

Etudes épidémiologiques

Date de la mise à jour : 31.12.2006

Les bases de l'épidémiologie

Définition

L'**épidémiologie** est la science des problèmes de santé dans une population.

Initialement centrée sur l'étude des maladies infectieuses, et en particulier sur celle des épidémies, l'épidémiologie s'est intéressée à partir des années 1950 à d'autres domaines, en particulier celui des cancers ou encore des maladies cardio-vasculaires.

Stricto sensu, l'épidémiologie est l'étude de la distribution des problèmes de santé (maladies, accidents, ...) et de leurs déterminants dans les populations humaines, cette étude ayant pour but la prévention des problèmes de santé.

Cette définition fait donc appel à plusieurs notions.

✓ **La distribution** : il s'agit de la fréquence du problème, de sa variation en fonction de différents paramètres : temps, espace, caractéristiques socio-démographiques (âge, sexe, etc.)

✓ **Les problèmes de santé** : ils peuvent correspondre à des maladies (infectieuses ou non, aiguës ou chroniques), des handicaps, des traumatismes, des suicides, des indicateurs cliniques (tension artérielle, poids) ou des indicateurs biologiques (hémoglobinurie, cholestérolémie)

✓ **Les déterminants** : selon les cas, il s'agit de facteurs environnementaux (agents physiques : température, radiations, bruit ; agents chimiques : polluants de l'air, de l'eau), liés aux habitudes de vie (vie professionnelle, alimentation, alcool, tabac, ...), génétiques, etc.

✓ **Les populations humaines** : l'épidémiologie s'intéresse au groupe et non à l'individu. C'est une grande différence par rapport à la médecine clinique. Ces populations sont définies d'après un critère géographique, temporel, socio-démographique, ...

✓ **La prévention** : elle repose sur la mise en place d'interventions, correspondant à de

nouvelles thérapeutiques, des campagnes de vaccination, d'information, de dépistage, ...

Grâce à une meilleure connaissance des problèmes de santé, la finalité de l'épidémiologie est donc de favoriser une meilleure santé des populations, la santé étant non pas seulement l'absence de maladie, mais « un état de complet bien-être physique, mental et social », selon la définition de l'Organisation Mondiale de la Santé (OMS).

Les domaines de l'épidémiologie

L'épidémiologie comprend plusieurs domaines théoriques. Généralement, on en distingue trois :

✓ **Epidémiologie descriptive** : son but est de décrire un phénomène de santé sur une population humaine en fonction du temps, de l'espace et des caractéristiques de cette population.

L'épidémiologie descriptive est indispensable pour appréhender l'importance d'un problème de santé, pour guider la planification sanitaire et les programmes de santé, et pour générer des hypothèses sur les facteurs de risque des maladies.

Ex : on souhaite étudier l'évolution des tumeurs cérébrales au cours du temps en France.

✓ **Epidémiologie analytique** : elle recherche les causes des maladies, établit les lois et conditions de leur propagation et recherche les moyens de contrôle et de lutte. L'identification des facteurs de risque des maladies est l'un des aspects les plus importants de l'épidémiologie. Elle permet de déterminer les groupes de population à haut risque et donc de cibler les interventions sanitaires (campagnes d'information, de dépistage...). L'épidémiologie analytique repose toujours sur la comparaison de plusieurs groupes.

Ex : on souhaite étudier si l'utilisation des téléphones mobiles est un facteur de risque des tumeurs cérébrales.

✓ **Epidémiologie évaluative** : elle vise à évaluer les actions de traitement ou de prévention.

Etudes épidémiologiques

En fonction de l'objectif recherché, plusieurs types d'études sont utilisés en épidémiologie.

