

Les bases de l'épidémiologie

Décembre 2019

Qu'est-ce que l'épidémiologie ?

Par définition, l'épidémiologie est l'étude de la **distribution** des événements de santé et de leurs **déterminants** dans les populations humaines. L'application de ses résultats sert de base aux interventions de prévention, de prise en charge des personnes concernées et d'évaluation des réponses sanitaires et sociales aux problèmes de santé. Il est important de noter que cette discipline, à la différence de la pratique clinique, s'intéresse à l'état de santé des **populations** (population générale ou groupes spécifiques) et non à l'individu.

L'épidémiologie peut poursuivre différents objectifs (les indicateurs seront présentés dans la suite du document) :

Décrire	Décrire et quantifier des événements de santé dans une population Indicateurs : prévalence, incidence
Analyser (étiologie)	Rechercher et mesurer la relation entre une exposition et la survenue d'un événement Indicateurs : risque relatif, odds-ratio
Evaluer	Evaluer l'efficacité d'une intervention (prévention, traitement, etc.) à remplir ses objectifs

La finalité de l'épidémiologie, et de la recherche en santé publique de manière générale, est de tirer profit de ces connaissances pour proposer des interventions et des politiques, fondées sur des **preuves scientifiques**, ou des éléments probants, permettant d'améliorer la santé et le bien-être, et de réduire les inégalités de santé.

Quelques éléments de vocabulaire

Événement (de santé) : phénomène étudié, que ce soit une maladie (morbidité), un dysfonctionnement de l'organisme, un décès (mortalité), ou encore un comportement à risque.

Déterminant : facteur ou exposition (biologique, comportemental, social, environnemental) déterminant la santé des populations. On parle de facteur de risque lorsque celui-ci augmente le risque d'apparition d'un événement de santé, et de facteur protecteur lorsque celui-ci diminue le risque.

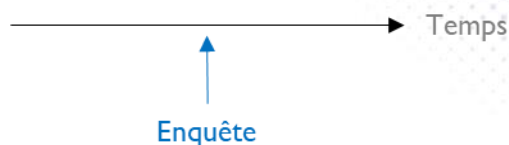
Les différents types d'études épidémiologiques

1) Les études observationnelles

Il s'agit d'études où l'investigateur observe la relation entre un événement et une exposition **telle qu'elle existe** dans la population.

- Les études transversales (études de prévalence)

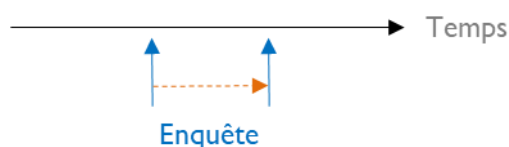
Ce sont des études **descriptives** visant à observer la population de façon ponctuelle. Elles permettent de mesurer l'exposition ou la présence d'un événement **à un instant donné**. Elles produisent une coupe transversale ou « photographie » de l'état de la situation, au moment de l'étude, dans la population étudiée.



Exemple : l'enquête OPPIDUM¹, visant à surveiller la consommation de substances psychoactives par des sujets présentant une pharmacodépendance, est une étude transversale où les données sont recueillies ponctuellement chaque année, au mois d'octobre.

- Les études d'incidence (longitudinales non comparatives)

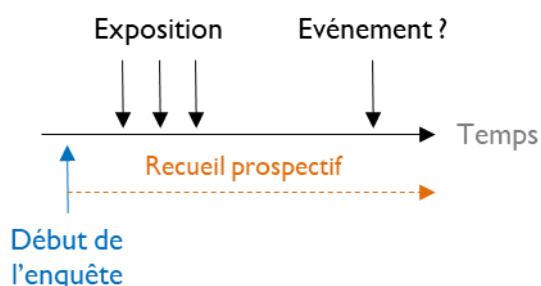
Ce sont des études **descriptives** visant à observer la population de manière répétée pendant un **intervalle de temps donné**. Cette observation récurrente permet de suivre l'évolution d'une exposition ou d'un événement au cours du temps.



