

Repères pour évaluer l'efficacité des actions de prévention en santé

Jean-Claude Desenclos,
chargé de missions scientifiques,
Santé publique France.

Les sciences biomédicales, humaines et sociales améliorent la compréhension des processus complexes à l'œuvre dans la genèse du fardeau¹ en santé publique et permettent de concevoir des interventions pour améliorer la santé des populations. Cependant, leur mise en œuvre en contexte de vie implique une phase de conception, de développement et d'évaluation rigoureuse. La validité scientifique, l'efficacité et l'innocuité de l'intervention doivent être étayées.

Répondre à des besoins de santé publique nécessite de définir des objectifs d'amélioration et

la mobilisation de ressources pour matérialiser les interventions préventives ou de promotion de santé. Censées produire des résultats positifs, elles peuvent aussi avoir des effets indésirables avec *in fine* un impact sur la santé de la population qu'il faut apprécier. Le diagramme proposé par Rozenn Desplatz et Marc Ferraci [1] (voir figure 1) permet de situer plusieurs critères selon la mise en relation respective des effets des différentes étapes. Les critères d'analyse ou d'évaluation peuvent être qualitatifs, quantitatifs, ou les deux (évaluation mixte).

La documentation de l'efficacité – et de l'efficience – des interventions est un enjeu important dans les pays qui investissent des ressources publiques dans l'amélioration de la santé des populations. Dans un

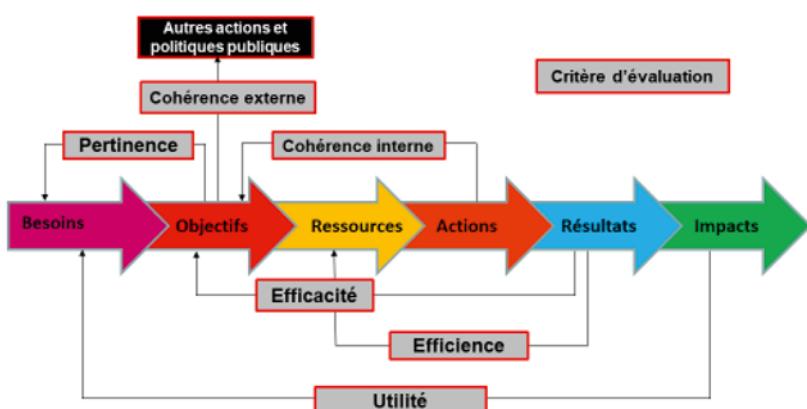
L'ESSENTIEL

■ **L'efficacité des interventions en prévention doit être démontrée en contexte de vie pour garantir qu'elles apportent à la population les bénéfices attendus. Plusieurs méthodes d'évaluation existent, avec différents niveaux de preuves, permettant d'estimer le gain de santé apporté par une action. L'évaluation quantitative se distingue d'autres approches, comme la capitalisation des expériences et les méthodologies qualitatives. Ces dernières apportent un éclairage légitime sur le contexte de mise en œuvre d'un programme et les conditions de son impact optimal.**

contexte de défiance, une démarche objective d'évaluation de leur efficacité et de leur impact est nécessaire. Elle doit être distinguée de l'évaluation portant sur d'autres critères légitimes comme la pertinence, la cohérence, la mise en œuvre ou l'utilité (voir figure 1).

L'efficacité quantifie la capacité de l'intervention à améliorer un ou des indicateur(s) de santé pertinent(s) et défini(s) *a priori* dans la population-cible et en un temps donné. On distingue l'efficacité (*efficacy*) qui est l'effet spécifique de l'intervention dans son indication spécifique et indépendamment des autres facteurs d'hôte et d'environnement, de l'efficacité en vie réelle (*effectiveness*) qui est évaluée dans le contexte de son déploiement. L'efficacité en vie réelle est inférieure à l'efficacité, sauf

Figure 1 – Déploiement d'une action de santé publique.



Source : Desplatz R., Ferraci M. (2016).

Légende : Les étapes de la mise en œuvre d'une action de santé publique sont schématisées par les flèches. Selon que l'on met en relation certaines de ces étapes entre elles, on peut définir des critères d'évaluation qualitatifs (pertinence, cohérence interne, cohérence externe) et quantitatifs (efficacité, efficience et impact) [1].

