région Ile-de-France a lancé en 2013, l'évaluation de ce dispositif afin d'en améliorer la qualité et d'apprécier son impact sur la réduction des inégalités sociales et territoriales de santé (ISTS).

Méthode. L'évaluation a été réalisée en deux temps : (1) analyse de la structuration des 18 CLS (revue de littérature et analyse documentaire) ; (2) analyse des dynamiques des CLS pouvant influencer les ISTS (39 entretiens conduits dans 5 sites).

Processus. Le comité de pilotage associait l'ARS et l'équipe universitaire. Il y a eu une collaboration étroite avec les chargés de mission et référents CLS au niveau local. Le rapport a été diffusé et les résultats ont été communiqués lors de colloques et par le biais d'articles.

Résultats. L'évaluation a montré une meilleure prise de conscience de l'enjeu de réduction des ISTS, le développement des missions d'animation territoriale de l'ARS et une politique plus concertée. En revanche, le niveau de priorisation de l'action sur le thème reste limité, des difficultés à mettre en œuvre une démarche globale de santé demeure et la participation des habitants reste faible. Il convient de faire de la lutte contre les ISTS un objectif plus explicite et d'adopter des stratégies adaptées.

Utilisations. L'évaluation a permis la révision des CLS de 2ème génération, la réorganisation du dispositif de pilotage régional et des modes d'organisation dans les délégations territoriales. Elle a contribué à l'élaboration d'un référentiel d'autoévaluation des CLS à destination des collectivités. Les acquis capitalisés sont ainsi exploitables pour d'autres projets.

Questions en débat. Des sujets, tels que la participation des habitants, l'action intersectorielle et la qualité des partenariats engagés, restent à approfondir. Si l'exploitation des résultats de l'évaluation reste relativement aisée au niveau régional, les difficultés actuelles sont liées à ce jour, à l'avenir incertain du dispositif CLS.

Encadré 4. Les enjeux méthodologiques de l'évaluation des interventions complexes

(D'après l'intervention du Pr. François ALLA, professeur à l'Université de Lorraine)

Depuis le début des années 2000, on assiste à une prise de conscience de la complexité croissante des interventions dans le domaine de la santé ; l'évaluation s'y adapte en développant des méthodes qui prennent en considération cette complexité. Face à la complexité définie par Edgar Morin comme « le tissu d'évènements, actions, interactions, rétroactions, déterminations, aléas qui constituent notre monde phénoménal », il s'agit de comprendre le tout d'une action (ici de santé publique) mais également chacune de ses composantes et d'identifier comment l'intervention se conduit dans un réel qui est complexe.

En santé, l'évaluation est longtemps restée régie par un modèle, celui du médicament, où l'on compare un produit à un placebo ou bien un produit A à un produit B. En regard, la complexification apparaît désormais dans les objets évalués (par exemple une organisation de soins aux acteurs multiples en interaction mutuelle constante) comme dans l'obligation de la prise en compte du réel. Dans le modèle classique de la recherche clinique, on évalue une intervention définie à l'avance en prenant garde que « les choses ne bougent pas » tandis qu'une intervention de santé publique est, par nature, en évolution continue et en adaptation au contexte, lequel est considéré comme un élément faisant partie de l'intervention). Dès lors, une intervention dont l'efficacité est démontrée dans un contexte donné peut être applicable dans un autre contexte sans y être pour autant efficace : la possibilité d'une transférabilité n'est pas acquise. Ceci bouleverse le paradigme méthodologique dérivé de la recherche clinique qui ferait de l'essai contrôlé randomisé (ECR) individuel en double aveugle, le modèle de référence pour l'évaluation : avec l'ECR, l'évaluateur tente d'ôter tout ce qui peut contaminer les caractéristiques des deux groupes comparés ; mais dans des interventions où le contexte constitue un élément du résultat obtenu, on ne peut plus « décontextualiser ». L'évaluation ne peut donc s'intéresser seulement aux résultats : une analyse des processus est nécessaire pour ouvrir la « boîte noire », pour tenter d'identifier pourquoi la même intervention a des effets différents dans des contextes différents ou pourquoi telle intervention s'est révélée inefficace (l'échec est-il imputable au contenu même de l'intervention ou aux modalités de sa mise en œuvre ?). L'évaluation de processus peut être utilisée pour évaluer la fidélité et la qualité de la mise en œuvre, ou pour clarifier les mécanismes de causalité et identifier les facteurs contextuels associés à la variation des résultats.

Les méthodes alternatives aux ECR peuvent aller d'une adaptation (faire des « essais pragmatiques »), à des analyses des processus et mécanismes pour ne pas seulement savoir si une intervention est efficace mais aussi déterminer pourquoi et comment cette efficacité apparaît. La théorie de programme permet notamment de clarifier les hypothèses sur la façon dont les interventions fonctionnent : sachant que toute intervention repose sur des hypothèses (souvent implicites) quant aux causes du problème et quant aux leviers du changement, il importe de mettre au jour le « modèle logique », i. e. « la représentation schématique des relations entre les ressources, les activités d'une intervention et les résultats attendus ». Les méthodes sont nombreuses mais non interchangeables. Il faut en fait « avoir pour perspective la complémentarité d'études qui répondent à des questions différentes et qui contribuent ensemble à la démarche ».

Enfin, les interventions complexes du domaine de la santé doivent contribuer à rapprocher évaluation et recherche. C'est d'ailleurs ce que formule la feuille de route de la Stratégie nationale de santé en ces termes : « Il convient d'accroître les contributions de la recherche en matière de définition et d'évaluation des politiques de sécurité sanitaire et environnementale et d'apporter des éléments aux débats sur l'amélioration de l'efficacité et de l'efficience du système de santé et des politiques de santé publique et de prévention ».