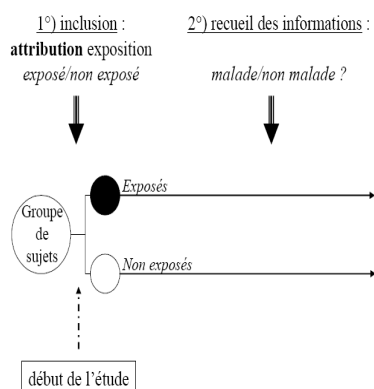
Toutes les études épidémiologiques peuvent être classées selon qu'il s'agit d'une expérimentation ou non. Cette distinction est importante car elle a des implications sur la mise en place de l'étude elle-même, son analyse, mais aussi l'interprétation et la portée des résultats.

✓ **Les études expérimentales** : l'équipe réalisant la recherche, maîtrise les conditions d'exposition des sujets (qui sera exposé, à quoi, comment, quand, où). Il s'agit d'une expérimentation. Les groupes exposés et non exposés à la caractéristique seront suivis puis comparés en ce qui concerne l'incidence de l'événement étudié. L'affectation d'un sujet dans l'un ou l'autre des groupes comparés peut être tirée au sort ou non. Il peut s'agir des essais thérapeutiques qui permettent d'évaluer l'efficacité ou la tolérance de traitements médicamenteux. L'exposition étant, dans ce cas, la prise d'un médicament.

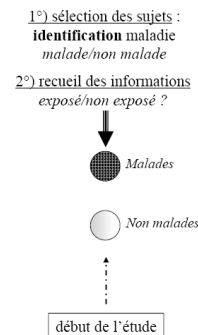
✓ **Les études non expérimentales** : l'équipe réalisant la recherche « observe » la réalité, sans intervenir sur les conditions d'exposition des sujets étudiés. C'est la raison pour laquelle ces études sont également appelées **études d'observation**.

Classiquement, on distingue :

- **Les études de cohortes** : ces études sont également appelées études **exposés/non exposés** car deux groupes de sujets (au moins), l'un exposé, l'autre non exposé à la caractéristique, ne présentant pas initialement la maladie étudiée, sont suivis puis comparés pour l'incidence (nombre de nouveaux cas) de la maladie



- **Les études cas-témoins** : les deux groupes étudiés sont, d'une part, des sujets **malades** (les « cas ») et, d'autre part, des sujets **non malades** (les « témoins »). Ces deux groupes seront comparés pour leur exposition antérieure à une ou plusieurs expositions.



Les études cas-témoins ne permettent pas de calculer l'incidence de la maladie parmi les sujets étudiés puisque le nombre de cas et de témoins est fixé par l'équipe réalisant la recherche.

- **Les études transversales**

Leur principe est simple : il s'agit de mesurer chez les individus d'une population donnée un ensemble de caractéristiques **au même moment**. Ce sont des études de courte durée, ne permettant d'estimer que des mesures de prévalence (nombre de cas à un moment donné).

Elles sont utilisées généralement pour des objectifs descriptifs. Elles sont également souvent utilisées à visée analytique, mais les résultats obtenus sont plus difficilement interprétables car l'exposition à la caractéristique et la présence de la maladie sont mesurées au même moment.

- **Les études écologiques**

Elles étudient l'association (la corrélation) entre les variations d'indicateurs **collectifs** d'exposition et de santé, lorsque l'on ne peut disposer de données à l'échelon individuel. Elles sont donc différentes des autres types d'étude utilisés en épidémiologie analytique car elles ne tiennent pas compte de l'exposition ou de l'état de santé de **chaque individu** appartenant aux groupes comparés, mais du **niveau moyen** d'exposition et de l'état de santé des populations, ce qui se traduit par l'emploi du terme « écologique ». On peut étudier la même population à différentes périodes (variations temporelles) ou plusieurs populations de différentes zones géographiques au cours de la même période (variations géographiques).

Ce schéma d'étude a beaucoup été utilisé pour étudier la relation entre la proximité d'antenne relais et la mortalité par cancer autour de ces antennes ([cf. fiche de synthèse stations de base](#)).

