

# Interprétation d'une enquête épidémiologique : type d'enquête, notion de biais, causalité (72)

Docteur José LABARERE  
Mars 2004 (Mise à jour mai 2005)

## Pré-Requis :

- Test de comparaison de proportions ( $\chi^2$ )
- Programme biostatistiques PCEM 2

## Mots-clés :

Epidémiologie, enquête, facteurs de risque, biais.

## Références :

- Czernichow P, Chaperon J, Le Coutour X. Epidémiologie : connaissances et pratiques. Paris : Masson, 2001.
- Bastuji-Garin. Interprétation d'une enquête épidémiologique. Type d'enquête, notion de biais, causalité. Revue du Praticien 1996 ;46 :997-1003.
- Touzet S, Colin C. Interprétation d'une enquête épidémiologique. Type d'enquête, notion de biais, causalité. Revue du Praticien 1999 ;49 :1797-1804.

## 1. Introduction

L'épidémiologie est une discipline scientifique dont l'objet est l'étude de la distribution des états de santé et de leurs déterminants dans les populations humaines.

L'épidémiologie procède par enquêtes d'observation : l'investigateur ne contrôle pas l'exposition aux facteurs d'exposition par opposition aux essais cliniques où l'investigateur contrôle l'allocation du traitement.

## 2. Types d'enquêtes épidémiologiques

### 2.1. Les enquêtes descriptives

Les enquêtes descriptives décrivent la distribution des états de santé et des facteurs de risque dans les populations (par exemple : prévalence du SIDA, incidence de l'infarctus du myocarde en population générale, mortalité par cancer du sein en Isère, prévalence du tabagisme chez les 15-24 ans...).

Elles se répartissent en :

#### 2.1.1. Enquêtes de prévalence

Ce sont des **enquêtes transversales** qui estiment le nombre de cas présents dans une population à un instant donné. (Exemple : enquête de prévalence « un jour donné » des infections nosocomiales en établissements de santé).

Le taux de prévalence est le rapport du nombre de cas recensés sur l'effectif de la population à un instant donné :

$P = M/N$  (compris entre 0 et 1)

M = nombre de cas (malades)

N = nombre total de sujets (malades + non-malades)

### 2.1.2. Enquêtes d'incidence

Ce sont des **enquêtes longitudinales** qui estiment le nombre de nouveaux cas de maladie dans une population, pendant une période donnée. (Exemple : registres des cancers présents dans certains départements).

Le taux d'incidence est le rapport du nombre de nouveaux cas de maladie recensés au nombre de personnes susceptibles d'être atteintes dans la population, pendant une période donnée.

$TI = M/PT$

M = nombre de nouveaux cas (malades) pendant la période d'étude

PT = nombre de personnes-temps (personnes-années, le plus souvent)

On parle, par exemple, d'un taux d'incidence annuel de 10 cas pour 100 000 personnes-années.

## 2.2. Les enquêtes étiologiques (ou analytiques)

Les enquêtes étiologiques analysent les relations entre l'exposition à un facteur de risque et un état de santé. Les enquêtes étiologiques sont toujours **comparatives** (elles comparent deux groupes différant soit sur la présence de la maladie, soit sur la présence du facteur de risque). Un facteur de risque est une caractéristique associée à une probabilité plus élevée de maladie. C'est une définition statistique qui ne permet pas d'affirmer le rôle causal.

### 2.2.1. Enquêtes exposé / non-exposé

Les enquêtes exposé / non-exposé consistent à **comparer la proportion de malades** (ou de décès) observée entre un groupe de sujets exposés à un facteur de risque et un groupe de sujets non exposés à ce facteur de risque, initialement indemnes de la maladie. L'enquête exposé / non-exposé peut être **prospective** (les groupes sont constitués au début de l'étude et on organise le suivi des sujets au cours des années qui suivent leur inclusion) ou **rétrospective** (l'inclusion des sujets se fait à partir d'une date du passé suffisamment lointaine pour que la maladie ait eu le temps de se développer et que la période de suivi soit écourtée (date d'embauche des employés d'une usine par exemple)). L'enquête exposé / non-exposé s'adresse de préférence aux facteurs d'exposition rares.



[Annexe : Schéma d'une enquête exposé / non exposé](#)  
(J. Labarere)

L'enquête exposé / non-exposé permet d'estimer le risque relatif de la maladie (RR).

