

# Les biais dans les études d'observation<sup>1</sup>

Peter Kleist

GlaxoSmithKline AG, Münchenbuchsee

## Quintessence

- Les études d'observation (par exemple les études de cohorte ou les études de cas-témoins) sont soumises par principe aux multiples risques liés au biais statistique.
- Si un biais de confusion est anticipé, on peut le contrôler par des méthodes statistiques. Par contre, un biais de sélection et/ou d'information endommage une étude d'observation de manière irrémédiable.
- Le lecteur de cet article sera confronté à diverses formes d'expression du biais: de cette façon, il apprendra à poser les bonnes questions lui permettant de mieux évaluer une étude d'observation.
- Lorsque c'est possible, il faut toujours donner la préférence aux études randomisées contrôlées, car les études d'observations ne constituent pas une solution de remplacement appropriée.
- Les études non randomisées ne se prêtent pas à l'évaluation de mesures de dépistage.

**La vie est l'art de tirer des conclusions valables à partir d'hypothèses erronées.**

*Samuel Butler (1612–1680), auteur satirique anglais*

## Problématique

Le domaine d'application des études d'observation (c.-à-d. des essais non interventionnels) s'étend à l'identification de facteurs de risque et d'indicateurs pronostiques, ainsi qu'à la détermination de l'étiologie des maladies. L'étude d'observation intervient également si un essai randomisé contrôlé est impossible à réaliser ou inacceptable sur le plan éthique. C'est pourquoi l'importance des études d'observation, particulièrement celle des études de cohorte et des études cas-témoin (ou études rétrospectives), est incontestable en épidémiologie et en médecine clinique.

Cependant, depuis de longues années, on discute de la question de savoir si l'étude d'observation pourrait se substituer partiellement à l'étude randomisée contrôlée. Pour y répondre, il faut comparer la consistance des résultats des deux types d'études. Il est vrai que deux méta-analyses ont su mettre en évidence de larges recoupements entre les résultats des études d'observation et ceux des études randomisées contrôlées [1, 2]; mais la crédibilité de ces méta-analyses est à son tour mise en doute [3]. Dans certains cas où les résultats des deux types d'études ont fortement divergé, les études d'observation ont généralement surestimé l'effet de la

mesure examinée [4]. Les exemples que nous connaissons tous sont celui des vitamines anti-oxydantes destinées à prévenir les tumeurs et celui du traitement de substitution hormonale post-ménopausique censé prévenir les maladies cardiovasculaires: les premières études d'observation laissaient supposer que ces deux traitements exerçaient un effet préventif, puis des études ultérieures, randomisées, ont montré que l'effet nocif de ces traitements surpassait leur effet bénéfique. Comment expliquer ces écarts? Les études d'observation non randomisées sont fondamentalement exposées à de multiples risques de biais entraînant une déformation systématique des résultats à des niveaux de l'étude très différents: choix des personnes, saisie des données, évaluation ou interprétation [5]. Autrement dit, il existe d'autres facteurs susceptibles d'influencer les résultats d'un essai.

Cet article cherche à attirer l'attention sur les formes les plus importantes d'expression de biais dans les études d'observation, et il voudrait permettre au lecteur d'avoir un regard critique sur les résultats de ces essais.

## Biais de sélection – quels éléments compare-t-on?

Tout d'abord, il faut examiner sur quelle base s'effectue le choix des participants à l'étude. Par exemple, dans une étude d'observation prospective, une partie des participants fut recrutée avec beaucoup de soin, et l'autre partie fut constituée de volontaires ayant répondu positivement à une seule prise de contact écrite [6]. Il s'est avéré que le groupe des volontaires était composé en majorité de femmes, que le degré de formation y était plus élevé et qu'il réagissait mieux aux tests cognitifs – et que six à huit ans plus tard, le taux de mortalité y était nettement plus faible que dans l'autre groupe.

Un tel biais, dit «biais de volontariat», prend une portée particulière dans les études d'observation, au sein desquelles les cohortes ou groupes de cas proviennent fréquemment de différentes sources ou ont été formés à partir de différents types de mesures. Par exemple, les fumeurs réagissent moins souvent aux questionnaires que les non-fumeurs et forment ainsi un groupe particulier dans le cadre des études d'observation [7]. En conséquence, vous devez observer la manière dont la sélection des participants s'est opérée, et vous demander si les différences de procédures de recrutement

<sup>1</sup> Cette formation continue est également accréditée par l'Association Suisse de Médecine Pharmaceutique (ASMP).



Peter Kleist

L'auteur certifie qu'aucun conflit d'intérêt n'est lié à cet article.

ou de réponse aux demandes étaient susceptibles d'influencer le résultat d'une étude.

Le choix du groupe de contrôle dans les études de cas-témoins constitue une forme particulière de biais de sélection (*control group bias*). Prenons un exemple hypothétique, avec la question suivante (tirée de la réf. [8]): la prise d'antirhumatisme non stéroïdiens inhibe-t-elle la formation d'un carcinome colorectal? Les «cas» sont les patients atteints d'un carcinome colorectal, et les témoins, les patients non atteints de cette affection. Si les témoins sont choisis dans une section de rhumatologie alors que les cas ne le sont pas, l'étude est fortement biaisée. En effet, l'usage bien plus prononcé de médicaments antirhumatisme dans le groupe témoin

**Le lecteur d'une étude d'observation devrait prêter une attention toute particulière aux caractéristiques des participants**

ferait pencher les résultats en direction d'un risque nettement plus faible des personnes sous traitement.

