

LES ETUDES DE COHORTE

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 de cohorte

*en
pratique*

Traitement statistique
DES ETUDES CLINIQUES

NF EN ISO 9001



INTRODUCTION

Parmi les études épidémiologiques, les études de cohorte appartiennent à la catégorie des études étiologiques.

Les études de cohorte cherchent à associer "exposition à un facteur de risque" et "survenue d'une maladie", l'objectif étant de démontrer que le risque de contracter la maladie considérée est plus important chez les sujets exposés au facteur de risque étudié que chez les sujets non exposés.

Les premières études de cohorte, initiées pour mettre en évidence un facteur de risque, sont apparues au cours des années 1950 aux États-Unis et en Grande-Bretagne. Elles avaient pour objectif pour l'une, de démontrer le rôle de la consommation de tabac sur l'étiologie des cancers broncho-pulmonaires (cohorte de médecins britanniques, Doll et Hill, 1954-1956) et pour l'autre (Case et al., 1954), d'évaluer le rôle des amines aromatiques dans la survenue du cancer de la vessie.

1

Définitions

Une étude de cohorte est une étude au cours de laquelle on compare la morbidité (ou la mortalité) observée dans un ou plusieurs groupes d'individus (cohorte), indemnes de la maladie à l'origine, exposés ou non à un facteur de risque soupçonné d'être à l'origine de la maladie. L'apparition éventuelle, au cours du temps, de la maladie étudiée est traquée chez chaque individu de la cohorte.

La cohorte peut être définie comme **un ou plusieurs groupes de sujets suivis au cours du temps et définis à partir de caractéristiques personnelles communes** : sexe, milieu ou origine géographique, milieu professionnel, ...

La cohorte peut être également prise au sens démographique du terme : «ensemble d'individus ayant vécu un événement semblable pendant la même période de temps».

2

Les différents types d'études de cohorte

L'étude de cohorte peut être **prospectiv e, historique** ou **une étude exposé/non exposé**.

- Une étude de cohorte est dite **prospectiv e** si l'inclusion dans l'étude et le recueil initial des informations sur l'exposition au facteur de risque sont réalisés lors de la mise en place de l'étude. La cohorte est suivie de façon prospectiv e au cours des années qui suivent l'inclusion.

L'étude de cohorte prospectiv e est utilisée lorsque **le nombre de sujets exposés** est **important** dans la population étudiée, **l'apparition de la maladie fréquente** et **le délai d'apparition** relativement **court**. Ce type d'étude est bien adapté à la surveillance des travailleurs d'un secteur industriel et à la mise en évidence des effets de faibles doses.

- Une étude de cohorte est dite **historique** si l'inclusion se fait à partir d'une date choisie dans un passé suffisamment lointain pour que la maladie étudiée ait eu le temps de se développer et que la durée de suivi de la cohorte soit ainsi raccourcie.

On fait appel à une étude de cohorte lorsque **le délai de survenue de la maladie a près exposition** au facteur de risque est **très important**, ce qui est le cas notamment pour les cancers ou les maladies cardiovasculaires.

- Un troisième type d'études de cohorte concerne **les études exposé/non exposé**. Il s'agit d'**études comparativ es** dans lesquelles deux cohortes de sujets sont constituées de façon distincte : une cohorte est exposée au facteur de risque, l'autre ne l'est pas. On compare l'incidence de la

maladie entre le groupe de sujets exposés et le groupe de sujets non exposés.

Comme pour les autres types d'études de cohorte, les deux cohortes sont suivies pendant une période au cours de laquelle est notée l'apparition de la maladie.

Les études exposé/non exposé nécessitent **moins de sujets** que les deux autres types d'études de cohorte, mais les cohortes constituées ne sont **pas représentatives de la population générale** et l'incidence de la maladie dans la population générale ne peut pas être estimée. L'étude exposé/non exposé est utilisée lorsque **l'exposition au facteur de risque est rare**.

3

Constitutions des cohortes

n SOURCES

La population peut être identifiée à partir de **différentes sources** :

- Archives d'entreprises ou de secteurs d'activité pour les maladies professionnelles.
- Mutuelles ou compagnies d'assurance pour l'étude des risques liés aux comportements alimentaires, à la consommation d'alcool ou de tabac.
- Archives de consultations médicales pour l'étude du risque lié à un antécédent médical ou thérapeutique.



n CRITÈRES D'INCLUSION

Pour éviter les biais de sélection, les critères d'inclusion doivent être **extrêmement rigoureux** :

- L'exposition au risque doit être définie avec précision de même que la durée et l'intensité d'exposition.
- L'intensité d'exposition doit être variable si l'on veut établir une relation dose-effet.
- L'information concernant le statut vital doit être accessible pendant toute la durée du suivi : seront ainsi exclus de la cohorte les sujets les plus difficiles à suivre, en évitant cependant d'exclure des sujets pour lesquels le risque est connu comme, par exemple, les travailleurs temporaires dans l'industrie nucléaire.

n MODE D'INCLUSION

Le recrutement d'une cohorte peut être **fixe, dynamique ou mixte**.

