

LES ETUDES EXPOSÉS-NON EXPOSÉS

Spécialisé dans le traitement statistique des études cliniques,
le Département Biométrie de FOVEA vous propose
au travers de ce fascicule quelques notions simples
concernant les études exposés-non exposés.

*en
pratique*

Traitement statistique
DES ETUDES CLINIQUES

NF EN ISO 9001 Ed. 2000



N°0071-2003-FR-COFRAC-QUA



INTRODUCTION

Les études épidémiologiques relèvent soit de **l'épidémiologie non expérimentale**, soit de **l'épidémiologie expérimentale** également appelée épidémiologie d'intervention.

A l'épidémiologie non expérimentale appartiennent les **études épidémiologiques descriptives** et les **études épidémiologiques étiologiques**.

Les études étiologiques visent à rechercher les **causes** et les **facteurs de risque des maladies**. Ce sont des études **compara-**

tives, non randomisées, comportant **2 groupes de populations** constitués de façon distincte. Elles peuvent être soit **prospectives**, soit **rétrospectives**.

L'objectif d'une étude prospective est de **rechercher l'information sur l'exposition ou non à un facteur de risque avant d'observer une éventuelle apparition de la maladie**. C'est à cette catégorie d'études qu'appartiennent les études exposés-non exposés.

1

Définition

L'étude exposés-non exposés est **un type particulier d'étude de cohorte**.

Il s'agit d'une **étude étiologique, comparative, observationnelle, longitudinale, prospective**, comparant 2 cohortes de sujets constituées de façon distincte, mais **non randomisées**.

Le principe d'une étude exposés-non exposés est de comparer la morbidité ou la mortalité dans deux ou plusieurs groupes d'individus initialement indemnes de la maladie et définis en fonction de leur exposition à un facteur de risque soupçonné de cette maladie.

L'étude exposés-non exposés nécessite **une période de suivi** entre l'inclusion d'un sujet et la mesure de son altération de santé.

Le terme de cohorte est uti-

lisé pour désigner les groupes de sujets suivis au cours du temps. L'inclusion dans l'étude et le recueil initial des informations sur l'exposition sont réalisés au moment de la mise en place de l'étude. Les cohortes sont suivies de façon prospective au cours des années qui suivent l'inclusion.

Deux groupes de sujets sont constitués : **un groupe exposé au facteur étudié** et **un groupe non exposé**. **Le facteur d'exposition est le facteur contrôlé** au départ, **le facteur maladie est le facteur aléatoire**. Ainsi, l'étude exposés-non exposés permet de **comparer** entre les 2 groupes **la fréquence de sujets malades**, mais pas la fréquence de l'exposition.

2

Choix d'une étude exposés-non exposés

Quand choisir une étude

exposés-non exposés ? Les avantages et les inconvénients sont les suivants :

■ Avantages

- Réduction du nombre de sujets à inclure
- Etude adaptée lorsque l'exposition au facteur de risque est rare
- Limitation des biais grâce au choix des exposés / non exposés au démarrage de l'étude
- Surveillance de la survenue de la maladie dans chacun des groupes
- Analyse possible de plusieurs maladies au travers de la même étude

■ Inconvénients

- Absence de randomisation des groupes
- Impossibilité d'estimer l'incidence de la maladie dans la population générale en raison de la non représentativité des groupes
- Suivi de longue durée engendrant des perdus de vue
- En cas de maladies rares, séries trop importantes à constituer
- Coût financier élevé

3

Choix de la population

Comme dans toute étude épidémiologique, la planification d'une étude exposés-non exposés commence par le choix de la population qui fera l'objet de l'étude. C'est une phase essentielle dont dépendra la qualité de l'étude.

On constitue une cohorte de sujets exposés et une cohorte de sujets non exposés au facteur de risque en limitant au maximum les biais de sélection.

Ces cohortes ne sont cependant **pas représentatives de la population générale**. Pour réaliser une étude sur un échantillon représentatif, il faudrait tirer au sort un échantillon de la population que l'on veut étudier (exemple : fumeuses enceintes pendant une période donnée), ce qui est souvent très difficile à réaliser.

■ Constitution de la cohorte " exposés "

Dans le groupe " exposés ", l'**exposition** doit être **clairement définie** et **d'un niveau suffisant** pour qu'il soit licite d'étudier un lien avec un risque de maladie. D'où le problème de la **mesure de l'exposition** à un facteur **en intensité** et **en durée**.

Il peut être nécessaire de ne considérer que les expositions ayant **une durée minimum**, c'est-à-dire de fixer **un seuil d'exposition cumulée** en deçà duquel les sujets ne sont pas considérés comme étant " exposés ". **L'intensité de l'exposition** doit être **suffisamment variable** si l'on souhaite définir **des catégories d'exposition contrastées** et rechercher **une relation " dose-effet "**.

