

# 21

## Méthodes d'étude épidémiologique

La principale distinction entre les méthodes d'études *in vitro* et *in vivo* et les méthodes d'étude épidémiologique réside dans sa nature observationnelle, par opposition au caractère expérimental de la toxicologie. La nature observationnelle de l'épidémiologie implique que l'exposition n'est pas randomisée (attribuée aléatoirement à chaque sujet), mais qu'elle est constatée, et est le fait de sa situation personnelle (lieu de vie, comportements, caractéristiques sociodémographiques). La conséquence est que la répartition des facteurs de risque de la maladie (autres que l'exposition considérée) peut différer entre les groupes de sujets comparés, ce qui peut entraîner ce que les épidémiologistes appellent un biais de confusion (voir plus bas).

### Principales approches de l'épidémiologie environnementale

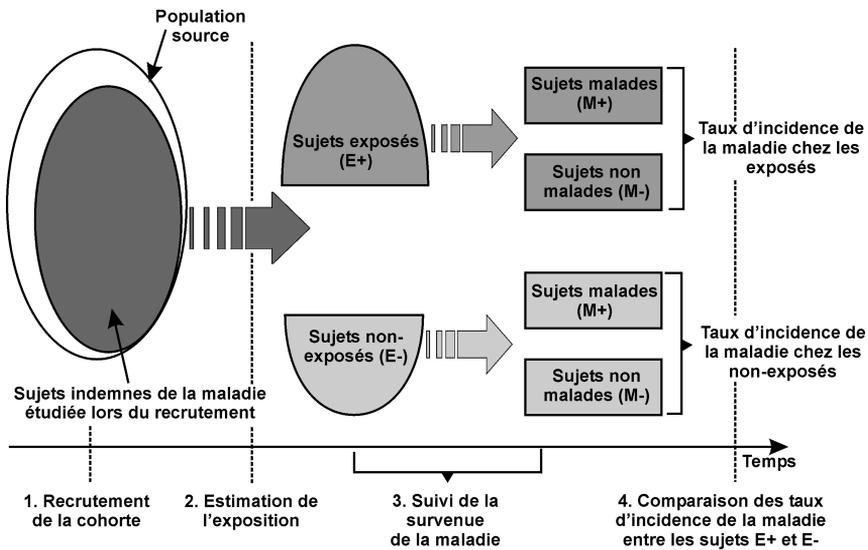
Les deux principaux types d'études épidémiologiques en population à visée étiologique (c'est-à-dire visant à identifier les facteurs de risque des événements de santé) sont l'étude de cohorte et l'étude cas-témoins. À cela il faut ajouter d'autres approches, moins fréquemment utilisées, telles que les approches écologiques ou celles reposant sur des séries de cas, et deux approches n'impliquant pas de recrutement de sujets mais s'appuyant sur des données déjà disponibles : la méta-analyse et l'étude de risque sanitaire.

#### Étude de cohorte

L'approche de cohorte consiste à recruter des sujets indemnes de la pathologie d'intérêt et de les suivre au cours du temps pour identifier au niveau individuel la survenue de cette pathologie. Généralement, l'exposition au facteur environnemental considéré est estimée lors de l'inclusion dans la cohorte, cette mesure pouvant être actualisée au cours du suivi. On peut en pratique recruter plusieurs cohortes suivies en parallèle, définies par exemple par leur niveau d'exposition : une cohorte de travailleurs exposés à un composé chimique donné, et une cohorte « non exposée » d'employés de la même entreprise mais ne travaillant pas à des postes exposés. Souvent, une unique cohorte est recrutée, incluant des sujets exposés à des niveaux divers.

Au cours du suivi, en plus d'une mesure (ou estimation) de l'exposition au(x) facteur(s) environnemental(taux) d'intérêt, des informations sur les facteurs de confusion potentiels sont recueillies ; diverses mesures de paramètres biologiques (marqueurs d'inflammation, polymorphismes génétiques...) peuvent aussi être réalisées.

L'analyse statistique consiste à comparer la fréquence de survenue de la maladie entre les différents groupes définis par leur niveau d'exposition, en corrigeant les différences de fréquence de la maladie qui seraient dues aux facteurs de confusion (figure 21.1). Cette analyse peut permettre d'étudier et de prendre en compte des facteurs modifiant l'effet éventuel des perturbateurs endocriniens, tels que des caractéristiques génétiques (Cantor et coll., 2010) ou sociodémographiques.



**Figure 21.1 : Principe de l'étude de cohorte prospective**

Du point de vue de l'épidémiologiste, l'approche toxicologique classique consistant à exposer des groupes d'animaux exposés à différents niveaux d'un composé ou d'un mélange de composés chimiques et à suivre la fréquence d'apparition de certains troubles de la santé (ou de suivre les variations de paramètres biologiques) au cours du temps pour la mettre en relation avec le niveau d'exposition au composé considéré correspond à une étude de cohorte dans laquelle l'exposition aurait été attribuée aléatoirement aux animaux.

### Étude cas-témoins

L'étude cas-témoins consiste à recruter d'une part des personnes présentant l'événement de santé d'intérêt, et d'autre part un groupe de personnes comparables mais ne présentant pas l'événement de santé (ou témoins). L'exposition

aux facteurs environnementaux considérés, ainsi que le niveau des facteurs de confusion potentiels, sont estimés chez chacun des sujets.

Si l'étude cas-témoins est réalisée sur les cas incidents (ceux survenant dans une zone ou un lieu de recrutement donné dans une période donnée), l'étude cas-témoins est en principe équivalente à une étude de cohorte réalisée à partir de la même population source (les habitants de la zone considérée ou les patients susceptibles d'aller consulter sur le lieu de recrutement choisi), mais dans laquelle on ne recruterait les sujets que lorsqu'un cas survient, et non pas avant la survenue de la maladie. Au lieu de suivre tous les sujets qui ne développeront pas la maladie dans la population source, on ne suit qu'un sous-groupe aléatoire de ceux-ci, beaucoup plus petit, recruté par exemple à raison de quelques témoins juste après la survenue de chaque nouveau cas. L'*odds-ratio* de maladie estimé par cette étude cas-témoins aura la même interprétation que le risque relatif instantané (*hazard ratio* en anglais) qui aurait été estimé à partir de l'ensemble de la cohorte.

Le critère de choix des témoins est crucial pour la validité de l'étude. En principe, ceux-ci ne doivent pas forcément être représentatifs de la population générale dans son ensemble, mais ils doivent être représentatifs des sujets qui, s'ils développaient la maladie, seraient recrutés comme cas. Si par exemple les cas correspondent aux sujets chez qui un cancer du sein a été diagnostiqué en 2010 dans un service hospitalier donné, les témoins doivent correspondre à un échantillon aléatoire des femmes qui, si elles développaient la maladie, seraient diagnostiquées dans ce même service. On peut envisager de recruter ces témoins dans d'autres services du même hôpital, à condition que ces services aient la même zone de recrutement (et le même bassin sociodémographique) que le service où sont recrutés les cas.

