



L'épidémiologie pour tous

Nombre de sujets nécessaire pour une enquête descriptive à visée quantitative

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

En épidémiologie descriptive, on peut avoir un objectif soit *quantitatif*, c'est-à-dire d'estimer (de quantifier) la fréquence (ou [prévalence](#)) d'une maladie dans une [population](#) donnée, soit *qualitatif*, c'est-à-dire de savoir si la maladie est présente ou non dans la population. La présente fiche traite le cas d'une enquête à visée quantitative.

Nature des sujets

En pathologie animale, il faut tout d'abord préciser la nature des « sujets » à introduire dans l'échantillon, c'est-à-dire l'unité épidémiologique d'intérêt. En effet, il peut s'agir soit d'unité « *animal* », soit d'unité « *élevage* » selon que l'on souhaite estimer le pourcentage de prévalence de la maladie chez les animaux (unité « animal » au numérateur et au dénominateur) ou dans les élevages (unité « élevage » au numérateur et au dénominateur).

Si l'unité épidémiologique d'intérêt est l'animal, la base de sondage pour le tirage au sort doit comporter l'ensemble des animaux ; si c'est l'élevage, elle doit lister l'ensemble des élevages.

Principe du calcul du nombre de sujets nécessaire

Il existe une interdépendance entre trois paramètres :

- le *pourcentage de prévalence* de la maladie dans la population d'intérêt,

- la *précision relative* de l'estimation du pourcentage de prévalence,
- et le *nombre de sujets* dans l'échantillon.

Le pourcentage de prévalence de la maladie dans la population d'intérêt est souvent mal connu, mais doit obligatoirement être estimé *a priori* pour déterminer le nombre de sujets à introduire dans l'échantillon (c'est la prévalence supposée ou attendue), l'enquête ayant pour objectif de le préciser.

Le choix de l'un des deux autres paramètres fixe automatiquement le niveau de l'autre. Ainsi, si pour une prévalence attendue donnée, on fixe le nombre de sujets dans l'échantillon, on détermine de ce fait le niveau de précision du résultat de l'enquête. A l'inverse, si on décide du niveau jugé nécessaire de précision du résultat, ceci détermine le nombre de sujets à introduire dans l'échantillon.

Compte tenu de l'interdépendance entre ces trois paramètres, pour connaître le nombre de sujets à introduire dans l'échantillon, il faut donc :

- estimer *a priori* le pourcentage de prévalence de la maladie étudiée (prévalence supposée ou attendue)
- et déterminer le niveau de précision souhaité.

Estimation *a priori* du pourcentage de prévalence de la maladie

Aussi paradoxal que cela puisse paraître, il faut avoir une idée de l'ordre de grandeur du pourcentage de prévalence de la maladie dans la région d'intérêt pour

pouvoir calculer le nombre de sujets nécessaire, afin d'obtenir le niveau choisi de précision du résultat. Pour ce faire, on peut se fonder sur des enquêtes antérieures ou, à défaut, sur diverses informations ou données provenant de la région.

Si l'on **surestime** le pourcentage de prévalence (c'est-à-dire si la prévalence réelle, connue à la suite de l'enquête, est inférieure à celle supposée *a priori*), l'inconvénient est d'avoir une précision insuffisante du résultat car inférieure à celle désirée.

Si, au contraire, on **sous-estime** le pourcentage de prévalence, l'inconvénient est d'avoir dépensé trop d'argent pour l'enquête par rapport au niveau de précision du résultat, jugé suffisant.



*Dans le cas où il est très **difficile d'avoir une estimation** pertinente a priori de la prévalence, il convient de connaître la **nature du risque que l'on prend : défaut de précision** si la maladie est plus rare que supposé ; **excès de dépense** si la maladie est plus fréquente qu'on ne le croyait.*

L'analyse des résultats en temps réel au cours de l'enquête peut permettre d'ajuster la taille de l'échantillon tiré au sort, en fonction du niveau de prévalence trouvé, par rapport à celui supposé.

Détermination du niveau de précision souhaité

L'obtention d'un résultat très précis lors d'une enquête descriptive dépend de la taille de l'échantillon utilisé. La contrepartie de cette précision est le coût de l'enquête qui augmente proportionnellement à la taille de l'échantillon. Il est donc important, avant de préparer le protocole qui définit la taille de l'échantillon, de bien réfléchir à la précision qui paraît indispensable : plus ce niveau sera élevé, plus l'enquête sera coûteuse.



