

# Épidémiologie

---

Jean Bouyer, Sylvaine Cordier, Patrick Levallois

La référence bibliographique de ce document se lit comme suit:

Bouyer J, Cordier S, Levallois P (2003)

Épidémiologie.

In : Environnement et santé publique - Fondements et pratiques, pp.89-118.

Gérin M, Gosselin P, Cordier S, Viau C, Quénel P, Dewailly É, rédacteurs.

Edisem / Tec & Doc, Acton Vale / Paris

Note : Ce manuel a été publié en 2003. Les connaissances ont pu évoluer de façon importante depuis sa publication.

# Épidémiologie

---

Jean Bouyer, Sylvaine Cordier, Patrick Levallois

- 1. Introduction**
- 2. Domaine de l'épidémiologie**
- 3. Mesures utilisées en épidémiologie**
  - 3.1 Prévalence
  - 3.2 Taux d'incidence
  - 3.3 Risque cumulé de maladie (ou incidence cumulée)
  - 3.4 Mesures d'association
- 4. Types d'enquêtes épidémiologiques**
  - 4.1 Études expérimentales
  - 4.2 Enquêtes descriptives
  - 4.3 Enquêtes étiologiques
- 5. Validité et précision**
  - 5.1 Introduction
  - 5.2 Biais de sélection
  - 5.3 Biais de classement
  - 5.4 Facteurs de confusion
- 6. Analyse des données**
  - 6.1 Estimation et tests
  - 6.2 Analyse
- 7. Interprétation des résultats. Causalité**
  - 7.1 Jugement de signification
  - 7.2 Jugement de causalité
  - 7.3 Critères de causalité
- 8. Conclusion**

## 1. INTRODUCTION

L'épidémiologie est une des disciplines qui concourent de façon importante aux progrès des connaissances en santé environnementale et en santé au travail. Son atout majeur est d'étudier les relations entre environnement et santé à un niveau global. Cette globalisation est faite selon deux «dimensions». En premier lieu, les résultats des études épidémiologiques portent sur des groupes de sujets, définis par exemple par leur exposition à telle ou telle condition environnementale ou de travail. Ce niveau d'analyse permet de dégager des moyennes et des tendances stables que l'observation individuelle ne rend pas perceptibles en raison de la variabilité importante entre les individus, qu'elle soit d'origine «biologique» ou qu'elle résulte de conditions d'exposition variables. Le second niveau de globalisation se situe au sein de l'individu lui-même. L'épidémiologie ne cherche pas à étudier ni à définir les mécanismes d'action des expositions sur l'organisme humain. Elle mesure leur effet «intégratif» par la survenue de pathologies ou, de façon plus générale, d'événements de santé. C'est le côté «boîte noire» de l'épidémiologie qui a été beaucoup débattu (Savitz, 1994; Skrabanek, 1994), et dont on voit bien les limites, mais aussi les avantages puisqu'à la fois il masque la connaissance des mécanismes biologiques fins et la rend non indispensable à la progression des connaissances des effets de l'environnement sur l'Homme.

L'épidémiologie n'est bien sûr pas seule et, comme dans la plupart des domaines scientifiques, les avancées en santé environnementale ont été et seront le fruit des résultats conjoints de plusieurs disciplines. C'est ainsi que la toxicologie apporte les «preuves» expérimentales de la nocivité de certaines substances que l'épidémiologie ne peut pas donner, mais elle les apporte sur l'animal ou dans des conditions «idéales», souvent éloignées de la réalité des expositions humaines. À l'inverse, l'épidémiologie est proche des conditions réelles d'exposition, mais a parfois du mal à séparer les effets d'expositions survenant de façon conjointe.

Ce chapitre présente les méthodes épidémiologiques en mettant l'accent sur les mesures utilisées en épidémiologie et les principaux types d'étude. Il permettra aussi d'aborder succinctement les principes de l'analyse des données épidémiologiques et de leur interprétation. Il

devrait permettre de comprendre les résultats des études épidémiologiques en ayant un regard critique sur leur méthodologie.

## 2. DOMAINE DE L'ÉPIDÉMOLOGIE

L'épidémiologie est une discipline plutôt récente. Son champ d'intérêt s'accroît d'années en années, et sa méthodologie est encore en pleine évolution. Même si on a coutume de rappeler que déjà Hippocrate 400 ans avant Jésus-Christ s'intéressait aux déterminants de la maladie, nous devons reconnaître que l'épidémiologie comme science a commencé à voir le jour au XIX<sup>e</sup> siècle (en particulier en Angleterre et en France) et qu'elle s'est véritablement développée au XX<sup>e</sup> siècle en particulier, et de plus en plus à l'aide des statisticiens et de la révolution informatique.

L'épidémiologie est classiquement définie comme l'étude de la distribution des maladies et de leurs déterminants dans les populations humaines (Bouyer et coll., 1993; Rothman et coll., 1998). Cependant, son champ s'est rapidement étendu pour couvrir l'étiologie de l'ensemble des problèmes de santé ainsi que leur contrôle (Last, 1983). Les définitions modernes de l'épidémiologie incluent même l'évaluation des interventions et le support aux politiques de santé. Cependant, dans ce chapitre, nous nous concentrerons sur sa définition primaire, soit l'étude de l'apparition de la maladie dans les populations humaines et son apport à l'évaluation des risques environnementaux.

## 3. MESURES UTILISÉES EN ÉPIDÉMOLOGIE

La définition de l'épidémiologie montre qu'il faut s'intéresser à deux types de mesures. D'une part, celles qui permettent de caractériser la distribution des maladies; il s'agit des mesures de risque et d'incidence (nous verrons la prévalence, le taux d'incidence et le risque cumulé). D'autre part, celles qui permettent de quantifier le lien entre une exposition et la maladie; il s'agit des mesures d'association (nous verrons principalement le risque relatif et l'odds ratio).

Les mesures qui caractérisent la distribution des maladies englobent des mesures de risque au sens strict, c'est-à-dire les probabilités d'être ou de devenir malade, et les mesures d'incidence

qui indiquent la «vitesse d'apparition» des cas de maladie (pour le taux d'incidence) ou de décès (pour le taux de mortalité). Une probabilité est un nombre sans unité de mesure, compris entre 0 et 1, alors qu'un taux d'incidence est pourvu d'une unité et peut être supérieur à 1.

### 3.1 Prévalence

La façon la plus «naturelle» de mesurer la fréquence d'une maladie dans une population est de calculer la proportion de malades *présents* dans la population à un instant donné. Cette mesure est dénommée la prévalence, notée P, et définie par  $P = \frac{M}{N}$  où M est le nombre de malades et N le nombre total de sujets (malades et non malades) de la population.

La prévalence intègre deux dimensions différentes de la fréquence de la maladie. D'une part, la durée de maladie (ou, du moins, la durée de la présence d'un malade dans la population). D'autre part, la «vitesse d'apparition» de nouveaux cas de maladie au sein de la population (c'est-à-dire le taux d'incidence qui est défini plus bas).

La prévalence est surtout utile en santé publique lorsqu'on s'intéresse à la planification des ressources de santé nécessaires dans une population. En recherche étiologique, cet indice est rarement utilisé, sauf dans quelques domaines particuliers, comme la périnatalité, ou dans le cas de pathologies fréquentes et sous-diagnostiquées (la dépression, par exemple).

L'estimation de la prévalence sur un échantillon est notée  $p_0$ , et l'intervalle de confiance correspondant est, lorsque la taille n de l'échantillon est assez grande,

$$p_0 \pm z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}^*$$

### 3.2 Taux d'incidence

Par définition, le taux d'incidence (TI) de la maladie est la «vitesse de production» de nouveaux cas au cours d'un intervalle de temps. Il est égal au nombre de nouveaux cas survenus dans cet intervalle de temps divisé par la taille de la population à risque. 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 à l'étude et sur l'ensemble de la durée de suivi, pendant laquelle les sujets sont

susceptibles d'être enregistrés comme de nouveaux cas. La population peut être ouverte, c'est-à-dire que des sujets peuvent y entrer ou en sortir au cours de la période de suivi. Dans ce cas, les durées cumulées correspondent aux périodes où le sujet est présent dans la population. L'unité de mesure la plus fréquente en épidémiologie est la personne-année, au point que le terme «personnes-années» est souvent employé comme terme générique à la place de «personnes-temps». Cependant, si l'unité de mesure du temps est le mois, la semaine ou le jour, on peut être amené à compter en personnes — mois — semaines ou jours.

La définition formelle du taux d'incidence est  $TI = \frac{m}{PA}$  où m est le nombre de nouveaux cas pendant la période  $[t, t+\Delta t[$  et PA le nombre de personnes-années cumulé sur la période  $[t, t+\Delta t[$ . Le taux d'incidence n'est pas une probabilité. En particulier, on exprime sa valeur en nombre de cas par personnes-temps. Par exemple, si l'unité de temps est l'année,  $TI = 10^{-4}$  se lit 1 cas pour 10 000 personnes-années (ou 10 cas pour 100 000 personnes-années).

Lorsqu'on estime le taux d'incidence sur un échantillon, son intervalle de confiance est donné par :

$$TI \pm z_{\alpha/2} \sqrt{\frac{TI}{PA}}$$

La plupart des calculs concernant le taux d'incidence sont sous-tendus, sur le plan mathématique, par l'existence d'une variable T qui mesure la date de survenue de l'événement étudiée (par exemple, la maladie ou le décès), et par la notion d'incidence instantanée (notée  $\lambda(t)$ ) qui lui est associée. La variable T a la propriété d'être «censurée» pour certains sujets, ceux pour lesquels l'événement n'est pas survenu pendant la période de suivi et pour lesquels la seule information est que T est supérieure à leur durée de suivi. Ces notions dépassent le cadre de ce livre et peuvent être trouvées ailleurs (par exemple: Hill et coll., 1990; Bouyer et coll., 1993; Estève et coll., 1993) sous le nom de «données de survie» (en référence à l'étude de la mortalité qui a été leur premier champ d'application) ou de «données censurées».

\*  $z_{\alpha/2}$  est le percentile  $(1-\alpha/2)$  de la loi normale centrée réduite. Le plus souvent, on prend  $\alpha = 5\%$  et on a alors  $z_{\alpha/2} = 1,96$ .

Notons enfin que, lorsque la population est stationnaire — c'est-à-dire lorsque aucune des caractéristiques de la maladie (telle que taux d'incidence ou prévalence) n'évolue au cours du temps —, la prévalence et le taux d'incidence sont liés par la relation  $P = \frac{TI d}{1 + TI d}$  où  $d$  est la durée moyenne de la maladie. Dans le cas, fréquent, où  $TI d$  est petit, cette relation est approchée par  $P = TI d$  qui illustre bien les deux dimensions de la prévalence indiquée précédemment.

### 3.3 Risque cumulé de maladie (ou incidence cumulée)

Par définition, le risque cumulé est la probabilité de devenir malade au cours d'une période fixée. Cela nécessite donc de préciser la durée de la période considérée. Le calcul est facile dans une population fermée (sans entrée ni sortie de sujets) et sans sujets perdus de vue: il suffit de diviser le nombre de nouveaux cas par le nombre de sujets non malades au début de la période. Sinon (population ouverte ou sujets perdus de vue), le risque cumulé de maladie pendant la période  $\Delta t$  est donnée par  $R(\Delta t) = 1 - \exp\{-TI \Delta t\}$  (Bouyer et coll., 1093). Il s'agit alors d'une probabilité conditionnelle (à l'absence de censure) pendant la période à l'étude.

Si le taux d'incidence est petit (ou plus précisément si  $TI \Delta t$  est petit), et seulement dans ce cas, cette expression est approchée par  $R(\Delta t) = TI \Delta t$ . On voit donc que, si  $\Delta t = 1$  (1 an si l'unité est l'année), on obtient  $R = TI$ , ce qui est commode à retenir, mais source de confusion entre les notions de taux d'incidence et d'incidence cumulée.

Lorsqu'on estime le risque cumulé sur un échantillon, l'intervalle de confiance est donné par  $[1 - \exp(-a \Delta t); 1 - \exp(-b \Delta t)]$ , où  $[a; b]$  est l'intervalle de confiance de  $TI$ .

La formule permettant de calculer  $R(\Delta t)$  en fonction de  $TI$  qui vient d'être donnée nécessite que  $TI$  soit constant sur l'intervalle de temps  $\Delta t$ . Lorsqu'on veut calculer le risque cumulé de maladie sur une longue période, cette hypothèse n'est, en général, plus satisfaite. On doit tenir compte, par exemple, du fait que l'âge augmen-

tant, l'incidence de la maladie augmente aussi. On est alors conduit à découper la période sur laquelle on veut calculer le risque cumulé de maladie en sous-périodes au sein desquelles on peut supposer le taux d'incidence constant. Nous noterons  $p$  le nombre de sous-périodes,  $\Delta t_k$  la durée de la  $k^e$  sous-période (les  $\Delta t_k$  ne sont pas nécessairement tous égaux) et  $TI_k$  le taux d'incidence correspondant. On montre alors que le risque de maladie pendant l'ensemble de la période est\*

$$R = 1 - \prod_{k=1}^p (1 - R_k) = 1 - \prod_{k=1}^p (\exp\{-TI_k \Delta t_k\}) = 1 - \exp\left\{-\sum_{k=1}^p TI_k \Delta t_k\right\}.$$

Si les taux d'incidence  $TI_k$  sont petits (ou plus précisément si les  $TI_k \Delta t_k$  sont petits), cette expression se simplifie en  $R = 1 - \prod_{k=1}^p (1 - TI_k \Delta t_k)$ .

### Cas particulier de la mortalité

Lorsqu'on s'intéresse à la survie, l'incidence de décès est appelée mortalité, et les mêmes mesures que précédemment peuvent être décrites. Les plus connues sont le *taux de mortalité* qui est le taux d'incidence de décès et la *létaleté* qui est l'incidence cumulée ou le risque de décès parmi les personnes atteintes d'une maladie au cours d'une période donnée.

### 3.4 Mesures d'association

L'étude de l'association entre une exposition  $E$  (ou facteur de risque) et la maladie  $M$  est une des étapes majeures de la recherche des facteurs étiologiques des maladies. Plusieurs questions complémentaires se posent. D'une part, celle de l'existence même d'un lien statistique entre l'exposition et la maladie. D'autre part, celle de la mesure de la force du lien entre  $E$  et  $M$  qui permet de quantifier l'accroissement du risque en fonction de l'exposition au facteur de risque; cela nécessite de choisir une mesure d'association. Enfin, bien sûr, celle de l'interprétation de l'association lorsque celle-ci a été établie. De façon générale, une mesure d'association est une mesure descriptive: elle permet de mesurer l'association statistique entre deux variables (une exposition et la fréquence d'une maladie), mais ne permet pas directement de savoir s'il y a un lien de cause à effet entre elles. Nous reviendrons sur cette question à la fin du chapitre. Il nous arrivera, cependant, pour alléger la présen-

\* Dans les expressions qui suivent,  $\prod_{k=1}^p (1 - R_k)$  indique le produit des  $(1 - R_k)$  pour  $k$  variant de 1 à  $p$ , et  $\sum_{k=1}^p$  indique la somme.

tation, d'employer le terme «effet de E» au lieu de «mesure de l'association entre E et M». Sauf indication explicite du contraire, cela n'aura pas de sens causal.