Facteurs de risque

Une des principales questions qui se posent à l'épidémiologiste concerne la recherche des **facteurs de risque** des maladies, c'est-à-dire la recherche de facteurs qui expliquent une incidence plus élevée (ou moins élevée) du problème de santé chez les sujets exposés que chez les sujets non exposés. Cela fait appel à l'épidémiologie analytique.

Cette recherche repose sur l'étude de l'existence d'une **association** (ou d'un lien, d'une relation) entre l'exposition à cette caractéristique et la maladie. Il est toujours nécessaire de **comparer** deux groupes (au moins).

Les études de cohortes et les études cas-témoins sont les plus couramment utilisées pour la recherche d'une association. Quelle que soit le schéma d'étude utilisé, leurs résultats doivent pouvoir être présentés dans un **tableau de contingence**, classant les individus selon :

- leur statut vis-à-vis de la **maladie** (malade/non malade)
- leur exposition au **facteur de risque** étudié (exposé/non exposé).

Tableau 1 : répartition des sujets d'une population selon le statut malade/non malade et exposé/non exposé

	Malades	Non malades	Total
Exposés	a	b	E_1
Non exposés	c	d	E_0
Total	M_1	M_0	T

La mesure d'association

La mesure d'association repose toujours sur la comparaison de deux groupes pour la mesure observée. Deux possibilités s'offrent à l'épidémiologiste.

- Soit on compare un groupe **exposé** à une caractéristique, et un autre **non exposé** à cette caractéristique (groupe de référence). Il s'agit d'**études de cohorte**. On calcule le nombre de nouveau cas (incidence) de la maladie étudiée dans chacun des deux groupes, et on écrit le **risque relatif (RR)**

- Soit on compare un groupe de sujets malades (« cas ») et un groupe de sujets ne présentant pas (et n'ayant jamais présenté) cette maladie (groupe de référence : « témoins »). Il s'agit d'**études cas-témoins**

La mesure d'association la plus fréquemment utilisée est alors le **rapport de cotes** d'exposition, RC (odds ratio en anglais, OR). Pour le calculer, il faut d'abord déterminer la cote d'exposition chez les cas et chez les témoins. La cote d'exposition

est le rapport entre la probabilité d'être exposé et la probabilité de n'être pas exposé.

En pratique, dans une population où l'événement est rare (prévalence < 5 %), le rapport de cote est une bonne approximation du risque relatif. L'interprétation du rapport de cote est alors similaire à celle du risque relatif.

Interprétation des résultats

Théoriquement, l'étendue des valeurs du RR et du RC peut aller de 0 à plus l'infini.

- **RR ou RC = 1,00**. Il n'y a pas d'association entre le facteur de risque et la maladie.

- **RR ou RC > 1,00**. L'association est positive : le risque de maladie est plus fort chez les sujets exposés que chez les sujets non exposés. On parle aussi de **facteur de risque**.

Ex : RR = 1,30. Un sujet exposé a un risque d'être malade 1,30 fois plus élevé (ou bien encore a un risque supérieur de 30 % d'être malade) qu'un sujet non exposé

- **RR ou RC < 1,00**. L'association est inverse. Le risque de maladie est moins fort chez les sujets exposés que chez les sujets non exposés. On parle aussi de **facteur protecteur**.

*Ex : RR = 0,65. Cela signifie que le risque d'un sujet exposé est 0,65 fois celui d'un sujet non exposé ou encore un sujet exposé a un risque plus faible de 35 % [(1 - 0,65) * 100] d'être malade qu'un sujet non exposé.*

ATTENTION : la mesure d'association en elle-même ne suffit pas à déterminer s'il existe une association entre une exposition et une maladie, mais qu'il faut également prendre en compte d'autres notions, telles que **l'intervalle de confiance** de la mesure d'association et de la **valeur du test statistique**, pour déterminer si l'association observée peut être due au seul hasard, ou, au contraire, est stable, c'est-à-dire «statistiquement significative».

