



L'épidémiologie pour tous

Epidémiologie descriptive : Qualité et interprétation des résultats d'une enquête à visée quantitative

Rédigée par B. Toma, B. Dufour, J. Rivière et J.J. Bénet ; validée par le bureau de l'AEEMA

Une enquête descriptive à visée quantitative a pour objectif de connaître la [prévalence](#) d'une maladie dans une [population](#). Elle repose souvent sur l'emploi d'un test de [dépistage](#). Compte tenu de l'importance de la taille de la population, elle est réalisée, en général, sur un [échantillon](#) de celle-ci.



La qualité des résultats de l'enquête dépend, entre autres, des caractéristiques du test de dépistage utilisé (cf. Fiches [Sensibilité d'un test de dépistage](#) et [Spécificité d'un test de dépistage](#)) et de l'échantillonnage.

Test de dépistage

En fonction de la [sensibilité](#) et de la [spécificité](#) du test utilisé (cf. Fiches [Sensibilité d'un test de dépistage](#) et [Spécificité d'un test de dépistage](#)) ainsi que de la prévalence réelle de la maladie dans la population, la fréquence des résultats faussement négatifs ou faussement positifs est variable et, donc, la [prévalence apparente](#) (total des résultats positifs), plus ou moins éloignée de la [prévalence réelle](#) (inconnue, sauf à soumettre à un test parfait la totalité de la population).



*Si l'on dispose d'informations fiables sur la sensibilité et la spécificité du test utilisé dans l'enquête, il est souhaitable d'estimer la **prévalence probable**, plus proche de la **prévalence réelle**, à partir de la **prévalence apparente** (cf. ci-dessous).*

Echantillonnage

Il existe différents types d'échantillons. Une distinction fondamentale est à faire entre les [échantillons représentatifs](#) (cf. Fiche [Représentativité d'un échantillon](#)) de la population d'origine et les [échantillons non représentatifs](#).



*Les qualités intrinsèques des résultats de l'enquête ([exactitude](#) et [précision](#)) sont **indépendantes** et ne dépendent pas de la même caractéristique de l'échantillon.*

- **L'exactitude du résultat dépend de la représentativité de l'échantillon** : la [représentativité](#) de l'échantillon garantit l'[exactitude](#) du résultat alors qu'un échantillon non représentatif peut conduire à un résultat plus ou moins erroné ([biaisé](#)) ; un échantillon représentatif permet d'estimer la prévalence de la maladie dans la population d'origine, à partir des

résultats obtenus sur l'échantillon, alors qu'un échantillon non représentatif ne le permet pas ;

- La **précision** du résultat dépend de la taille de l'échantillon : le résultat est d'autant plus précis que la **taille** de l'échantillon est importante.

Interprétation des résultats

En tenant compte de l'échantillonnage

Echantillon non représentatif

La description de l'échantillon, étape indispensable à toute étude épidémiologique, permet de le comparer à la population dont il provient et à laquelle on veut appliquer le résultat. En cas d'écart significatif au plan statistique, on peut affirmer que l'échantillon n'est pas représentatif.

Dans une telle situation, ou dans celle résultant de l'absence d'une base de sondage ou d'une connaissance insuffisante de la population (faune sauvage, par exemple), il n'est pas opportun d'estimer une prévalence dans une population à partir des résultats obtenus sur ce type d'échantillon car le risque peut être élevé que ce dernier soit biaisé et, donc, le résultat, erroné (par rapport à la prévalence réelle dans la population).



Dans ces conditions, il est préférable de n'exprimer le résultat que pour l'échantillon étudié, en attirant l'attention sur l'impossibilité de l'inférer à la population d'intérêt, tout en essayant d'estimer le sens de l'erreur possible (surestimation ou sous-estimation).

Echantillon représentatif



Si la comparaison de la description de l'échantillon à celle de la population permet de conclure que l'échantillon est bien représentatif de la population, il est possible d'estimer la valeur de la prévalence probable de la maladie dans la population d'origine à partir des résultats de prévalence obtenus sur l'échantillon.

A condition que le nombre de sujets dans l'échantillon soit suffisant (cf. livre *Epidémiologie appliquée*, page 194), une formule permet de calculer une fourchette de valeurs dans laquelle il existe 95 % de chance pour que la valeur de la prévalence réelle soit située. En l'absence de biais de représentativité dans le protocole et dans son application au cours de l'enquête, la probabilité est élevée (95 %) pour que la prévalence réelle soit située au sein de la fourchette, et le risque, faible (5 %) pour qu'elle soit à l'extérieur de la fourchette.

Le résultat s'exprime alors de la façon suivante :

Pourcentage de prévalence de la maladie dans la population : 20 % ; IC_{95 %} [18 % ; 22 %]

IC_{95 %} signifie : intervalle de confiance (ou fourchette) à 95 % de probabilité. Le pourcentage de prévalence trouvé dans l'échantillon est de 20 % et le pourcentage de prévalence dans la population a 95 % de probabilité d'être compris entre 18 % et 22 %.



La précision du résultat est d'autant meilleure que la fourchette est de faible amplitude.



Par rapport à la précision attendue, ayant servi à calculer le nombre de sujets nécessaire dans l'échantillon, la **précision obtenue** du résultat est **supérieure** si la valeur de la **prévalence probable** est **supérieure** à celle de la **prévalence supposée** (et inférieure, dans le cas contraire).

En tenant compte des qualités du test de dépistage

La prise en compte de la sensibilité et de la spécificité du test conduit à transformer le résultat obtenu sur l'échantillon (**prévalence apparente**) en une valeur plus proche de celle de la **prévalence réelle**, à savoir la prévalence probable dans l'échantillon. Ceci est possible grâce à une formule mathématique (page 135 du livre *Epidémiologie appliquée*). La valeur de la prévalence probable est parfois assez différente de celle de la prévalence apparente.

Exemple : pour un test de dépistage de sensibilité 90 % et de spécificité 95 %, une prévalence apparente de 13,5 % sur un échantillon de 1 000 sujets correspond à une prévalence probable de 10 % dans l'échantillon.



L'estimation de la prévalence d'une maladie dans une population nécessite la **prise en compte** préalable des **performances du test de dépistage** sur les résultats obtenus sur les échantillons.

En tenant compte des risques de biais

Différents **biais** peuvent survenir au cours de la réalisation d'une enquête : non respect de certains points du protocole (nombre de prélèvements par sujet enquêté ; attitude différente de certains enquêteurs dans la manière de collecter les informations, etc.), refus de participer de la part de sujets ou de propriétaires de sujets tirés au sort, etc.



L'interprétation des résultats d'une enquête doit passer par la **recherche de ces différents risques de biais**, leur analyse et l'estimation du **sens** (surestimation ou sous-estimation) de leur effet sur les résultats.



Toma B. et al. : *Epidémiologie appliquée*, 2018, AEEMA éditeur, 614 p.

Terminologie : des liens hypertextes (en bleu souligné) vous permettent d'accéder aux définitions sur le site de l'AEEMA

Fiches AEEMA : [Sensibilité \(individuelle ou collective\) d'un test de dépistage](#) ;
[Spécificité, \(individuelle ou collective\) d'un test de dépistage](#) ;
[Représentativité d'un échantillon](#).