Dans ce qui suit, le facteur de risque et la maladie sont caractérisés par des variables dichotomiques (ayant uniquement deux valeurs: présence ou absence). Pour le facteur de risque, les deux catégories sont notées E+ pour les sujets exposés et E- pour les sujets non exposés. Pour la maladie, les malades sont notés M+ et les non-malades sont notés M-. Nous nous intéresserons principalement au cas où l'on mesure le *risque* ou probabilité d'être atteint, c'est-à-dire la probabilité d'être malade à un moment donné (ou prévalence) ou la probabilité de le devenir au cours d'une période fixée. Ce risque sera noté R. Nous donnerons quelques éléments concernant le cas où on considère non pas le risque lui-même, mais un taux d'incidence ou une incidence instantanée.

### Modèle additif et modèle multiplicatif

Soit  $R_1$  le risque de maladie chez les sujets exposés au facteur de risque et  $R_0$  le risque de maladie chez les sujets non exposés.

Les deux principaux types d'indices pour quantifier l'association entre la maladie et le facteur de risque sont l'excès de risque (modèle additif) :  $\Delta = R_1 - R_0$ ; et le risque relatif (modèle multiplicatif) :  $RR = \frac{R_1}{R_0}$

Bien entendu, d'autres indices sont envisageables. Certains, tout en donnant une expression numérique particulière de la relation entre l'exposition et la maladie, se ramènent finalement à l'un des deux précédents. C'est le cas de la différence relative  $\frac{R_1 - R_0}{R_0}$  qui apporte la même information que le risque relatif puisqu'elle est égale à  $RR - 1$ . Cette dernière est aussi appelée excès de risque relatif; il s'agit d'une mesure du risque fréquemment utilisée en évaluation de risque (Krewski et coll., 1999). D'autres sont de nature plus différente. Le plus utilisé est l'odds ratio défini par  $OR = \frac{R_1 / (1 - R_1)}{R_0 / (1 - R_0)}$ . Il s'agit du rapport de la quantité  $\frac{R}{1 - R}$  calculée chez les exposés à sa valeur chez les non-exposés. La quantité  $\frac{R}{1 - R}$  est appelée «odds» en anglais;

d'où le nom d'odds ratio donné à la dernière mesure définie. En français, certains auteurs reprenant l'analogie des courses de chevaux, où la probabilité de victoire est présentée sous la forme de «cote» similaire à l'odds, utilisent le terme «rapport de cotes» (Bernard et Lapointe, 1987). Nous expliquerons plus loin les raisons qu'il peut y avoir à utiliser l'OR.

Le risque relatif ou l'OR peuvent varier de 0 à l'infini. Un risque relatif de 2 signifie que le risque de maladie est le double dans le groupe exposé par rapport au groupe non exposé. L'absence de relation entre la maladie et le facteur de risque correspond à un excès de risque égal à 0, un risque relatif égal à 1 ou encore un odds ratio égal à 1. Si l'on s'en tient à la seule mesure de l'association entre E et M dans une population donnée, il y a peu de raisons de choisir l'un des trois indices ( $\Delta$ , RR ou OR) plutôt qu'un autre. Sur le plan qualitatif, ils sont équivalents: ils indiquent en même temps l'existence ou l'absence d'une relation entre exposition et maladie, puisque les égalités  $\Delta = 0$ ,  $RR = 1$  et  $OR = 1$  sont équivalentes. Sur le plan quantitatif, leurs valeurs numériques sont différentes, mais s'ordonnent de façon identique. Si, pour deux expositions  $E_1$  et  $E_2$ , on a  $\Delta_1 > \Delta_2$ , on a aussi  $RR_1 > RR_2$  et  $OR_1 > OR_2$ . Le choix entre modèle additif et multiplicatif repose sur ce que l'on connaît de la nature du lien entre l'exposition et la maladie. Les arguments directs fondés sur le mécanisme détaillé de la maladie sont rarement disponibles. Parmi les exceptions notoires, citons la cancérogenèse où un modèle multi-étapes de développement des tumeurs a été proposé par plusieurs auteurs et s'est montré en accord avec de nombreuses observations (Armitage et Doll, 1961; Breslow et Day, 1980). Ce modèle spécifie que, pour qu'une cellule donne naissance à une tumeur décelable, plusieurs étapes distinctes, ordonnées et irréversibles sont nécessaires. Chaque étape a une certaine probabilité de se produire dans un délai fixé. L'exposition à des cancérogènes modifie ces probabilités. On peut montrer que ce modèle est souvent en accord avec le modèle multiplicatif; par exemple, si deux facteurs de risque distincts agissent sur deux étapes différentes, leurs effets conjoints se multiplient.

En dehors de cas particuliers tels que le modèle multi-étapes, c'est en général pour des raisons empiriques — le constat que tel ou tel modèle présente une meilleure adéquation aux

observations — que l'on préfère un modèle plutôt qu'un autre. On choisira ainsi le modèle multiplicatif si, lorsque l'on change de population d'étude,  $\frac{R_1}{R_0}$  reste stable et le modèle additif si c'est  $R_1 - R_0$  qui reste stable. En pratique, «changer de population d'étude» signifie, par exemple, qu'on s'intéresse à l'association entre une même exposition et une même maladie dans des populations de différents pays ou de différentes régions ou, au sein d'une même population, dans différentes classes d'âge, et plus généralement dans différents sous-groupes de cette population. Dans la plupart des cas qui ont été examinés, c'est le modèle multiplicatif qui apparaît le plus conforme aux observations (Breslow et Day, 1980).

Dans ce chapitre, nous suivons ces constatations empiriques en nous intéressant essentiellement au modèle multiplicatif. L'association entre exposition et risque sera donc mesurée par le risque relatif ou par l'odds ratio.

### Risque relatif et odds ratio

Le risque relatif et l'odds ratio ont des qualités différentes et symétriques sur le plan de l'interprétation des résultats et sur celui de leurs propriétés statistiques. Le risque relatif permet d'exprimer l'association entre l'exposition et la maladie de façon facilement interprétable: c'est le facteur par lequel le risque de maladie est multiplié en présence de l'exposition. Les résultats présentés à l'aide du risque relatif se trouvent ainsi aisément explicables et compréhensibles pour un public non spécialisé en épidémiologie. L'odds ratio, lui, exprime la relation entre  $R_0$  et  $R_1$  de façon moins immédiate. À l'inverse, l'odds ratio peut être estimé dans tous les types d'enquête, alors que ce n'est pas le cas du risque relatif. On montre de plus que l'odds ratio a de meilleures propriétés mathématiques que le risque relatif lorsque la fréquence de la maladie est caractérisée par un risque. En particulier, les modèles multivariés tels que le modèle logistique permettent d'estimer l'OR et non le RR. On préfère donc souvent l'odds ratio pour les calculs statistiques.

Heureusement, dans de nombreuses situations rencontrées en épidémiologie, OR est peu différent de RR, de sorte que l'on peut à la fois satisfaire les exigences des statistiques en estimant l'odds ratio et celles d'une meilleure

présentation des résultats en interprétant sa valeur comme on le ferait pour celle du risque relatif. En effet, l'odds ratio et le risque relatif sont liés par la relation  $OR = RR \frac{(1 - R_0)}{(1 - R_1)}$ . Lorsque la maladie est rare ( $R_0$  et  $R_1$  petits), les valeurs numériques du risque relatif et de l'odds ratio sont donc proches.

D'un point de vue pratique, lorsque l'exposition est caractérisée par une variable dichotomique ( $E^+$  pour les exposés et  $E^-$  pour les non-exposés) et que la maladie est caractérisée par un *risque* (c'est-à-dire par la probabilité d'être malade ou de le devenir au cours d'une période donnée), les données obtenues sur un échantillon d'enquête peuvent se présenter sous la forme d'un tableau à quatre cases où  $M^+$  désigne les sujets malades ou devenus malades au cours de la période d'étude et  $M^-$  les sujets non malades:

	$E^+$	$E^-$	
$M^+$	a	b	$m_1$
$M^-$	c	d	$m_0$
	$n_1$	$n_0$	

Alors, quel que soit le type de l'enquête, l'estimation de l'odds ratio est  $OR = \frac{ad}{bc}$

Pour calculer l'intervalle de confiance, il existe plusieurs formules. Elles sont toutes approchées et donnent des résultats très proches sur le plan numérique. Une des méthodes couramment utilisée consiste à calculer d'abord l'intervalle de confiance de  $\ln OR$  grâce à la formule:

$$[B_i; B_s] = \ln OR \pm z_{\alpha/2} \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

On obtient ensuite l'intervalle de confiance de OR en prenant l'exponentielle des bornes:  $[\exp(B_i); \exp(B_s)]$ .

Dans les enquêtes de cohorte (avec population fermée), l'estimation du risque relatif (rapport des incidences cumulées) est donnée par  $RR = \frac{a/(a+c)}{b/(b+d)}$ . Le calcul de l'intervalle de confiance passe aussi par celui du logarithme:  $\ln RR \pm z_{\alpha/2} \sqrt{\frac{c}{n_1 a} + \frac{d}{n_0 b}}$  dont on prend l'exponentielle des bornes. Dans les enquêtes transversales, on peut aussi calculer l'équivalent, soit le ratio de prévalences, mais on a tendance à utiliser de plus en plus l'odds ratio (Rothman et Greenland, 1998).

### Mesures d'association fondées sur le taux d'incidence ou l'incidence instantanée

À partir du taux d'incidence ou de l'incidence instantanée d'une maladie chez des sujets exposés ( $TI_1$  ou  $\lambda_1(t)$ ) et non exposés ( $TI_0$  ou  $\lambda_0(t)$ ), on peut également définir deux mesures d'association: l'excès de risque [différence  $TI_1 - TI_0$  ou  $\lambda_1(t) - \lambda_0(t)$ ] et le risque relatif [rapport  $TI_1/TI_0$  ou  $\lambda_1(t)/\lambda_0(t)$ ]. Notons cependant que l'odds ratio ne peut pas être défini, car une incidence n'est pas une probabilité.

### Mesure d'impact potentiel

Le risque relatif ou l'odds ratio mesurent l'intensité du lien entre le facteur de risque et la fréquence de la maladie au sein d'une population, mais ils ne permettent pas d'évaluer l'impact du facteur de risque au niveau de l'ensemble de la population, c'est-à-dire la proportion des cas qui pourraient être évités si le facteur de risque était absent.

Cette information est donnée par la fraction de risque attribuable, qui mesure la proportion des cas «attribuable» à l'exposition étudiée (Coste et Spira, 1991) et qui est notée RA. Il faut bien sûr insister sur le fait que la notion de cas «attribuables» à l'exposition n'a de sens que si l'exposition est un agent causal de la maladie. Nous ferons dans un premier temps cette hypothèse pour calculer la fraction risque attribuable; nous reviendrons ensuite sur l'interprétation pratique de cet indice.

La fraction de risque attribuable dans la population est définie, lorsque le risque relatif est supérieur à 1, par  $RA = \frac{P_E(RR-1)}{P_E(RR-1)+1}$ , où RR est le risque relatif et  $P_E$  la proportion de sujets exposés dans la population. Son estimation et son intervalle de confiance dépendent du type d'enquête dans lequel elle est calculée. Le calcul de la fraction de risque attribuable s'étend au cas où l'exposition comporte plus de deux classes (Bouyer et coll., 1993).

Deux autres mesures d'impact potentiel sont parfois utilisées: la fraction préventive et la fraction étiologique.

La fraction préventive, notée FP, est la proportion de cas de maladie évités par la présence d'un facteur dans le cas où l'exposition est un facteur protecteur pour la santé ( $RR < 1$ ). C'est effectivement l'équivalent de la fraction de risque attribuable pour les facteurs protecteurs. On montre que  $FP = P_E(1-RR)$ .

La fraction étiologique FE est, par définition, la proportion de cas attribuables à l'exposition *parmi les sujets exposés*. Pour la calculer, il suffit de reprendre l'expression de la fraction de risque attribuable et de l'appliquer à une population où 100 % des sujets sont exposés, c'est-à-dire  $P_E = 1$ . On obtient  $FE = \frac{RR-1}{RR}$ . On remarque que la FE n'apporte pas d'information supplémentaire par rapport au risque relatif, puisque l'un se déduit de l'autre. Elle permet cependant d'exprimer la relation entre l'exposition et la maladie par un indice qui fournit un renseignement important pour la prévention. En effet, pour un groupe de sujets exposés au facteur de risque, la fraction étiologique fournit directement la proportion de cas qui disparaîtrait en l'absence du facteur. Ainsi, même pour un risque relatif assez petit, par exemple  $RR = 2$ , FE est égale à 50 %. Autrement dit, la suppression du facteur de risque conduit à une diminution de 50 % du risque de maladie chez les sujets exposés (si, comme nous le supposons depuis le début de ce paragraphe, la relation entre l'exposition et la maladie est causale).

Revenons pour conclure sur l'interprétation de la fraction de risque attribuable (RA). Elle a été construite pour mesurer la proportion des cas qui disparaîtrait en l'absence de l'exposition. Cela suppose que le lien entre l'exposition et la maladie soit bien causal. Si, par exemple, on indique que la fraction de risque attribuable de maladie coronarienne associée au tabac est égale à 27 %, c'est parce que l'on considère que la consommation de tabac influence de façon causale le risque de maladie coronarienne et donc que 27 % des cas pourraient disparaître si personne ne fumait dans cette population. Il arrive cependant qu'on s'écarte de cette rigueur. La notion de fraction de risque attribuable est en effet particulièrement importante en santé publique, car elle permet d'estimer la proportion de cas évitables si des actions sont entreprises pour supprimer l'exposition. Cela conduit, en pratique, à étendre l'utilisation de la fraction de risque attribuable à des situations où la causalité, tout en étant plausible, n'est pas démontrée. Cette pratique est très discutable et nécessite une très grande prudence dans l'interprétation de ses résultats. Par ailleurs, comme plusieurs facteurs peuvent concourir de façon conjointe à l'apparition de la maladie, la frac-

tion de risque attribuable «brute» présentée ci-dessus tend à surestimer l'impact du facteur de risque dans la population. En particulier, la somme des fractions de risque attribuables à différents facteurs de risque est souvent supérieure à 100 % (Rothman et Greenland, 1998). Il est possible de corriger, au moins partiellement, ce problème en calculant une fraction de risque attribuable ajustée sur d'autres facteurs de risque, ou une fraction de risque attribuable à l'effet conjoint de plusieurs facteurs de risque (Bruzzi et coll., 1985; Coste et Spira, 1991).

#### 4. TYPES D'ENQUÊTES ÉPIDÉMIOLOGIQUES

Les enquêtes épidémiologiques se divisent en plusieurs grandes catégories comme le schématise la figure 4.1. Une première division sépare les enquêtes d'observation des études expérimentales. Les enquêtes d'observation se divisent ensuite en enquêtes descriptives et étiologiques qui ont des objectifs différents comme leur nom l'indique. Enfin, les enquêtes étiologiques se divisent en trois catégories principales selon leur méthodologie: étude de cohorte, étude cas-témoins et étude transversale. Comme toute classification, celle que nous présentons comporte une part d'arbitraire, et certaines enquêtes ont du mal à y trouver leur place. D'autres types d'enquêtes peuvent être décrits; il s'agit en fait d'enquêtes «hybrides» par rapport aux trois précédentes (Kleinbaum et coll., 1982).