L'information sur le statut initial doit être **accessible** pendant toute la durée du suivi. Les sujets du groupe " exposés " ne doivent **pas** être **atteints de la maladie ni d'aucune de ses manifestations précoces**.

Le recrutement de la cohorte peut être **fixe** ou **dynamique**.

- Dans une cohorte fixe, seuls les sujet " exposés " **présents lors de la constitution de la cohorte** sont inclus dans l'étude.

- Dans une cohorte dynamique, **tous les nouveaux sujets répondant aux critères d'inclusion** depuis la date de la constitution de la cohorte sont recrutés.

■ Constitution de la cohorte de référence

Dans les études exposés-non exposés, l'inclusion dans le groupe de référence vise à constituer **un groupe a priori comparable** à celui des sujets " exposés ".

L'idéal serait de constituer le groupe " non exposés " avec des sujets en tous points

semblables aux sujets " exposés ", sauf en ce qui concerne le facteur d'exposition. En effet, on cherche à s'assurer que l'observation d'une incidence différente dans le groupe " exposés " n'est pas imputable au mode de sélection des sujets.

Le groupe de référence doit être **comparable** au groupe " exposés " pour **des caractéristiques générales** comme l'âge, le sexe, la catégorie socioprofessionnelle, les antécédents médicaux, le terrain... Si cela est nécessaire, il est possible de **stratifier** ce groupe sur certaines caractéristiques.

Ce groupe doit pouvoir être suivi avec les mêmes méthodes, le même soin et pendant la même durée que le groupe " exposés ".

Dans certaines études, les cohortes " exposés " et " non exposés " ne sont définies qu'après le recrutement, la base de recrutement garantissant que ces groupes sont suffisamment contrastés et de taille suffisante. C'est l'exemple des groupes de fumeurs et de non fumeurs de la cohorte des médecins britanniques de Doll et Hill.

Dans les études de natalité, la population nationale ou régionale est souvent utilisée comme population de référence par commodité et par souci d'économie. Cela peut cependant présenter des inconvénients. Ainsi, si le facteur de risque est trop fréquent dans la population générale, la proportion de sujets " exposés " classés

LES ÉTUDES EXPOSÉS-NON EXPOSÉS EN PRATIQUE

" non exposés " peut ne pas être négligeable.

Enfin, ce choix peut être à l'origine d'un biais de sélection si le groupe " exposés " diffère de la population générale par rapport à des caractéristiques comme l'âge, le sexe ou la catégorie socio-professionnelle dont l'influence sur le risque est fréquente.

4

Les biais

Différents types de biais peuvent survenir au cours de la vie de l'étude.

■ Au moment du choix des sujets

- **Biais de sélection** : les groupes comparés n'ont pas *a priori* le même risque d'exposition et/ou de survenue de l'événement ; les biais de sélection ne peuvent pas être corrigés.

■ Au moment du recueil des données

- **Biais dans le suivi des groupes** : les sujets sont inclus à différentes dates alors que le date du endpoint est identique pour tous ; ce biais peut être compensé par le calcul des unités " patient-temps ".

- **Biais de classement** : erreurs de mesure systématiques de l'exposition ou de la survenue de la maladie ; ils peuvent influencer la valeur

de l'Odds Ratio, soit en la sous-estimant, soit en la sur-estimant.

- **Biais dus à la durée d'exposition** pouvant complètement modifier les résultats.

- **Biais d'information** : le recueil des données d'exposition ou de survenue de l'événement diffère entre les " exposés " et les " non exposés " ; ce biais peut être évité par une évaluation à l'aveugle de l'exposition et de la survenue de l'événement.

■ Au moment de l'analyse des données

- **Biais de confusion** : l'ex-

	Malades	Non malades	Effectif total
Exposés	a	b	a+b
Non exposés	c	d	c+d
Effectif total	a+c	b+d	a+b+c+d

position et la survenue de l'événement sont tous deux liés à un facteur tiers. Les facteurs de confusion doivent être identifiés au début de l'étude et pris en compte dans l'analyse statistique.

5

Analyse des résultats

■ Nombre de sujets nécessaires

Le nombre de sujets nécessaires dans une étude exposés-non exposés cor-

respond au nombre de sujets assurant **une puissance supérieure ou égale à 80%**.

La puissance est d'autant plus importante que le nombre d'observations réalisées est important. Le nombre de sujets est soumis à la règle de 3 : il faut 3 fois plus de sujets que la fréquence de survenue de l'événement pour avoir 95% de chances d'observer un cas. Ainsi, pour observer un événement dont la fréquence de survenue est de 1/1.000, il faut recruter 3.000 sujets.

