

La santé publique est le champ de l'action publique visant à prévenir les maladies, prolonger la vie et promouvoir la santé à travers l'effort organisé de la société et de ses composantes. Dans la mesure où la santé conditionne les activités humaines, où les crises de santé publique peuvent avoir des répercussions sur de nombreux secteurs d'activité, bien au-delà du système de santé, et où la santé se construit et peut être altérée dans tous les environnements (l'environnement familial, professionnel, le système éducatif, les transports, etc.) et dans une multitude de situations (les loisirs, l'alimentation, etc.), la recherche en santé publique porte sur tous les domaines de la société, tous les environnements, bien au-delà de ce qui est identifié comme le système de santé ; elle concerne toutes les pathologies et paramètres biologiques et sanitaires, tous les âges et groupes, ainsi que bien entendu les inégalités entre ces groupes. Cette recherche permet de répondre à de grandes questions sur la fréquence et la gravité des problèmes de santé, tant dans leurs dimensions individuelles que collectives, leurs causes, les mécanismes (biologiques, sociologiques, etc.) sous-jacents et les solutions possibles, qu'elles passent par des interventions centrées sur l'individu, les communautés, l'environnement urbain, physique, chimique, socioculturel, le système de santé au sens strict, les acteurs économiques, le droit, etc. À ce titre, elle englobe toutes les échelles de la molécule à la planète et s'appuie sur de nombreuses disciplines qualitatives et quantitatives, incluant sciences de l'environnement, de la nature (biologie fondamentale, épidémiologie), de la société (sociologie, économie, histoire, droit, etc.) et mathématiques. Ses approches sont expérimentales (essais

randomisés contrôlés, études d'intervention randomisées, etc.) ou observationnelles (cohortes, études cas-témoins, études de prévalence ou de surveillance), quantitatives ou qualitatives, comme c'est plus souvent le cas dans les sciences humaines et sociales.

Introduction

Finalités de la recherche

La finalité de la recherche en santé publique est de fournir les connaissances sous-tendant la question de la préservation et l'amélioration de la santé des populations, incluant la réduction des inégalités de santé, et permettant de mener les politiques en lien avec ces objectifs (**Figure 12.1**). Dans une vision un peu restrictive, on peut la voir comme la *science des causes externes des maladies et de leur contrôle*, ou encore la science de la prévention sanitaire et de l'amélioration de la santé. On peut distinguer parmi ses objectifs spécifiques, ceux de :

- décrire la fréquence des troubles de la santé, l'histoire « naturelle » de la maladie, ainsi que la distribution des déterminants connus ou suspectés de la santé (démographie, facteurs environnementaux, sociaux, comportementaux, etc.) : c'est le *volet descriptif* ;
- comprendre les causes (principalement extérieures) de ces troubles de la santé ou, de façon équivalente, les effets des facteurs environnementaux au sens large (facteurs physiques, chimiques, psychosociaux, biologiques, comportementaux, etc.) : on parle parfois ici d'identification des dangers ou de *volet étiologique* ;

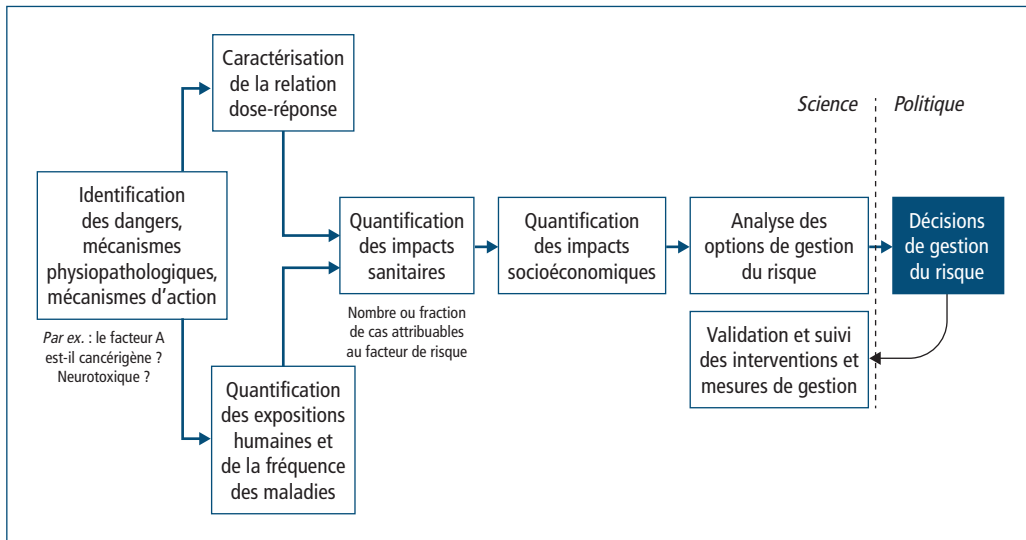


Figure 12.1. Les grands objectifs de la recherche en santé publique et leur contribution théorique aux décisions de gestion du risque. Noter qu'il s'agit d'un aperçu de ce qui est possible et qu'en pratique toutes ces étapes ne sont pas nécessaires pour toutes décisions publiques, qui peuvent aussi être prises par exemple sur la base de résultats limités à l'identification des dangers par exemple (adapté de [1]).

- contribuer à la compréhension des mécanismes sous-jacents ; ceux-ci peuvent être de nature biologique, psychologiques, sociaux (comme lorsqu'il s'agit de comprendre la réticence vaccinale), économiques ou politiques (pour comprendre les freins à la mise en œuvre de certaines politiques de santé publique) : c'est le *volet mécanistique* ;
- quantifier les impacts sociétaux correspondants ; le terme d'impact est entendu comme le nombre (ou la proportion) des cas de pathologie (ou d'années de vie en bonne santé perdues) attribuables aux dangers ou facteurs de risque identifiés ou suspectés et est synonyme de risque ; cet impact est la concrétisation dans un contexte social donné du danger évoqué ci-dessus et correspond au *volet d'évaluation* ou quantification du risque ou des impacts ;
- identifier les *interventions efficaces pour préserver ou améliorer la santé* (y compris en termes de réduction des inégalités sociales ou territoriales, et y compris, mais pas seulement, concernant le système de santé, que ces interventions soient incitatives ou contraignantes), et comprendre les mécanismes

(sociaux ou autres) déterminant l'efficacité de ces interventions ; cette finalité inclut toute la recherche sur le système de santé (structure, efficacité, financement, équilibre entre prévention et thérapie, etc.).

Quelques jalons de la recherche passée

Attribuer à un champ de recherche une découverte spécifique procède toujours un peu de l'arbitraire. Malgré tout, on peut reconnaître au moins une *contribution* de la recherche en santé publique dans les connaissances suivantes :

- la démonstration de la transmission du choléra par les eaux usées (John Snow, vers 1850) ;
- les importantes inégalités de mortalité et de santé avec les classes sociales (Villermé et ses études comparant la mortalité infantile selon les milieux sociaux, 1820-1840, ayant permis de restreindre – très progressivement – le travail infantile) ;
- l'identification de l'impact de la fumée de tabac sur la mortalité, le cancer du poumon et de nombreuses autres pathologies [2] ;

- le rôle des comportements alimentaires, de l'activité physique, de la consommation d'alcool et de tabac sur les maladies cardiovasculaires (cohorte de Framingham et autres études à partir des années 1950) ;
- la démonstration des effets de la pollution atmosphérique (particules fines notamment) sur la santé, à des doses toujours plus faibles [3] ;
- la démonstration des effets du plomb sur le neurodéveloppement et la santé en général, ainsi que l'étude des inégalités sociales d'exposition au plomb ;
- l'établissement du lien entre température et mortalité et l'estimation de l'impact présent et attendu du changement climatique sur la mortalité due à ce lien...