##### 4.1 Études expérimentales

De façon générale, on qualifie d'étude expérimentale toute enquête où l'attribution de l'ex-

position est contrôlée par l'investigateur, c'est-à-dire que ce dernier a pu choisir quels sujets sont exposés et lesquels ne le sont pas, ainsi que le type d'exposition. De façon générale, l'étude expérimentale s'apparente à une étude de cohorte dans laquelle l'exposition est sous le contrôle du chercheur. La capacité à montrer que les associations mesurées sont de nature causale est meilleure dans les situations expérimentales que dans les enquêtes d'observation. L'avantage, comparativement aux méthodes d'observation, est alors de pouvoir isoler l'exposition étudiée des autres facteurs de risque (facteurs de confusion) afin que tout changement, dans l'effet observé, puisse être attribué uniquement à l'exposition. Cette attribution causale est d'autant plus aisée qu'il y a tirage au sort et que la comparabilité des groupes a été maintenue tout au long de l'enquête. On parle alors d'essai «randomisé».

Comme la méthode entraîne une exposition choisie par le chercheur et non par le participant, ce type d'étude nécessite que des conditions soient réunies pour garantir la sauvegarde des droits de la personne. En particulier, les participants doivent être clairement informés des avantages et inconvénients résultants de leur participation, ils doivent signer un formulaire de *consentement* confirmant leur volonté à participer et aussi avoir la possibilité en tout temps d'abandonner l'étude sans aucun préjudice. En fait, compte tenu de sa nature (visant à exposer volontairement des sujets), cette méthode est utilisée en épidémiologie principalement pour étudier l'effet d'une intervention à visée préventive.

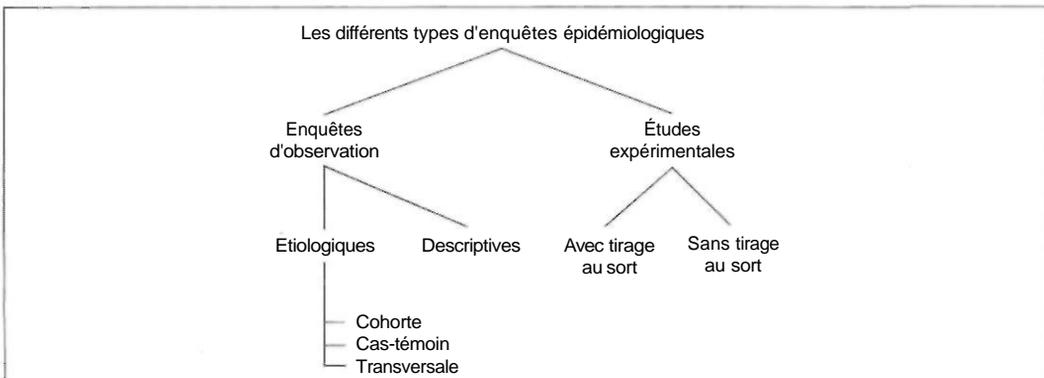


Figure 4.1 Les différents types d'enquêtes épidémiologiques

### Essai clinique randomisé

Il s'agit de la méthode de référence pour évaluer l'efficacité d'un traitement, qu'il soit médicamenteux ou non. Afin d'isoler l'effet de la composante active du traitement (par rapport à l'effet psychologique attendu, appelé communément «effet placebo»), il est courant d'utiliser dans le groupe ne recevant pas le traitement une exposition de type *placebo* ou intervention simulée\*. Puisque la connaissance du traitement utilisé peut influencer le diagnostic de la maladie fait par le médecin, mais aussi le respect du protocole par le patient, il est usuel que ni le médecin et ni le patient n'aient connaissance du véritable traitement. Cette méthode est ainsi parfois appelée *méthode en double aveugle*. On trouvera plus de détails sur ce type d'études dans des ouvrages spécialisés (Laplanche et coll., 1986; Bouvenot et Vray, 1999).

En contrepartie de ses avantages pour évaluer l'efficacité d'un traitement, la grande standardisation des essais randomisés diffère des situations généralement rencontrées dans la réalité où les expositions ne sont pas distribuées au hasard ni de façon constante au cours du temps. Le recours à des expériences randomisées est en fait assez peu fréquent en épidémiologie environnementale. Elles se rencontrent principalement dans les essais thérapeutiques destinés à étudier l'efficacité des traitements. À titre d'exemple, l'essai clinique a été utilisé pour évaluer l'efficacité des traitements (médicamenteux ou non) de l'intoxication au plomb (Lanphear et coll., 1999; O'Connor et Rich, 1999). Ce type d'étude est aussi parfois utilisé pour évaluer l'effet possible d'une exposition environnementale à des niveaux inférieurs aux normes en vigueur, particulièrement dans le cas de courte exposition. Ainsi, on peut citer le cas de l'effet des expositions aux champs magnétiques d'extrêmes basses fréquences (Selmaoui et coll., 1996).

### Essai préventif

Il s'agit d'une étude expérimentale s'adressant généralement à des personnes bien portantes, dans le but de prévenir l'apparition de maladies éventuelles. Ce type de méthode est utilisé fréquemment en médecine préventive afin d'évaluer certaines interventions préventives. C'est

le cas en particulier de l'évaluation de l'efficacité des vaccins. Peu d'essais préventifs à notre connaissance ont été réalisés en santé environnementale. On peut cependant citer l'exemple de l'étude de l'efficacité de suppléments de vitamines pour la prévention du cancer du poumon (Omenn et coll., 1996).

### Essai communautaire

Le traitement ou l'exposition se fait alors au niveau d'une communauté (ville, village, école, etc.) plutôt qu'au plan individuel. L'intervention est habituellement de nature préventive, et l'étude vise à évaluer son efficacité pour réduire l'apparition de problèmes de santé. L'essai peut être véritablement de nature expérimentale (l'exposition étant choisie par le chercheur et son allocation étant faite de façon aléatoire). L'essai peut aussi être de nature *quasi expérimentale*. Dans ce dernier cas, le choix de l'exposition peut être décidé selon différents critères de faisabilité. Cependant, on essaie toujours d'avoir un groupe sans traitement le plus comparable possible au groupe avec traitement. Le consentement de participation est habituellement donné par les responsables politiques ou administratifs de la communauté sous étude. Des procédures doivent cependant être établies pour que les individus participants soient informés des objectifs de l'étude. On retrouve dans la littérature plusieurs exemples de l'application de cette méthode en santé environnementale, dont celui historique de la fluoruration des eaux de consommation (Arnold et Dean, 1956).

### 4.2 Enquêtes descriptives

La distinction généralement faite entre études descriptives et étiologiques comporte une large part d'arbitraire. Les secondes doivent fournir des arguments essentiels en faveur ou à l'encontre de l'hypothèse d'un rôle étiologique des facteurs de risque étudiés alors que les premières ont pour objectif principal de fournir des statistiques permettant de connaître l'état sanitaire de la population (fréquence de la maladie, tendances temporelles ou géographiques) sans le mettre explicitement en rapport avec des facteurs de risque.

\* Cependant, lorsqu'un traitement reconnu efficace est disponible, l'utilisation d'un placebo sera considérée comme non éthique, le traitement reconnu sera alors généralement utilisé comme traitement de référence (dans le groupe *contrôle*) afin de comparer son efficacité à celle du nouveau traitement.

En pratique, il y a souvent un passage continu des unes aux autres. Très peu d'enquêtes descriptives se limitent à présenter la fréquence d'une maladie à un moment donné. Le plus souvent, les variations temporelles ou géographiques sont décrites et peuvent être le point de départ d'hypothèses sur les causes de la maladie.

Traditionnellement, la production périodique de statistiques de santé, par exemple les taux d'incidence ou de mortalité par sexe et groupe d'âge, permet la surveillance de l'état de santé des populations sur une base régionale ou nationale (Thacker et Stroup, 1998). L'étude plus précise des variations de ces taux selon des caractéristiques telles que l'âge, le sexe ou la catégorie socio-professionnelle, l'étude de leur évolution dans le temps ou dans l'espace a souvent permis de générer des hypothèses sur des facteurs étiologiques potentiels liés à ces variations. La recherche plus formelle de corrélations au niveau des groupes entre les paramètres de santé et leurs déterminants potentiels est l'objet des études écologiques.

### **Études des variations selon les caractéristiques de personnes**

L'étude des variations de l'incidence d'une maladie selon le sexe, l'âge ou le niveau socio-économique est souvent une première étape dans toute étude descriptive et permet souvent de soulever certaines hypothèses étiologiques. Ainsi, l'examen attentif des variations de fréquence en fonction du sexe permet souvent d'évoquer des hypothèses sur le rôle possible de facteurs médicaux liés au sexe (hormones) ou de facteurs environnementaux: habitudes de vie (consommation de tabac ou d'alcool, par exemple) ou encore d'expositions professionnelles, agissant séparément ou en combinaison. La fréquence de la plupart des pathologies a tendance à augmenter avec l'âge, mais l'observation de pics d'incidence à des âges plus jeunes (pour les cancers de l'enfant, par exemple) permet d'identifier des groupes d'étiologies *a priori* différentes. Les catégories socio-professionnelles réunissent un ensemble de déterminants sociaux (revenus, niveau d'études, accès aux soins) définissant des groupes qui présentent un risque inégal vis-à-vis de la maladie. Les études de l'INSEE en France (Institut national de la statistique et des études économiques) ont mis en évidence dès les années 1980 - surtout chez les

hommes - des disparités de mortalité entre groupes sociaux. Ainsi, les cadres et professions libérales sont les mieux protégés, suivis des professions intermédiaires, des agriculteurs, des artisans et commerçants, puis des employés et enfin des ouvriers. Dix ans après, les différences persistent et se retrouvent également au niveau de la morbidité (Desplanches et coll., 1996).

### **Variations géographiques et temporelles**

Les Systèmes d'informations géographiques (SIG) permettent d'enregistrer, d'indexer et de retrouver des informations recueillies sur une base géographique, qu'il s'agisse de paramètres de santé, de données démographiques ou de mesures d'exposition environnementales, par exemple. L'analyse géographique des cartes ainsi produites permet de nombreuses applications sur le plan de la santé publique allant de la production de données administratives (caractéristiques des populations et de leur recours aux soins, localisation des structures de soins existantes, etc.) à la génération d'hypothèses (identification de clusters, hypothèse sur un modèle de propagation de maladies infectieuses, ou sur les causes possibles des variations géographiques observées pour les maladies non transmissibles). Les SIG peuvent aussi servir de soutien aux études étiologiques (Neutra, 1999). En effet, en épidémiologie environnementale, l'exposition est souvent définie sur une base écologique à partir des lieux de résidence des sujets (pollution de l'air, résidence à proximité d'une source de pollution), et le développement de ces systèmes permet des estimations rapides et plus précises des expositions individuelles, y compris intégrées dans le temps (Beyea et Hatch, 1999). C'est ainsi que, dans une étude sur les relations entre l'asthme infantile et le trafic routier, les SIG ont été utilisés pour préciser l'exposition de cas et témoins: les adresses des enfants ont été saisies sur une base géographique («géocodées») et un cercle d'environ 150 m a été défini autour de chaque adresse; la liste des rues et voies à grande circulation traversant les différents cercles, la longueur des segments correspondants et le trafic sur ces voies, exprimé en nombre moyen de voitures par jour de semaine, ont été obtenus à partir d'un autre SIG municipal (English et coll., 1999).

Traditionnellement, l'étude des effets de la pollution de l'air s'appuie surtout sur les variations temporelles des phénomènes. La mesure

périodique des polluants dans l'air (particules, oxydes d'azote, oxydes de soufre) a permis l'étude des effets aigus, le plus souvent respiratoires ou cardiaques, des pics de pollution grâce à l'étude conjointe des variations chronologiques des mesures et des taux de morbidité à quelques jours d'intervalle.

### Études écologiques

La mise en évidence de variations géographiques ou temporelles de la fréquence de maladies conduit naturellement à vouloir mettre en relation des variations concomitantes des fréquences d'exposition à des facteurs de risque potentiels, obtenues collectivement sur les mêmes populations. Il s'agit d'*études de corrélation écologique* ou *études écologiques*. Le terme «écologique» est employé ici, par opposition à individuel, pour signifier le fait que les unités statistiques de ces études ne sont pas des individus mais des populations. On cherchera à corrélér, par exemple, les niveaux moyens de contamination des eaux par les pesticides obtenus par commune avec les fréquences de pathologies potentiellement associées, obtenues dans les mêmes communes. Le caractère écologique d'une exposition est particulièrement vrai dans les études environnementales dès que les variations géographiques ou temporelles des niveaux d'exposition excèdent les variations individuelles. C'est le cas bien sûr de la pollution de l'air. En ce qui concerne la pollution des eaux, les habitudes individuelles de consommation (usage d'eau embouteillée en particulier) introduisent une composante individuelle plus forte.

Même si ces études ont des limites importantes (liées en particulier à l'absence d'information et de contrôle sur les variables agissant au niveau individuel), elles ont connu récemment un regain de popularité en santé publique (Susser, 1994a et 1994b). En fait, il est des situations où non seulement ce type d'études est le seul possible, mais aussi le plus adéquat. C'est en particulier le cas des évaluations d'interventions communautaires telles que celles que nous avons mentionnées dans la section études expérimentales (essai communautaire). L'objectif étant alors d'évaluer véritablement l'effet d'une intervention sur le plan d'une communauté, incluant ainsi tous les aspects de faisabilité et d'interaction. Pour plus de détails sur les types d'études écologiques et leurs problèmes

methodologiques spécifiques, nous renvoyons le lecteur à des articles plus spécialisés (Walter, 1991a et 1991b; Morgenstern, 1998).

### Investigation d'agrégats

Inévitablement, il arrive périodiquement que des cas de maladies non transmissibles apparaissent de façon concentrée dans le temps et dans l'espace. Ce regroupement ou agrégat de cas peut parfois être le signe d'une cause environnementale commune à l'ensemble de ces cas, mais le plus souvent n'être que le résultat de fluctuations statistiques. Les départements de santé publique sont souvent interpellés sur l'existence de tels agrégats et ont été amenés à développer un certain nombre de règles de décision sur les investigations à entreprendre (Rothenberg et Thacker, 1992; De Wals et Niyonsenga, 1996). En fait, l'investigation d'agrégats est fréquemment décevante, et le plus souvent aucune cause précise ne peut expliquer l'agrégat observé, même lorsque la pathologie étudiée se développe rapidement après l'exposition, ce qui est le cas par exemple des anomalies de la reproduction (Kallen, 1988).

La situation est plus favorable lorsque la surveillance des taux d'incidence permet de mettre en évidence une sur-incidence localisée et persistante dans le temps. C'est ainsi que la cartographie de l'incidence des cancers au Canada avait révélé un taux élevé de cancer de la vessie chez les hommes dans la région du Saguenay au Québec (Wigle, 1977). Le fait que cet excès ait été surtout visible chez les hommes a permis de suspecter l'existence d'un risque professionnel spécifique à la région. De fait, les investigations complémentaires mirent en cause les expositions aux produits de combustion dans le département de l'électrolyse de l'industrie de l'aluminium (Thériault et coll., 1984).

Plus récemment, une incidence élevée de cancers respiratoires en Nouvelle-Calédonie était expliquée, à l'issue d'investigations épidémiologiques et environnementales, par l'utilisation par les Mélanésiens d'un badigeon à base d'amiante de type trémolite pour recouvrir le mur des maisons (Goldberg et coll., 1995; Luce et coll., 2000).

En résumé, si de nombreux agrégats ou pics d'incidence sont le fait du hasard, certains d'entre eux sont la conséquence d'une cause environnementale, déjà connue ou nouvelle. Ces

«alarmes» sont d'autant plus susceptibles d'aboutir à l'identification d'une cause, que la population étudiée est stable, que le facteur de risque est fréquent dans la population locale (et peu fréquent ailleurs!) et que le risque relatif est élevé.