Une étude cas-témoins est par définition rétrospective : le recrutement s'effectue après la survenue de la maladie. Ceci peut poser un problème dans le cas où l'exposition au facteur environnemental d'intérêt est difficile à estimer rétrospectivement, comme c'est typiquement le cas pour les perturbateurs endocriniens. Dans ce cas, d'importantes erreurs de classement sur l'exposition sont attendues, qui pourront biaiser l'association estimée entre exposition et survenue de la maladie (voir plus bas, estimation des expositions). Une exception notable correspond à l'étude cas-témoins nichée dans une cohorte (voir ci-dessous).

### **Cohorte ou cas-témoins ?**

En principe, l'approche de cohorte (idéalement, avec une randomisation de l'exposition, rarement possible pour des raisons éthiques si l'exposition est potentiellement néfaste) est considérée comme l'approche reine. C'est elle qui, par sa nature prospective, permet d'aller vers une mesure de chaque facteur dans la période pertinente, de façon prospective, et, en reconstituant précisément la chronologie des expositions et événements de santé, d'éviter

les biais de causalité inverse. Toutefois cette approche peut dans certains cas se révéler très lente et coûteuse à réaliser, et, dans certaines situations, l'étude cas-témoins apportera en un temps et pour un coût plus limités une information de qualité quasi équivalente.

Si on ne peut donner de règles générales quant à l'utilisation d'un *design* d'étude dans une situation particulière, il existe des situations typiques dans lesquelles chaque approche va se révéler plus pertinente : si le taux d'incidence de la maladie est élevé, si le délai de latence supposé entre le moment de l'exposition et celui de la survenue de la maladie est relativement court (de quelques semaines à quelques mois, voire un petit nombre d'années), une approche de cohorte peut se révéler relativement peu coûteuse et mérite d'être considérée. Si la maladie est rare, une étude cas-témoins peut se révéler judicieuse. Dans le cas du cancer du testicule, dont l'incidence chez l'homme adulte est de l'ordre de 10 cas pour 100 000 personnes-années, il faut suivre 100 000 hommes pendant 10 ans pour recruter 100 cas. Une étude cas-témoins à partir des cas diagnostiqués dans plusieurs grands centres hospitaliers pourra en un à deux ans recruter ces cas, et environ 300 témoins, et fournir une information similaire (mais peut-être davantage de biais) dans un délai bien plus bref et avec un effectif total de 400 sujets. Toutefois, cette approche cas-témoins risque de ne pas être pertinente si l'exposition au facteur environnemental considéré ne peut être estimée que par le dosage d'un biomarqueur peu persistant dans l'organisme. Dans ce cas, le dosage réalisé au moment du diagnostic du cas a peu de chances de fournir une bonne estimation de l'exposition dans la fenêtre temporelle biologiquement pertinente, plusieurs années auparavant. Dans le cas où on s'intéresse aux effets à long terme des expositions subies durant la vie intra-utérine, expositions souvent très difficiles à estimer rétrospectivement, il y a peu d'alternatives pertinentes à la mise en place d'une cohorte de femmes enceintes, chez qui différents prélèvements biologiques sont réalisés, avec un suivi à long terme de leurs enfants. De telles cohortes ont été mises en place dans différents pays, par exemple la cohorte Moba en Norvège (incluant environ 100 000 enfants), la *Danish National Birth Cohort* (environ 90 000 enfants) ou, en France, les cohortes Pélagie (3 500 enfants bretons) ou Eden (1 900 enfants nés à Nancy et Poitiers).

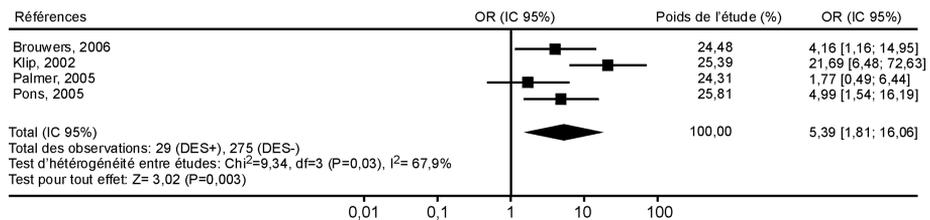
### **Cohorte et cas-témoins**

Dans le cas particulier d'une étude cas-témoins nichée dans une cohorte, les cas et les témoins sont recrutés à partir d'une unique cohorte (en général, en recrutant tous les cas de la pathologie d'intérêt déclarés dans la cohorte, et un sous-groupe des sujets n'ayant pas développé la maladie). L'intérêt de l'approche est que les informations sur les caractéristiques des sujets et leur niveau d'exposition ont généralement été recueillies lors du recrutement dans la cohorte, c'est-à-dire avant la survenue de la maladie (prospectivement). En particulier, si des prélèvements biologiques ont été réalisés à l'inclusion dans

la cohorte et congelés, ceux-ci peuvent être utilisés après la survenue des cas pour doser un biomarqueur d'exposition. Si la maladie est rare, le coût total est bien plus faible que si les dosages avaient été réalisés sur l'ensemble de la cohorte.

## Méta-analyse

La méta-analyse consiste à synthétiser les résultats de plusieurs études sur un sujet donné pour obtenir une estimation plus précise et fiable de l'effet du facteur considéré sur le risque de maladie (figure 21.2). L'unité d'analyse n'est donc plus le sujet ou le patient, mais une étude elle-même. Cette approche peut permettre de discuter l'homogénéité des études sur un sujet donné, et peut s'avérer particulièrement utile pour des maladies ou des facteurs d'exposition rares, pour lesquels chaque étude individuelle risque d'être peu informative. C'est une forme standardisée de synthèse de la littérature scientifique sur un sujet, qui est plus formalisée que la simple revue de la littérature. Sur le principe, l'*odds-ratio* issu de la méta-analyse peut être interprété comme une moyenne pondérée des *odds-ratios* issus de chaque étude individuelle, le poids donné à chaque étude dépendant de sa précision.



**Figure 21.2 : Exemple de méta-analyse, reprenant les résultats des études concernant l'association entre l'exposition féminine au distilbène (DES) durant la vie intra-utérine et le risque de survenue d'hypospadias dans la descendance**

Le losange noir indique la valeur de l'*odds-ratio* méta-analytique et ses extrémités gauche et droite indiquent les limites de l'intervalle de confiance à 95% de cet OR méta-analytique (d'après Slama et Cordier 2010).

## Étude d'impact sanitaire

L'étude d'impact sanitaire vise à quantifier l'impact sanitaire d'un facteur donné, à l'échelle d'une communauté. Elle consiste, connaissant les niveaux d'exposition dans la communauté et l'effet de l'exposition sur le risque de maladie au niveau individuel (relation dose-réponse estimée à partir d'études épidémiologiques par exemple), à quantifier le risque de maladie, c'est-à-dire le nombre de cas de la maladie attribuables (attendus) du fait de l'installation industrielle ou du facteur considéré. Une telle étude n'implique pas de mesurer directement la fréquence des maladies dans la communauté concernée,

comme on le ferait dans une étude épidémiologique de cohorte ou de type cas-témoins, ni de suivre à long terme d'importantes populations.