Cette liste n'est bien sûr pas exhaustive et vise avant tout à illustrer la diversité des facteurs de risque et troubles de santé considérée.

Toute aussi diverse est la nature des approches sur lesquelles repose cette recherche, que nous allons maintenant passer en revue en insistant sur celles relevant de l'épidémiologie, plus formalisées, ce qui ne signifie pas que les autres disciplines telles que celles relevant des sciences humaines et sociales ne contribuent pas de façon essentielle à la santé publique.

La description donnée ci-dessous des grandes approches utilisées pour répondre aux différents objectifs de la recherche en santé publique pourra être complétée par des présentations plus détaillées [4, 5].

Aperçu des méthodes de la recherche en santé publique

Décrire les maladies, l'environnement, les comportements

La description des maladies à l'échelle des populations passe par des études d'incidence (renseignant sur le nombre de nouveaux cas d'une pathologie, indépendamment de ses causes), prévalence (identifiant les cas présents à un

moment donné, qu'ils soient survenus récemment ou pas) et auprès des patients (pour renseigner l'histoire de la maladie une fois survenue). Une option méthodologiquement simple mais très lourde logistiquement est d'identifier tous les cas incidents de la maladie survenant dans une zone donnée : c'est la logique des registres, dont des exemples sont le registre des causes médicales de décès de l'Inserm (CépiDc), les deux registres nationaux des cancers pédiatriques ou la quinzaine des registres départementaux français des cancers de l'adulte. Il est aussi bien sûr possible d'obtenir une estimation précise de la fréquence de l'événement d'intérêt sans exhaustivité. Méthodologiquement, le cadre est celui de la *théorie des sondages* [6], qui permet de définir à quelles conditions on peut faire une inférence correcte sur la fréquence de la maladie (et les incertitudes associées) en n'étudiant qu'un sous-groupe de la population. Le plan de sondage élémentaire correspond à l'échantillonnage aléatoire simple (qui implique de disposer d'une base de sondage, ce qui est loin d'aller de soi), mais il existe des plans plus complexes, tels que l'échantillonnage stratifié et les sondages à probabilités d'inclusion inégales (utiles si l'on souhaite à moindre coût bien décrire un groupe de petite taille par rapport aux autres considérés, qui sera surreprésenté dans l'échantillon recruté par rapport aux autres groupes). Un exemple d'un tel plan de sondage complexe est celui de l'étude EpiCov, réalisée durant la crise Covid-19 à partir de 2020 pour décrire la séroprévalence de la maladie, et qui a consisté à représenter chaque département français avec un même effectif, et surreprésenter les populations socialement défavorisées [7]. On peut rappeler que les sondages par quotas appréciés des instituts de sondage d'opinion ne font pas partie des sondages probabilistes et ne permettent pas de contrôler rigoureusement les biais et estimer les incertitudes.

Ces méthodes peuvent aussi bien sûr être utilisées pour décrire les comportements (cas des baromètres renseignant sur la consommation de tabac ou des premières études sur les comportements sexuels du Pr Kinsey aux États-Unis en 1948 et 1953, ou de l'équipe du Pr Spira en France au début de l'épidémie de VIH [8]), les facteurs environnementaux (méthodes

d'échantillonnage de l'air, de l'eau pour décrire leur pollution, ou encore de l'alimentation, dans le cas de l'enquête de l'alimentation totale réalisée par l'ANSES¹) et la contamination de l'organisme par des facteurs chimiques, *via* les études de biosurveillance telles que NHANES aux États-Unis [9], ou Esteban [10] et maintenant Albane en France².

Le développement du SNDS (système national des données de santé) suscite d'importants espoirs du point de vue de la recherche étiologique et la surveillance. À l'heure actuelle, sa base principale ne contient pas de données systématiques sur les diagnostics réalisés mais seulement sur les prescriptions médicales (qui peuvent permettre de prédire plus ou moins précisément la pathologie ayant donné lieu à la prescription), les coûts hospitaliers et les causes médicales de décès.

Les études descriptives telles que celles de biosurveillance peuvent simultanément fournir un tableau de la santé et de certains facteurs de risque. Toutefois, la méthodologie des études d'incidence simples, de par sa nature transversale (renseignement sur la maladie et ses facteurs de risque simultanément, sans suivi temporel), n'est pas la plus efficace pour identifier les effets sanitaires des facteurs de risque. Une exception est celle des études descriptives concernant les maladies infectieuses, dont les déterminants sont souvent plus proches du déclenchement de la maladie et pour lesquelles des études descriptives d'analyse de cluster peuvent permettre de comprendre certains facteurs de risque.

Comprendre les causes des maladies-l'identification des dangers

■ Essais d'intervention randomisés

Abordons l'étude des facteurs de risque des pathologies par l'essai randomisé (*randomised controled trial* [RCT]), dans lequel on assigne aléatoirement à un groupe de sujets indemnes de l'événement de santé d'intérêt chacun des

traitements ou facteurs dont on souhaite comparer l'influence sur un paramètre biologique ou événement de santé prédéfini. Ce tirage au sort garantit que, en espérance, les groupes comparés sont identiques du point de vue des autres facteurs de risque de la maladie. Si possible, ni les volontaires ni l'équipe en charge du suivi ne savent quel traitement a été attribué à chacun (essai en double aveugle), afin de limiter l'effet placebo et certaines erreurs de mesure et diagnostic. Ainsi, toute différence observée dans la probabilité de développer la maladie (ou d'en guérir, si on part de patients atteints d'une même pathologie) ne peut être attribuée qu'au traitement lui-même, et pas à une différence dans les facteurs de risque de la maladie ou la mesure de sa survenue. Le fait de prédéfinir un critère de jugement principal (l'événement de santé d'intérêt) permet aussi de contrôler les associations fortuites dues aux comparaisons multiples.

Dans le détail, plusieurs modalités et options existent (voir **Figure 12.3**), notamment :

- *bras parallèles versus essai croisé (crossover)* : plutôt que de comparer un groupe de sujets exposés à un facteur ou traitement donné à un autre groupe exposé à un autre facteur (éventuellement un placebo), correspondant à une approche avec des bras de traitement parallèles, dans l'essai croisé (ou en *crossover*), on compare un sujet considéré à un moment donné avec lui-même à une autre période durant laquelle le traitement diffère ; ce qui est randomisé n'est plus le traitement reçu mais l'ordre selon lequel chacun reçoit chaque traitement. Une telle approche croisée n'est adaptée qu'à l'étude de traitements ayant des effets à plutôt court terme sur des paramètres de santé. Citons aussi les essais de plateforme adaptative, qui visent à considérer de façon plus souple et rapide des traitements multiples et analyses en sous-groupe [11] ;
- *le niveau de randomisation, individuel ou collectif* : l'essai randomisé peut aussi être réalisé au niveau de communautés ou groupes (en particulier s'il y a un risque de *contamination* de l'intervention entre les sujets), à l'échelle desquels une exposition est appliquée et la fréquence d'un événement de santé

1. <https://www.anses.fr/fr/content/les-etudes-de-l'alimentation-totale-eat>

2. <https://www.enquete-albane.fr/>

suivie et comparée entre les groupes ayant reçu chaque traitement (*essais en cluster*, ou *cluster randomized trials* [12]) ; un exemple consiste à sélectionner un certain nombre de quartiers et à y mettre en œuvre diverses approches (marketing social, développement d'équipements sportifs) pour promouvoir l'activité physique afin de quantifier leur bénéfice ;

- la finalité de l'essai peut être de mettre en évidence la supériorité d'un traitement sur un autre (*essai de supériorité*, le plus fréquent) ou de fournir des éléments en faveur d'une efficacité similaire entre deux traitements (*essai de non-infériorité*, *essai d'équivalence*) ; cette distinction a des conséquences sur l'effectif nécessaire notamment [13].