L'intervalle de confiance

En effet, un RR de 1,2 ou de 0,9 ne veut pas forcément dire que dans le premier cas le facteur est un facteur de risque (augmentant de 20% la maladie) et dans le 2^{ème} cas, un facteur protecteur (diminuant de 10% la maladie)

Il est indispensable de construire à partir de la valeur trouvée et de la taille de l'étude (nombre de sujet), un intervalle de confiance avec l'idée simple qui est que si cet intervalle contient la valeur 1, il n'existe pas de relation statistique.

Ainsi, un RC = 1,30 avec un intervalle de confiance compris entre 0,5 et 2,00 (noté [0,5 ;

2,00]) et comprenant la valeur 1, signifie que bien qu'on ait mis en évidence une augmentation du risque, celui-ci n'est pas statistiquement significatif.

En revanche, un RR ou RC = 1,30 avec un intervalle comprenant la valeur 1 comme par exemple [1,10 ; 1,90], signifie qu'il existe une augmentation du risque **statistiquement significative**.

La puissance de l'étude

Comme on vient de le voir pour l'intervalle de confiance, la taille de l'échantillon va conditionner la puissance de l'étude

Si l'étude est trop peu puissante, on risque de ne pouvoir mettre en évidence une relation statistique même si elle est vraie

Or dans le cas de la téléphonie mobile, les risques évoqués sont, s'ils sont vrais, très faibles individuellement, il faut donc de très grosses études pour les mettre en évidence ce qui paradoxalement conduit souvent les scientifiques qui ne montrent pas de relation à conclure : il y a peut-être une relation mais nous ne l'avons pas mise en évidence, une étude plus puissante pourrait éventuellement le faire... et le doute s'installe.

Les biais

Si une relation statistique est démontrée, cela ne veut pas dire que l'exposition la cause de survenue de la maladie.

Il y a entre la relation statistique et l'imputation causale tout un faisceau d'erreurs possibles, on parle alors de biais qu'il est important de prendre en compte avant de conclure de manière définitive.

Schématiquement, on distingue 3 grands types de biais :

✓ **Les biais de sélection** : il s'agit de biais provenant d'une distorsion lors de la constitution de l'échantillon d'étude.

Ex : Le choix des témoins : dans le cas d'études cas-témoins (très utilisées pour étudier l'exposition aux téléphones mobiles), il est très délicat de choisir de « bons témoins » et l'on peut être amenés à des erreurs importantes si on n'y prend pas garde.

Imaginons par exemple 100 cancers du cerveau choisis dans des hôpitaux urbains qui ont un certain âge et un certain profil socio-économique. Si on choisit comme témoin des gens sans cancers du cerveau mais très jeunes ou encore avec un

statut socio-économique faible ou vivant à la campagne, on risque de trouver une relation entre les radiofréquences et le cancer entièrement expliquée par la différence d'âge ou le statut socio-économique entre les cas et les témoins.

✓ **Les biais d'information**

Ces biais concernent l'ensemble des informations recueillies pour la comparaison des groupes de l'étude. Il s'agit donc des erreurs commises en classant les sujets selon le statut exposé/non exposé et selon le statut malade/non malade.

Ex : La recherche de l'exposition aux téléphones mobiles : dans le cadre de la téléphonie mobile, il n'est pas toujours simple de retrouver rétrospectivement par le seul interrogatoire les précisions concernant l'exposition (fréquence, durée d'utilisation du téléphone mobile) ce d'autant que l'on sait que les malades ont toujours une meilleure « mémoire » que les non malades et ont même souvent tendance à exagérer leur exposition. Ainsi on peut s'attendre à ce que des personnes atteintes d'un cancer du cerveau exagèrent leur utilisation de mobiles alors que les témoins pourraient au contraire les négliger.

✓ **Les biais de confusion** : un biais de confusion survient lorsqu'un **facteur externe**, encore appelé **tiers facteur**, entraîne une distorsion de l'estimation de l'effet de l'exposition sur la maladie.