Par exemple, une étude d'impact sanitaire a estimé en 2001 que la pollution atmosphérique liée au trafic routier était susceptible d'entraîner environ 18 000 décès par an chez les adultes de plus de 30 ans en France (intervalle de confiance à 95 %, de 10 700 à 24 700 décès). Cette étude (Kunzli et coll., 2000) a été réalisée en estimant la distribution des niveaux d'exposition aux polluants atmosphériques de la population française, et en appliquant des modèles mathématiques (relations dose-effet) obtenus à partir de plus petites populations et quantifiant le nombre de décès attendus dans chaque groupe en fonction de son niveau d'exposition.

### Étude transversale

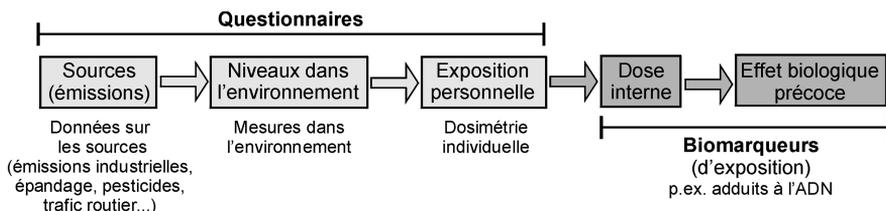
Si l'échantillonnage transversal correspond bien à un mode d'échantillonnage classique utilisé en théorie des sondages, en épidémiologie descriptive ou en science des expositions pour décrire la prévalence d'une pathologie, d'un critère biologique (études de séroprévalence par exemple) ou d'une exposition dans une population à un moment donné, il n'y a pas au sens strict d'étude étiologique transversale. En revanche, des analyses statistiques sont fréquemment réalisées à partir d'échantillons transversaux de population chez qui la présence d'un événement de santé est recherchée et l'exposition à un facteur environnemental estimé. Ceci peut être vu comme une version dégradée de l'étude de cohorte, avec un recueil simultané de l'information sur l'exposition et l'événement de santé (alors que dans une cohorte, l'information sur l'exposition est normalement recueillie avant celle sur les paramètres de santé d'intérêt), et une analyse portant sur les cas prévalents de la maladie plutôt que sur les cas incidents. Dans le cas de facteurs environnementaux dont l'exposition est estimée à partir de biomarqueurs peu persistants dans l'organisme (comme c'est le cas pour certains perturbateurs endocriniens, comme le bisphénol A), cette approche est notamment limitée, tout comme l'approche cas-témoins, car elle ne permet pas d'estimer directement l'exposition dans la fenêtre biologiquement pertinente, qui peut remonter à plusieurs années avant la survenue du cas dans le cas d'événements de santé pour lesquels une programmation à long terme est probable. Une étude réalisée dans le Massachusetts fournit un exemple de cette approche : à partir d'un échantillon transversal d'hommes consultant dans une clinique d'infertilité, des échantillons de sperme et d'urine ont été recueillis le même jour. L'échantillon d'urine a permis de doser la concentration de bisphénol A, qui a été mise en relation avec la concentration spermatique, en faisant l'hypothèse que la concentration urinaire de bisphénol A pouvait être considérée comme un bon marqueur du niveau moyen de bisphénol A dans les mois précédents (Meeker et coll., 2010).

## Autres types d'études épidémiologiques

D'autres approches sont utilisées en épidémiologie environnementale. Il s'agit en particulier des approches écologiques : de type temporel, correspondant à l'approche des séries temporelles (Bell et coll., 2004), ou spatial, dans lesquelles l'unité d'observation est la population dans son ensemble et non pas l'individu, des approches *case-crossover* (Maclure 1991). Ces approches n'ayant pour l'instant pas été appliquées à des polluants tels que les perturbateurs endocriniens ne seront pas détaillées ici.

## Estimation des expositions en épidémiologie environnementale

On peut considérer que c'est à la suite d'une chaîne d'événements plus ou moins complexes selon les polluants que ceux-ci finissent par arriver en contact avec l'Homme (figure 21.3) : émission du ou des polluants à partir de certaines sources (d'origine anthropique ou naturelle), dispersion/répartition dans les différents grands compartiments de l'environnement (eau, air, sol, chaîne alimentaire, médicaments, habitat...), où les polluants sont transférés et parfois transformés, pour arriver jusqu'aux micro-environnements en contact direct avec l'Homme ; internalisation d'une partie des polluants dans l'organisme (par inhalation, ingestion, absorption dermique, irradiation...), métabolisme (figure 21.3), stockage/excrétion. Les grandeurs correspondant à chacune de ces étapes sont différentes : les polluants émis dans l'environnement correspondent à une quantité (totale ou par unité de temps) ; les niveaux mesurés dans l'environnement à une concentration (par volume d'air, d'eau ou unité de masse végétale ou animale) ; l'exposition humaine, qui correspond au sens strict à la quantité du polluant qui a été en contact avec les surfaces d'échanges de l'organisme pendant une période donnée ; la dose (interne) correspond à la quantité de polluant qui a pénétré dans l'organisme (quantité ou concentration par litre de fluide biologique ou kg de tissu).



**Figure 21.3 : Représentation des grandes familles d'approches utilisées en épidémiologie environnementale pour estimer l'exposition aux polluants environnementaux**

Le type d'étude choisi ne détermine généralement pas de façon univoque l'approche pouvant être utilisée pour estimer l'exposition. Cette approche

pourra reposer, de façon non exclusive, sur des questionnaires (aux sujets de l'étude, leurs proches, leur personnel soignant...), des modèles environnementaux déjà disponibles, des mesures environnementales (par exemple à proximité des lieux de vie des sujets), des mesures d'exposition personnelle (à partir de dosimètres portés par les sujets dans une période définie), des dosages de biomarqueurs d'exposition à partir de prélèvements biologiques réalisés chez les sujets (sang, urine, lait...).

L'exposition est souvent résumée sous forme d'une quantité, mais il est crucial d'identifier aussi la fenêtre d'exposition considérée dans l'étude, c'est-à-dire la période temporelle (définie de façon calendaire, ou plus pertinemment par rapport à l'âge ou au stade de développement du sujet) au cours de laquelle cette exposition est quantifiée.

D'autres grandeurs liées à l'exposition peuvent être aussi considérées : débit de dose, caractère aigu ou chronique de l'exposition, dépassement d'un certain niveau d'exposition...