## Calcul du risque relatif (enquête exposé / non-exposé)

	Malades	Non-malades
Exposés	a	b
Non-exposés	c	d

$Re = a / (a+b)$   
 $Rne = c / (c+d)$

**Risque relatif = RR = Re / Rne**

[Annexe : Calcul du risque relatif à partir du tableau de contingence dans une enquête exposé / non-exposé](#)  
(J. Labarere)

L'interprétation du risque relatif se fait de la façon suivante :

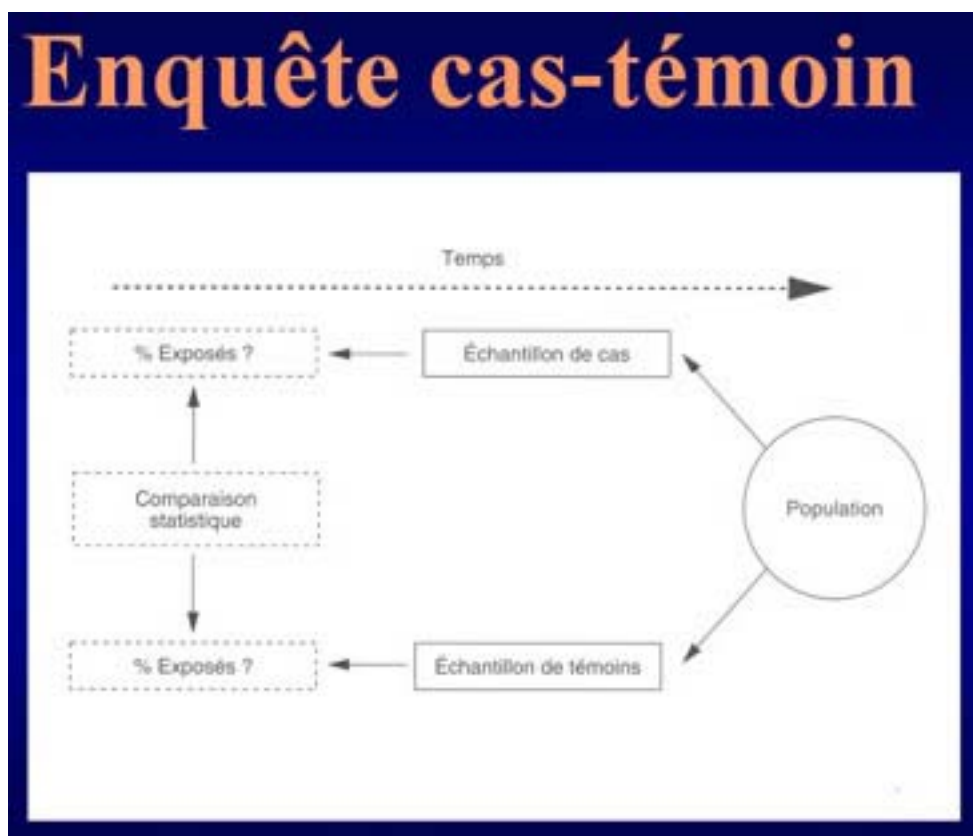
- RR = 1 : absence de relation entre le facteur de risque et la maladie
- RR > 1 : risque accru de maladie (facteur de risque)
- RR < 1 : risque réduit de maladie (facteur protecteur)

La p-value du test du  $\chi^2$  effectué sur le tableau de contingence ou l'intervalle de confiance à 95% [IC95%] du risque relatif (qui donne la précision de l'estimation du risque relatif) permettent de porter un jugement de signification statistique :

- $RR > 1$ ,  $px^2 < 0.05$  ou [IC95%] excluant la valeur 1 : risque significativement accru
- $RR < 1$ ,  $px^2 < 0.05$  ou [IC95%] excluant la valeur 1 : risque significativement réduit
- $px^2 > 0.05$  ou [IC95%] incluant la valeur 1 : on ne met pas en évidence de relation statistiquement significative entre le facteur étudié et la maladie.

### 2.2.2. Enquêtes cas / témoins

Les enquêtes cas-témoins consistent à **comparer la fréquence d'exposition** antérieure à un (ou plusieurs) facteur(s) de risque dans un groupe de « cas » atteints de la maladie étudiée, et dans un groupe de « témoins » indemnes de celle-ci. La mesure du facteur de risque est effectuée rétrospectivement. L'enquête cas / témoin s'adresse de préférence aux maladies rares.