### 4.3 Enquêtes étiologiques

Dans les enquêtes étiologiques, la méthode utilisée consiste à comparer des groupes de sujets pour mettre en évidence l'association entre une exposition et une maladie ou pour connaître de façon précise les modalités de cette association. On distingue trois types principaux d'enquêtes étiologiques: les enquêtes de cohorte, les enquêtes cas-témoins et les enquêtes transversales. Dans un schéma temporel où l'exposition E précède la maladie M, les trois types d'enquête se distinguent par le moment d'inclusion des sujets et par le type d'information recueilli. En outre, les biais potentiels les plus fréquents ne sont pas les mêmes selon le type d'enquête, ce qui conduit à des modalités et des précautions différentes dans leur protocole.

#### Enquêtes de cohorte

Le terme «cohorte» est utilisé pour désigner un groupe de sujets ayant une caractéristique commune. Il peut s'agir d'une génération, définie par une même période de naissance, d'un corps de métier (les salariés de l'industrie du caoutchouc ou l'ensemble des médecins d'un pays), de personnes ayant subi une exposition particulière (la cohorte des personnes irradiées à Nagasaki ou Hiroshima ou celle des femmes traitées pour un cancer du sein), de porteurs d'une caractéristique génétique (comme la trisomie 21) ou encore de sujets réunis sur une base géographique (cohorte des habitants de Framingham) (Feinleib, 1985).

Les enquêtes de cohorte consistent à étudier la morbidité (ou la mortalité) observée dans une ou des cohortes ayant différents niveaux d'exposition au facteur sous étude. On se limite bien sûr aux sujets initialement indemnes de la maladie, et, très souvent, les sujets sont séparés en fonction de leur exposition connue à un facteur de risque, soit parce qu'ils constituent dès le départ des cohortes distinctes, soit parce que cette séparation est effectuée au sein de la cohorte initiale. Dans la définition que nous adoptons ici, les enquêtes de cohorte nécessitent

une période de suivi entre l'inclusion d'un sujet et la mesure de l'altération de santé. On les appelle également parfois «enquêtes longitudinales», même si initialement cette dénomination était réservée aux études comportant des mesures répétées au cours du temps chez un même individu. Lorsque l'exposition est dichotomique et que l'on compare l'incidence de la maladie d'un groupe exposé à celle d'un groupe non exposé, on parle ainsi parfois d'enquête exposé-non exposé.

L'inclusion dans l'étude et le recueil initial des informations sur l'exposition peuvent se faire au moment de la mise en place de l'enquête. La cohorte est alors suivie de façon *prospective* au cours des années qui suivent l'inclusion; on parle donc de cohorte prospective. La durée totale de l'enquête (c'est-à-dire le délai séparant la date d'inclusion des sujets et la date où l'on arrête de les suivre et où les résultats sont analysés) varie d'une enquête à l'autre, essentiellement selon la nature de la maladie et le temps de latence entre l'exposition et la maladie. L'inclusion peut également se faire à partir d'une date déterminée du passé (date d'embauche dans une entreprise, date de premier cancer), suffisamment lointaine pour que la maladie étudiée ait eu le temps de se développer et que le temps de suivi de la cohorte soit écourté. On parle alors de *cohorte historique*. Le début de l'étude à proprement parler peut se situer à un moment intermédiaire entre l'exposition et la maladie. L'échantillon d'enquête doit être reconstitué *a posteriori* de façon à retrouver celui qu'on aurait pris pour une enquête prospective. Une enquête historique présente l'avantage de raccourcir la durée effective de l'enquête, mais la difficulté de retrouver tous les sujets devant faire partie de l'échantillon peut être accrue. Par ailleurs, la mesure des expositions peut être moins précise (alimentation, par exemple). Une des premières études de cette sorte a été publiée en 1954 par Case et coll. pour préciser le rôle respectif de différentes amines aromatiques dans la survenue du cancer de la vessie. L'enquête avait été précédée par la publication, depuis la fin du XIX<sup>e</sup> siècle, de plusieurs séries de cas de cancers de la vessie survenus dans l'industrie des colorants et mettant en cause l'aniline comme agent cancérigène. Puis, au cours des années 1940, on avait mis en évidence par expérimentation chez le chien un effet cancérigène de la bêta-naphtylamine, non retrouvé

pour la benzidine. L'étude de Case et coll. a eu une efficacité spectaculaire, en montrant que l'exposition à la benzidine, à l'alpha-naphtylamine et à la bêta-naphtylamine multipliait par 30 environ le risque de cancer de la vessie, alors que l'exposition à l'aniline n'était pas elle-même cancérigène (ce qui n'a toujours pas été infirmé à ce jour). Elle a entraîné la mise en place de mesures de protection dès la fin des années 1950 et conduit, par extrapolation, à une étude plus systématique du risque associé aux amines aromatiques, notamment dans l'industrie du caoutchouc. Le temps de latence élevé du cancer de la vessie et l'importance visible du risque imposaient ici une stratégie plus rapide que l'attitude prospective, et l'approche par enquête cas-témoins en population générale était tout à fait inadaptée du fait de la rareté de l'exposition.

Les premières enquêtes de cohortes construites pour mettre en évidence un facteur de risque de maladie sont apparues au cours des années 1950 aux Etats-Unis et en Grande-Bretagne. Elles ont d'abord été consacrées à l'étude d'associations déjà fortement suggérées par des enquêtes cas-témoins ou des études de cas, comme la cohorte des médecins britanniques de Doll et Hill (1954 et 1956) conçue pour évaluer le rôle de la consommation de tabac sur la mortalité par cancer, ou celle de R. Case et coll. (1954), sur le rôle des amines aromatiques dans le cancer de la vessie. Les enquêtes cas-témoins étaient encore mal maîtrisées et inspiraient de la méfiance, notamment parce que le fait que la connaissance du statut exposé-non exposé soit postérieure à la connaissance du statut malade-non malade peut constituer une sérieuse source de biais. À cet égard, les enquêtes de cohorte se rapprochaient davantage d'un schéma expérimental. C'est aussi ce qui a poussé Doll et Hill, au début des années 1950, à renforcer la crédibilité des résultats de leur enquête cas-témoins montrant l'effet du tabac sur les risques de cancer du poumon par une enquête de cohorte chez les médecins britanniques.

Lors de la mise en œuvre d'une enquête de cohorte, les points auxquels il faut prêter particulièrement attention sont les suivants:

- La population source doit être choisie de façon à ce que le niveau d'exposition soit suffisant et suffisamment variable, que la durée de suivi puisse être suffisamment

longue et surtout que l'information sur l'état de santé des sujets ou sur leur statut vital accessible pour toute la durée du suivi et de façon semblable pour tous les sujets.

- Les sujets exposés doivent être comparés à un groupe de référence (non exposé) qui peut être soit interne à la cohorte soit externe. Dans le second cas, on peut comparer les exposés à la population nationale ou régionale. Cela présente l'avantage d'être assez facile à mettre en œuvre et d'économiser les moyens nécessaires. En revanche, il ne faut pas que le facteur de risque étudié soit trop fréquent, il est pratiquement impossible d'ajuster sur des facteurs de confusion en dehors des facteurs démographiques (âge, sexe, période, catégorie sociale, région), et il faut penser aux biais de sélection possibles (comme le «*healthy worker effect*» ou «*effet du travailleur en bonne santé*»). Si un groupe de référence est suivi en même temps que les exposés, il doit être *a priori* comparable à la cohorte exposée en tout point sauf sur l'exposition (âge, sexe, catégorie socio-professionnelle) et avoir des possibilités de suivi identiques à celles du groupe exposé.
- L'inclusion des sujets peut consister à prendre tous ceux qui sont présents à la date de constitution de la cohorte (on parle de cohorte fixe) ou tous les nouveaux sujets (par exemple les nouveaux embauchés dans une entreprise) depuis la création de la cohorte (on parle de cohorte dynamique). Dans le premier cas, le recrutement est plus rapide, mais on risque des biais de sélection, le maintien dans la population suivie est lié à la santé ou à l'exposition.
- Le recueil de l'information concerne la survenue de la maladie et la mesure des expositions. Le critère étudié peut être la mortalité. Ce sont les enquêtes de cohorte les plus répandues, car l'information sur les décès est plus standardisée et plus accessible que l'information sur les maladies, et les valeurs de référence de la population générale sont connues. La mortalité intègre cependant à la fois la survenue de la maladie et la capacité à la traiter. C'est pourquoi les enquêtes de cohorte actuelles s'intéressent de plus en plus à l'incidence de la maladie dont les résultats sont plus faciles à interpréter.

L'analyse des enquêtes de cohorte consiste essentiellement à comparer le taux d'incidence de maladie observé chez les sujets exposés à celui d'un groupe de référence, le rapport des deux taux étant le risque relatif associé à l'exposition. Dans les enquêtes n'incluant pas de groupe de référence, on utilise le taux d'incidence des statistiques nationales ou régionales généralement standardisés sur l'âge et le sexe pour calculer un rapport comparatif d'incidence (*Standardized Incidence Ratio* ou SIR) ou un rapport comparatif de mortalité (*Standardized Mortality Ratio* ou SMR).

### Enquêtes cas-témoins

Le principe d'une enquête cas-témoins est de comparer la fréquence d'exposition antérieure à un facteur de risque dans un groupe de «cas», atteints de la maladie étudiée, et dans un groupe de «témoins» n'ayant pas cette maladie. Les sujets sont inclus dans l'étude au moment de la survenue de la maladie. Le recueil de l'information concerne l'exposition ainsi que les co-facteurs; il est toujours rétrospectif (sauf si l'enquête cas-témoins est «nichée» dans une cohorte où l'exposition ou certains co-facteurs ont déjà été collectés). Par exemple, dans une étude sur les risques de cancer associés aux herbicides, des auteurs suédois ont comparé la fréquence d'exposition aux herbicides au cours de leur vie professionnelle dans un groupe de malades atteints de lymphome (cas) et dans un groupe de taille égale, tiré au sort dans la population suédoise (témoins). Une fréquence plus élevée parmi les cas indique un lien possible entre herbicides et cancer. Un autre exemple, non plus dans la population d'une région, mais à l'intérieur de l'industrie de l'aluminium au Canada: l'exposition cumulée aux goudrons parmi les cas de cancer de vessie apparus dans l'industrie est comparée à celle d'un groupe témoin constitué de travailleurs de la même industrie n'étant pas atteints de cancer de la vessie.

On voit donc, dès maintenant, que le protocole d'une enquête cas-témoins n'est pas figé, et de nombreuses variantes sont imaginables en fonction des définitions données aux groupes des cas et des témoins; toutefois, un certain nombre de principes devront être respectés qui seront présentés plus loin.

Parmi les mesures d'association décrites précédemment, seules quelques-unes sont

estimables par une enquête cas-témoins. En effet, le principe consiste à comparer des fréquences d'exposition, et donc aucune des mesures décrites s'appuyant sur des fréquences de maladie ne pourra être utilisée. La mesure d'association utilisée est l'odds ratio (odds ratio de l'exposition qui est en fait égal à l'odds ratio de la maladie).

Une des difficultés importantes des enquêtes cas-témoins est le choix d'un groupe témoin approprié. En fait, le choix des témoins est très lié à la façon dont les cas ont été définis, et nous allons donc indiquer brièvement les éléments qui guident la définition et le choix des groupes que l'on compare.

#### *Définition et sélection des cas*

Les critères et méthodes diagnostiques permettant de caractériser les cas doivent être bien définis, objectifs et réalisés sur tous les sujets. Il faut éviter que le diagnostic soit fait à l'occasion d'un examen motivé par une exposition particulière, par exemple lors de la surveillance de travailleurs occupant certains postes en médecine du travail. On risquerait d'avoir un échantillon de cas où les sujets exposés seraient sur-représentés. Il faut enfin choisir entre maladie primitive ou secondaire.

Les cas peuvent être incidents (les nouveaux cas au fur et à mesure qu'ils se présentent) ou prévalents (tous les cas encore présents enregistrés dans les années précédant le début de l'enquête). Les seconds permettent un recrutement plus rapide des sujets d'enquête, mais ils sont sous-représentés en sujets dont la durée de survie est courte après le diagnostic. On préfère donc prendre des cas incidents pour éviter ce biais de sélection. On peut aussi envisager de prendre tous les cas diagnostiqués depuis une date fixée, ce qui implique d'inclure dans l'enquête les cas décédés depuis cette date. Il n'y a pas alors de biais de sélection, mais cela pose des problèmes importants de recueil de l'information concernant les cas décédés et de choix des témoins.

La population où les cas seront recrutés doit être choisie de façon à ce que l'exposition soit suffisamment fréquente et de niveaux variés pour que la comparaison des fréquences d'exposition ait une puissance statistique suffisante.

Il y a trois «sources» principales de cas: les malades figurant dans un registre de pathologie, par exemple, registre de cancers ou registre de

malformations congénitales (il s'agit alors de l'ensemble des malades de la région ou du pays couvert par le registre), les malades hospitalisés dans un ou plusieurs services de traitement (on parle de cas hospitaliers) ou les malades recensés dans le cadre d'une enquête de cohorte.

#### *Définition et sélection des témoins*

De façon générale, les témoins devraient être constitués d'un échantillon représentatif des non-malades de la population dont sont issus les cas. La définition de cette population ne pose pas de problème lorsque les cas sont issus d'un registre ou d'une cohorte: il s'agit respectivement de la population générale de la zone du registre et de l'ensemble des sujets de la cohorte. C'est beaucoup moins simple pour des cas hospitaliers; la population des témoins pourrait être définie par l'ensemble des sujets qui, s'ils sont malades, seraient traités dans les services dont sont issus les cas. On voit que cela dépend du mode de recrutement de ces services hospitaliers, c'est-à-dire notamment de leur position géographique et du choix des sujets eux-mêmes. On prend en général des témoins hospitaliers non atteints par la pathologie étudiée, mais atteints d'une autre pathologie. Dans ce cas, il faut faire attention que les maladies des témoins n'aient pas de facteurs de risque connus qui soient communs avec la maladie des cas.

Pour que les populations d'origine des cas et des témoins soient les plus proches possibles, on utilise souvent l'appariement. Il permet à la fois de prendre en compte un facteur de confusion (ou plusieurs) et de limiter les biais de sélection en réunissant des témoins venant d'une population semblable à celle des cas. Par exemple, on constate qu'un appariement sur l'âge, la zone de résidence et le pays d'origine a tendance à rendre aussi les deux groupes équilibrés pour le statut socio-économique et le niveau de scolarisation. Cependant, les limites de l'appariement ont été mises en évidence à plusieurs reprises et incluent la non-possibilité d'étudier le risque associé aux variables d'appariement et la possibilité de confusion résiduelle si l'effet des variables d'appariement n'est pas contrôlé au moment de l'analyse (Clayton et Hills, 1993; Rothman et Greenland, 1998). En pratique, on se limite de plus en plus à appairer les cas et les témoins pour un ou deux facteurs très fortement associés à la maladie et pour lesquels le contrôle de la confusion sera amélioré à l'aide de l'appariement.

Le recueil des informations, fait de façon rétrospective, se fait le plus souvent par questionnaire (auto-administré ou par enquêteur). Le problème principal est celui des biais de classement dû aux erreurs de mémoire. Il faut donc s'efforcer que les informations recueillies soient «objectives», que les questions soient fermées pour limiter les différences d'expression entre les sujets et que l'interrogatoire soit à l'aveugle, ce qui est souvent difficile s'agissant de malades et de témoins. Il faut que le recueil de l'information soit le plus semblable possible entre les cas et les témoins si on veut éviter qu'une différence constatée entre ces deux groupes soit attribuable au mode de recueil plutôt qu'à la maladie. Par exemple, s'il y a plusieurs enquêteurs, il faut répartir les cas et les témoins entre les différents enquêteurs. De même, si le choix a été fait d'inclure des cas décédés, il est recommandé d'avoir des témoins décédés aussi, pour que le recueil de l'information soit de même nature. Enfin, la codification des réponses doit être précise et à l'aveugle.