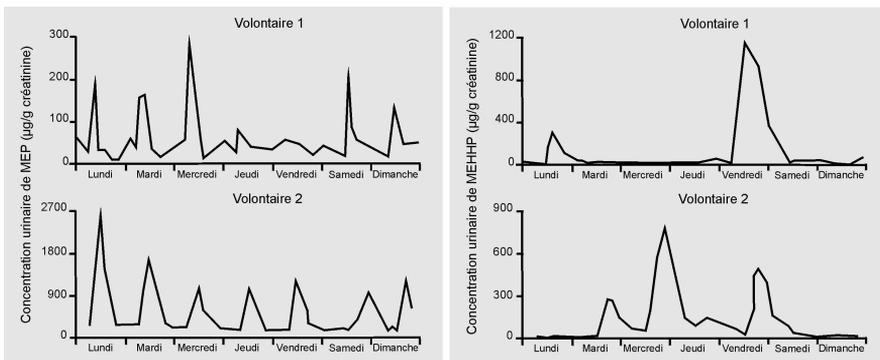
### **Biomarqueurs d'exposition**

Dans le cas (très fréquent aujourd'hui) où des prélèvements biologiques sont réalisés, tous les outils de la biochimie et de la biologie moléculaire peuvent être utilisés si les conditions de prélèvement et de stockage le permettent. C'est le cas entre autres des outils de protéomique et de métabolomique. En complément de ces approches globales, une approche polluant par polluant est possible, et reste la plus fréquemment utilisée : dans ce cas, un composé ou un petit nombre de composés d'une même famille sont dosés simultanément (par exemple, le PFOA et le PFOS). Si l'effectif est suffisant, l'analyse statistique peut permettre de prendre en compte simultanément ces différents polluants afin de chercher à identifier leur effet propre. Par exemple, dans une étude concernant l'effet des perturbateurs endocriniens sur les caractéristiques spermatiques, les auteurs ont cherché à isoler l'association entre l'exposition aux parabènes avec les caractéristiques spermatiques de l'association entre l'exposition au bisphénol A avec les caractéristiques spermatiques (Meeker et coll., 2011).

La problématique des mélanges est parfois abordée sous un autre angle ; dans le cas d'une étude concernant l'impact de perturbateurs endocriniens sur le risque de cancer du sein, un indice global du fardeau de xéno-œstrogènes dans l'organisme a été estimé à partir de prélèvements de tissu adipeux et son association avec le risque de cancer du sein quantifiée (Fernandez et coll., 2004 ; Ibarluzea et coll., 2004).

Si l'approche par biomarqueur d'exposition peut sembler séduisante au premier abord car elle fait la jonction entre l'épidémiologie et les outils de la biologie moléculaire et de la biochimie, et également par son caractère objectif, elle peut se révéler très limitée dans différentes situations : c'est en particulier le cas

quand on connaît mal le métabolisme chez l'humain des composés étudiés (auquel cas ses principaux métabolites peuvent ne pas être connus), quand le métabolisme varie fortement avec la dose, et quand le composé parent et ses principaux métabolites ont une demi-vie très courte (typiquement, quelques heures) dans l'organisme. Dans ce dernier cas (figure 21.4), il est peu probable qu'un unique prélèvement biologique suffise à donner une bonne estimation de l'imprégnation ou même de la hiérarchie des sujets en termes d'imprégnation, et d'autant moins que la fenêtre d'exposition biologiquement pertinente est longue. Une alternative consiste à recueillir plusieurs échantillons biologiques répartis durant l'ensemble de la fenêtre d'exposition biologiquement pertinente, ce qui peut se révéler logistiquement ou éthiquement difficile à grande échelle, ou dans le cas de prélèvements invasifs.



**Figure 21.4 : Variations de la concentration urinaire de deux métabolites de phtalates au cours de la semaine chez des volontaires sains (d'après Preau et coll., 2011)**

Chaque figure correspond au profil du métabolite concerné au cours de la semaine chez un volontaire en situation d'exposition « normale ».

## Erreur et biais

Les 4 grandes sources d'erreur en épidémiologie sont l'erreur aléatoire, les biais de confusion, les biais de classement et les biais de sélection. La notion d'erreur indique un écart entre la « vraie » valeur (par exemple le poids de naissance) et la valeur estimée (le poids de naissance mesuré par la sage-femme avec une balance donnée). Si cette erreur a une composante systématique, c'est-à-dire si la valeur estimée sur un grand nombre d'estimations de la valeur d'intérêt diffère de la valeur réelle (comme cela se produirait si la balance était faussée et avait tendance par exemple à sous-estimer de 100 g la masse réelle), on parle d'erreur systématique ou de biais. Si la moyenne des valeurs estimées correspond à la valeur réelle (comme cela surviendrait si la balance avait tendance alternativement à sous-estimer puis à surestimer de 100 g la masse réelle), on parle d'erreur aléatoire et non pas de biais.

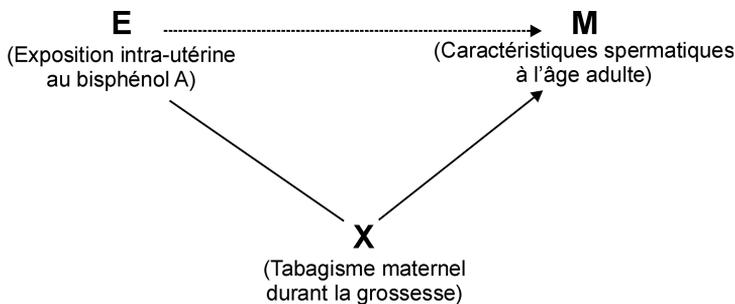
## Erreur aléatoire

C'est la composante non systématique (c'est-à-dire sans direction privilégiée, ou nulle en moyenne) de l'erreur. Dans une vision déterministe, l'erreur aléatoire est due en partie à notre méconnaissance des facteurs influençant la survenue de l'événement d'intérêt : plus on diminue cette méconnaissance, plus les variations aléatoires s'amenuisent.

Une autre composante de l'erreur aléatoire est due aux fluctuations d'échantillonnage (ou erreur d'échantillonnage). Augmenter la taille de l'échantillon d'étude (sans que cela se fasse au détriment de la précision et la validité des mesures faites sur les sujets) permet de limiter les fluctuations d'échantillonnage et donc l'erreur aléatoire.

## Biais de confusion

On parle de biais de confusion quand il existe un facteur extérieur à la chaîne causale reliant l'exposition à la maladie considérée qui fausse l'estimation de l'association entre exposition et maladie. Ceci peut typiquement survenir si ce facteur extérieur est à la fois associé à l'exposition et à la probabilité de survenue de la maladie (sans en être une cause). Si par exemple on s'intéresse à l'effet possible de l'exposition intra-utérine au bisphénol A provenant du régime alimentaire maternel sur la fertilité à l'âge adulte, le tabagisme maternel pendant la grossesse est susceptible de constituer un facteur de confusion (figure 21.5) : il peut être associé au régime alimentaire, et donc à l'exposition intra-utérine au bisphénol A. D'autre part, il est susceptible d'influencer la fertilité de la descendance à l'âge adulte (Jensen et coll., 2005).



**Figure 21.5 : Facteur de confusion : diagramme causal d'une étude hypothétique visant à caractériser l'impact de l'exposition intra-utérine au bisphénol A sur les caractéristiques spermatisques à l'âge adulte**

S'il est associé au régime alimentaire maternel (influençant l'exposition au bisphénol A du fœtus) et à la fertilité à l'âge adulte, le tabagisme maternel pendant la grossesse est un facteur de confusion potentiel.

Les biais de confusion dus à des facteurs connus peuvent être corrigés par différentes approches. La plus fréquemment utilisée est de nature statistique et

consiste à mesurer le facteur correspondant et à le prendre en compte par ajustement dans le modèle de régression estimant l'association entre exposition et maladie. L'association ainsi estimée sera équivalente à celle qu'on aurait observée si la fréquence du facteur d'exposition était la même entre le groupe exposé et le groupe non exposé, c'est-à-dire si on avait supprimé l'association entre l'exposition et le facteur de confusion.