Exemple : dans le cas des maladies à déclaration obligatoire, les études réalisées à partir de registres de morbidité (relatifs à un ou plusieurs événements de santé), comme par exemple le système de surveillance du VIH en France qui recense en continu depuis 2003 les nouveaux diagnostics d'infection par le VIH.

- Les études de cohorte

Ce sont des études longitudinales **analytiques**, qui consistent à suivre dans le temps un groupe d'individus partageant un certain nombre de caractéristiques communes, afin d'identifier la survenue d'un événement. Le suivi individuel des individus permet de recueillir prospectivement l'exposition à plusieurs facteurs au cours du temps. L'objectif est alors de comparer l'évolution de la survenue de l'événement **entre individus exposés et non-exposés** et de mettre en évidence des facteurs de risque ou de protection vis-à-vis de l'événement étudié.



Exemple : la cohorte « CSAPA sentinelle », qui va être mise en place à l'ANPAA dans neuf CSAPA (10, 18, 21, 30, 33, 58, 66, 71, 89). Elle va permettre de suivre dans le temps les consultants, et d'identifier les facteurs individuels et structurels favorables à une amélioration de leur santé et de leur accompagnement.

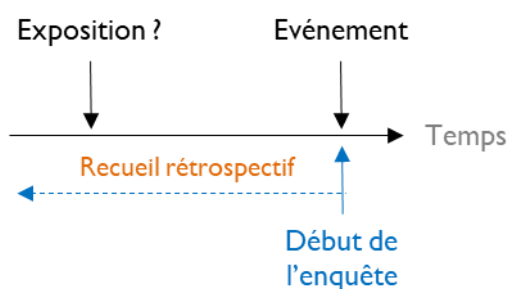
A noter qu'il existe des cohortes historiques où l'exposition a déjà eu lieu au moment du début de l'enquête.

¹ « Observation des Produits Psychotropes Illicites ou Détournés de leur Utilisation Médicamenteuse », enquête réalisée par le réseau des centres d'évaluation et d'information sur la pharmacodépendance (CEIP)

Exemple : la cohorte Gazel mise en place en 1989 par l'Inserm en coopération avec plusieurs services d'EDF-GDF qui permet de suivre l'état de santé de 20 000 salarié.es.

- Les études cas-témoins

Ce sont des études longitudinales **analytiques** où un groupe de cas (présentant l'événement) et un groupe de témoins (ne présentant pas l'événement) sont recrutés. Des données concernant leur exposition sont alors recueillies de manière rétrospective. Dans ce type d'étude, il est important que les témoins soient représentatifs de la population source d'où sont issus les cas, c'est-à-dire qu'ils soient comparables aux cas, à l'exception de l'événement étudié. L'objectif est alors de comparer le niveau des facteurs d'exposition **entre les cas et les témoins** et de mettre en évidence des facteurs de risque ou de protection vis-à-vis de l'événement étudié.



Exemple : une étude-cas témoins réalisée par l'Inserm avait pour objectif d'évaluer si les troubles anxieux étaient davantage fréquents chez les femmes présentant des troubles du comportement alimentaire (TCA)². Dans cet exemple, les troubles anxieux ont été mesurés chez des femmes présentant des TCA (cas) et chez des femmes n'en présentant pas (témoins).

2) Les études interventionnelles

Il s'agit d'études où l'investigateur **intervient** sur le statut d'exposition (nature, moment) des individus. La survenue de l'événement est alors comparée entre le groupe exposé et le groupe contrôle (sans exposition).

- Les études expérimentales

Les études expérimentales sont **randomisées**, c'est-à-dire que la répartition des individus dans les groupes est réalisée par tirage au sort. Ces études ou essais randomisés peuvent être qualifiés :

- d'ouverts, lorsque le statut d'exposition est connu de l'individu ;
- en simple aveugle, lorsque le statut d'exposition est inconnu de l'individu ;
- en double aveugle, lorsque le statut d'exposition est inconnu de l'individu et de l'investigateur.