Dossier

Miser sur des interventions efficaces pour améliorer la prévention

quand l'action concerne les maladies infectieuses, du fait de l'effet externe des infections (prévenir la survenue d'infections réduit indirectement les infections secondaires du fait de leur nature transmissible).

L'impact est la réduction d'un problème de santé ou l'amélioration de facteurs positifs à la santé qui résultent d'une intervention en santé publique. L'impact dépendra de l'*effectiveness*, de la proportion de la population qui y est exposée, de son observance et d'un éventuel effet externe. L'efficience est la capacité de l'intervention à produire les résultats avec un niveau de ressource raisonnable. L'utilité est un concept plus large, car elle intègre par exemple le bien-être, l'absence de souffrance, la satisfaction des besoins fondamentaux, les préférences des citoyens, etc.

L'évaluation repose sur des essais d'intervention communautaire

L'évaluation de l'impact d'une intervention de santé publique doit se concevoir sur la base d'éléments argumentés. L'intervention devra avoir été suffisamment conceptualisée et codifiée : rationnel, innocuité, faisabilité, acceptabilité (individuelle, sociale et par les professionnels), etc. Cette phase de préparation implique une revue multidisciplinaire de la littérature et des expériences, voire des études pilotes incluant des approches participatives. Elle aboutit à la formulation de l'intervention dont on juge qu'elle apportera une amélioration de la santé de la population. L'intervention pourra être individuelle (programme de dépistage ou d'éducation sanitaire dispensé à chaque personne-cible par des professionnels de santé ou par des intervenants sociaux) ou collective, ciblant des groupes socialement définis (programme de sevrage tabagique sur le lieu de travail, programme d'éducation sanitaire scolaire, etc.).

L'objectif de l'évaluation est alors d'estimer son efficacité et d'en déduire l'impact en population, en établissant la situation qu'aurait connue la population sans l'intervention, ce qui permet de façon contre-factuelle d'en estimer le gain de santé en la comparant à celle observée en

présence de l'intervention. On a pour cela recours à des essais d'intervention communautaire (EIC), dont la perspective varie. Il peut s'agir d'établir la preuve de concept d'une action innovante par une évaluation optimale et maîtrisée (on évalue ici l'efficacité au sens des Anglo-Saxons). Quand une intervention a déjà fait l'objet d'évaluation(s) positive(s), avant d'envisager son implantation dans un autre contexte ou dans un autre système de santé, il s'agit de s'assurer de la validité externe de l'évaluation ou des évaluations déjà conduite(s) dans le nouveau contexte. Enfin, une autre perspective est celle du déploiement – ou de la mise à l'échelle –, par exemple au niveau national ou dans plusieurs pays et dans des contextes territoriaux et sociaux variables. Dans ces deux cas de figure, on évalue non plus l'efficacité, mais l'efficacité en vie réelle.

Le niveau de preuve (validité interne) d'un EIC variera en fonction du choix méthodologique. Il sera maximal si les participants reçoivent l'intervention au hasard (randomisation), mais au prix d'une faisabilité plus complexe et d'une validité externe (extrapolation à d'autres contextes) moindre. Le niveau d'application de l'intervention est à considérer (individus ou groupes définis [grappes/clusters] de personnes), de même que la mise en place ou non d'un groupe de comparaison, et le fait de disposer ou non de mesures avant ou après l'intervention. On distingue ainsi plusieurs types d'EIC [2 ; 3] :

- *EIC expérimentaux avec randomisation de l'intervention* dans lesquels l'échantillon expérimental et l'échantillon témoin sont par nature comparables, hormis l'intervention. L'application de l'intervention (et la randomisation) peut être individuelle ou au niveau de groupes de personnes ;
- *EIC quasi expérimentaux avec un groupe de comparaison géographique non aléatoire*. On parle d'approche quasi expérimentale, car l'intervention est mise en œuvre par les investigateurs dans le cadre de l'étude (contrairement aux « expériences naturelles » où l'évaluation est réalisée au vu de la mise en œuvre ou non de l'intervention) ;