## Biais de classement

La réalisation d'une étude épidémiologique peut être vue comme un exercice de métrologie, dans lequel on mesure (ou estime) chez des sujets recrutés selon un protocole défini et validé un niveau d'exposition, la survenue d'un événement de santé, et la présence de facteurs de confusion potentiels. Pour des raisons logistiques et aussi liées à la complexité des paramètres étudiés, chacune de ces mesures est susceptible de se faire avec une certaine erreur.

L'existence de ces erreurs de mesure (ou de classement) n'entraîne pas automatiquement de biais et n'invalide donc pas forcément l'étude. En particulier, si l'erreur de mesure sur un facteur donné (par exemple, l'exposition à un composé perfluoré estimée à partir de dosages plasmatiques réalisés dans une fenêtre biologiquement pertinente par rapport à l'événement de santé étudié) est totalement aléatoire, c'est-à-dire indépendante de toutes les caractéristiques des sujets y compris de leur exposition réelle, cette erreur aura pour principale conséquence de diminuer la précision de l'association estimée entre exposition et maladie. Il n'est pas sûr qu'un biais survienne, mais s'il y en a un celui-ci correspondra généralement à une sous-estimation de l'association entre exposition et maladie. En d'autres termes, l'intervalle de confiance associé au risque relatif quantifiant cette exposition sera plus large que si l'exposition avait été mesurée sans erreur, et la puissance statistique de l'étude sera diminuée. Le tableau 21.I indique par exemple que si l'exposition estimée est corrélée à l'exposition réelle avec un coefficient de corrélation de 0,6 et que l'effet de l'exposition réelle correspond à un risque relatif de 3, il faudra recruter 2 à 3 fois plus de sujets dans l'étude que si l'exposition était mesurée sans erreur, et le risque relatif estimé sera en moyenne de 1,9 (doublement du risque en cas d'exposition, alors qu'il est en réalité triplé)(de Klerk et coll., 1989).

Dans d'autres cas, l'erreur de classement pourra entraîner une surestimation de l'association entre exposition et maladie, ou une sous-exposition (Jurek et coll., 2005).

L'erreur de classement peut aussi bien concerner l'exposition d'intérêt que la mesure de l'événement de santé ou d'un des facteurs de confusion potentiels. Dans ce cas, l'ajustement sur ce facteur mal mesuré est susceptible de ne pas permettre d'éliminer l'ensemble du biais de confusion (on parle de biais de confusion résiduel). Un tel biais de confusion résiduel peut aussi survenir si le

codage de la variable représentant le facteur de confusion dans le modèle de régression n'est pas approprié (Slama et Werwatz, 2005).

**Tableau 21.1 : Conséquence de l'erreur de mesure sur l'exposition sur le nombre de sujets nécessaires dans une étude cas-témoins pour garantir une puissance statistique satisfaisante (80 %) et sur le risque relatif caractérisant l'association entre exposition et survenue de la maladie**

	Nombre de sujets nécessaires	Risque relatif estimé
Sans erreur de mesure	106 cas, 106 témoins=212	3,0
Avec erreur de mesure ( $r=0,64$ )	286 cas, 286 témoins=572	1,9
Avec erreur de mesure ( $r=0,20$ )	3 072 cas, 3 072 témoins = 6 144	1,2

Les valeurs sont données pour différents niveaux de corrélation ( $r$ ) entre la valeur réelle de l'exposition et son estimation avec erreur utilisée dans l'étude épidémiologique (d'après de Klerk et coll., 1989).

La correction des erreurs de classement est plus complexe à mettre en œuvre que la correction des biais de confusion, mais elle est possible. Elle repose généralement sur la quantification de l'erreur de mesure (par exemple en mesurant avec différentes approches le facteur susceptible d'être mal estimé) et sur sa prise en compte dans des analyses de sensibilité pouvant reposer sur des simulations (Lash et Fink, 2003).

### Biais de sélection

Il faut noter qu'il n'y a pas de nécessité pour la population étudiée d'être représentative de la population générale du pays. En revanche, la population source d'où est issue la cohorte doit être clairement identifiée, et c'est vis-à-vis de cette population que la représentativité doit être atteinte, afin de limiter les biais de sélection éventuels. Cette représentativité peut être discutée à la lumière du taux de participation et d'une comparaison entre les sujets ayant accepté et ceux ayant refusé de participer. D'autres types de biais de sélection existent, notamment dans les études de cohorte, où un taux important de sujets perdus de vue au cours du suivi peut entraîner un biais de sélection.

### Interprétation des résultats d'une étude épidémiologique

Généralement, les résultats d'une étude épidémiologique sont sous la forme d'un paramètre quantifiant l'association entre l'exposition considérée et la caractéristique de santé. Si le paramètre de santé est de nature continue (concentration d'une hormone stéroïdienne, par exemple), le paramètre correspondra à la variation de cette caractéristique entre le groupe exposé et le groupe non-exposé (ou entre 2 groupes ayant des niveaux d'exposition différents). Si la caractéristique de santé est binaire (de type malade/non malade),

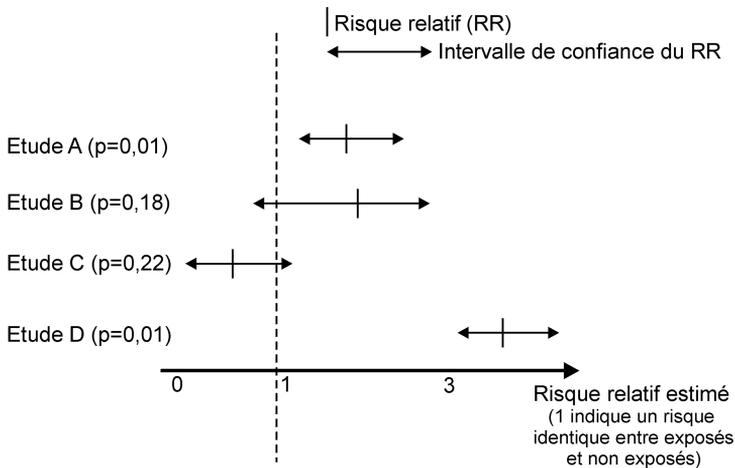
le paramètre quantifiant l'association entre exposition et maladie correspondra généralement au rapport entre la probabilité (ou fréquence) de la maladie chez les exposés, par rapport au groupe non exposé. Il s'agira selon l'étude et le modèle statistique d'un risque relatif, risque relatif instantané, ou *odds-ratio*. Un risque relatif de 1,5 signifie que la fréquence de la maladie est augmentée de 50 % chez les sujets exposés, par rapport aux sujets non exposés. Plus rarement, d'autres paramètres sont estimés, tels que la fraction de risque attribuable, ou la proportion de cas de la maladie qu'on peut attribuer à l'exposition considérée. Le paramètre quantifiant l'association entre exposition et maladie est corrigé de l'effet de facteurs de confusion pris en compte, c'est-à-dire qu'il caractérise l'association qu'on observerait s'il n'y avait pas de biais de confusion dû aux facteurs mesurés dans l'étude. Ce paramètre est une mesure d'association statistique mais, en l'absence de biais, elle est souvent interprétée comme une mesure d'impact, ou au moins comme une estimation de l'impact de l'exposition sur la fréquence de la maladie.