Critères de causalité

Enfin, lorsque la relation statistique est prouvée, des arguments classiques de causalité doivent être pris en compte. Ce sont les suivants :

Critères internes à l'étude

- ✓ Séquence dans le temps (l'exposition à l'agent causal précède l'effet)
- ✓ Force de l'association
- ✓ Spécificité de la cause et de l'effet (cause présente seulement chez les malades et pour cette maladie)
- ✓ Relation dose effet entre exposition et maladie
- ✓ Cohérence interne (prise en compte des biais de classement, de sélection et des facteurs de confusion).

Critères externes à l'étude

- ✓ Constance de l'association et reproductibilité dans d'autres études
- ✓ Plausibilité biologique
- ✓ Cohérence avec les connaissances générales

Lorsque l'on cherche si le téléphone mobile, ou plutôt les radiofréquences émises par ces téléphones sont un risque pour la santé, on peut mettre en place différentes études épidémiologiques.

- **Etude de cohorte** : on peut suivre dans le temps le groupe exposé aux radiofréquences (RF+) et le groupe non exposé (RF-) et observer le nombre de nouveaux cas (incidence) de cancer du cerveau. On peut alors calculer le risque relatif et son intervalle de confiance.

Cette méthode est quasi idéale sauf que bien sûr, elle est très longue (et donc très coûteuse) car elle peut nécessiter plusieurs années et elle impose un certain nombre de garanties (minimum de perdus de vue, difficultés lorsqu'un individu change de statut passant de non exposé aux RF- à exposé aux RF ou l'inverse, etc.).

- **Etude cas-témoins** : dans ce cas on peut imaginer d'interroger un certain nombre de malades sur leur exposition aux téléphones mobiles et faire de même pour les non malades. Comme l'échantillon n'est pas représentatif, on ne peut calculer le risque relatif. Celui-ci est alors approximé par ce qu'on appelle l'odds Ratio (ou rapport de cote) cité souvent dans les articles sous la forme OR et que l'on peut (même si ce n'est pas tout à fait exact !) assimiler au risque relatif.

Cette technique épidémiologique a évidemment le très grand avantage d'être assez courte, n'imposant aucun suivi ce qui lui donne un immense avantage sur l'étude de cohorte.

Ces études cas-témoins sont d'ailleurs les plus souvent utilisées dans le cadre de la téléphonie mobile comme l'étude INTERPHONE par exemple ([cf. fiche étude interphone](#)) ou les autres études sur téléphone mobile et tumeurs du cerveau ([cf. fiche tumeurs de la tête et du cou](#))

MAIS, cette méthode est beaucoup moins « fiable » que la méthode prospective et elle est exposée à de nombreuses erreurs (appelés aussi biais).

On retient 2 biais majeurs que nous l'avons vu précédemment :

- Le choix des témoins
- La recherche de l'exposition

On comprend bien que ces 2 biais majeurs font que l'épidémiologie préfère les enquêtes de cohortes

qui malheureusement sont beaucoup plus rares et difficiles à mettre en place.

Il faut encore insister sur un point.

➤ L'impossible réfutation

Les études épidémiologiques évoquées ci-dessus sont toutes comparatives et l'analyse statistique qui fournira la réponse peut montrer une différence avec un certain risque d'erreur contrôlé (le risque de 1^{ère} espèce) et dans ce cas la relation statistique est prouvée.

Cependant, si la différence n'existe pas (par exemple entre le groupe exposé et non exposé), le statisticien ne pourra jamais démontrer cette inexistence.

Tout au plus pourra-t-il conclure qu'il n'a pu mettre en évidence une relation ce qui n'est en rien synonyme de « une non relation est démontrée » !

Là est toute l'ambiguïté et la difficulté du problème surtout sous le dogme de principe de précaution

- Si quelque chose existe, on le prouve (si on peut).
- Si quelque chose n'existe pas..., on dit que « cela reste possible ». et ce doute devient facilement un début de preuve !