- *EIC quasi expérimentaux sans groupe de comparaison géographique, mais avec mesure temporelle avant et après*. La mesure peut être ponctuelle ou reposer sur des séries chronologiques suffisamment longues avant et après l'intervention. La modélisation permet alors de prévoir la dynamique en l'absence d'intervention et de la comparer avec celle observée avec intervention. On ne peut cependant garantir l'absence d'une évolution spontanément favorable ou expliquable par des facteurs de confusion, notamment sociaux ;
- *EIC implicites* qui ne disposent que d'une mesure faite après l'intervention et sans groupe de comparaison, par exemple une automesure ponctuelle explorant une amélioration du fait de l'intervention. L'évaluation de l'efficacité ou de l'impact n'est alors pas possible.

Aspects méthodologiques, éthiques et opérationnels

La mise en œuvre d'EIC soulève des questions méthodologiques multiples et spécifiques [2 ; 3]. Les questions éthiques vont au-delà de la recherche biomédicale habituelle de par la dimension populationnelle² des EIC. La déclaration d'Ottawa propose 15 recommandations spécifiques à passer en revue pour assurer leur valeur éthique [4]. Les aspects opérationnels ne sont pas moindres et impliquent une préparation minutieuse en termes de ressources et de financement, d'organisation, de planification, d'implication des communautés concernées, d'information et de communication, d'assurance qualité et de monitorage. Parmi ces questionnements, citons :

- *Approche randomisée ou quasi expérimentale ?* La perspective est ici déterminante. S'il s'agit d'une preuve de concept, la question de la randomisation est explicitement posée sous réserve de sa faisabilité et des aspects éthiques. S'il s'agit d'une implantation ou du déploiement d'une intervention, on est avant tout intéressé par l'efficacité et l'impact en vie réelle pour maximiser la validité externe des résultats : l'approche quasi expérimentale ou par série chronologique, sans groupe de comparaison géographique, est alors à considérer.

● *EIC individuels ou en grappes ?*

L'approche en grappes permet d'intégrer la nature collective d'une intervention. Elle permet de gérer le « biais de contamination » entre individus, de prendre en compte l'effet externe éventuel et d'apprécier l'*effectiveness* dans une perspective de mise en œuvre en routine. Elle soulève néanmoins des questions éthiques spécifiques [4] et implique un effet dit « grappe », qui augmente le nombre de sujets à recruter. Choisie à bon escient, la nature de la grappe retenue anticipe la future mise à l'échelle. La nature des grappes retenues est fonction de l'intervention et de la population ciblée. Elle peut être :

- géographique (régions, ville, village, quartier) ;
- démographique (foyer familial) ;
- sociale (lieu de travail, école, collectivité de vie, établissement d'hébergement pour personnes âgées dépendantes – Ehpad) ;
- sanitaire (services de santé, hôpitaux, patientèle, centres de dépistage, de prévention).

● *L'EIC en grappes en biseau ou escalier (stepped-wedge), dans lequel toutes les grappes reçoivent l'intervention, chacune y étant exposée séquentiellement et aléatoirement réconcilie le besoin d'une méthode robuste (randomisation) et les questions éthiques et logistiques afférentes [5].*

● *L'adhésion et l'observance des personnes dans le groupe exposé à l'intervention ?* C'est une question récurrente. Une absence d'efficacité peut refléter une observance insuffisante plutôt qu'un manque réel d'efficacité, qu'il est nécessaire d'anticiper lors du calcul du nombre de sujets à inclure et par le monitorage précis de l'intervention pour prendre en compte le niveau d'observance lors de l'analyse qui introduit cependant des biais avec des tailles d'échantillon souvent insuffisantes.

● *L'efficacité de l'intervention est-elle modulée par certaines caractéristiques, notamment sociales ?* L'efficacité en pratique courante est souvent moindre dans les groupes les moins affluents socialement. Une approche stratifiée en distinguant plusieurs niveaux sociaux ou territoriaux pertinents permet d'évaluer cette hétérogénéité avec des échantillons

plus grands. Une alternative est de circonscrire l'évaluation dans un contexte de population socialement défavorisée.