#### **Enquêtes transversales**

Dans les enquêtes transversales, l'échantillon est issu de l'ensemble de la population sans être sélectionné sur l'exposition (comme dans les enquêtes de cohorte) ni sur la maladie (comme dans les enquêtes cas-témoins). Les sujets inclus dans l'étude sont tous ceux qui sont présents au moment de l'enquête, et on mesure classiquement l'exposition ou l'état de santé à ce même moment. Ainsi définies, elles ont principalement un objectif descriptif, visant à évaluer la fréquence d'une exposition ou d'une maladie ou la distribution d'un paramètre de santé. Le choix de la population source et la représentativité de l'échantillon sont les principaux éléments qui conditionnent alors la qualité de l'étude.

La facilité de mise en œuvre de ce type d'enquêtes a conduit à étendre leur utilisation à la recherche étiologique. Il est alors fréquent que l'interrogatoire des sujets porte aussi sur leurs expositions et leurs maladies passées. Une enquête transversale est alors caractérisée par son type de recrutement plus que par le type d'informations recueillies. C'est ainsi qu'elles sont largement utilisées en épidémiologie des risques professionnels, dans le domaine particulier de la recherche et de l'évaluation de mar-

queurs fonctionnels ou biologiques, indicateurs précoces d'une atteinte infraclinique.

Les enquêtes transversales posent principalement deux types de problèmes. D'une part, la relation temporelle entre le facteur de risque étudié et la pathologie peut être difficile à établir (l'exposition précède-t-elle ou non la survenue de la pathologie?). D'autre part, le recrutement transversal exclut les sujets absents ou ayant quitté la population avant le début de l'enquête et peut être à l'origine de biais de sélection.

En épidémiologie étiologique, si les études de cohorte et cas-témoins constituent les principaux outils, un certain nombre d'hypothèses peuvent être raisonnablement abordées par une approche transversale dans la mesure où certaines précautions sont prises à l'étape du protocole et dans l'interprétation des résultats:

- Limitation des enquêtes transversales à l'étude de marqueurs de l'état de santé pour lesquels la prévalence est un bon indicateur et dont la survenue ne conduit pas à modifier le comportement vis-à-vis de l'exposition étudiée, afin d'éviter les biais de sélection.
- Prudence dans l'interprétation des résultats en raison des problèmes liés à la temporalité de la relation exposition-maladie. Toutefois, on a vu que, si l'échantillonnage dans les enquêtes est par définition transversal, le recueil d'informations peut être à la fois transversal et rétrospectif, ce qui permet de tenir compte de certains paramètres temporels.

### Choix du type d'enquête

Lorsqu'on organise une enquête épidémiologique, il est assez rare qu'un seul type d'enquête soit possible. Le plus souvent, plusieurs types d'enquête sont envisageables, voire tous. Il est cependant pratique d'utiliser des critères qui orientent le choix, sachant que le choix final est un compromis entre des contraintes contradictoires. Nous parlerons essentiellement des enquêtes de cohorte et cas-témoins. Les enquêtes transversales ont, pour les critères que nous envisageons dans ce paragraphe, des caractéristiques proches de celles des enquêtes de cohorte. Les principaux critères de choix sont la fréquence de la maladie et celle de l'exposition, l'intérêt principal pour l'étiologie de la maladie ou pour les conséquences de l'exposition, les biais, la temporalité entre l'exposition et la ma-

ladie, et le temps de latence, la durée et le coût de l'enquête.

Une enquête de cohorte est surtout utilisée dans deux types de situations. La première est celle de l'étude des risques associés à une exposition peu fréquente et donc exceptionnellement rencontrée dans une enquête cas-témoins (le risque de cancer de la vessie associé à l'exposition aux amines aromatiques au cours de la fabrication des colorants, le risque de second cancer associé au fait d'avoir subi une chimiothérapie, par exemple). Les études de cohorte prospectives sont bien adaptées à la surveillance des travailleurs d'un secteur industriel et à la mise en évidence des effets à faible dose. Les enquêtes de cohorte historiques sont particulièrement intéressantes pour les maladies dont le temps d'induction est élevé comme les cancers ou les maladies cardio-vasculaires, et permettent d'exploiter l'information sur les expositions passées. Les enquêtes historiques de mortalité sont *a priori* d'autant plus intéressantes que le taux de mortalité de la pathologie étudiée est plus proche de son taux d'incidence. La seconde situation est celle de l'étude des risques associés à des expositions très fréquentes pour lesquelles on veut une mesure d'une précision incompatible avec l'évaluation rétrospective ou dont la mesure peut être sérieusement influencée par le statut cas-témoin (vitamines et cancers, expositions environnementales et démence sénile).

De façon générale, les enquêtes cas-témoins sont bien adaptées à l'étude de maladies rares pour lesquelles le délai depuis le début de l'exposition est relativement long, situation dans laquelle une enquête de cohorte nécessiterait une population de grande taille, suivie sur plusieurs années. Par ailleurs, elles permettent de prendre en compte d'autres facteurs de risque qu'il n'est parfois pas possible de mesurer dans les enquêtes de cohorte historique en particulier. Elles sont également bien adaptées à l'étude d'expositions dispersées pour lesquelles la constitution d'une cohorte serait difficile. En outre, lorsqu'elles se déroulent en population générale, elles reflètent les caractéristiques des expositions «en moyenne», en termes d'intensité et de fréquence d'exposés, et permettent donc d'avoir une idée exacte de la part des différents facteurs dans l'étiologie de certaines maladies dans la population générale, par l'intermédiaire du calcul du pourcentage de risque attribuable.

Dans le cas où l'enquête se déroule à l'intérieur d'une industrie (enquête cas-témoins «nichée» au sein d'une cohorte), elle permet un gain d'efforts important, en ne s'intéressant qu'aux cas intervenus dans la cohorte et à des témoins en nombre limité choisis dans la même cohorte.

Leurs inconvénients sont liés aux biais possibles dus en particulier à la constitution d'un groupe témoin et au recueil rétrospectif parfois lointain des expositions (ou des co-facteurs). Malgré cela, leur rapidité d'exécution comparée aux enquêtes de cohorte explique leur développement actuel dans de nombreux domaines de l'épidémiologie.

Les différents points qui viennent d'être discutés montrent clairement qu'il n'y a pas un type d'enquête idéal de façon absolue. La progression des connaissances sur la nature d'une association entre une exposition et une maladie n'est en fait possible que grâce à la variété des enquêtes qui sont réalisées sur le sujet, chacune ayant ses propres faiblesses et qualités. Il convient donc d'adapter le choix du type d'enquête à chaque situation particulière en tenant compte des critères que nous avons énumérés ci-dessus.

### Plan du protocole d'une enquête étiologique

Le protocole décrit les différentes phases du déroulement d'une enquête qui vont de l'exposé explicite de la question posée et du cadre général dans lequel elle se situe jusqu'au plan d'analyse des résultats. Il faut insister sur l'importance, pour de multiples raisons, de *rédiger* l'ensemble du protocole.

L'ordre de présentation que nous avons adopté est celui que l'on utilise en général pour la rédaction finale, mais il est clair que, au moment de la discussion du protocole, les choix faits sur certains points peuvent remettre en cause ou modifier des choix faits sur des points antérieurs. Le protocole n'est finalement achevé qu'après suffisamment «d'allers et retours» pour que l'ensemble soit cohérent.

En donnant ci-dessous un plan type, nous n'avons pas d'autre ambition que de fournir un guide pour la discussion et la rédaction du protocole. Il doit être bien compris que ce plan peut et doit être adapté à chaque enquête: certains points peuvent être non pertinents (s'il n'y a pas d'appariement, par exemple) et d'autres peuvent être ajoutés si certaines précisions méritent d'être indiquées.

### Plan type du protocole d'une enquête étiologique

- 
- A. Définition du problème posé et des objectifs de l'enquête
  - B. «Traduction» du problème en hypothèses de recherche vérifiables empiriquement et réfutable par des tests statistiques
  - C. Définition des populations cibles et source
  - D. Définition de l'échantillon
  - E. Questionnaire et mode de recueil des données
  - F. Réalisation pratique de l'enquête
  - G. Principes généraux de l'analyse statistique
  - H. Publication des résultats
- 

## 5. VALIDITÉ ET PRÉCISION

### 5.1 Introduction

Les notions de *validité* et de *précision* interviennent lorsqu'il s'agit de généraliser les mesures (de fréquence ou d'association) obtenues sur un *échantillon* d'étude à une population plus grande (population cible).

La mesure obtenue sur un échantillon est une *estimation* de la vraie mesure définie au niveau de la population (RR, OR). Cette estimation serait très vraisemblablement différente si l'on étudiait un autre échantillon provenant pourtant de la même population. Ces variations, dites *aléatoires*, résultent du phénomène des *fluctuations d'échantillonnage*. La connaissance des lois statistiques qui régissent ces fluctuations permet donc d'assortir la mesure obtenue d'une précision statistique, un intervalle de confiance, par exemple. Elle permet également de conclure à l'existence d'une différence significative entre une estimation et une valeur attendue, compte tenu des fluctuations d'échantillonnage, à l'aide de *tests statistiques* appropriés (Armitage et Berry, 1971; Scherrer, 1984; Bouyer, 1996).

On dira qu'une estimation est *biaisée* (ou non valide) si un certain nombre d'erreurs *systématiques* dans la constitution de l'échantillon (*sélection*), dans le *classement* des sujets ou résultant de la présence de tiers-facteurs (*ou facteurs de confusion*) font que l'estimation obtenue au

niveau de l'échantillon est une mesure incorrecte de l'indice défini au niveau de la population. Elle ne pourra être corrigée en partie que par la prise en compte des erreurs au moment de l'analyse ou de la discussion des résultats sur le sens probable du biais.

## 5.2 Biais de sélection

Il y a un biais de sélection lorsque la population dont est extrait l'échantillon d'étude (population source) est différente de la population à laquelle on souhaite généraliser les résultats (population cible) ou lorsque les groupes de comparaison (exposés/non-exposés ou cas/témoins) ne sont pas comparables. Nous passerons successivement en revue différents types de situations courantes en épidémiologie, dans lesquelles des biais de sélection sont systématiquement ou potentiellement présents.

### Biais dans la sélection de l'échantillon d'étude

Il s'agit de situations dans lesquelles il n'est possible de recenser qu'une partie de la population source pour en extraire un échantillon d'étude. Par exemple, une population en activité professionnelle à une date donnée alors que nous nous intéressons à toutes les personnes ayant ou ayant eu une activité professionnelle, ou encore la population des malades hospitalisés, alors que nous nous intéressons à l'ensemble des personnes atteintes de la maladie.

Une étude transversale a pour particularité un recrutement «instantané» des sujets, c'est-à-dire que ne seront inclus que les sujets présents au moment de l'étude. Ceci exclut par définition tous les sujets qui satisfont par ailleurs tous les autres critères d'inclusion, mais qui, pour diverses raisons (maladie, mobilité), ne sont pas présents au moment de l'enquête. Il peut alors y avoir un biais de sélection si les raisons d'absence sont directement ou indirectement liées à la présence du facteur de risque. C'est le cas par exemple pour une exposition professionnelle entraînant le retrait du poste de travail.

### Biais dans le choix du groupe de référence

Par groupe de référence, nous entendons le groupe témoin dans les études cas-témoins ou le groupe non exposé dans les autres types d'étude.

Dans une étude cas-témoins, idéalement, le groupe témoin doit être représentatif de la population source dont sont issus les malades

(cas). Lorsque les cas ont été identifiés à l'aide d'un registre, la population source est par définition la population de la région couverte par le registre, et le choix d'un groupe témoin représentatif de cette population ne pose alors théoriquement pas de problème. Dans le cas où les témoins sont recrutés en milieu hospitalier, la population source est plus difficile à définir: en théorie, il s'agit de l'ensemble des sujets qui, s'ils étaient atteints de la maladie étudiée, seraient hospitalisés dans l'un des services où ont été identifiés les cas. Il s'agit donc de caractériser cette population en termes socio-démographiques, c'est-à-dire de rechercher les déterminants (lieu de résidence, niveau socio-économique) qui ont abouti à l'hospitalisation dans l'un des services inclus dans l'étude. L'absence de base de sondage correspondante entraînera le plus souvent la constitution d'un groupe de référence à partir, par exemple, de malades admis dans le même hôpital pour une autre cause, ayant des caractéristiques démographiques semblables aux cas (appariement). Le choix de la ou des causes d'hospitalisation «témoins» est aussi une source possible de biais de sélection, car idéalement la fréquence d'exposition au facteur de risque dans ce groupe doit être semblable à celle de la population source. Certains facteurs de risque, comme le tabac, étant à l'origine de plusieurs maladies, le choix peut être difficile. Par exemple, le groupe des malades atteints de maladies cardio-vasculaires n'est pas un bon groupe témoin dans l'étude de l'association entre consommation de tabac et cancer de vessie, dans la mesure où le tabac étant également un facteur de risque de maladies cardio-vasculaires, la fréquence d'exposition au facteur de risque sera nettement surestimée dans le groupe témoin et en conséquence la mesure de l'association tabac-cancer de vessie sera sous-estimée.

Dans les études de cohorte, le groupe exposé est souvent comparé à la population générale. Ceci est vrai en particulier dans les études de mortalité, puisque la plupart des pays publient régulièrement des statistiques de mortalité par âge, sexe et région pour chaque cause de décès. La fréquence de mortalité pour certaines causes n'est donc calculée que pour le groupe exposé et comparée ensuite aux statistiques établies dans la population générale correspondante. Là, encore, il s'agit de déterminer les caractéris-

tiques de la population source du groupe exposé de façon à définir la population source non exposée. Dans le cas d'une cohorte professionnelle, par exemple, le groupe exposé est souvent composé d'hommes, résidant d'une région donnée, dont l'état de santé leur a permis d'être employé souvent dans des postes pénibles. La population générale masculine de la même région est composée en partie de personnes dont l'état de santé est plus précaire, et globalement le taux de mortalité de cette population est égal ou supérieur à celui de la cohorte étudiée. Paradoxalement, on observe donc souvent que, malgré la présence d'une exposition professionnelle particulière ayant motivé la mise en place d'une étude, la mortalité de la cohorte étudiée est moins élevée que celle de la population de référence. Ce phénomène est appelé "*Healthy Worker Effect*" dans la littérature anglo-saxonne. Il n'exclut pas toutefois que, pour certaines causes spécifiques plus particulièrement liées à l'exposition, la mortalité soit plus élevée dans la cohorte. L'interprétation des différences observées dans ce type d'étude doit donc toujours se faire en gardant à l'esprit cette différence de construction entre les deux groupes de comparaison.

### 5.3 Biais de classement

Un biais de classement peut intervenir si une erreur systématique affecte la mesure du facteur de risque ou de l'état de santé. Si, comme dans la situation la plus simple présentée ici, les sujets sont classés de façon dichotomique pour l'exposition (exposé/non-exposé) et pour l'état de santé (malade/non-malade), les erreurs de classement aboutiront, par exemple, à considérer non-exposés des sujets qui sont exposés en réalité et vice versa, ou à considérer non-malades des sujets malades en réalité (et vice versa). Les considérations présentées ici s'appliqueraient également si le facteur de risque ou l'état de santé étaient mesurés de façon quantitative.