Exemple : un essai thérapeutique permettant d'évaluer l'effet d'un traitement, l'exposition étant soit le traitement, soit le placebo.

- Les études quasi-expérimentales

Les études quasi-expérimentales sont **non randomisées**, c'est-à-dire que l'attribution des groupes est soit réalisée par l'investigateur, soit fondée sur l'autosélection (choix des participants).

Les enquêtes **avant-après** permettent de comparer les situations avant et après une intervention. Dans certains cas, chaque individu peut être son propre témoin. Les enquêtes **ici-ailleurs** permettent de

² <https://epidemiologie-france.aviesan.fr/en/epidemiologie-france/fiches/case-control-study-on-patients-with-addictive-behaviour/fre-fr>



comparer, au même moment, les situations dans deux groupes (exposé et non exposé) distincts géographiquement.

Exemple : l'évaluation à l'ANPAA du programme d'éducation thérapeutique du patient « Choizitaconso© » en RdRD en alcoologie est une étude quasi-expérimentale où le devenir d'un groupe ayant reçu l'intervention (CSAPA d'Avignon) est comparé avec celui d'un autre groupe ne l'ayant pas reçue (CSAPA de Digne-les-Bains et clinique Saint-Barnabé à Marseille), et ceci avant et après l'intervention.

Les mesures en épidémiologie

1) La prévalence

La prévalence d'un événement correspond au **nombre de cas** dans une population **à un moment donné**, indépendamment de l'ancienneté de l'événement. Elle est mesurée lors d'enquêtes transversales et permet de mesurer la fréquence d'un événement dans une population à un moment précis.

Exemple : les résultats du Baromètre de Santé Publique France indiquent qu'en 2018, la prévalence du tabagisme quotidien des 18-75 ans était de 25,4% (28,2% des hommes et 22,9% des femmes)³.

2) L'incidence

L'incidence d'un événement correspond au nombre de **nouveaux cas** pour un événement dans une population donnée **pendant une période donnée**. Elle est mesurée lors d'enquêtes longitudinales et permet de mesurer la proportion d'individus présentant l'événement sur une période de temps spécifiée.

Exemple : en 2015, l'incidence des cancers attribuables à la consommation d'alcool a été estimée à 7,9% en France⁴.

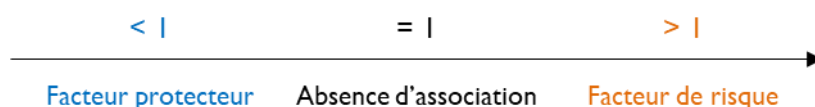
3) Le risque relatif et l'odds-ratio

Le risque relatif et l'odds-ratio (ou rapport des côtes) permettent de mesurer l'**association** entre un facteur d'exposition et un événement. Ces mesures permettent de quantifier la relation mais ne garantissent pas que celle-ci soit causale.

Risque relatif (RR) : la survenue de l'événement est comparée entre les individus exposés et les individus non-exposés. Cet indicateur est estimable dans les études de cohorte.

Odds-ratio (OR) : la proportion d'individus exposés est comparée entre les cas et les témoins. Cet indicateur est estimable dans tous les types d'enquêtes, notamment dans les études cas-témoins.

Ces deux indicateurs peuvent alors s'interpréter en fonction de leur valeur (le détail de leur calcul est disponible à la fin du document) :



³ Andler R, Richard JB, Guignard R, Quatremère G, Verrier F, Gane J, Nguyen-Thanh V. Baisse de la prévalence du tabagisme quotidien parmi les adultes : résultats du Baromètre de Santé publique France 2018. Bull Epidémiol Hebd. 2019;(15):271-7. http://beh.santepubliquefrance.fr/beh/2019/15/2019_15_1.html

⁴ Shield KD et al. New cancer cases in France in 2015 attributable to different levels of alcohol consumption. Addiction. 2018 Feb;113(2):247-256. doi: 10.1111/add.14009.