- Dans une cohorte fixe sont recrutés **uniquement les sujets présents au moment de la constitution de la cohorte**, notamment si celle-ci concerne l'étude des conséquences d'un événement unique (*exemples : explosions atomiques d'Hiroshima et de Nagasaki, nuage de Tchernobyl*).
- Dans une cohorte dynamique sont inclus, entre la date de constitution de la cohorte et la date de fin de suivi, **tous les nouveaux sujets répondant aux critères d'inclusion**. La cohorte dynamique élimine les biais de recrutement induits, dans le cas d'un facteur de risque ayant des répercussions rapides, par l'élimination des sujets les plus fragiles et la sélection des sujets les plus résistants ou les moins exposés.
- Une cohorte mixte associera les deux modes de recrutement décrits ci-dessus.

4

Analyse des études de cohorte

n EFFECTIFS NÉCESSAIRES

L'unité statistique des études de cohorte est la personne-temps. Le calcul des effectifs nécessaires est réalisé non pas sur la base du nombre de sujets mais sur celle du nombre de personnes-temps (années, mois, semaines, jours, ...) générés par les sujets constituant la cohorte.

n PERSONNE -TEMPS

Chaque sujet génère un nombre de personnes-temps qui est le **nombre d'unités de temps** (années, mois, ...) pendant lesquelles le sujet a été **exposé au facteur de risque avant d'être malade** ou encore **la durée de sa présence dans l'étude** si, à l'issue de celle-ci, il n'a pas présenté la maladie considérée.

n TAUX D'INCIDENCE

Par définition, le taux d'incidence de la maladie (T_i) est égal au nombre de nouveaux cas (m) par unité de temps divisé par la taille de la population.

La taille de la population se mesure en «personnes-temps». Il s'agit de la somme des durées, cumulées sur l'ensemble de la population considérée, pendant laquelle les sujets sont susceptibles d'être enregistrés comme des nouveaux cas. Soit m le nombre de nouveaux cas pendant la période $[t, t + \Delta t[$ et PA le nombre cumulé de personnes-années sur la période $[t, t + \Delta t[$, le taux d'incidence T_i est défini par :

$$T_i = \frac{m}{PA}$$

Le taux d'incidence s'exprime en cas par personnes-années (si l'unité de mesure du temps est l'année).

n RISQUE RELATIF

L'analyse des études de cohorte

consiste essentiellement à **comparer le taux d'incidence de la maladie observé chez les sujets exposés au facteur de risque** (ou supposé comme tel) (T_1) à **celui observé chez les sujets non exposés** (T_0).

Le rapport de ces deux taux étant le risque relatif associé à l'exposition :

$$RR = \frac{T_1}{T_0}$$

La valeur du risque relatif RR permet de dire que **le risque de présenter la maladie est RR plus important chez les sujets exposés que chez les sujets non exposés**.

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

Si $RR < 1$, le facteur est dit «**protecteur**» et diminue le risque d'être atteint de la maladie.

Le calcul de l'**intervalle de confiance du risque relatif** permet d'indiquer la **précision de la mesure** 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**.

n FACTEURS DE CONFUSION

Pour s'assurer qu'un facteur de risque est responsable à 100% de l'apparition d'une maladie, il faut s'assurer que d'autres facteurs de risque, non testés, n'interviennent pas.

Ces facteurs de risque «cachés» s'appellent **facteurs de confusion**. Ils sont **liés à la fois à la maladie et au facteur de risque étudié**. Ils augmentent ou diminuent l'intensité de l'association entre le facteur de risque étudié et la maladie. Ainsi, une fausse relation entre consommation de café et maladie coronarienne était due au fait que les buveurs de café ont tendance à être fumeurs, le tabac étant fortement associé à la maladie coronarienne.

LES ETUDES DE COHORTE EN PRATIQUE

En isolant l'effet de la consommation de tabac, la relation entre consommation de café et maladie coronarienne disparaît.

Les facteurs de confusion doivent être identifiés puis pris en compte au début de l'étude ainsi qu'au moment de l'analyse statistique. Lors de l'analyse, la population peut être stratifiée selon les modalités du facteur de confusion et l'ajustement réalisé par la méthode de Mantel-Haenzel, l'analyse multivariée pouvant également être utilisée.

En résumé, sachant que la population d'une étude de cohorte se répartit ainsi :

	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

les résultats d'une étude de cohorte peuvent s'exprimer de la façon suivante :

- Incidence de la maladie

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

- Fréquence de l'exposition

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

- Risque absolu

â Chez les sujets exposés

$$Re = \frac{a}{(a+b)}$$

â Chez les sujets non exposés

$$Ro = \frac{c}{(c+d)}$$

- Risque relatif = rapport des risques absolus

$$RR = \frac{Re}{Ro}$$

- Excès de risque = différence des risques absolus

$$ER = Re - Ro$$

5 Conclusion

Les études de cohorte sont destinées à **explorer les risques encourus par une population exposée**.

Leur **méthodologie** est relativement **lourde** en raison de la **rareté des pathologies étudiées** et des **mo yens nécessaires au suivi de cohortes importantes**. Cependant les études de cohorte peuvent conduire à **une analyse plus fine des indices de santé**.

6 Bibliographie

1. **Bouyer J., Hemon D., Cor dier S., Der riennic 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