[Annexe : Schéma d'une enquête cas / témoin](#)  
(J. Labarere)

Le risque relatif de maladie ne peut pas être estimé dans une enquête cas-témoin. On comprend qu'un échantillon composé, par choix de l'investigateur, de 50% de malades (les « cas ») ne permet pas d'estimer le risque de maladie. Une enquête cas / témoin permet de calculer l'odds ratio qui est une bonne approximation du risque relatif lorsque la maladie est rare dans la population (<5%, soit la majorité des cas en épidémiologie).

**Calcul de l'odds ratio  
(enquête cas / témoin)**

	FDR +	FDR -
Cas	a	b
Témoins	c	d

**Odds ratio = (a x d) / (b x c)**

(attention à l'ordre des lettres)

[Annexe : Calcul de l'odds ratio dans une enquête cas / témoin](#)  
(J. Labarere)

L'odds ratio s'interprète comme le risque relatif.

### 3. Notion de biais

Le biais désigne une erreur systématique dans l'estimation d'un paramètre (prévalence, odds ratio, risque relatif...). Il doit être distingué des fluctuations aléatoires d'échantillonnage qui représentent seulement un défaut de précision de l'estimation.

Il existe trois types de biais :

#### 3.1. Biais de sélection

Les biais de sélection affectent la constitution de l'échantillon d'enquête, c'est à dire le processus par lequel les sujets sont choisis au sein de la population. Il sont à craindre chaque fois que l'échantillon d'enquête n'est qu'une sélection de la population d'étude.

Les principales sources de biais de sélection sont :

- la constitution d'un échantillon par un autre moyen que le tirage au sort (sujets volontaires) ;
- les non-réponses à une enquête ;
- le recrutement de témoins en milieu hospitalier qui ne sont pas représentatifs de la population générale (biais de Berkson) ;
- les sujets perdus de vue dans les enquêtes exposé / non-exposé.

#### 3.2. Biais de classement

Le biais de classement désigne une erreur systématique de mesure de l'exposition ou de la maladie. Ils conduisent à mal classer les sujets en « malades / non malades » ou en « exposés / non exposés ».

Il peut s'agir d'un biais de classement différentiel entre les deux groupes :

- d'un biais d'enquêteur ou de subjectivité : un enquêteur cherchera avec plus d'acuité une exposition passée chez un cas, s'il a connaissance du statut malade du sujet ;
- d'un biais de mémorisation : un sujet atteint d'une pathologie (cas) se souvient davantage de ces expositions passées qu'un sujet indemne de la pathologie (« témoin »).

Il peut aussi s'agir d'un biais de classement non différentiel (tensiomètre mal étalonné par exemple).

### 3.3. Biais de confusion

Un biais de confusion désigne une erreur systématique dans l'estimation d'une mesure d'association (odds ratio ou risque relatif) entre le facteur étudié et la maladie, du fait d'un défaut de prise en compte d'un facteur de confusion. Un facteur de confusion est un facteur lié à la fois à l'exposition et à la maladie étudiées.



Annexe : Schéma d'un facteur de confusion  
(J. Labarere)

## 4. Notion de causalité

La question de la causalité d'une relation entre un facteur et une pathologie intervient **après l'objectivation d'une association statistiquement significative**. Il s'agit de savoir si l'association mise en évidence correspond à une relation de cause à effet. Une enquête d'observation ne permet pas de conclure à une relation causale (contrairement à un essai randomisé bien conduit). Il faut donc examiner les arguments qui plaident en faveur de la causalité de la relation :

### 4.1. Critères internes à l'étude

- Séquence temporelle : l'exposition doit précéder l'apparition de la pathologie ;
- force de l'association : risque relatif ou odds ratio élevé ;
- relation de type « dose-effet » entre l'exposition et la fréquence de la pathologie ;
- spécificité de la cause et de l'effet : la relation est spécifique si le facteur de risque est présent chez presque tous les malades et seulement chez eux, et ce pour cette seule maladie ;

- cohérence interne de l'étude : prise en compte la plus complète possible des facteurs de confusion et minimisation des biais de sélection ou de classification.

#### 4.2. Critères externes à l'étude

- Constance de l'association et reproductibilité dans diverses situations (différentes périodes, régions ou populations) ;
- cohérence des résultats avec les résultats d'études publiées ;
- plausibilité biologique : modèles expérimentaux chez l'animal ;
- cohérence avec les connaissances générales et les hypothèses qui ont conduit à la réalisation de l'enquête ;
- parallélisme de la distribution (dans l'espace et dans le temps) du facteur de risque et de la pathologie.