Des approches qualitatives complémentaires

L'amélioration de la santé des populations via des interventions à caractères collectif et social en prévention et promotion de la santé implique, en amont de leur mise en œuvre, une évaluation rigoureuse qui peut être complexe, coûteuse en ressources et en temps. Elle nécessite une expertise approfondie avec une forte interaction avec la recherche. L'évaluation quantitative de l'efficacité et de l'impact fait l'objet de critiques : elle ne prendrait pas suffisamment en compte les contextes dans lesquels les interventions sont mises en place ; elle évaluerait l'efficacité « toutes choses étant égales par ailleurs », paradigme qui ne collerait pas à la réalité [6]. Cette dernière appréciation est à moduler, car il est tout à fait possible d'intégrer les caractéristiques (sociales, territoriales, genre, éducation, etc.) dont on a des raisons de penser qu'elles modifieraient l'efficacité.

D'autres approches, qualitatives ou mixtes, dites « réalistes », sont ainsi proposées [7]. Y est en particulier promu le savoir expérientiel, défini comme une vérité apprise par

l'expérience personnelle et collective de ceux qui la vivent (les pairs) [7]. La capitalisation des expériences apparaît alors importante pour la mise en œuvre des interventions et surtout pour leur transférabilité. En effet, si l'évaluation quantitative éclaire l'impact d'une action dans une population, elle n'informe que très partiellement sur la façon dont les acteurs de terrain élaborent ou adaptent leurs stratégies d'action à leur contexte et à ses évolutions [6 ; 7]. Si ces approches ne quantifient pas directement l'efficacité ou l'impact, elles éclairent légitimement sur les conditions, les processus et les contextes de mise en œuvre et indirectement sur les conditions d'un impact positif voire optimal. ■

1. Le fardeau de la maladie (*Burden of Disease, BoD*) : méthode scientifique visant à quantifier de manière systématique et comparative l'ampleur de la dégradation ou de la perte de la santé liées aux maladies, aux traumatismes et aux facteurs de risque par âge, sexe et zone géographique à un moment donné (NDLR). En ligne : <https://www.santepubliquefrance.fr/les-actualites/2024/estimation-du-fardeau-environnemental-et-professionnel-de-la-maladie-en-france-quels-sont-les-couples-maladie-facteur-de-risque-a-priorise>

2. Une des questions est que l'on assigne des groupes complets à l'intervention ou non, sans demander en amont l'avis des personnes les constituant.

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] Desplatz R., Ferraci M. *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens*. Paris : France Stratégie, coll. Évaluation, 2016 : 72 p. En ligne : https://www.strategie.gouv.fr/files/files/Publications/2016%20SP/guide_methodologique_20160906web.pdf
- [2] Desenclos J.-C., Dabis F. Méthodes épidémiologiques de l'évaluation d'une intervention de santé publique. In F. Dabis & J.-C. Desenclos, (éds.), *Épidémiologie de terrain. Méthodes et applications*. Arcueil : John Libbey Eurotext, 2016 : 792 p.
- [3] Hayes R., Moulton L.H. *Cluster Randomised Trials*. Boca Raton (Florida) : CRC Press, 2009 : 315 p.
- [4] Weijer C., Grimshaw J. M., Eccles M. P., McRae A. D., White A., Brehaut J. C. et al. The Ottawa statement on the ethical design and conduct of cluster randomized trials. *PLoS Medicine*, 2012, vol. 9, n° 11. En ligne : <https://pubmed.ncbi.nlm.nih.gov/23185138/>
- [5] Hemming K, Haines T. P., Chilton P. J., Girling A. J., Litford R. J. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ*, 6 février 2015, vol. 350. En ligne : <https://www.bmjjournals.org/content/350/bmj.h391>
- [6] Alla F., Cambon L., Ridde V. *La Recherche interventionnelle en santé des populations. Concepts, méthodes et applications*. Éditions IRD, 2023.
- [7] Soudier B., Ferron C., Laurent A. Capitalisation des expériences en promotion de la santé : enjeux, apports et méthode. *La Santé en action*, 2021, vol. 456 : p. 9-11. En ligne : <https://www.santepubliquefrance.fr/docs/capitalisation-des-experiences-en-promotion-de-la-sante-enjeux-apports-et-methode>