Le paramètre quantifiant l'association entre exposition et maladie est toujours assorti d'un intervalle de confiance (généralement à 95 %), et souvent d'un degré de signification, généralement noté « p ». L'intervalle de confiance peut être interprété (un peu abusivement) comme un intervalle qui a de bonnes chances de contenir la vraie valeur du paramètre quantifiant l'association entre exposition et maladie. Il donne une idée de l'amplitude de l'erreur aléatoire, ou encore de la précision de l'étude : si l'intervalle de confiance est large, on peut considérer que l'étude est peu précise, ou encore peu informative (un vaste éventail de valeurs sont possibles pour le paramètre quantifiant l'association entre exposition et maladie) ; si l'intervalle de confiance est étroit, on peut considérer l'étude comme informative. Le degré de signification donne une idée de la probabilité qu'on aurait d'observer une valeur estimée du paramètre quantifiant l'association entre exposition et maladie sous l'hypothèse (notée  $H_0$ ) où il n'y a pas en réalité d'association entre exposition et maladie, c'est-à-dire si on suppose que le « vrai » risque relatif vaut 1. Plus le degré de signification est faible, moins le résultat observé aurait eu de chances d'être observé s'il n'y avait pas d'association réelle. En conséquence, plus le degré de signification est faible, plus on a tendance à rejeter cette hypothèse  $H_0$  d'une absence d'association entre exposition et maladie : on va donc être tenté de considérer qu'il y a bien une association. Si le degré de signification est élevé, l'association observée est compatible avec l'hypothèse d'une absence d'association, et pourrait être expliquée par des fluctuations aléatoires. En conséquence, on ne peut pas rejeter l'hypothèse  $H_0$ . Toutefois, cela ne signifie pas qu'on doive accepter cette hypothèse, et que les résultats indiquent qu'il n'y a pas d'association réelle. Si le degré de signification est élevé, en toute rigueur, on ne peut tirer de conclusion forte dans un sens ou l'autre, et l'étude est peu informative.

Une erreur d'interprétation fréquemment commise est de considérer que deux études ayant des degrés de signification très différents (par exemple valant

0,01 et 0,20, considérés respectivement comme « statistiquement significatif » et « statistiquement non significatif ») sont contradictoires. Le fait que plusieurs études sur une même question aient parfois des degrés de signification inférieurs à 5 % et parfois supérieurs à 5 % ne correspond pas forcément à une situation où la littérature doit être considérée contradictoire. Cette situation est même attendue si l'exposition a un effet réel sur le risque de maladie et si les études réalisées n'ont pas toutes une puissance statistique très élevée. Comparer les « p » n'est donc pas très informatif.

En revanche, une comparaison des intervalles de confiance associés à un même paramètre et tirés de différentes études peut s'avérer très informatif (figure 21.6). En schématisant, si les intervalles de confiance se recoupent assez largement, on peut considérer que la littérature est assez cohérente ou homogène. Il est alors utile de réaliser une méta-analyse de ces études (cf. figure 21.2) qui donnera une estimation de l'association la plus probable, correspondant approximativement à une moyenne des paramètres quantifiant l'association dans les différentes études, ainsi que de l'intervalle de confiance.



**Figure 21.6 : Exemples hypothétiques de risques relatifs (et leur intervalle de confiance et degré de signification) d'études quantifiant toutes une même association entre un facteur d'exposition donné et le risque de survenue d'une maladie (études supposées sans biais)**

Le risque relatif issu de l'étude A est statistiquement significatif, celui issu de l'étude B ne l'est pas, et pourtant, considérées simultanément les études A et B sont cohérentes et fournissent l'intervalle de valeurs 1-3 comme le plus probable pour la valeur réelle du risque relatif. Les études B et C ont toutes deux des degrés de signification « non significatifs » au seuil de 5 %, et pourtant elles semblent peu compatibles (l'une indique que le RR est probablement supérieur à 1, l'autre qu'il est probablement inférieur à 1). De même, les études A et D (toutes deux significatives) ne sont pas très cohérentes quant à la valeur la plus probable du RR. Ceci illustre que la cohérence (ou du moins l'homogénéité) de la littérature ne peut être jugée à partir des degrés de signification des études, mais plutôt des intervalles de confiance des mesures d'association.

Si en revanche un grand nombre d'intervalles de confiance sont disjoints (typiquement, un intervalle de confiance allant de 0,2 à 0,9, un autre allant de 1,5 à 7,2), alors on pourra parler d'hétérogénéité des résultats de la littérature, voire de « contradiction ». Tout ceci suppose qu'on considère des études qui sont toutes dépourvues de biais...

## Causalité

On peut définir la cause d'une maladie comme étant un événement ou une caractéristique antérieur à la survenue de cette maladie et qui était nécessaire pour que la maladie survienne au moment où elle est survenue, toutes les autres conditions étant fixées. Si la cause avait été absente, la maladie ne serait pas survenue, ou elle serait survenue à un autre moment (Rothman et coll., 2008). Cette définition ne fait pas de distinction entre un facteur qui initierait la survenue de la maladie et un autre qui ne ferait que précipiter sa survenue : tous deux sont considérés comme des causes de la maladie.

Les épidémiologistes utilisent (souvent implicitement) différents modèles de causalité. Un modèle de causalité assez fréquemment considéré est celui des causes suffisantes composées, proposé par Rothman (Rothman 1976 ; Rothman et coll., 2008). Dans ce modèle, chaque cause unique de la maladie n'est généralement ni nécessaire ni suffisante (on peut fumer sans mourir des effets du tabac, ou développer un cancer du poumon sans fumer ni être exposé à la fumée de tabac). En revanche, le modèle considère qu'il existe des causes composées (ensemble de facteurs et caractéristiques), qui sont elles suffisantes pour entraîner la survenue de la maladie chez certains sujets. Pour une maladie donnée, il existe a priori plusieurs causes composées pouvant la déclencher, n'incluant pas toujours les mêmes causes élémentaires, ce qui est cohérent avec l'observation qu'il existe rarement des causes nécessaires à la survenue des maladies complexes.

Une caractéristique importante de la survenue des maladies humaines complexes est leur caractère multicausal. Très peu de maladies ont des causes uniques. Même une maladie comme la phénylcétonurie (arriération mentale résultant d'un trouble du métabolisme d'un acide aminé, la phénylalanine, lui-même dû à une mutation génétique), considérée comme génétique, peut être vue comme environnementale : en effet, une intervention alimentaire appropriée (régime pauvre en phénylalanine) permet d'éviter la manifestation de l'arriération mentale ; le régime alimentaire est donc aussi une des causes de la maladie (Rothman et coll., 2008).