Dans l'exemple du risque relatif, une valeur supérieure à 1 indique que l'incidence de l'événement est supérieure chez les individus exposés par rapport à celle calculée chez les individus non-exposés. Le facteur d'exposition augmente ainsi le risque de survenue de l'événement, on parle alors de facteur de risque. A noter que plus la valeur du RR ou de l'OR est éloignée de 1, plus l'association est forte.

Exemple : les résultats de l'étude Méthaville sur la faisabilité et l'efficacité de l'initiation d'un traitement par méthadone en médecine de ville ont montré que la méthadone était associée à une diminution de l'intensité de la douleur à 6 mois (OR = 0,29) et 12 mois (OR = 0,30) de suivi, alors que le risque de suicide était associé positivement à la douleur (OR = 5,90)⁵.

L'incertitude

L'épidémiologie passe le plus souvent par l'analyse de données recueillies à partir d'un **échantillon** de la population source, plus ou moins vaste selon l'objet de la recherche. Cet échantillon doit être **représentatif** de la population source (critères socio-démographiques à définir, le plus souvent âge, sexe, origine géographique, etc.).



1) La variabilité

Les mesures de risque (prévalence, incidence) et d'association (RR, OR) sont des **estimations** obtenues à partir de ces échantillons. Si ces indicateurs étaient mesurés à partir d'autres échantillons, les valeurs obtenues seraient alors très certainement différentes, bien que proches et se répartissant aléatoirement autour de la « bonne » valeur. C'est ce qu'on appelle les **fluctuations d'échantillonnage** (la taille de 30 sujets est une référence minimale le plus souvent).

Il convient alors de donner ces estimations avec leur **intervalle de confiance à 95%**, c'est-à-dire l'intervalle dans lequel la vraie valeur a de fortes chances (95%) de se situer. Un intervalle de confiance à 95% correspond à un risque d'erreur de 5%, ce qui est classiquement utilisé en épidémiologie. A noter que plus la taille de l'échantillon est grande, plus le risque d'erreur diminue.

2) Les biais

Un biais est une **erreur systématique**, se répétant sur chaque échantillon, qui fausse les résultats (une association entre une exposition et un événement de santé par exemple) dans un sens donné. Les études présentant souvent des biais, il est important de les anticiper dès la conception de l'étude afin de les contrôler. A noter que la taille de l'échantillon n'aura aucun impact sur la minimisation d'un biais.

- Les biais de sélection

Ces biais résultent de la manière de **choisir** et de **suivre** les individus de l'échantillon. C'est le cas par exemple d'une étude visant à mesurer l'impact d'une campagne de prévention sur les infections sexuellement transmissibles (IST), et où les individus sont recrutés dans des cabinets de médecins généralistes. Le nombre d'individus présentant des IST sera plus important qu'en population générale car les « malades » ont tendance à plus consulter.

⁵ Nordmann S et al. Pain in methadone patients: Time to address undertreatment and suicide risk (ANRS-Methaville trial). PLoS One. 2017 May 17;12(5):e0176288. doi: 10.1371/journal.pone.0176288.

- Les biais de mesure (ou de classement)

Ces biais sont dus à une erreur dans le « **classement** » (nomenclature imprécise ou défaut de spécification) des expositions ou des individus présentant l'événement. Il peut s'agir par exemple d'un biais de mémorisation dans les études cas-témoins où les individus atteints d'une pathologie (cas) peuvent avoir tendance à se souvenir davantage des expositions que les individus ne présentant pas la pathologie (témoins). Un biais lié à la subjectivité de l'enquêteur peut également mener à une mesure erronée de l'exposition si celui-ci cherche avec plus d'acuité une exposition passée chez un cas que chez un témoin.