On distingue en général deux types d'erreurs de classement dont les conséquences sont différentes sur les estimations obtenues pour la mesure d'association: d'une part, les erreurs *différentielles* qui affectent différemment les catégories comparées, la proportion de sujets considérés exposés à tort est différente chez les malades et les non-malades ou encore la

fréquence de sujets classés malades à tort est plus grande chez les exposés que chez les non-exposés, par exemple, et d'autre part les erreurs *non différentielles* affectant également les deux catégories comparées. Dans le cas d'erreurs différentielles, le biais peut conduire à une augmentation ou une diminution de l'estimation par rapport à la vraie valeur selon les erreurs rencontrées. Dans le cas d'erreurs non différentielles, on peut montrer que, quelles que soient l'ampleur et la direction de l'erreur affectant de façon égale les deux groupes comparés, l'estimation obtenue sera biaisée vers l'absence d'association; en d'autres termes, la force de l'association sera sous-estimée (Bouyer et coll., 1993).

Les types courants d'études épidémiologiques donnent de nombreux exemples d'erreurs différentielles. Dans une étude cas-témoins, des sujets ayant subi un événement de santé grave (cas) sont comparés à des sujets en général non atteints (témoins). Lorsque l'exposition antérieure au facteur de risque est mesurée par un interrogatoire du sujet lui-même, sur ses habitudes de vie, ses expositions professionnelles passées, il est probable que les expositions potentiellement nocives, pouvant avoir un lien avec l'état de santé actuel, soient remémorées avec plus d'acuité par les cas que par les témoins, et qu'ils y soient même aidés par des enquêteurs zélés! On aboutit donc à une sous-estimation de la fréquence d'exposition chez les témoins et donc à une surestimation de l'OR. Ce biais est minimisé par l'utilisation d'un groupe témoin malade, des sujets atteints d'autres cancers témoins d'une étude sur les cancers de vessie, des enfants atteints d'autres malformations témoins d'une étude sur les spina bifida, etc. Mais, si dans ce cas les erreurs de classement sont probablement diminuées, d'autres types de biais sont probables. Les études exposés/non-exposés dans lesquelles les sujets exposés bénéficient d'un meilleur suivi médical aboutissent à une sous-estimation de la fréquence de malades chez les non-exposés et donc là encore à une probable surestimation de la mesure d'association (RR, par exemple).

On voit donc qu'il existe de nombreuses situations pouvant aboutir à une surestimation de la mesure d'association et donc à conclure à tort à l'existence d'un lien entre un facteur de risque et une maladie. Les épidémiologistes sont donc extrêmement vigilants pour éviter ce type d'erreur, utilisant différents types de techniques

(interrogatoire ou examen médical à l'aveugle, choix de groupes de comparaison plus proches) permettant de minimiser ce biais.

Les erreurs non différentielles sont celles qui continuent à affecter la mesure de l'exposition ou de l'état de santé, de façon irréductible, même si tous les efforts ont été mis en œuvre pour que les groupes comparés soient traités de façon identique. Par exemple, dans une étude de mortalité, des erreurs de classement sur la cause de mortalité sont inévitables si l'on utilise les certificats de décès. La qualité d'une étude sur les anomalies radiologiques sera subordonnée à la qualité technique des radiographies à interpréter. Dans les deux situations, les erreurs sur l'évaluation de l'état de santé, dans la mesure où elles sont non différentielles, aboutiront à sous-estimer la mesure d'association et donc éventuellement à empêcher de conclure à l'existence d'une vraie relation. La discussion des résultats doit tenir compte de ce biais.

En conclusion, une étude épidémiologique aboutira à des conclusions informatives si de bonnes garanties existent qu'un biais de classement différentiel a été évité et si l'ampleur des erreurs non différentielles n'est pas telle qu'elles aboutissent à une forte sous-estimation de la force de l'association. De nombreuses techniques existent permettant d'atteindre ces deux objectifs; quelques exemples ont été donnés, on en trouvera d'autres dans Bouyer et coll. (1993).

#### 5.4 Facteurs de confusion

Prenons immédiatement un exemple de facteur de confusion (Bouyer et coll., 1993).

Soit une population de 11 000 sujets répartis de la façon suivante:

	E <sup>+</sup>	E <sup>-</sup>
M <sup>+</sup>	184	680
M <sup>-</sup>	816	9320
Total	1000	10 000

où E<sup>+</sup> = exposés au facteur de risque étudié;

E<sup>-</sup> = non-exposés; M<sup>+</sup> = malades;

M<sup>-</sup> = non-malades.

On obtient  $RR = 18,4 \% / 6,8 \% = 2,7$ .

Si l'on répartit la même population en strates d'âge, on obtient les tableaux suivants:

	Strate 1		Strate 2		Strate 3	
	< 25 ans		25-40 ans		> 40 ans	
	E <sup>+</sup>	E <sup>-</sup>	E <sup>+</sup>	E <sup>-</sup>	E <sup>+</sup>	E <sup>-</sup>
M <sup>+</sup>	16	200	48	240	120	240
M <sup>-</sup>	184	4800	252	2760	380	1760
Total	200	5000	300	3000	500	2000
	RR1 = 2		RR2 = 2		RR3 = 2	

On voit donc que les valeurs du risque relatif (RR) dans chaque strate d'âge (RR1 = RR2 = RR3 = 2) sont différentes du risque relatif avant stratification, ou risque relatif brut (RR=2,7). On dira que l'âge a agi ici comme facteur de confusion dans la relation entre E et M. La vraie relation entre E et M est mesurée par le risque relatif par strate, c'est-à-dire une fois éliminée l'influence de l'âge sur la mesure de cette relation. L'écart entre le risque relatif brut et les risques relatifs par strate résulte de deux phénomènes:

- d'une part, la fréquence de la maladie augmente avec l'âge: elle est de 4,2 % dans la strate 1, de 8,7 % dans la strate 2 et de 14,4 % dans la strate 3;
- d'autre part, les exposés sont en moyenne plus âgés que les non-exposés: on voit en effet que 50 % des exposés ont plus de 40 ans, contre seulement 20 % des non-exposés.

Ainsi, une partie de l'augmentation du risque de maladie attribuée à l'exposition lorsqu'on considère le risque relatif brut sur l'ensemble de la population est en fait attribuable à l'âge. On voit donc qu'un facteur F est facteur de confusion s'il est à la fois facteur de risque de la maladie étudiée et lié à l'exposition E.

Il faut noter que, dans le cas où les estimations des risques relatifs ne seraient pas identiques d'une strate à l'autre, on serait en présence d'une *interaction* et non plus d'un biais de confusion, c'est-à-dire la situation dans laquelle l'effet du facteur de risque varie selon les valeurs que prend le facteur de confusion.

Les biais liés aux facteurs de confusion ont l'avantage par rapport aux autres types de biais qu'il est possible de les prendre en compte jusqu'à les éliminer au moment de la mise en place d'une enquête et de les corriger au moment de l'analyse statistique. Nous présen-

terons ici les différentes techniques permettant la prise en compte d'un facteur de confusion au moment de la mise en place d'une enquête, les techniques statistiques utilisables au moment de l'analyse (*ajustement*) seront exposées plus loin dans la section «Analyse de données».

La méthode de choix pour la prise en compte des facteurs de confusion au moment de la mise en place de l'enquête épidémiologique est le tirage au sort (*voir les études expérimentales*). Il permet en effet d'allouer au hasard les sujets provenant d'une même population source à l'un des groupes de comparaison (traité ou non traité dans le cas d'essais thérapeutiques). Le tirage au sort permet *a priori* d'équilibrer les caractéristiques des deux groupes, caractéristiques dont certaines sont des facteurs de confusion potentiels. Toutefois, si le tirage au sort reste la méthode de référence, il ne s'applique pas à la plupart des études épidémiologiques d'observation. Dans ce cas, les groupes exposés et non exposés sont déjà constitués au moment de l'étude, et les populations sources diffèrent par un certain nombre de facteurs, certains étant des facteurs de risque de la maladie étudiée et donc facteurs de confusion possibles. Une solution consiste donc à équilibrer la distribution des facteurs de confusion entre les groupes comparés (exposés/non-exposés ou cas/témoins). Cet équilibrage peut se faire par appariement, soit au niveau de strates du facteur de confusion, on appelle alors ce procédé *stratification* (les deux groupes ont la même proportion d'hommes et de femmes, par exemple), soit par *appariement individuel* (à un sujet exposé est apparié un sujet non exposé de même âge et de même sexe, par exemple). Dans les deux cas, l'une des conditions nécessaires à l'existence d'une confusion (lien entre F et E, ou lien entre F et M) n'est plus remplie, et le biais de confusion disparaît. Il faut noter que, dans le cas d'une étude cas-témoins, cet appariement ne contrôle pas complètement le biais de confusion qui doit être complété par une prise en compte au niveau de l'analyse. Dans le choix des variables d'appariement, il faut veiller également à ne pas sur-apparier, c'est-à-dire à ne pas restreindre la variabilité du facteur de risque étudié, parce qu'il était en partie lié au facteur d'appariement. Par exemple, dans une étude cas-témoins sur le rôle possible d'expositions environnementales, un appariement sur le lieu de résidence (destiné

à équilibrer les origines socio-démographiques des recrutements en milieu hospitalier, par exemple) va entraîner une trop grande similitude sur les expositions des cas et des témoins et risque de faire disparaître l'association recherchée.

Les facteurs de confusion à considérer dans une enquête, soit au moment de l'appariement soit au moment de l'analyse, sont en général choisis parmi les facteurs de risque connus de la maladie étudiée ou ses facteurs de variation (comme l'âge, le sexe ou la catégorie socio-professionnelle).

## 6. ANALYSE DES DONNÉES

### 6.1 Estimation et tests

Nous avons évoqué plus haut l'existence de fluctuations d'échantillonnage et de lois statistiques qui les régissent. Ainsi, il est possible, connaissant le pourcentage de malades dans une population, de savoir dans quel intervalle de valeurs se situeront 95 % des pourcentages observés sur différents échantillons de même taille tirés de cette même population. L'estimation d'un paramètre sur une population (fréquence, risque relatif) consiste à effectuer le chemin inverse: connaissant la valeur observée sur un échantillon, quelle est la valeur la plus probable de ce paramètre dans la population (estimation ponctuelle) et quel est l'intervalle dans lequel se situe très probablement (à 95 %, par exemple) la vraie valeur du paramètre sur la population (intervalle de confiance)?

L'estimation ponctuelle préconisée pour ses bonnes propriétés statistiques pour la plupart des indices utilisés en épidémiologie est obtenue par la méthode du maximum de vraisemblance; cette estimation est le plus souvent égale à la valeur observée sur l'échantillon. Lorsque l'estimateur (la fonction ou formule mathématique permettant d'obtenir l'estimation à partir des valeurs observées sur l'échantillon) a une distribution normale, l'intervalle de confiance de l'estimation est alors obtenu par la formule générale  $A \pm z_{\alpha/2} \sqrt{\text{var}(A)}$  où A représente l'estimation de l'indice (p, TI, RR, OR), var(A), la valeur de la variance de l'estimateur pour l'échantillon et  $z_{\alpha/2}$  la valeur limite de la loi centrée réduite correspondant au risque  $\alpha$  (par exemple, pour  $\alpha = 5\%$ , on a  $z_{\alpha/2} = 1,96$  et on parle d'intervalle de confiance à 95 %).

Certaines formules permettant le calcul de la variance (et donc de l'intervalle de confiance) ont été indiquées dans la section 3 consacrée aux mesures utilisées en épidémiologie. Pour les autres, nous renvoyons à des livres spécialisés (Scherrer, 1984; Bernard et Lapointe, 1987; Bouyer et coll., 1993).

Une autre situation dans laquelle les fluctuations d'échantillonnage doivent être prises en compte est par exemple celle, courante en épidémiologie, dans laquelle on compare un pourcentage de malades (ou fréquence de maladie) chez des exposés et des non-exposés (ce qui revient aussi à comparer un RR à 1) ou un pourcentage d'exposés chez des malades et des non-malades (ce qui revient à comparer un OR à 1). La comparaison ne peut pas se faire, bien sûr, à partir des seules estimations, mais en tenant compte des fluctuations d'échantillonnage, à l'aide de *tests statistiques* permettant de conclure à l'existence ou non d'une certaine association avec un certain risque d'erreur. Par exemple, l'existence d'une association significative entre E et M sera testée en comparant deux pourcentages observés à l'aide d'un test. L'épidémiologie utilise un sous-ensemble de tests statistiques adaptés au type de variables (souvent qualitatives) et d'hypothèses testées. L'exposé détaillé de ces tests et de leurs conditions d'application peut être trouvé dans des ouvrages d'épidémiologie plus complets (Breslow et Day, 1980, 1987; Kleinbaum et coll., 1982; Rothman et Greenland, 1988; Bouyer et coll., 1993).

## 6.2 Analyse

Dans l'étude de la relation entre un facteur de risque (E) et un état de santé (M), il est rare que le déroulement de l'étude soit suffisamment simple pour que le résultat final puisse se résumer à estimer la force de l'association entre E et M, avec son intervalle de confiance, et de tester la signification statistique de l'estimation obtenue. Le plus souvent, en effet, les différents biais évoqués plus haut (sélection, classement, confusion) altèrent (biaisent) la vraie valeur de l'estimation. Parmi eux, le biais de confusion peut être corrigé au moment de l'analyse statistique. Cette analyse consiste à estimer une valeur  $RR_a$  ou  $OR_a$  mesurant l'association entre E et M ajustée sur un (ou plusieurs) facteur de confusion F. Différentes techniques disponibles

vont être évoquées: il s'agit d'une part de la méthode d'ajustement de Mantel-Haenszel et d'autre part de l'analyse multivariée.

### Méthode d'ajustement de Mantel-Haenszel

Le principe de la méthode de Mantel-Haenszel consiste à estimer d'abord l'association entre E et M pour chaque classe (ou strate) de valeurs du facteur de confusion F (donc à F constant), puis à calculer une valeur moyenne pondérée des estimations par strate qui sera la valeur ajustée. Cette valeur ajustée sera assortie d'un intervalle de confiance, et sa signification statistique sera testée au moment du test de Mantel-Haenszel.

Dans l'exemple utilisé au début du chapitre sur les biais de confusion, le facteur de confusion (âge) avait été divisé en trois classes ou strates: <25 ans, 25-40 ans, >40 ans. De façon générale, si on appelle k le nombre de strates, on peut estimer k risques relatifs conditionnels  $RR_i$ , (ou odds ratios  $OR_i$ ). Le risque relatif ajusté RR sera obtenu après transformation logarithmique par la formule suivante:

$$\text{Ln}(RR_a) = \frac{\sum w_i \text{Ln}(RR_i)}{\sum w_i}$$

$$\text{avec } w_i = \frac{1}{\text{var}(\text{Ln}(RR_i))}$$

si dans chaque strate on utilise les notations suivantes:

	strate F <sub>i</sub>		
	E <sup>+</sup>	E <sup>-</sup>	
M <sup>+</sup>	a <sub>i</sub>	b <sub>i</sub>	m <sub>1i</sub>
M <sup>-</sup>	c <sub>i</sub>	d <sub>i</sub>	m <sub>0i</sub>
Total	n <sub>1i</sub>	n <sub>0i</sub>	n <sub>i</sub>

où a<sub>i</sub> = nombre de malades exposés; b<sub>i</sub> = nombre de malades non exposés; m<sub>1i</sub> = nombre de malades dans la strate i;...

on a  $\text{var}(\text{Ln}(RR_i)) = \frac{c_i}{n_{1i}a_i} + \frac{d_i}{n_{0i}b_i}$  pour chaque strate i.