Ceci illustre qu'il n'est pas pertinent d'opposer les causes environnementales et génétiques des maladies : il faut plutôt les voir comme des facteurs intervenant à des niveaux causaux différents, éventuellement emboîtés les uns dans les autres mais en aucun cas opposés ou complémentaires. Une mutation

génétique entraînant la survenue d'une maladie dans la génération  $n$  peut avoir été causée par une exposition environnementale à la génération  $n-1$ . Et cette exposition environnementale peut elle-même avoir une origine psychosociale. Comme le souligne Rothman, à un certain niveau d'observation, la quasi-totalité des causes peuvent être considérées comme étant de nature génétique, alors qu'à un autre niveau les causes d'une même maladie peuvent être considérées comme étant essentiellement de nature environnementale, sans que ceci soit contradictoire. À ce titre, ainsi que pour d'autres raisons, si on considère que 40 % des cas d'une maladie sont attribuables à des facteurs génétiques, il serait erroné de considérer que 60 % des cas sont attribuables à l'environnement.

Il n'existe pas de recette simple à appliquer pour déterminer si une association rapportée dans une étude épidémiologique est de nature causale. Si on entend parfois parler de « critères de causalité de Bradford Hill » dont il suffirait de vérifier s'ils sont vérifiés par une étude donnée, ceci fait référence à une liste de standards (Hill 1965) qui, s'ils ont, pour certains de la logique (par exemple le « critère » de temporalité, selon lequel la cause doit précéder l'effet), n'ont pas de réel fondement scientifique. Pour reprendre des arguments développés notamment par Rothman (2002), on peut citer le critère de spécificité d'une association : il existe beaucoup de polluants environnementaux ayant des effets sanitaires variés, donc il semble difficile de justifier qu'un facteur apparemment associé à une seule pathologie soit une cause de cette pathologie plus vraisemblable qu'un autre facteur associé à plusieurs pathologies. Un autre exemple, le critère de relation dose-effet (strictement) monotone, qui exclut les associations avec seuil ou les relations en U, qui peuvent correspondre à des effets réels. S'il n'existe pas de critères de causalité aisément applicables, il existe bien des arguments sur la validité interne à une étude et, de manière liée, des niveaux de preuves différents apportés par différents types d'étude. Ces arguments de validité concernent l'absence de biais, ou plus précisément l'absence de biais susceptibles de créer une association ayant l'amplitude observée dans l'étude. Une étude pour laquelle des biais résiduels existent, mais susceptibles d'entraîner une sous-estimation de l'amplitude de l'association, pourra être prise en compte. Concernant le niveau de preuve apporté par différents types d'étude, on considère généralement qu'il décroît quand on va des études randomisées aux études d'observation avec des informations et une analyse au niveau individuel (en privilégiant les cohortes prospectives sur les études rétrospectives), puis jusqu'aux études écologiques spatiales (les études écologiques temporelles, telles que les séries temporelles, peuvent apporter un niveau de preuve élevé, comparable aux études avec des données individuelles). Les études de cas sont considérées comme apportant le niveau de preuve le plus faible.

Si on peut admettre, comme on l'entend souvent, que l'épidémiologie ne permet pas d'établir de relation causale, il ne faut pas oublier de compléter cette affirmation en remarquant que c'est le cas de chacune des disciplines

impliquées dans la santé environnementale. De même qu'une unique étude épidémiologique, aussi bien faite soit elle, ne permet pas d'établir avec certitude la causalité d'une association (ne serait-ce que parce que les associations fortuites dues aux fluctuations aléatoires, si elles deviennent moins probables à mesure que la taille de l'échantillon et la précision des mesures augmentent, ne peuvent jamais être écartées avec certitude), une étude toxicologique ou une expérimentation *in vitro* ne le permettent pas (ce qui ne signifie pas qu'elles ne permettent pas de caractériser le mécanisme d'action du composé) : d'une part, ces autres approches sont elles aussi bien souvent soumises aux fluctuations aléatoires, mais surtout, elles sont limitées au modèle expérimental sur lesquelles elles reposent. La généralisation des résultats à d'autres modèles ou espèces, et notamment à l'espèce humaine, ne relève pas d'un processus scientifique et est généralement entachée d'incertitude : même si on admet qu'une expérience de toxicologie permet d'établir qu'un composé est hépatotoxique chez le rat Wistar, elle ne permettra pas d'établir l'existence d'une telle relation causale dans l'espèce humaine, et il y a même fort à parier que l'effet du composé différerait chez le rat Sprague-Dawley...

En pratique, dans le contexte de la recherche en santé environnementale et de la gestion du risque sanitaire lié aux facteurs environnementaux, il apparaît pertinent de considérer l'établissement de la causalité comme étant essentiellement le fruit d'une expertise scientifique impliquant de nombreuses disciplines. En confrontant les travaux de ces différentes disciplines, les experts « décident » de la plausibilité de l'effet d'un composé sur un paramètre de santé. Ceci peut être vu comme l'élaboration d'une théorie scientifique concernant l'effet de ce composé, qui inclut le métabolisme, les organes, tissus, récepteurs cibles de ce composé, les effets phénotypiques observés et la relation dose-effet attendue à l'échelle de la population. Plutôt qu'une réponse en oui ou non, ce processus est susceptible d'apporter des éléments sur la causalité de façon graduée, indiquant si une association donnée est très peu probable (peu d'éléments en faveur d'un effet causal), probable, très probable... Ce type de classification est similaire dans sa logique à celle utilisée par le Centre international de recherche sur le cancer dans son classement des substances susceptibles d'être cancérogènes.

## **Perturbateurs endocriniens et santé humaine : vers l'étude épidémiologique idéale**

Il n'y a pas d'approche épidémiologique unique pour répondre à une question aussi complexe que celle de l'impact des perturbateurs endocriniens sur un ensemble très varié de paramètres biologiques et de santé. On peut toutefois tirer certaines conclusions sur les méthodes à privilégier à partir des grandes caractéristiques du problème posé.

Ces caractéristiques sont :

- du point de vue des polluants considérés, une multiplicité de composés de différentes natures chimiques, ayant chacun des sources et des voies d'exposition multiples ; des synergies ou antagonismes possibles entre différents polluants. Ces polluants très nombreux ont chacun différents (parfois beaucoup de) métabolites, pour lesquels il est encore rarement possible d'identifier un unique métabolite actif. Beaucoup de ces métabolites sont très peu ou peu persistants dans l'organisme (demi-vie de quelques heures à quelques semaines) ;
- du point de vue de la fenêtre d'exposition, l'existence possible de différentes fenêtres de sensibilité au cours de la vie, avec des hypothèses fortes concernant la fenêtre développementale (vie intra-utérine et premières années de vie) ; une période de latence pouvant être longue (plusieurs années voire dizaines d'années), et des effets pouvant dans certains cas persister ou apparaître dans la descendance de la génération exposée ;
- des mécanismes biologiques complexes, probablement multiples, pouvant impliquer le système endocrinien, d'autres systèmes (système immunitaire ou nerveux par exemple) et, à une autre échelle, l'épigénétique ;
- du point de vue de la relation dose-effet, l'existence possible de relations non strictement monotones et une difficulté de se servir des relations dose-effets obtenues aux doses élevées pour prédire les effets des doses plus faibles ;
- des événements et paramètres biologiques et sanitaires très variés pouvant être influencés par les perturbateurs endocriniens.