L'essai randomisé contrôlé au niveau individuel est l'approche principale utilisée pour valider les médicaments et dispositifs médicaux, ainsi que celle de la toxicologie *in vivo* (dans les autres espèces que l'humain en général). En pratique, toutefois, il n'est pas toujours possible ou souhaitable de contrôler l'exposition, ou d'inclure un nombre de volontaires suffisant pour étudier des maladies relativement rares. Une autre option est la mise en place d'une cohorte.

■ Étude de cohorte

L'étude de cohorte ressemble à l'essai randomisé à la différence près que l'exposition au facteur de risque est *observée* et non pas *contrôlée* par l'équipe de recherche (notons que l'expression de *cohorte* n'exclut pas rigoureusement que l'exposition soit randomisée, même si beaucoup préfèrent utiliser l'expression d'essai randomisé contrôlé dès qu'il y a contrôle de l'exposition au niveau individuel). Ces études peuvent être réalisées sur des effectifs bien plus importants (au-delà de 100 000 sujets), ce qui permet d'étudier de nombreuses pathologies simultanément – en principe toutes celles dont la fréquence est telle qu'il y aura dans la cohorte un nombre conséquent de cas au bout d'une certaine durée de suivi. Un effectif important peut aussi permettre d'étudier différents facteurs de risque avec des prévalences variables, possibilité étendue par le développement des biomarqueurs.

L'approche consiste à recruter des sujets indemnes de la ou des maladies d'intérêt, à caractériser leurs facteurs de risque et les déterminants potentiels considérés, puis à suivre les sujets afin d'identifier les cas incidents des maladies (ou le suivi de l'évolution de paramètres biologiques) d'intérêt. L'absence de randomisation ne garantit pas la comparabilité des groupes comparés, entraînant de possibles biais de confusion³. Leur contrôle peut se faire par des approches statistiques telles que le redressement ou, plus souvent, l'ajustement *via* un modèle de régression⁴, parfois en ayant recours à un score de propension. L'analyse statistique (par un modèle de survie tel que celui de Cox) permet d'estimer l'association entre chaque facteur de risque et la force de survenue de la maladie, ajustée sur les facteurs pertinents (**Figure 12.2**). Cet ajustement statistique est à la cohorte ce que la randomisation est à l'essai randomisé – il vise à garantir la comparabilité des groupes ayant différents niveaux d'exposition. Contrairement à ce qu'on entend parfois, l'étude de cohorte n'est pas censée, pour estimer de façon rigoureuse l'association entre exposition et survenue de la maladie, être « représentative » d'une population donnée [15]. Par exemple, de nombreuses cohortes impliquant un suivi lourd sont réalisées sur des populations plutôt favorisées, disposant de temps libre pour s'investir dans l'étude, ou sur-représentant les professions médicales ; si cette restriction peut limiter l'aptitude à étudier les interactions avec des paramètres sociaux, elle ne limite pas forcément la validité des résultats concernant l'effet possible de l'exposition. En effet, d'une part cette restriction peut limiter les biais de confusion en accroissant l'homogénéité de la population d'étude et, d'autre part,

3. Un facteur de confusion est défini comme un facteur *influençant* à la fois le déterminant et l'événement de santé d'intérêt. Il ne s'agit donc ni de conséquences possibles de l'exposition d'intérêt (qui peuvent constituer des facteurs intermédiaires, ou médiateurs, dans l'association entre exposition et maladie), ni des conséquences de la maladie, dont le contrôle est susceptible d'être nuisible, causant un biais de sélection [14].

4. Pour diverses raisons, les essais randomisés sont eux aussi généralement ajustés sur certains facteurs de risque de l'événement d'intérêt.

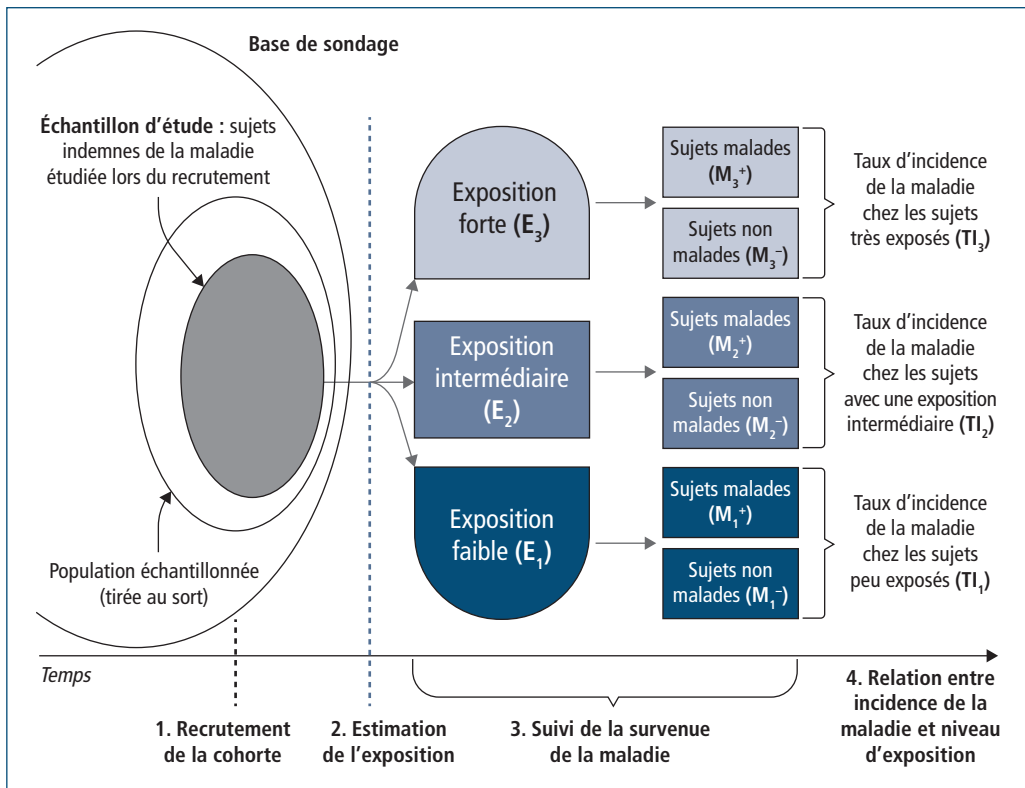


Figure 12.2. Principe de l'étude de cohorte. L'analyse consiste en principe à comparer le taux d'incidence de la maladie dans les groupes définis par leur niveau d'exposition au facteur d'intérêt, ce qui en pratique peut se faire grâce à un modèle de régression tel que le modèle de Cox. L'exposition peut bien sûr avoir plus de trois niveaux ou être considérée comme une variable continue, et l'analyse peut être ajustée sur différents facteurs de confusion (d'après [4]).

la généralisation des résultats ne s'obtient pas en répétant l'étude dans toutes les populations (groupes sociaux, pays, etc.) possibles mais plutôt par des arguments extérieurs sur les mécanismes d'action [15].

La France est riche en études de cohortes, qu'elles aient été constituées en milieu clinique ou en population générale. Cette approche est puissante, rigoureuse, permet souvent, par son approche prospective, de mesurer avec une bonne finesse les paramètres biologiques, pathologies, expositions et facteurs de confusion ; en revanche elle peut être lente à mettre en œuvre – la durée de suivi nécessaire étant définie par le nombre d'années permettant qu'un nombre « suffisant » de cas de la maladie surviennent dans la population d'étude.