Une formule équivalente existe pour OR en remplaçant  $RR_i$  par  $OR_i$  dont la variance du logarithme est  $\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}$  (Bouyer et coll., 1993).

L'intervalle de confiance du risque relatif ajusté s'obtient suivant la formule générale suivante:

$$\text{Ln}(RR_a) \pm z_{\alpha/2} \sqrt{\frac{1}{\sum w_i}}$$

avec  $w_i = (\text{var } \text{Ln}(RR_i))^{-1}$  comme plus haut.

Le test de Mantel-Haenszel permet donc de tester l'existence d'une différence statistiquement significative entre  $RR_a$  (ou  $OR_a$ ) et 1.

Il s'écrit sur le même principe qu'un test classique  $\chi^2$

$$\chi^2_{MH} = \frac{\left( \sum \left[ a_i - \frac{n_i m_i}{n_i} \right]^2 \right)}{\sum \left[ \frac{m_{0i} m_{1i} n_{0i} n_{1i}}{n_i^2 (n_i - 1)} \right]^2}$$

Les limites de signification sont celles du  $\chi^2$  à 1 degré de liberté. Les conditions d'application (Bouyer et coll., 1993) doivent être vérifiées.

L'intérêt de l'analyse stratifiée de Mantel-Haenszel est qu'elle permet une visualisation pas à pas des estimations obtenues pour chaque strate des différents facteurs de confusion, et la mise en évidence concrète des interactions éventuelles ou des incohérences possibles. La mesure globale ajustée résumant toutes les strates n'est en toute rigueur interprétable qu'une fois fait cet examen des données. En effet, bien que cet aspect ne soit qu'effleuré dans ce chapitre, il est important de mentionner que, si un effet (quantifié ici par le RR) est principalement présent dans une strate ou varie de façon importante selon la valeur des différentes strates du co-facteur, on dira que le co-facteur est un facteur modifiant ou qu'il agit en interaction avec le facteur exposition. La présentation d'une mesure globale d'effet est alors d'intérêt discutable, et les données seront plutôt présentées selon les strates d'intérêt. Cet aspect d'interaction ou d'effet modifiant est testé par un test d'homogénéité permettant de statuer si la variation de l'effet selon les strates peut être attribuable au hasard.

Les limites de l'analyse stratifiée tiennent au nombre maximum de strates informatives qu'il est possible de constituer avec un nombre de sujets donné. Il est en effet théoriquement possible d'étudier simultanément plusieurs facteurs de confusion: dans ce cas, le nombre total de strates sera égal au produit des nombres de classes des différents facteurs étudiés, par exemple vouloir étudier comme facteurs de confusion le sexe et l'âge (regroupé en trois classes) donnera six strates (femmes <25 ans; hommes <25 ans; femmes 25-40 ans, etc.). Les petits effectifs de chaque strate limiteront alors fortement la précision de l'analyse. C'est pourquoi, avec le développement des logiciels d'analyse statistique, les méthodes permettant l'analyse simultanée des relations entre de nombreuses variables (multivariées) sont largement utilisées aujourd'hui.

## Analyse multivariée

L'analyse multivariée consiste à modéliser la relation entre une variable Y caractérisant l'état de santé et le facteur de risque principalement étudié E, et différentes autres variables X facteurs de confusion potentiels.

La formulation générale est donc  $Y = f(E, X_1, X_2, \dots, X_i, \dots)$ .

S'il s'agit d'une fonction f linéaire, on aura, par exemple,

$Y = \alpha + \beta_0 E + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \dots$  dans laquelle  $\alpha, \beta_0, \beta_1, \dots, \beta_i, \dots$ , sont des coefficients estimés sur la population à partir des données observées sur l'échantillon. On a ainsi modélisé la relation entre Y et E en ajustant sur les variables  $X_i$  facteurs de confusion potentiels, qui ont été incluses dans le modèle.

Y et f varient suivant le type de données analysées:

- si l'état de santé est caractérisé par une mesure quantitative (poids de naissance, dosage biologique), Y, variable à expliquer, sera la valeur moyenne de cette mesure. La fonction sera en général une fonction linéaire, sur des variables éventuellement transformées (log, racine carrée), et la méthode sera la régression linéaire multiple;
- si l'état de santé est caractérisé en oui/non (malade/non-malade), Y est le risque de maladie R ou une fonction de ce risque. Dans la régression logistique, on aura, par exemple,  $Y = \text{Ln} \frac{R}{1-R}$ , et f sera une fonction linéaire;
- si l'état de santé est caractérisé en oui/non (malade/non-malade), mais qu'on s'intéresse à la date de survenue de la maladie, Y est l'incidence instantanée  $\lambda(t)$  ou une fonction de celle-ci. Dans le modèle de Cox, on aura par exemple  $Y = \text{Ln} \lambda(t)$  et f sera une fonction linéaire.

En pratique, seul un petit nombre de modèles est utilisé en épidémiologie analytique dont le choix repose essentiellement sur la nature de la variable utilisée pour mesurer la maladie comme nous l'avons indiqué ci-dessus. Il existe cependant des circonstances en biostatistique où le problème principal est justement la recherche du bon modèle. Par exemple, l'étude des effets des faibles doses de radiations ionisantes ou les analyses de séries spatio-temporelles.

Si nous considérons donc que le choix du type de modèle est facile à régler, parce qu'il y a peu de types de modèles en concurrence, il reste à choisir les variables qui doivent figurer dans le modèle. Il faut tout d'abord insister sur la nécessité d'adopter une stratégie pour choisir les variables. En effet, même lorsque le type de modèle a été choisi et que les variables qui peuvent éventuellement y être incluses ainsi que leur codage ont été déterminés, le nombre de possibilités reste considérable. Prenons, par exemple, la situation où il y a 7 variables  $X_i$ , toutes en 0/1. Le modèle contenant toutes ces variables ainsi que toutes leurs interactions\* a  $2^7 = 128$  coefficients. On dit que c'est un modèle saturé, car il a autant de coefficients que de catégories de sujets définies par les valeurs prises pour les sept variables. Le nombre de sous-modèles possibles est gigantesque ( $2^{128}$  qui est supérieur à  $10^{38}$ ). Il est bien sûr exclu de comparer ces modèles différents les uns avec les autres pour choisir le meilleur. Il faut donc adopter une stratégie pour choisir le modèle jugé le meilleur.

L'alternative est la suivante: soit choisir le modèle saturé qui est celui qui permet de représenter parfaitement les données, car il a autant de coefficients que de catégories de sujets, mais alors chaque coefficient estimé a une grande variance en raison du grand nombre de variables et, surtout, il est fréquent que ce modèle soit difficile à interpréter; soit choisir un sous-modèle, souvent plus facile à interpréter, dont les coefficients, moins nombreux, auront une variance plus faible, mais qui peut laisser place à un biais de confusion résiduel ou à une inadéquation aux données par non-prise en compte de certaines variables.

Le choix entre ces deux possibilités dépend pour une part du point de vue où on se place. Si l'objectif est la prédiction la meilleure de  $Y$ , la question d'une bonne prise en compte de la confusion n'est pas décisive: n'importe quel bon ensemble de prédicteurs (avec ou sans contusion) convient, même si on doit s'attacher à sélectionner les facteurs de risque causaux dont les effets ont plus de chance d'être reproductibles d'une population à l'autre. En revanche, si l'objectif est d'estimer au mieux la relation entre une exposition et la maladie, le problème essentiel est l'élimination de tout

phénomène de confusion et, souvent, une bonne description de la relation dose-effet entre l'exposition et la maladie.

Ce second point de vue (étude de la relation entre une exposition et la maladie) est le plus fréquent en épidémiologie étiologique. Le choix des variables  $X_i$  que l'on va finalement retenir comprend trois étapes principales (Bouyer et coll., 1993). En premier lieu, le choix des variables que l'on va sélectionner initialement comme variables pouvant être prises en compte dans le problème étudié, puis le choix de leur codage et des interactions que l'on va retenir. On aboutit ainsi à un ensemble de variables qui sont susceptibles d'entrer dans le modèle. Il faut alors choisir, au sein de ces variables, celles que l'on va finalement conserver. Dans le processus conduisant au choix du modèle final, il ne faut pas perdre de vue l'objectif initial. Dans le cas de l'épidémiologie étiologique, il s'agit de quantifier et de tester la relation entre l'exposition et la maladie, après l'avoir «débarrassée» des effets de confusion dus aux autres variables. Il est, d'autre part, important d'aboutir à un modèle que l'on sache interpréter sur le plan épidémiologique.

On trouvera un exposé plus détaillé de ces questions capitales dans les ouvrages spécialisés (Breslow et Day, 1980; Kleinbaum et coll., 1982; Bouyer et coll., 1993).

## 7. INTERPRÉTATION DES RÉSULTATS. CAUSALITÉ

Même si l'épidémiologie permet de décrire de nombreuses associations impliquant diverses maladies et facteurs de risque potentiels, elle a fréquemment de la difficulté à établir un lien de causalité. Ceci est particulièrement vrai pour l'épidémiologie d'observation. À titre d'exemple, mentionnons le temps qu'il aura fallu pour que la cigarette soit reconnue comme facteur causal dans l'augmentation de l'incidence du cancer du poumon: la cigarette ne fut reconnue comme facteur étiologique du cancer du poumon que dans les années 1960, alors qu'elle était déjà soupçonnée dans les années 1940 (Stolley et Lasky, 1995).

L'interprétation des résultats d'une enquête étiologique comprend deux étapes principales:

\* C'est-à-dire les interactions entre deux variables  $X_i$  et  $X_j$ , mais aussi celles entre trois variables et jusqu'à celle entre les 7 variables.

un jugement de signification qui s'appuie sur l'analyse statistique des résultats de l'enquête et un jugement de causalité qui fait souvent appel à des résultats extérieurs à l'enquête.

### 7.1 Jugement de signification

Une enquête étiologique permet de tester l'existence d'une association entre une exposition et une maladie. Le jugement de signification consiste à conclure, sur la base de l'analyse statistique, que l'on rejette ou que l'on ne rejette pas l'hypothèse d'absence d'association. Cette conclusion ne peut être émise avec une crédibilité suffisante qu'à la suite d'une analyse statistique correctement conduite. Cela signifie, bien sûr, que les tests adéquats ont été utilisés, mais aussi que les facteurs de confusion potentiels ont été pris en compte et que les risques d'erreurs ont été contrôlés. Ce dernier point est peut-être trop souvent négligé malgré son importance. Sa discussion est différente selon que le résultat est significatif ou ne l'est pas.

Dans le cas d'un résultat non significatif, on doit s'interroger sur la puissance de l'enquête. Si elle est suffisante, l'enquête apporte une connaissance scientifique utile. Même si elle ne permet pas d'accepter l'hypothèse d'absence d'association, elle permet de conclure que la force de l'association entre l'exposition et la maladie (mesurée par le risque relatif, par exemple) est inférieure à une certaine valeur. Si la puissance est trop faible, on ne peut que se limiter au pauvre constat que l'instrument utilisé pour tester l'existence d'une association entre l'exposition et la maladie n'était pas de bonne qualité, de sorte que la conclusion porte sur les insuffisances de l'enquête en tant qu'outil plutôt que sur le fond du problème.

Dans le cas d'un résultat significatif, il faut s'assurer que le risque d'erreur a été correctement contrôlé, c'est-à-dire qu'on ne s'est pas mis en situation de conclure à l'existence d'une association avec un risque d'erreur supérieur aux 5 % requis. Cette question se pose dans la plupart des enquêtes épidémiologiques, car on est le plus souvent conduit à tester plusieurs associations dans la même enquête. Il faut alors considérer que le risque n'est réellement contrôlé que pour la ou les questions d'intérêt principal telles qu'elles ont été définies *a priori* dans le protocole. Pour les autres associations que l'on est amené à

étudier, soit parce que les données l'ont suggéré, soit parce qu'il était prévu d'explorer des directions secondaires, il faut considérer que le risque a été mal contrôlé et que les résultats trouvés doivent être confirmés par des études spécifiques.

### 7.2 Jugement de causalité

La question de la causalité intervient lors d'un résultat significatif. Il s'agit de savoir si l'association mise en évidence entre une exposition et une maladie correspond à une relation de cause à effet. Cette question a fait l'objet de nombreux débats philosophiques, et nous n'avons pas l'intention de discuter de ces différents courants de pensées. Le lecteur intéressé à la discussion de ces aspects appliqués à l'épidémiologie pourra se référer à l'excellent recueil de textes assemblés par Rothman (1988). Mentionnons, toutefois, qu'avec les meilleurs critères d'interprétation, il est toujours possible de conclure de façon erronée en regard des résultats d'une ou de plusieurs études, en rejetant par exemple une association comme causale alors qu'elle l'est.

Tout en sachant qu'une conclusion causale n'est pas accessible à partir d'une seule enquête d'observation, il faut examiner les éléments qui militent en faveur d'une telle conclusion et ceux qui la contredisent. Les principaux critères de causalité sont discutés brièvement ci-dessous.

### 7.3 Critères de causalité

Plusieurs auteurs ont proposé des critères pour faciliter l'interprétation des résultats d'une ou de plusieurs études épidémiologiques. Les critères les plus reconnus sont ceux de Hill (1965). Ces critères ne peuvent être utilisés qu'après avoir écarté des biais importants comme explication possible aux résultats significatifs observés. Ces critères sont les suivants :

- *Force de l'association*: La force de l'association est mesurée par le risque relatif (RR). Plus le RR est élevé, plus l'association est forte. L'association est généralement considérée comme faible si le RR est inférieur à 1,5 et forte, s'il est supérieur à 3. À titre d'exemple, dans le cas de l'association tabac-cancer du poumon, l'association est considérée comme particulièrement forte avec un RR d'environ 10. Le risque relatif associé à la fumée secondaire (fumée inspirée par les non-fumeurs)

est généralement inférieur à 1,5 et donc considéré comme faible. Plus l'association est forte, moins il y a probabilité que le RR observé puisse être expliqué par la confusion résiduelle et donc plus l'association est probablement causale. Ceci dit, il faut se rappeler que, si l'association est véritablement causale, un faible RR ne peut être négligé. Il peut en fait être responsable d'une proportion de cas importante dans la population (si l'exposition est fréquente ou l'incidence de la maladie chez les non-exposés importante).

- *Constance*: Plus l'association est observée fréquemment dans plusieurs études, effectuées dans divers pays et si possible avec des outils différents, plus l'association paraîtra crédible. Il y a en fait peu de chance que les mêmes erreurs puissent se reproduire de façon similaire dans différents contextes et avec différents outils de recherche. Cependant, il faut aussi reconnaître qu'à l'inverse la présence de facteurs contributeurs (facteurs intervenant dans le mécanisme causal) chez certaines populations puisse modifier la relation observée, et donc faire varier le RR.
- *Spécificité*: Une relation sera spécifique lorsque la même cause donnera toujours le même effet. En fait, comme déjà mentionné par Hill, ce critère est un facteur intéressant lorsqu'il est présent, mais son absence ne permet certainement pas de rejeter la causalité. En effet, la plupart des expositions environnementales (agissant le plus souvent de façon systémique) entraîneront divers problèmes de santé. Les intoxications au plomb ou à l'arsenic en sont un bon exemple.
- *Temporalité*: L'exposition doit toujours être présente avant l'apparition de la maladie. Nous ajoutons que l'exposition doit être présente pendant la période d'induction potentielle de la maladie, soit la période précédant la période de latence. En d'autres termes, comme on s'attend à ce que le diagnostic d'un cancer soit précédé d'une période de latence de plusieurs années (variant selon le type de cancer), il sera improbable qu'une exposition récente, survenue dans les mois précédant le diagnostic, puisse être responsable de l'apparition de ce cancer. Cependant, l'exposition récente peut être un

indicateur des habitudes passées, et, si un facteur agit au niveau de la progression d'un cancer, il pourra agir dans les derniers mois de son évolution.