■ Présentation des résultats

Les résultats d'une étude exposés-non exposés se présentent de la façon suivante :

La cote de maladie est :

- a/c chez les sujets exposés
- b/d chez les sujets non exposés

■ Mesures d'association

La mesure de l'association entre exposition et maladie peut être exprimée en :

- Risque relatif (RR)
- Odds Ratio (OR)
- Risque attribuable (RA)
- Fraction étiologique du risque (FE)

● Risque relatif

$$RR = \frac{\text{Probabilité de survenue de l'événement si "exposé"}}{\text{Probabilité de survenue de l'événement si "non exposé"}}$$

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

● **Odds Ratio = Rapport de cotes**

Utilisé à la place du risque relatif lorsque l'échantillon n'est pas représentatif de la population ou lorsque la maladie est rare (prévalence < 10%), l'odds ratio (OR) est une bonne estimation du risque relatif.

$$OR = \frac{a/c}{b/d}$$

ou

$$OR = \frac{a \times d}{b \times c}$$

Le risque relatif (RR) et l'odds ratio (OR) mesurent l'association entre le facteur de risque et la maladie.

La valeur X du RR ou de l'OR permet de dire que le risque de présenter la maladie est X fois plus important chez les sujets exposés que chez les sujets non exposés.

Si RR ou OR > 1, le facteur augmente le risque d'être atteint de la maladie.

Si RR ou OR < 1, le facteur est dit " protecteur " et diminue le risque d'être atteint de la maladie.

Le calcul de l'intervalle de confiance à 95% permet d'indiquer la précision de RR ou d'OR et de tester sa significativité. Si cet intervalle de confiance ne comprend pas la valeur 1, on conclut alors que l'association entre le facteur de risque et la maladie est statistiquement significative. Dans le cas contraire, la relation n'est pas démontrée.

● **Risque attribuable pour la population**

Le risque attribuable (RA) mesure la proportion de cas de la maladie que l'on peut attribuer au facteur de risque dans l'ensemble de la population, y compris chez les sujets " non exposés ".

En termes de Santé Publique, cet indice permet de mesurer l'impact global du facteur de risque sur la population.

Le risque attribuable pour la population dépend à la fois de la valeur du risque relatif (RR) et de la fréquence de l'exposition dans la population (Pe).

$$RA = \frac{Pe(RR - 1)}{[Pe(RR - 1)] + 1}$$

sachant que Pe est donnée par la formule :

$$Pe = \frac{a+b}{a+b+c+d}$$

● **Fraction étiologique du risque = Risque attribuable pour les exposés**

La fraction étiologique du risque (FE) dépend du risque relatif :

$$FE = \frac{RR - 1}{RR}$$

Il est possible de remplacer la valeur du risque relatif (RR) par celle de l'odds ratio (OR) dans le calcul du risque attribuable (RA) ainsi que dans celui de la fraction étiologique du risque (FE).

6

Conclusion

Les études exposés-non exposés sont des études étiologiques, prospectives, longitudinales, observationnelles comparant 2 cohortes, l'une de sujets exposés, l'autre de sujets non exposés.

Elles sont destinées à mettre en évidence l'association entre l'exposition à un facteur de risque et la survenue d'une maladie.

LES ÉTUDES EXPOSÉS-NON EXPOSÉS EN PRATIQUE

Pour **interpréter le risque en termes de causalité**, plusieurs conditions doivent être réunies :

- liaison statistiquement significative entre facteur de risque et maladie
- recherche et prise en compte des facteurs de confusion
- recherche de biais susceptibles de modifier la relation
- réunion d'un certain nombre de facteurs de causalité

Enfin, **la nature causale** de la relation **ne peut être mise en évidence à partir d'une seule étude.**

Les études exposés-non exposés trouvent leur indication lorsque **l'exposition au facteur de risque est rare**. Ce type d'étude peut porter sur **un nombre de sujets relativement faible**.

En revanche, elles ne permettent **pas d'estimer l'incidence de la maladie étudiée, leur délai de réalisation est souvent long et le suivi des sujets difficile.**

7

Bibliographie

1. Bouyer J., Hemon D., Cordier S., Derriennic F., Stücker I., Stengel B., Clavel J.

Epidémiologie – Principes et méthodes quantitatives
Les Editions INSERM 1993

2. Rumeau-Rouquette C., Bréart G., Padieu R.

Méthodes en épidémiologie
Flammarion
Médecine-Sciences

3. Schwartz D.

Méthodes statistiques à l'usage des médecins et des biologistes
Flammarion 1993

4. Touzet S., Colin C.

Interprétation d'une enquête épidémiologique
La Revue du Praticien ; 49, 1999

5. Moore N.

Université d'Eté
Pharmaco-Epidémiologie,
septembre 2004