■ Étude cas témoins

L'étude cas-témoins permet de contourner cette limitation, en recrutant d'emblée des cas de la maladie d'intérêt, ce qui peut être fait rapidement s'il existe des approches permettant de les identifier – par exemple un service hospitalier où se fait leur diagnostic. Son principe est d'échantillonner séparément des cas de la maladie et une population « comparable » indemne de la maladie (les témoins⁵), et de renseigner rétrospectivement les niveaux du

5. Une confusion est parfois faite autour du groupe témoins, qui, en toxicologie désigne un groupe non exposé à un facteur de risque d'intérêt – une substance chimique par exemple – alors que dans l'approche épidémiologique cas-témoins, les témoins représentent le groupe indemne de la maladie.

ou des facteurs de risque d'intérêt et des facteurs de confusion potentiels dans ces deux groupes. L'approche est donc rétrospective. Si les sites de recrutement diagnostiquent beaucoup de cas, on peut ainsi, en un an, recruter des centaines de cas, alors que, dans le cas d'une maladie dont l'incidence est de 10 cas pour 100 000 personnes par années, recruter 100 cas nécessite de suivre 10 000 sujets initialement indemnes de la maladie durant 10 ans. Si l'étude cas-témoins est réalisée sur des cas incidents, elle est en théorie, dans les estimations qu'elle fournit, équivalente (aux erreurs de mesure près) à une étude de cohorte – c'est-à-dire que l'*odds-ratio* issu de l'étude cas-témoins incidente est équivalent au risque relatif (*hazard rate*) du modèle de Cox d'une étude de cohorte [5].

Les limites de l'approche cas-témoins sont liées à sa nature rétrospective, qui entraîne davantage d'erreurs par rapport à une cohorte dans la caractérisation des facteurs de risque (liées aux biais de mémorisation, à l'impossibilité, généralement, de caractériser rétrospectivement un paramètre biologique avec une forte variabilité temporelle), au fait que l'étude est généralement restreinte à une unique pathologie, à la difficulté d'identification et recrutement du groupe témoin pertinent. Concernant ce dernier, il est censé correspondre à un échantillon représentatif des sujets qui, s'ils avaient développé la maladie dans la période considérée, auraient été éligibles dans le groupe des cas. Si les cas sont ceux recrutés dans un service hospitalier donné pour un diagnostic de cancer du poumon en 2025, les témoins devraient ainsi correspondre aux sujets qui, s'ils avaient développé un cancer du poumon en 2025, auraient été diagnostiqués dans ce même service (ce qui implique d'avoir une bonne idée de la zone de recrutement du service et de disposer d'une base de sondage permettant d'en identifier un sous-groupe aléatoire).

Il est possible de réaliser des études cas-témoins sans recruter de témoins, en utilisant chaque cas comme son propre témoin – ce qui n'a de sens que pour les facteurs de risque variant rapidement dans le temps, et pour caractériser leurs effets à court terme

seulement – c'est l'étude cas-croisé (*case crossover*). On peut aussi, pour contourner les biais dus à la mesure rétrospective de l'exposition, nicher l'étude cas-témoins dans une cohorte réalisée précédemment ; si des prélèvements biologiques y ont été réalisés prospectivement, ceux-ci peuvent être utilisés pour caractériser l'exposition de façon prospective, mais uniquement chez les cas et un sous-échantillon des autres sujets, limitant le coût des dosages par rapport à une analyse de cohorte conventionnelle (cas-témoins nichés dans la cohorte, approche cas cohorte).

■ Études écologiques spatiales

De même qu'on peut comparer la fréquence d'un événement de santé entre des sujets présentant des contrastes d'exposition (c'est l'étude de cohorte évoquée plus haut), il est possible de le faire à l'échelle de groupes et communautés : il s'agit dans ce cas d'une étude écologique. Les différences principales avec la cohorte sont que les données sur la maladie et l'exposition sont agrégées à l'échelle de groupes, et que, souvent, expositions et événement de santé sont caractérisés pour les mêmes périodes (l'analyse est transversale). Le premier point a des conséquences importantes, car l'agrégation des données ne permet pas un contrôle des biais de confusion aussi efficace que dans les études individuelles (cohortes, cas-témoins) : c'est ce qu'on appelle le biais écologique [16]. Pour cette raison notamment, on considère que le niveau de preuve apporté par les études écologiques *spatiales* est bien plus faible que celui provenant d'études sur données individuelles caractérisant les facteurs de confusion potentiels précisément.

On peut rapprocher les *études de cluster* des études écologiques spatiales. Ces études consistent à identifier des zones définies spatialement pour lesquelles l'incidence de la maladie dans une période donnée est plus élevée qu'attendu étant donné la taille de la zone. Si cette approche peut permettre d'identifier des facteurs de risque ayant un effet très fort et à court terme (cas d'agents infectieux ou, par le passé, d'expositions professionnelles fortes), elle se révèle souvent peu informative pour des

maladies chroniques telles que le cancer [17] et des facteurs de risque ayant un effet faible ou décalé dans le temps. Une façon plus efficace de remonter à la cause d'une pathologie en partant des cas est de s'appuyer sur l'approche cas-témoins déjà évoquée, qui inclut généralement bien plus de cas qu'une investigation de cluster (permettant d'étudier des facteurs de risque sans supposer qu'ils ont un très « fort » effet sur le risque de maladie), formule des hypothèses *a priori* sur les facteurs de risque et ne fait pas l'hypothèse que la localisation spatiale est très fortement liée au(x) facteur(s) de risque. Tout ceci ne s'applique bien sûr pas aux maladies infectieuses, pour lesquelles les enquêtes sur les foyers de cas peuvent être très efficaces pour identifier un pathogène ou son mode de transmission, ou limiter la progression d'une épidémie.

■ Études écologiques temporelles (et séries temporelles)

Une autre façon de comparer des groupes est de s'appuyer sur les contrastes temporels, et non pas spatiaux, d'exposition ; il est dans ce cas possible de comparer le nombre de nouveaux cas d'une pathologie dans une même communauté à différentes périodes – par exemple la mortalité avant, pendant et après un fort épisode de pollution, ou, symétriquement, une période durant laquelle la pollution a décliné, par exemple du fait de mesures de restriction du trafic routier durant les jeux olympiques. Si ces périodes sont proches, on peut faire l'hypothèse que la communauté a peu varié du point de vue de la plupart des facteurs de risque de la maladie (ex. : tabac, obésité ou facteurs génétiques) et que les biais de confusion sont moins susceptibles d'exister que pour une étude écologique spatiale, dans laquelle on va par exemple comparer deux villes. L'expression d'*expérience naturelle* est parfois utilisée pour certaines études observationnelles pour lesquelles on peut considérer l'attribution de l'exposition similaire à ce qui serait observé si elle avait été tirée au sort [18], ce qui peut notamment être le cas pour des comparaisons à court terme des sujets d'une même communauté chez qui l'exposition

a brusquement changé (par exemple du fait d'une grève dans une usine émettant des polluants atmosphériques).

Les études écologiques temporelles peuvent considérer un nombre important d'unités temporelles, par exemple plusieurs années si l'unité d'analyse est le jour, et en rapprochant la variation du nombre de cas journaliers de l'exposition à ceux d'un facteur dont les niveaux varient peu spatialement mais fortement temporellement, et sont aussi suivis avec le même pas de temps (*analyse de séries temporelles* [19]). Cette approche s'est révélée très puissante pour mettre en valeur les effets à court terme de la pollution atmosphérique ou de la température sur la mortalité. Une limite de ces études écologiques temporelles est le fait qu'elles ne permettent que de caractériser les effets à court terme de l'exposition d'intérêt, sous-estimant généralement l'impact global de l'exposition par rapport à ce qui serait capturé par une étude de cohorte avec un suivi à long terme.