- *Dose-réponse* (appelé par Hill gradient biologique): Il s'agit du critère le plus important en toxicologie pour déterminer le caractère toxique d'une substance et permettre l'établissement de niveaux sécuritaires d'exposition. Généralement, l'effet doit augmenter progressivement avec la dose d'exposition, mais ce n'est pas toujours le cas. La courbe en U décrivant l'effet de la consommation d'alcool sur les maladies cardiovasculaires en est un bon exemple. Par ailleurs, compte tenu des difficultés rencontrées dans les études rétrospectives pour mesurer l'exposition antérieure, la courbe dose-réponse observée en épidémiologie peut être très faible ou inexistante à cause d'erreurs dans le classement des sujets exposés.
- *Plausibilité biologique*: Il s'agit d'un critère toujours important pour juger de la nature causale d'une association. Est-ce que des explications biologiques plausibles peuvent expliquer le résultat observé? Malheureusement, ce concept est interprété différemment selon les auteurs, ce qui prête à confusion (Weed et Hursting, 1998). En pratique, si la substance sous étude donne le même effet, ou un effet similaire, chez l'animal, on parlera de plausibilité biologique. Lorsque seulement des études effectuées sur des cellules sont positives (tests *in vitro* ou *in vivo*), les conclusions en regard de ce critère seront plus difficiles.
- *Cohérence*: La cohérence réfère à la compatibilité de l'observation avec les connaissances ou observations déjà acquises sur le même sujet. Il s'agit donc d'une notion plus large que celle de la plausibilité biologique. Il peut s'agir par exemple de la cohérence des résultats des études écologiques avec les études étiologiques ou encore de la cohérence des résultats observés en milieu résidentiel et en milieu de travail.
- *Preuve expérimentale*: Il s'agit de la preuve par expérimentation chez l'humain. Il s'agit principalement de la preuve expérimentale de la disparition de l'effet lorsque l'exposition disparaît. En fait, la preuve expérimentale chez

l'humain est difficile à établir, mais elle est parfois utilisée lors d'exposition brève et d'effets aigus (allergie).

- *Analogie*: Il s'agit de la notion générale que les propriétés toxiques d'une substance ou d'un agent peuvent être déduites à partir des connaissances de la toxicité d'autres agents ou substances de la même famille. Même si cette notion générale est *a priori* intéressante, elle s'avère en pratique parfois en défaut (HAP ou BPC, par exemple).

L'utilisation de ces critères a fait l'objet de nombreux débats; il ne s'agit certainement pas d'une panacée susceptible de solutionner tous les cas difficiles d'interprétation des associations épidémiologiques. Ils ne peuvent, en particulier, pallier une qualité médiocre des études effectuées sur un sujet donné (s'appliquer uniquement à des études écologiques, par exemple). Par ailleurs, aucun de ces critères n'est essentiel, sauf évidemment le critère de temporalité, pour déclarer une association causale. Cependant, les critères de constance, de dose-réponse ainsi que de plausibilité biologique sont souvent déterminants pour affirmer un lien causal.

Ces critères sont utilisés fréquemment en épidémiologie environnementale pour juger de nature d'une association. Citons, à titre d'exemple, l'évaluation du lien entre la leucémie de l'enfant et l'exposition aux champs électromagnétiques (Levallois, 1995) ou encore celui de l'évaluation de l'étiologie de la maladie des combattants de la guerre du Golfe (Joellenbeck et coll., 1998).

## 8. CONCLUSION

Ce chapitre a permis de réviser les principales notions d'épidémiologie en s'attardant principalement aux aspects de recherche étiologiques. Certains exemples tirés du domaine de la santé environnementale ont été donnés, mais - comme le lecteur a pu s'en rendre compte - la méthode épidémiologique n'est pas différente en santé environnementale et dans les autres domaines de la santé publique. Quelques aspects sont cependant spécifiques et rendent souvent difficile le travail des épidémiologistes environnementaux. Il s'agit en particulier des faibles niveaux d'exposition et de l'influence des autres facteurs concomitants (Hémon, 1995).

Par ailleurs, la nécessité de plus en plus ressentie de tenir compte des résultats des études épidémiologiques, même imparfaites, dans l'évaluation des risques environnementaux a entraîné une réflexion visant à améliorer l'interprétation de ces études. En particulier, des recommandations pour l'utilisation de la méta-analyse (analyse quantitative des résultats de plusieurs études) en épidémiologie environnementale ont été proposées (Blair et coll., 1995). Des recommandations ont aussi été formulées pour déterminer l'utilisation pouvant être faite des résultats d'études épidémiologiques environnementales en fonction de leur qualité (Hertz-Picciotto, 1995).

## Bibliographie

- Armitage, P. et G. Berry. *Statistical methods in medical research*, Blackwell Scientific Publications, Oxford, 1971.
- Armitage, P. et R. Doll. *Stochastic models for carcinogenesis*, Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability: Biology and Problems of Health, University of California Press, Berkeley, 1961, p. 19-38.
- Arnold, F. A. J. et H. T. Dean. «Effect of fluorurated water supplies on dental caries prevalence: 10th year of the Grand Rapids-Muskegon Study», *Publ Health Rep*, 71, 1956, p. 652-658.
- Bernard, P. M. et C. Lapointe. *Mesures statistiques en épidémiologie*. Presses de l'Université du Québec, 1987.
- Beyea, J. et M. Hatch. «Geographic exposure modeling: a valuable extension of Geographic Information Systems for use in environmental epidemiology» *Environ Health Perspect*, 107, 1999, p. 181-190.
- Blair A. et coll. «Guidelines for application of meta-analysis in environmental epidemiology», *Regul Toxicol Pharm*, 22, 1995, p. 189-197.
- Bouvenot, G. et M. Vray. *Essais cliniques. Théorie, pratique et critique*, Flammarion, Médecine-Sciences. Paris, 1999, p. 450.
- Bouyer, J. *Méthodes statistiques. Médecine - Biologie*, ESTEM, les Éditions INSERM, Paris, 1996.
- Bouyer, J., D. Hémon, S. Cordier, F. Derriennic, I. Stücker, B. Stengel et J. Clavel. *Epidémiologie - Principes et méthodes quantitatives*, INSERM, 1993.
- Breslow, N. E. et N. E. Day. *Statistical methods in cancer research*, volume I, «The design and analysis of case control studies», Oxford University Press, Lond, 1980.
- Breslow, N. E. et N. E. Day, *Statistical methods in cancer research*, volume II, The design and analysis of cohort studies, Oxford University Press, Londres, 1987.
- Bruzzi, P., S. B. Green, D. P. Byar, L. A. Brinton, et C. Schaire. «Estimating the population attributable risk for multiple risk factors using case-control data», *Am J Epidemiol*, 122, 1985, p. 904-914.
- Case, R. A. M., M. E. Hosker, D. B. McDonald et J. T Pearson. «Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the british chemical industry», Part 1, «The role of aniline, benzidine, alpha-naphtylamine and beta-naphtylamine», *Br J Ind Med*, 11, 1954, p. 75-104.
- Clayton, D. et M. Hills. *Statistical models in epidemiology*, Oxford Science Publications, Oxford University Press, 1993.
- Coste, J. et A. Spira. «La proportion de cas attribuable en Santé Publique: définition(s), estimation(s) et interprétation» [The proportion of attributable cases in Public Health: definition, estimation and interpretation]. *Rev Epidém Santé Publ*, 39, 1991, p. 399-411.
- De Wals, P. et T. Niyonsenga. «Investigation d'agrégats d'anomalies de la reproduction: théorie et pratique», *Arch Publ Health*, 1996, 54, p. 43-58.
- Desplanques, G., A. Mizrahi et A. Mizrahi. «Mortalité et morbidité par catégories sociales», *Solidarité Santé*, 4, 1996, p. 75-85.
- Doll, R. et A. B. Hill. «The mortality of doctors in relation to their smoking habits: A preliminary report», *Br Med J*, 2, 1954, p. 1451-1455.
- Doll, R. et A. B. Hill. «Lung cancer and other causes of death in relation to smoking: A second report on the mortality of British doctors», *Br Med J*, 2, 1956, p. 1071-1081.
- English, P., R. Neutra, R. Scalf, M. Sullivan, L. Waller et L. Zhu. «Examining associations between childhood asthma and traffic flow using a geographic information system», *Environ Health Perspect*, 107, 1999, p. 761-767.
- Estève, J., E. Benhamou et L. Raymond. «Statistical Methods in Cancer Research», vol. IV - *Descriptive Epidemiology*, International Agency for Research on Cancer (1994), Lyon, France (version française: «Méthodes statistiques en épidémiologie descriptive», INSERM, 1993).
- Feinleib, M. «The Framingham study: sample selection, follow-up, and methods of analyses», *Natl Cancer Inst Monogr*, 67, 1985, p. 59-63.
- Goldberg, P., D. Luce, M. A. Billon-Galland, P. Quenel, C. Salomon-Nekiriai, J. Nicolau, P. Brochard et M. Goldberg. «Rôle potentiel de l'exposition environnementale et domestique à la trémolite dans le cancer de la plèvre en Nouvelle-Calédonie», *Rev Epidém Santé Publ*, 43, 1995, p. 444-450.

- Hémon, D. «Recherche épidémiologique sur l'environnement et la santé: quelques aspects méthodologiques», *Rev Epidém Santé Publ*, 43, 1995, p. 395-411.
- Hertz-Picciotto, I. «Epidemiology and quantitative risk assessment: a bridge from science to policy», *Am J Publ Health*, 85, 1995. p. 484-491.
- Hill, A. B. «The environment and disease. Association or causation?», *Proceedings of the Royal Society of Medicine*, 58, 1965, p. 295-300.
- Hill, C, C. Com-Nougué, A. Kramar, T. Moreau, J. O'Quigley, R. Senoussi et C. Chastang. *Analyse statistique des données de survie*, INSERM — Flammarion, 1990.
- Joellenbeck, L. M., P. J. Landrigan et E. L. Larson «Gulf War veterans' illnesses: a case study in causal inference», *Env Res*, 79, 1998, p. 71-81.
- Kallen, B. *Epidemiology of human reproduction*, CRC Press, Boca Raton, Floride, 1988.
- Kleinbaum, D. G., L. L. Kupper et H. Morgenstern. *Epidemiologic Research. Principles and quantitative methods*, Lifetime Learning Publications, 1982.
- Krewski, D., E. Cardis, L. Zeise et V. J. Feron. «Empirical approaches to risk estimation and prediction», dans S. Moolgavkar, D. Krewski, L. Zeise, E. Cardis et H. Moller (rédacteurs) «Quantitative estimation and prediction of human cancer risks», *ARC Scientific Publications*, Lyon, 131, 1999, p. 131-178.
- Lanphear, B. et coll. «Primary prevention of childhood lead exposure: a randomized trial of dust control». *Pediatrics*, 103, 1999, p. 771-777.
- Laplanche, A., G Com-Nougué et R. Flamant. *Méthodes statistiques appliquées à la recherche clinique*, Flammarion, Paris, 1986.
- Last, J. *A dictionary of epidemiology*, Oxford Medical Publications, Oxford, 1983.
- Levallois, P. «Do Power Frequency Magnetic Fields Cause Leukemia in Children?», *Am J Prev Med*, 11, 1995, p. 263-270.
- Luce, D., I. Bugel, P. Goldberg, M. Goldberg, C. Salomon, M. A. Billon-Galland, J. Nicolau, P. Quénel, J. Fevotte et P Brochard. «Environmental exposure to tremolite and respiratory cancer in New Caledonia: a case-control study», *Am J Epidemiol*, 151, 3, 2000.
- Morgenstern, H. «Ecologic studies», dans Rothman et Greenland *Modern Epidemiology*, Little, Brown and Company, Boston 1998, p. 459-480.
- Neutra, R. R. «Computer Geographic Analysis: a commentary on its use and misuse in public health», dans A. Lawson, A. Biggeri, D. Bohning, E. Lesaffre, J. F. Viel et R. Bertollini (rédacteurs) *Disease mapping and risk assessment for public health*, Wiley, Chichester, 1999, p. 311-319.
- O'Connor, M. E. et D. Rich «Children with moderately elevated lead levels: is chelation with DMSA helpful?», *Clin Pediatr*, 38, 1999, p. 325-331.
- Omenn, G. S. et coll. «Effects of a combination of beta carotene and vitamine A on lung cancer and cardiovascular disease», *New Engl J Med*, 334, 1996, p. 1150-1155.
- Rothenberg, R. B. et S. B. Thacker. «Guidelines for the investigation of clusters of adverse health effects», dans P. Elliot, J. Cuzick, D. English et R. Stern (rédacteurs), *Geographical and environmental epidemiology: methods for small-area studies*, Oxford University Press, Oxford, 1992, p. 264-277.
- Rothman, K. J. *Causal inference, Epidemiology Ressources*, Chesnut Hill, 1988.
- Rothman, K. J. et S. Greenland. *Modern Epidemiology*, Little, Brown and Company, Boston, 1998.
- Savitz, D. A. «In defense of black box epidemiology», *Epidemiology*, 5, 1994, p. 550-552.
- Scherrer, B. *Biostatistique*, Gaëtan Morin Éditeur, Boucherville, 1984.
- Selmaoui, B., J. Lambrozo et Y. Touitou. «Magnetic fields and pineal function in humans: evaluation of nocturnal acute exposure to extremely low frequency magnetic fields on serum melatonin and urinary 6-sulfatoxy melatonin circadian rhythm», *Life Sci*, 58, 1996, p. 1539-1549.
- Skrabanek, P. «The emptiness of the black box», *Epidemiology*, 5, 1994, p. 553-555.
- Stolley, P. D. et T. Lasky. *Investigating Disease Patterns, the Science of Epidemiology*, Scientific American Library, New York, 1995.
- Susser, M. «The logic in ecological: I. The logic of analysis». *Am J Publ Health*, 84, 1994a, p. 825-829.
- Susser, M. «The logic in ecological: II. The logic of design». *Am J Publ Health*, 84, 1994b, p. 830-835.

- Thacker, S. B. et D. F. Stroup. «Public health surveillance», dans R. C. Browson et D. B. Petirti (rédacteurs) *Applied Epidemiology, Theory to Practice*, Oxford University Press, 1998, p. 105-135.
- Thériault, G., C. Tremblay, S. Cordier et S. Gingras. «Bladder cancer in the aluminium industry», *Lancet*, 8383, 1, 1984, p. 947-950.
- Walter, S. D. «The ecologic method in the study of environmental health. I. Overview of the method». *Environ Health Perspect*, 94, 1991a, p. 61-65.
- Walter, S. D. «The ecological method in the study of environmental health. II. Methodologic issues and feasibility», *Environ Health Perspect*, 94, 1991b, p. 67-73.
- Weed, D. L. et S. D. Hursting «Biologic plausibility in causal inference: current method and practice», *Am J Epidem*, 147, 1998, p. 415-425.
- Wigle, D. T «Bladder cancer: possible new high-risk occupation», *Lancet*, 2, 8028, 1977, p. 83-84.