Ces caractéristiques incitent à privilégier des approches de type cohorte prospective, adossées à d'importantes biobanques, qui ont l'avantage de permettre :

- d'étudier simultanément différents paramètres de santé ;
- de pouvoir réaliser des prélèvements biologiques répétés durant la période biologiquement pertinente, autorisant le dosage de paramètres biologiques pouvant éclairer les mécanismes en jeu ;
- de doser efficacement et de façon prospective l'exposition à des perturbateurs endocriniens.

Dans le cas où on s'intéresse à l'effet à court ou moyen terme d'expositions à l'âge adulte (par exemple sur des caractéristiques hormonales ou spermatozoaires), cette approche peut prendre la forme d'un échantillon de quelques dizaines à quelques centaines de sujets chez qui des prélèvements d'urine, éventuellement de sang ou de sperme, sont réalisés de façon répétée sur une période de plusieurs mois à quelques années (Bonde et coll., 1996).

Dans le cas où on s'intéresse à l'effet des expositions durant la vie intra-utérine, le *design* privilégié est probablement celui de la cohorte mère-enfants, consistant à recruter le plus tôt possible durant la grossesse un nombre important de femmes, de réaliser différents prélèvements biologiques répétés chez la mère durant la grossesse puis chez l'enfant après la naissance, et de suivre à moyen et long terme la santé de l'enfant. Un exemple intéressant est celui de

la *Danish National Birth Cohort*, dans laquelle environ 91 800 femmes enceintes ont été recrutées entre 1996 et 2002, et suivies avec leur enfant. Des prélèvements biologiques ont été réalisés en cours de grossesse, à partir desquels les concentrations de composés perfluorés ont été dosés dans un sous-échantillon de la cohorte. Cette étude a d'ores et déjà permis de fournir des informations importantes sur les conséquences possibles de l'exposition durant la grossesse aux composés perfluorés sur la fertilité (Fei et coll., 2009), le déroulement, l'issue de la grossesse (Fei et coll., 2007 et 2008) et le développement de l'enfant (Andersen et coll., 2010).

Pour certains événements de santé très rares (certains cancers notamment), une option serait de combiner les données des grandes cohortes existant dans le monde (de tels efforts sont en cours pour certaines pathologies, par exemple en relation avec les facteurs génétiques). Des études cas-témoins sont envisageables, surtout si elles sont nichées dans des cohortes comprenant une bibliothèque, pour permettre de disposer d'une estimation prospective de l'exposition.

## BIBLIOGRAPHIE

ANDERSEN CS, FEI C, GAMBORG M, NOHR EA, SØRENSEN TI, OLSEN J. Prenatal exposures to perfluorinated chemicals and anthropometric measures in infancy. *Am J Epidemiol* 2010, **172** : 1230-127

BELL ML, SAMET JM, DOMINICI F. Time-series studies of particulate matter. *Annu Rev Public Health* 2004, **25** : 247-280

BONDE JP, GIWERCMAN A, ERNST E. Identifying environmental risk to male reproductive function by occupational sperm studies: logistics and design options. *Occup Environ Med* 1996, **53** : 511-519

CANTOR K, VILLANUEVA CM, SILVERMAN DT, FIGUEROA JD, REAL FX, et coll. Polymorphisms in GSTT1, GSTZ1, and CYP2E1, Disinfection byproducts, and risk of bladder cancer in Spain. *Environ Health Perspect* 2010, **118** : 1545-1550

DE KLERK NH, ENGLISH DR, ARMSTRONG BK. A review of the effects of random measurement error on relative risk estimates in epidemiological studies. *Int J Epidemiol* 1989, **18** : 705-712

FEI C, MCLAUGHLIN JK, TARONE RE, OLSEN J. Perfluorinated Chemicals and Fetal Growth: A Study within the Danish National Birth Cohort. *Environ Health Perspect* 2007, **115** : 1677-1682

FEI C, MCLAUGHLIN JK, TARONE RE, OLSEN J. Fetal growth indicators and perfluorinated chemicals: a study in the Danish National Birth Cohort. *Am J Epidemiol* 2008, **168** : 66-72

FEI C, MCLAUGHLIN JK, LIPWORTH L, OLSEN J. Maternal levels of perfluorinated chemicals and subfecundity. *Hum Reprod* 2009, **24** : 1200-1205

FERNÁNDEZ MF, RIVAS A, OLEA-SERRANO F, CERRILLO I, MOLINA-MOLINA JM et coll. Assessment of total effective xenoestrogen burden in adipose tissue and identification

of chemicals responsible for the combined estrogenic effect. *Anal Bioanal Chem* 2004, **379** : 163-170

HILL AB. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine* 1965, **58** : 295-300

IBARLUZEA JM J, FERNÁNDEZ MF, SANTA-MARINA L, OLEA-SERRANO MF, RIVAS AM, AURREKOETXEA JJ, et coll. Breast cancer risk and the combined effect of environmental estrogens. *Cancer Causes Control* 2004, **15** : 591-600

JENSEN MS, MABECK LM, TOFT G, THULSTRUP AM, BONDE JP. Lower sperm counts following prenatal tobacco exposure. *Hum Reprod* 2005, **20** : 2559-2566

JUREK AM, GREENLAND S, MALDONADO G, CHURCH TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. *Int J Epidemiol* 2005, **34** : 680-687

KUNZLI N, KAISER R, MEDINA S, STUDNICKA M, CHANEL O, FILLIGER P, et coll. Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lancet* 2000, **356** : 795-801

LASH TL, FINK AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology* 2003, **14** : 451-458

MACLURE M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 1991, **133** : 144-153

MEEKER JD, EHRlich S, TOTHTL, WRIGHT DL, CALAFAT AM, TRISINI AT, et coll. Semen quality and sperm DNA damage in relation to urinary bisphenol A among men from an infertility clinic. *Reprod Toxicol* 2010, **30** : 532-539

MEEKER JD, YANG T, YE X, CALAFAT AM, HAUSER R. Urinary concentrations of parabens and serum hormone levels, semen quality parameters, and sperm DNA damage. *Environ Health Perspect* 2011, **119** : 252-257

PREAU JL JR, WONG LY, SILVA MJ, NEEDHAM LL, CALAFAT AM. Variability over 1 week in the urinary concentrations of metabolites of diethyl phthalate and di(2-ethylhexyl) phthalate among eight adults: an observational study. *Environ Health Perspect* 2010, **118** : 1748-1754

ROTHMAN KJ. Causes 1976. *Am J Epidemiol* 1995, **141** : 90-95, discussion 89

ROTHMAN KJ. *Epidemiology. An introduction*. New-York, Oxford University Press, 2002

ROTHMAN KJ, GREENLAND S, LASH TL. *Modern epidemiology*. Philadelphia, PA, Lippincott Williams & Wilkins, 2008

SLAMA R, WERWATZ A. Controlling for continuous confounding factors: non- and semi-parametric approaches. *Rev Epidemiol Sante Publique* 2005, **53** : 2S65-2S80

SLAMA R, CORDIER S. Environmental contaminants and impacts on healthy and successful pregnancies. In: *Environmental impacts on reproductive health and fertility*. WOODRUFF TJ, JANSSEN SJ, GUILLETTE LJ JR, GIUDICE LC (eds). *Cambridge University Press*, 2010