■ Études écologiques spatiotemporelles : la méthode des différences de différences

Cette approche (aussi appelée méthode des doubles différences) repose à la fois sur des comparaisons spatiales et temporelles : schématiquement, la tendance temporelle dans un paramètre de santé est estimée dans un groupe de sujets ou communautés considéré non exposé, et la comparaison (la différence) avec la tendance temporelle dans le même paramètre observé dans un groupe exposé permet d'estimer l'effet de cette exposition débarrassé des tendances temporelles dans la maladie non due à l'exposition [20, 21].

■ Études familiales

Évoquons aussi les études familiales, qui consistent à caractériser l'agrégation des cas d'une maladie au sein de familles, et peuvent permettre d'estimer la contribution de facteurs génétiques dans la survenue de la maladie (ex. : estimation de l'hérédité) et d'identifier ces facteurs si les sujets sont génotypés [22].

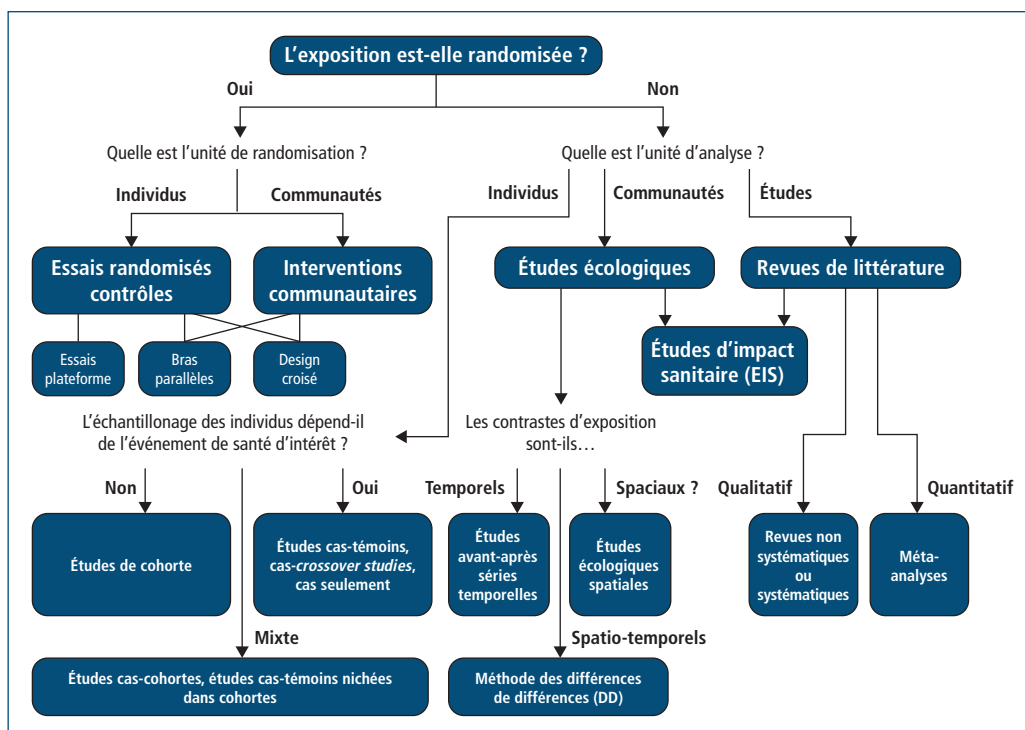


Figure 12.3. Principaux types d'études étiologiques en épidémiologie (adapté de [23]).

Comprendre les mécanismes des maladies

■ Mécanismes internes des maladies et l'ère de l'épidémiologie moléculaire

La diminution des coûts de nombreux dosages biologiques et des volumes d'échantillons biologiques nécessaires à leur caractérisation précise par les outils de spectrométrie de masse et chromatographie notamment, permettent désormais de quantifier un nombre important de paramètres biologiques chez chaque individu. Il peut s'agir de séquencer le génome, de caractériser des paramètres immunitaires, d'inflammation, hormonaux, l'épigénome (méthylation de l'ADN par exemple), les transcrits (ARN, micro-ARN, etc.), le protéome (niveaux circulants de différentes protéines), métabolome, microbiome et autres « omiques », etc. En génétique, ces travaux ont donné naissance aux GWAS (*genome-wide association studies*, ou études pangénomiques),

qui permettent de caractériser le lien entre la présence de polymorphismes génétiques et l'incidence ou la sévérité des maladies ou de traits spécifiques [24]. À part pour les caractéristiques génétiques des individus, qui varient peu dans le temps, ces études à visée mécanistique nécessitent des données prospectives et s'appuient généralement sur des cohortes. Ces études, à l'interface entre l'épidémiologie classique et la biologie des systèmes, posent de nombreux défis liés à la nécessité de mesurer les différents paramètres biologiques dans une fenêtre de temps pertinente (impliquant *a priori* des prélèvements biologiques répétés), à la dimension généralement grande des données recueillies et à l'utilisation d'approches statistiques adaptées au système biologique et au plan d'étude.

■ Analyses de médiation

Une situation qui a été clairement formalisée est celle où on suppose que l'effet d'un facteur E (par exemple une exposition environnementale ou une

mutation génétique) sur un paramètre de santé Y est en partie expliqué par l'effet de E sur un paramètre biologique intermédiaire M (le *médiateur*, par exemple l'expression d'une protéine ou un marqueur d'inflammation). Ceci correspond au cadre théorique de l'analyse de médiation [25]. Ce cadre permet de fournir une estimation de la part de l'effet de E sur le paramètre de santé due à l'effet de E sur le médiateur M (**Figure 12.4A**). En pratique, le médiateur potentiel peut correspondre à un ensemble de variables de dimension moyenne à haute (par exemple une couche de données sur la composition du microbiote intestinal ou les marques épigénétiques sur l'ensemble du génome, **Figure 12.4B**), ce qui pose des défis dans l'analyse statistique [26].

La question de l'*analyse des interactions* est plus complexe, notamment parce que ce terme est polysémique (interactions biologiques, chimique et statistique n'étant pas identiques), et parce que l'interaction statistique, qui peut être testée, dépend du modèle statistique utilisé. Il est généralement admis que les situations d'interaction pertinentes pour la santé publique (permettant par exemple d'identifier un sous-groupe à cibler pour une intervention si les moyens ne permettent pas d'agir sur toute la population) correspondent à celles mesurées sur une échelle additive (c'est-à-dire une modification de la mesure de l'effet additif des facteurs de risque considérés), qui n'est pas celle du modèle de régression logistique

souvent utilisé en épidémiologie ; des approches statistiques ont été proposées pour identifier ces interactions « additives » dans le contexte de tels modèles multiplicatifs [5, 27].

■ Mécanismes sociopolitiques

À l'échelle des populations, de nombreux autres mécanismes, extérieurs à l'organisme sont bien sûr aussi pertinents. On peut par exemple, avec les outils de la recherche qualitative ou quantitative de la sociologie et des sciences politiques, chercher à comprendre ce qui contribue à la réussite d'une campagne vaccinale ou comment certaines questions sont invisibilisées par des groupes y ayant intérêt (agnotologie, ou étude de la production culturelle de l'ignorance et du doute [28]).