*La question à se poser est de savoir **à quoi servira le résultat de l'enquête**, à quelles(s) décision(s), à partir de quel(s) seuil(s) de prévalence ?*

Détermination du nombre de sujets nécessaire

Après avoir défini l'unité épidémiologique, estimé le pourcentage de prévalence attendue et déterminé la précision relative souhaitée du résultat, il est possible de consulter un tableau (page 188 du livre *Epidémiologie appliquée*), qui permet d'obtenir le nombre nécessaire d'unités à introduire dans l'échantillon pour atteindre la précision désirée si la prévalence est proche de celle attendue.

Il faut alors vérifier si la situation correspond ou non à ce que l'on appelle une « **population infinie** ».



Lorsque le rapport (appelé « fraction de sondage ») de l'effectif nécessaire de sujets dans l'échantillon à l'effectif de la population à la population d'intérêt est faible, on considère qu'il s'agit d'une population infinie. Le seuil a été fixé à 10 %.

Si ce rapport est **inférieur à 10 %**, le nombre lu dans le tableau de la page 188 est celui des unités à introduire dans l'échantillon.

Si ce rapport est **supérieur à 10 %**, le nombre de sujets à introduire est inférieur à celui de ce tableau et, ce, d'autant plus que le pourcentage du rapport est élevé (c'est-à-dire que le nombre de sujets dans l'échantillon se rapproche de celui de la population totale) : il peut être obtenu à partir des tableaux des pages 190 à 192.

Le nombre d'unités nécessaire peut également être obtenu à l'aide du tableur de l'annexe 9 « [Taille d'échantillon aléatoire simple](#) » du livre *Epidémiologie appliquée* présent sur le site de l'AEEMA.

NB 1 : les valeurs trouvées dans les tableaux cités ci-dessus ont été calculées pour un test de dépistage de sensibilité parfaite. Si la sensibilité est inférieure à 1, il convient de corriger ces valeurs en les divisant par la valeur de la sensibilité.



Exemple : nombre de sujets nécessaire (issu du tableau) : 100 ; sensibilité du test : 0,9 ; nombre ajusté de sujets nécessaire en tenant compte de la sensibilité du test : $100 / 0,9 = 111$.

NB 2: lors d'enquête sur un échantillon « captif » (c'est-à-dire pour lequel les sujets ne peuvent pas refuser de participer à l'enquête), il n'est pas nécessaire d'augmenter le nombre nécessaire d'unités pour obtenir la précision désirée (si la prévalence est conforme à celle attendue). En revanche, dans bien des cas, des refus de participation de propriétaires d'animaux désignés par le tirage aléatoire risquent de compromettre

l'atteinte du niveau de précision désiré. Par précaution, il peut être utile de prévoir un nombre de sujets tirés au sort supérieur à celui correspondant à la précision souhaitée, et d'y recourir en fonction des refus enregistrés. Mais il conviendra de vérifier que ce refus de participation de certains sujets n'occasionne pas un biais de sélection qui pourrait altérer la représentativité de l'échantillon.

Les éléments conditionnant le nombre de sujets nécessaire dans une enquête descriptive, quantitative ou qualitative, sont indiqués dans le tableau 1.

Tableau 1. Éléments interdépendants en relation avec le nombre de sujets nécessaire pour une enquête descriptive, quantitative ou qualitative

<i>Approche quantitative</i>	<i>Approche qualitative</i>
Prévalence supposée de la maladie (ou prévalence attendue)	Pourcentage de <i>prévalence</i> à détecter (prévalence limite = seuil)
Précision relative du résultat	Niveau d' <i>erreur par défaut</i> acceptable (risque d'erreur bêta)
Nombre de sujets nécessaire	Nombre de sujets nécessaire



Pour une enquête descriptive à visée quantitative, le **nombre de sujets nécessaire** dans l'échantillon est conditionné, d'une part, par le **pourcentage de prévalence attendu** (ou présumé), d'autre part, par la **précision relative souhaitée du résultat**.

Plus la **prévalence attendue est faible**, plus élevé est le nombre de sujets nécessaire, pour un même niveau de précision.

De même, pour un niveau donné de prévalence attendue, le **nombre de sujets nécessaire est d'autant plus élevé que la précision relative jugée nécessaire est importante**.



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

Annexe : « [Taille d'échantillon aléatoire simple](#) »

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

Fiches AEEMA : [Nombre de sujets nécessaire pour une enquête descriptive à visée qualitative](#) ;

[Epidémiologie descriptive : qualité et interprétation des résultats d'une enquête à visée quantitative](#) ;

[Erreurs](#) ;

[Représentativité d'un échantillon](#)