- Les biais de confusion

Il existe un biais de confusion lorsque qu'un facteur est **simultanément associé à l'exposition et à l'événement**, pouvant être à l'origine d'une fausse association. C'est le cas par exemple du tabac dans l'association entre alcool et cancer du poumon. En effet, les fumeurs sont plus à risque de développer un cancer du poumon, mais les personnes consommant de l'alcool ont plus tendance à fumer que les personnes ne consommant pas d'alcool. A noter que ces biais peuvent être redressés à l'étape d'analyse des données.

L'interprétation des résultats

1) Lien de causalité

Seules les études expérimentales randomisées peuvent prouver que l'association entre l'exposition et l'événement est **causale**. Dans le cas des études observationnelles, il ne s'agit que d'une **présomption** de causalité. Des critères peuvent néanmoins être pris en compte afin d'évaluer la causalité, notamment :

- la temporalité : l'exposition doit précéder l'événement ;
- la force de l'association, mesurée par le risque relatif ou l'odds-ratio ;
- la relation dose-effet (relation entre le nombre de cigarettes fumées par jour et le risque de cancer du poumon par exemple) ;
- la reproductibilité des résultats : l'association est retrouvée dans d'autres études, notamment dans d'autres populations ;
- la plausibilité biologique de l'association.

Ainsi, affirmer un lien causal entre un facteur d'exposition et un événement nécessite **plusieurs études** qui pourront être comparées entre elles. Les difficultés méthodologiques et l'incertitude étant inhérentes à toute enquête épidémiologique, ce sont l'ensemble des résultats qui permettront d'accumuler suffisamment de preuves pour prouver qu'un facteur est la cause de la survenue d'un événement.

2) Généralisation des résultats

Les études épidémiologiques étant le plus souvent réalisées à partir d'un **échantillon**, il est important que celui-ci soit **représentatif** de la population source pour que les résultats puissent être **généralisés** à l'ensemble de la population. La représentativité de l'échantillon dépend essentiellement du mode de sélection des individus : si les individus sont tirés au sort, l'échantillon sera, en principe, représentatif. Ainsi, les résultats des études et leur généralisation seront d'autant plus fiables que l'échantillon est représentatif de la population source.

Quelques références

- L'épidémiologie sans peine, 2^{ème} édition. Marcel Goldberg. Frison Roche, 2000, 194 p. ISBN : 2876710390.
- Statistiques – Epidémiologie, 4^{ème} édition. Thierry Ancelle. Maloine, 2017, 352 p. ISBN : 2224035225.
- Vidéos : « Comprendre les données probantes issues de la recherche » réalisées par le Centre de collaboration nationale des méthodes et outils (Canada) : <https://www.nccmt.ca/fr/developpement-des-capacites/videos#url>

Pour aller plus loin

Calcul du risque relatif :

	Événement	Pas d'événement	
Exposé	a	b	a+b (nombre d'individus exposés)
Non-exposé	c	d	c+d (nombre d'individus non-exposés)
	a+c	b+d	

$$Risque\ relatif\ (RR) = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}} \rightarrow \begin{array}{l} \text{Incidence de l'événement chez les exposés} \\ \text{Incidence de l'événement chez les non-exposés} \end{array}$$

Calcul de l'odds-ratio :

	Événement	Pas d'événement	
Exposé	a	b	a+b
Non-exposé	c	d	c+d
	a+c (nombre de cas)	b+d (nombre de témoins)	

$$Odds - ratio\ (OR) = \frac{\frac{a}{(a+c)}}{\frac{b}{(b+d)}} \rightarrow \begin{array}{l} \text{Odds d'être exposé chez les cas} = \text{probabilité d'être exposé} \\ \text{divisé par la probabilité de ne pas être exposé chez les cas} \\ \text{Odds d'être exposé chez les témoins} = \text{probabilité d'être exposé} \\ \text{divisé par la probabilité de ne pas être exposé chez les témoins} \end{array}$$