Études d'impact sanitaire : quantifier les risques

Toutes les études étiologiques (d'association) décrites précédemment fournissent une estimation de l'effet d'un facteur de risque sur l'incidence ou la prévalence d'une maladie, souvent sous la forme d'un risque relatif ou *odds-ratio*, mesure qui, seule, est peu parlante en termes de santé publique ou pour la société. Une exposition associée à un risque relatif d'une maladie élevé (5 ou 10) peut entraîner peu de cas dans une population (si l'exposition est peu fréquente), alors qu'une autre dont le risque

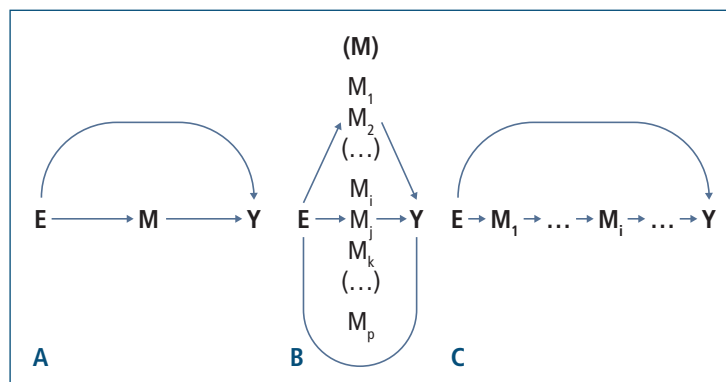


Figure 12.4. Analyses de médiation entre un facteur de risque E (exposition environnementale, comportement, etc.) et un événement de santé Y. **A.** Situation avec un unique médiateur M. **B.** Médiateur multidimensionnel. **C.** Ensemble de médiateurs s'influençant mutuellement (adapté de [26]).

relatif est bien plus faible (par exemple 1,5) peut entraîner bien plus de cas, si l'exposition est généralisée. Au-delà de la mesure d'association, l'estimation du nombre (ou de la proportion) de cas d'une pathologie attribuable à chaque exposition est donc pertinente. L'estimation de cette fraction de risque attribuable, concept ancien mais subtil de l'épidémiologie, implique de combiner les informations sur l'association entre exposition et maladie (une relation dose-réponse, qui peut être résumée par un risque relatif en cas de linéarité de cette relation sur une certaine échelle), la distribution de l'exposition et l'incidence moyenne [29]. La méthodologie est celle des études d'impact sanitaire [30], dont un exemple de résultats est donné en **figure 12.5**. D'autres approches d'évaluation des risques ou de la sécurité peuvent être utilisées, en particulier si les seules relations dose-réponse disponibles proviennent d'espèces autres que l'humain.

Insistons sur la différence avec les études étiologiques mentionnées plus haut : la plupart d'entre elles (cohortes, cas-témoins, séries temporelles, etc.) estiment une variation relative de la fréquence de la maladie associée à une variation donnée de l'exposition (donc celle attendue en comparant deux groupes dont l'exposition

diffère d'une unité ou d'une quantité donnée), et cette variation est estimée à partir de données individuelles ; les études d'impact sanitaire, elles, fournissent une estimation du nombre de cas de maladie attribuable à un facteur ou une intervention donnée à l'échelle de la population, en comparant les nombres de cas attendus entre une situation selon laquelle le facteur ou l'intervention est présent, potentiellement à des niveaux variables, et une autre situation, dite contrefactuelle, sans l'intervention ou le facteur, ou supposant que le facteur a été ramené en dessous d'un certain niveau (scénario dit contrefactuel) [30]. Ces études d'impact constituent une des opportunités pour intégrer les impacts sanitaires chez l'humain avec ceux concernant l'environnement, dans une vision de santé planétaire, comme l'illustrent les travaux sur les cobénéfices sanitaires du changement climatique.

De telles études d'impact peuvent être complétées en estimant le coût économique associé aux nombres de cas de maladies sous chacun des scénarios comparés, ou encore en rapportant les coûts d'une ou plusieurs interventions possibles aux bénéfices sanitaires attendus de ces interventions (étude coût-bénéfice), fournissant autant de paramètres pouvant être considérés par les décideurs devant trancher

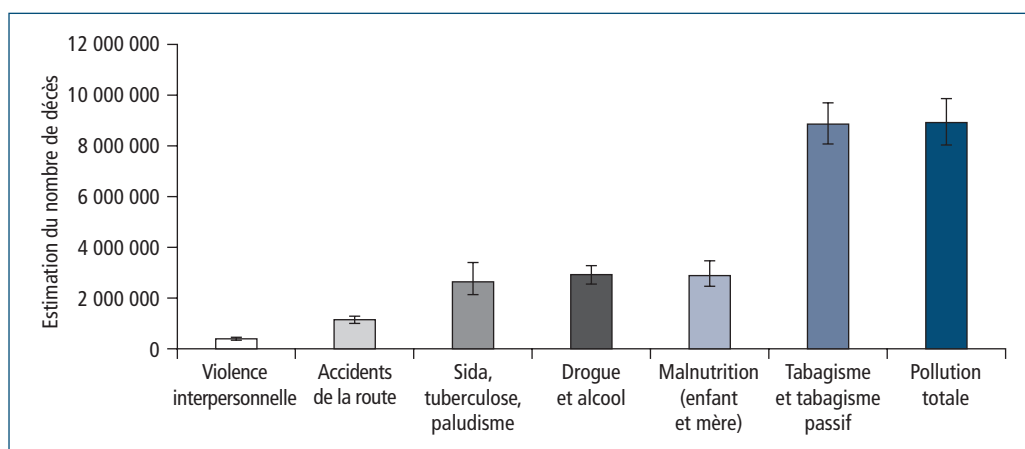


Figure 12.5. Estimation du nombre de décès attribuables annuellement (2019) à différents facteurs de risque à l'échelle de la planète, à partir des travaux du *Global Burden of Disease project* piloté par l'*Institute for Health Metrics and Evaluation* (IHME). La colonne « total pollution » inclut les décès attribuables à la pollution atmosphérique (particules fines et ozone), au plomb, à la qualité de l'eau de boisson et à certains carcinogènes en milieu professionnel (d'après [31]). (Merci de vérifier la traduction)

entre différentes options de gestion d'un risque identifié. Cette quantification économique impose d'attribuer un coût économique à la vie humaine ou aux années de vie perdues, et de recourir à des taux d'actualisation variables et parfois arbitraires pour prendre en compte les impacts futurs ; ces choix soulèvent des questions pratiques et surtout éthiques sur la valeur économique relative de la vie et de la santé entre pays et entre les époques ou générations [32].

Valider les interventions

Ces études *in silico* (car réalisées en comparant des paramètres généralement issus de la littérature, en particulier les relations dose-réponse) sur le bénéfice attendu d'interventions possibles, ne renseignent pas sur leur faisabilité pratique et conditions de réalisation optimales dans un contexte donné. C'est un des objets des études d'intervention, qui peuvent aussi permettre étudier les freins et facteurs favorisant l'efficacité d'interventions de santé publique à visée de prévention. Ces études peuvent être randomisées, à l'échelle individuelle ou du groupe comme discuté plus haut ; elles peuvent, en plus des traitements et dispositifs médicaux, concerner tout instrument (information, comme du marketing social ou un label tel que le Nutri-Score, incitation financière, norme, etc.) et type d'intervention, que ce soit au niveau des comportements individuels, de l'environnement (installation d'installations sportives, de distributeurs de seringues, verdissement des villes) ou de la structure et du fonctionnement du système de santé [33]. De telles études interventionnelles peuvent aussi bien être réalisées avant une décision de santé publique (pour étudier son efficacité potentielle), qu'après (**Figure 12.1**), même si les conditions et moyens le permettant sont rarement prévus par les décideurs.

Synthétiser les connaissances

Les résultats de ces différents travaux de recherche peuvent et doivent être synthétisés, exigence liée à la fois aux fluctuations aléatoires et, souvent, à la faible puissance de chaque étude. Cette synthèse peut aussi permettre de discuter et corriger une des sources de biais les

plus préoccupantes, le biais de publication [34]. Ce travail de synthèse, qui favorise la diffusion et l'utilisation des connaissances établies, est une étape essentielle avant l'action publique, en particulier dans le contexte du déficit de l'action publique de prévention par rapport à l'investissement dans la thérapie. Il peut permettre de dissiper le caractère apparemment incohérent de la littérature sur un sujet (qui est parfois une illusion due à l'erreur consistant à considérer contradictoires une étude rapportant une association entre un facteur et une pathologie donnés et une autre ne la rapportant pas⁶). Il pourrait sembler non spécifique à la santé publique, mais ce n'est pas le cas, du fait notamment que la nature de certains travaux, et notamment les études étiologiques, qui aboutissent à l'estimation d'un paramètre principal, se prêtent à des synthèses quantitatives, sous la forme de méta-analyses, une des formes principales des revues de la littérature [35] ; rappelons aussi que certaines approches de synthèse des connaissances ont émergé du champ de la recherche biomédicale, comme le champ de l'*evidence-based medicine*, qu'on pourrait définir et étendre à la santé publique comme le processus de recherche, d'évaluation systématique et formalisée et d'utilisation des résultats de la recherche comme base des décisions médicales ou de santé publique (définition adaptée de [36]). Une synthèse identifiant les interventions de santé publique dont l'efficacité a été démontrée peut, par exemple, être réalisée [37].

Établissement de la causalité en santé publique

De telles synthèses sont essentielles pour l'établissement du niveau de preuve concernant les connaissances de santé publique. Il est important pour la recherche en santé publique, si elle veut aider la société dans ses décisions, d'être capable d'identifier les associations causales, ou

6. Rappelons qu'une étude ne rapportant pas d'association (« statistiquement significative ») entre A et B ne démontre pas l'absence de lien entre A sur B ; simplement, elle ne permet pas d'exclure que A et B ne soient pas liés.

au moins celles dont le niveau de preuve est le plus élevé. Ce qui nous amène à la question de la causalité. Nous n'avons pas la place pour une présentation détaillée du concept de causalité dans les sciences biomédicales, mais rappelons brièvement à ce sujet quelques contresens parfois entendus :

- il n'existe pas de critères simples pour apprécier le niveau de preuve d'une étude épidémiologique ; en particulier, Bradford Hill n'en a pas publié. Il a proposé, dans un article de 1965, quelques aspects concernant une association, à considérer avant de pouvoir juger de sa nature causale, tout en rappelant que l'établissement de la causalité ne peut être ramené à un ensemble de règles. Ces aspects ont été discutés et, à l'exception de celui de temporalité (selon lequel la cause putative doit précéder la survenue de la maladie), jugés non pertinents par Rothman [5] ;
- rappeler que « corrélation n'est pas causalité » (*correlation is not causation*) dans le contexte d'études de santé publique est inutile, voire faux. En effet, la plupart de la recherche en santé publique ne s'appuie pas sur des corrélations mais sur des associations, généralement débarrassées de certains biais par le *design* de l'étude ou des approches statistiques (par exemple un ajustement sur les facteurs de confusion ou une pondération des données par un score de propension prédisant la probabilité d'exposition) ; c'est seulement dans le cas d'études expérimentales chez l'humain ou l'animal (cas de la toxicologie par exemple) qu'on peut se permettre de faire de l'inférence à partir de simples corrélations, sous certaines conditions ; or ces approches apportent généralement des éléments forts en faveur de la causalité.

Concernant spécifiquement la santé publique, le développement des travaux autour du concept d'inférence causale, par des auteurs comme J. Pearl, D. Rubin, J. Robins ou M. Hernán, au cours des dernières décennies, a permis de formaliser mathématiquement le concept de biais de confusion et d'identifier les approches, y compris à partir de données observationnelles, permettant de fournir une estimation de l'effet causal d'un facteur sur un événement de santé [38, 39]. À un niveau plus agrégé, celui de la

synthèse des connaissances sur un sujet, le niveau de preuve peut être caractérisé dans une approche pragmatique en synthétisant les connaissances issues de chacune des différentes disciplines ayant considéré le sujet, et notamment les travaux épidémiologiques, toxicologiques *in vivo*, et plus mécanistiques, puis en « intégrant » ces connaissances disciplinaires. Différentes approches permettent de combiner le niveau de preuve apporté par chaque champ disciplinaire [40, 41].

Quelques frontières et tensions actuelles de la recherche en santé publique

La recherche en santé publique est traversée de diverses lignes de tension ou structuration. Citons celle entre la recherche structurée à partir des expositions (*exposure-oriented*) et celle orientée à partir des maladies (*outcome-oriented*). La seconde tend à être plus proche de la recherche médicale sur la thérapie et la pathophysiologie, alors que la première est plus proche des sciences de l'environnement et peut-être aussi de la société – mais elles se retrouvent bien sûr et même si l'erreur de mesure n'a pas toujours de conséquences irrémédiables, il est *a priori* pertinent d'accorder tout autant de soin dans la caractérisation des expositions que dans l'identification des altérations biologiques ou la survenue des pathologies. Une autre tension, recouvrant en partie la première, est celle entre un pôle attiré par l'échelle moléculaire, avec le développement réel de l'épidémiologie moléculaire lors du premier quart du XXI^e siècle, la constitution de biobanques pour du phénotypage profond, du génotypage ou une caractérisation de mécanismes biologiques ou d'exposition dans une proportion importante d'études, situation qui était une exception au siècle précédent, et un pôle tourné vers l'échelle planétaire, avec le développement des concepts d'exposome, une seule santé, santé planétaire, santé globale [42] – sans que ces tendances soient opposées (la recherche sur l'exposome, par exemple, étant proche des questions de mécanisme d'action et de la toxicologie). L'épidémiologie moléculaire pose des

défis en termes de biais potentiels – la caractérisation des expositions par des biomarqueurs internes limitant *a priori* l'erreur de mesure mais accroissant les biais de confusion possibles par des caractéristiques individuelles, génétiques, biologiques ou comportementales [43] – et de prise en compte de la malédiction de la haute dimension, qui limite l'aptitude à extraire de l'information à mesure que le ratio entre le nombre de prédicteurs potentiels (souvent de l'ordre du million dans des études *omiques*) et le nombre de sujets croît [44].

Conclusion

Ce chapitre a fourni un aperçu de la diversité des objectifs de la recherche en santé publique, et aussi de ses approches, restreint à celles relevant de l'épidémiologie. Pour des raisons de place, certaines questions importantes n'ont pas été abordées, comme :

- les questions d'éthique et de réglementation, en lien notamment avec la numérisation de cette recherche (protection et partage des données et prélèvements biologiques par exemple) et le développement à venir de l'Espace européen des données de santé ;
- la distinction et l'articulation entre recherche et surveillance en santé publique ;
- le financement de la recherche, qui, en France, repose principalement sur l'IRESP (Institut pour la recherche en santé publique) et l'ANR (Agence nationale de la recherche), sur un nombre relativement faible d'institutions privées, en plus des fonds européens.

Concernant ce dernier point, il faut noter l'absence, à quelques exceptions notables près, de mécanisme liant les financements

disponibles pour la recherche en santé publique aux besoins (quelle que soit la façon dont ceux-ci sont identifiés), contrairement à ce qui se passe pour la recherche thérapeutique pour laquelle, en plus de financements non ciblés indispensables, un financement conséquent est adossé au budget de l'Assurance maladie (l'ONDAM, cas du programme hospitalier de recherche clinique). Il n'existe pas de logique similaire pour la recherche en prévention, malgré le fait qu'elle constitue le deuxième pilier de la préservation de la santé avec la thérapie (une exception notable étant celle du financement de la recherche sur les addictions, soutenue par l'IRESP *via* le fonds de lutte contre les addictions, qui est lui-même alimenté en particulier par les taxes sur la vente de tabac).

Pour conclure, la recherche en santé publique s'est fortement développée à partir de la seconde moitié du ^{xx}e siècle dans un dialogue avec les sciences des données, de l'environnement, du vivant (incluant toxicologie, biologie cellulaire et moléculaire et recherche en pathophysiologie) et de la société (qui, pour beaucoup d'entre elles, font intégralement partie des disciplines de la santé publique), ce qui lui permet de reposer aujourd'hui sur une palette de schémas d'études et d'outils permettant de décrire les pathologies, d'identifier leurs facteurs de risque à court moyen et long terme, de contribuer à comprendre les mécanismes sous-jacents, de quantifier le risque associé, de définir des interventions dans tous les groupes de population permettant de limiter le fardeau de maladie correspondant. Tout ceci est essentiel pour nourrir l'action publique en prévention, dans une vision qui peut aussi intégrer la préservation de l'environnement.

Références

1. Slama R. *Causes et conditions extérieures des maladies et de la santé*. Paris : Collège de France : Fayard, 2022 : 101 p. (Leçons inaugurales du Collège de France).
2. Doll R, Hill AB. Lung Cancer and Other Causes of Death in Relation to Smoking. *Br Med J* 1956 ; 2 : 1071-81.
3. Pope CA, Coleman N, Pond ZA, Burnett RT. Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ Res* 2020 ; 183 : 108924.
4. Slama R. *Le Mal du Dehors. L'influence de l'environnement sur la santé humaine*. Versailles : Quæ, 2017.

5. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. *Modern epidemiology*. Fourth edition. Philadelphia : Lippincott Williams & Wilkins, 2021 : 1 250 p.
6. Tillé Y. *Théorie des sondages : échantillonnage et estimation en populations finies*. 2^e éd. Malakoff : Dunod, 2019. (Sciences sup).
7. Warszawski J, Beaumont AL, Seng R, *et al.* Prevalence of SARS-Cov-2 antibodies and living conditions: the French national random population-based EPICOV cohort. *BMC Infect Dis* 2022 ; 22 : 41.
8. Spira A, Bajos N. *Les comportements sexuels en France*. Paris : la Documentation française, 1993. (Collection des rapports officiels).
9. Tyrrell J, Melzer D, Henley W, Galloway TS, Osborne NJ. Associations between socioeconomic status and environmental toxicant concentrations in adults in the USA: NHANES 2001-2010. *Environ Int* 2013 ; 59 : 328-35.
10. Fillol C, Oleko A, Saoudi A, *et al.* Exposure of the French population to bisphenols, phthalates, parabens, glycol ethers, brominated flame retardants, and perfluorinated compounds in 2014-2016: Results from the Esteban study. *Environ Int* 2021 ; 147 : 106340.
11. Adaptive Platform Trials Coalition. Adaptive platform trials: definition, design, conduct and reporting considerations. *Nat Rev Drug Discov* 2019 ; 18 : 797-807.
12. Dron L, Taljaard M, Cheung YB, *et al.* The role and challenges of cluster randomised trials for global health. *Lancet Glob Health* 2021 ; 9 : e701-710.
13. Herr M, Descatha A, Aegerter P. Essais de non-infériorité et d'équivalence : les points clés de leur méthodologie. *La Revue de Médecine Interne* 2018 ; 39 : 352-9.
14. Hernán MA, Hernández-Díaz S, Robins JM. A Structural Approach to Selection Bias. *Epidemiology* 2004 ; 15 : 615-25.
15. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013 ; 42 : 1012-4.
16. Greenland S, Robins J. Invited commentary: ecologic studies--biases, misconceptions, and counterexamples. *Am J Epidemiol* 1994 ; 139 : 747-60.
17. Goodman M, Naiman JS, Goodman D, LaKind JS. Cancer clusters in the USA: What do the last twenty years of state and federal investigations tell us? *Crit Rev Toxicol* 2012 ; 42 : 474-90.
18. Craig P, Katikireddi SV, Leyland A, Popham F. Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research. *Annu Rev Public Health* 2017 ; 38 : 39-56.
19. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. *Int J Epidemiol* 2013 ; 42 : 1187-95.
20. Wang G, Hamad R, White JS. Advances in Difference-in-differences Methods for Policy Evaluation Research. *Epidemiology* 2024 ; 35 : 628-37.
21. Dimick JB, Ryan AM. Methods for Evaluating Changes in Health Care Policy: The Difference-in-Differences Approach. *JAMA* 2014 ; 312 : 2401.
22. Duggal P, Ladd-Acosta C, Ray D, Beaty TH. The Evolving Field of Genetic Epidemiology: From Familial Aggregation to Genomic Sequencing. *Am J Epidemiol* 2019 ; 188 : 2069-77.
23. Slama R, Vernet C, Nassan FL, Hauser R, Philippat C. Characterizing the effect of endocrine disruptors on human health: The role of epidemiological cohorts. *C R Biol* 2017 ; 340 : 421-31.
24. Abdellaoui A, Yengo L, Verweij KJH, Visscher PM. 15 years of GWAS discovery: Realizing the promise. *Am J Hum Genet* 2023 ; 110 : 179-94.
25. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. *Annu Rev Public Health* 2016 ; 37 : 17-32.
26. Blum MGB, Valeri L, François O, *et al.* Challenges Raised by Mediation Analysis in a High-Dimension Setting. *Environ Health Perspect* 2020 ; 128 : 055001.
27. VanderWeele TJ, Knol MJ. A Tutorial on Interaction. *Epidemiologic Methods* 2014 ; 3.
28. Henry E. *La fabrique des non-problèmes : ou comment éviter que la politique s'en mêle*. Paris : Sciences po, les presses, 2021.
29. Poole C. A history of the population attributable fraction and related measures. *Ann Epidemiol* 2015 ; 25 : 147-54.
30. Rigaud M, Buckers J, Bessems J, *et al.* The methodology of quantitative risk assessment studies. *Environ Health* 2024 ; 23 : 13.

31. Fuller R, Landrigan PJ, Balakrishnan K, *et al.* Pollution and health: a progress update. *Lancet Planet Health* 2022 ; 6 : e535-547.
32. Clowney S. Environmental ethics and cost-benefit analysis. *Fordham Environmental Law Review* 2006 ; 18 : 105-50.
33. Alla F, Cambon L, Ridde V. *Recherche interventionnelle en santé des populations*. IRD Éditions, 2024.
34. Lin L, Chu H. Quantifying Publication Bias in Meta-Analysis. *Biometrics* 2018 ; 74 : 785-94.
35. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009 ; 26 : 91-108.
36. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *BMJ* 1995 ; 310 : 1122-6.
37. Van Der Vliet N, Suijkerbuijk AWM, De Blaeij AT, *et al.* Ranking Preventive Interventions from Different Policy Domains: What Are the Most Cost-Effective Ways to Improve Public Health? *Int J Environ Res Public Health* 2020 ; 17 : 2160.
38. Pearl J, Mackenzie D, Cohen A, Berland F. *Le livre des pourquoi: la nouvelle science des causes et des effets*. Paris : Hermann ; 2024.
39. Hernan M, Robins J. *Causal Inference: What If*. Boca Raton : Chapman & Hall/CRC ; 2020.
40. Hardy A, Benford D, Halldorsson T, *et al.* Guidance on the use of the weight of evidence approach in scientific assessments. *EFSA J* 2017 ; 15.
41. Colzin S, Crépet A, Wies B, *et al.* A plausibility database summarizing the level of evidence regarding the hazards induced by the exposome on children health. *Int J Hyg Environ Health* 2024 ; 256 : 114311.
42. Rabinowitz PM, Pappaioanou M, Bardosh KL, Conti L. A planetary vision for one health. *BMJ Glob Health* 2018 ; 3 : e001137.
43. Weisskopf MG, Webster TF. Trade-offs of Personal Versus More Proxy Exposure Measures in Environmental Epidemiology. *Epidemiology* 2017 ; 28 : 635-43.
44. Hastie T, Tibshirani R, Wainwright MJ. *Statistical learning with sparsity: the lasso and generalizations. First issued in paperback*. Boca Raton London New York : CRC Press, 2020 : 351 p. (Monographs on statistics and applied probability).