Dans une étude cas-témoin, réelle cette fois [9], cherchant à caractériser le risque de sida, on a effectué une première comparaison avec des patients d'une clinique pour maladies sexuellement transmissibles, et une seconde avec des personnes issues du voisinage de patients atteints de sida. L'Odds ratio par rapport au premier groupe témoin se situait à 2,9, et celui par rapport au deuxième, à 52. L'étendue de cette différence montre l'influence que peut exercer un biais résultant du mauvais choix du groupe témoin. En lisant une étude de cas-témoins, vous devez donc toujours vous demander si les témoins sont vraiment représentatifs au regard de la question posée. Les trois exemples qui suivent illustrent d'autres manifestations des biais de sélection dus au processus de sélection des participants:

Dans les études d'observation citées plus haut, les personnes ayant suivi un traitement de vitamines ou de substitution hormonale avaient de manière générale une conscience plus développée de leur santé que les témoins, et cette circonstance expliquait que leur risque d'atteinte de cancer ou de maladie cardiovasculaire ait été plus faible (biais des personnes adoptant un style de vie plus sain ou *healthy-user bias* [10, 11]).

Un autre type de biais de sélection peut apparaître en cas de risques se concurrençant dans une même étude, c.-à-d. lorsque deux événements possibles s'excluent mutuellement (biais dit de «risques concurrents», ou *competing risk bias*). Exemple: chez les personnes appréciant le vin rouge, le taux plus faible de maladies cardiovasculaires pourrait découler du seul fait que les gros buveurs décèdent plus rapidement pour d'autres causes, telles que l'insuffisance hépatique, et qu'ainsi ils échappent à l'exposition au risque de complications vasculaires.

Notre troisième exemple est le biais de confusion «incidence-prévalence», appelé également biais de survie sélective (ou biais de Neyman, en référence au mathématicien polonais Jerzy Neyman [12]), qui peut survenir dans les études transversales (*cross-sectional studies*) ou dans les études de cas-témoins. Si au lieu de relever le nombre de nouveaux cas dans un intervalle

de temps donné (incidence), on examine des cas déjà connus à un moment précis (prévalence), alors le facteur de risque présente un rapport avec le pronostic plutôt qu'avec l'étiologie de la maladie. Par exemple, le sous-groupe des patients décédant peu de temps après le diagnostic sera sous-représenté dans l'étude. Ainsi une augmentation du taux de mortalité chez les «cas» ne sera pas perçue, car la sélection s'est opérée sur des sujets dont le pronostic à long terme est plus favorable. Pour les mêmes raisons, on risque de sous-estimer l'influence de la surcharge pondérale sur les maladies cardiaques, car une partie des patients peut mourir pour d'autres causes liées aux comorbidités de l'obésité, comme par exemple les conséquences d'un diabète. A l'inverse, l'insuline pourrait passer à tort pour un facteur de risque du diabète, car elle prolonge la durée de la maladie, élevant ainsi la prévalence des patients sous traitement insulinaire.

Pour détecter l'origine d'un biais, il convient en conséquence de vous poser les questions suivantes: y a-t-il d'autres facteurs, par exemple la prise de vitamines, ou d'autres comportements susceptibles de combinaison avec la caractéristique étudiée? Existe-t-il des risques concurrents chez les participants de l'étude? A-t-on utilisé la prévalence plutôt que l'incidence d'une maladie, et se peut-il qu'un biais survienne parce que la caractéristique examinée exerce une influence sur la durée de la maladie ou la mortalité?

### Biais d'information – qu'examine-t-on?

Le «biais d'information» ou «biais de mesure» survient lorsque les méthodes de collecte des informations sur les participants de l'étude sont insuffisantes et que les erreurs de saisie ou les lacunes d'informations se répartissent inégalement entre les groupes d'examen. Quelques exemples: on questionne les patients du groupe «cas» directement et les témoins seulement par téléphone, ou alors on relève les informations sur les cas dans leurs dossiers médicaux et celles sur les témoins par renseignement oral.

Le biais d'information peut également se manifester si la présence d'une maladie précise amène l'enquêteur à insister davantage sur l'un des facteurs de risque possibles lors de sa recherche (*exposure suspicion bias*, ou «biais de suspicion d'exposition»), par exemple en procédant à une anamnèse approfondie sur l'alcool – ou inversement, si la connaissance d'une exposition antérieure ou d'un résultat d'examen l'incite à rechercher avec une plus grande intensité les signes d'une maladie (*diagnostic suspicion bias*, ou «biais de suspicion de diagnostic»), par exemple en effectuant une investigation plus poussée du carcinome de la prostate en cas de valeurs de PSA plus élevées.

C'est pourquoi vous devez vous poser d'abord les questions suivantes: a-t-on relevé les observations de la

**Il faut se demander si le relevé de données s'est effectué de la même façon chez tous les participants à l'étude**

même manière chez tous les participants? L'enquêteur a-t-il eu connaissance au préalable de l'appartenance des sujets examinés aux différents groupes d'étude resp. de leur statut d'exposition? Les personnes impliquées dans la saisie des données (enquêtes, évaluation des dossiers médicaux) doivent absolument ignorer la répartition en groupes et l'exposition antérieure de la personne examinée.

#### Existe-t-il d'autres facteurs ou d'autres comportements susceptibles d'influencer la caractéristique étudiée?

Une autre forme du biais d'information, le «biais de mémorisation», ou *recall bias*, apparaît essentiellement dans les études de

cas-témoins. Par exemple, les femmes ayant mis au monde un enfant handicapé se préoccupent davantage des causes possibles, et se souviennent ainsi plus facilement des médicaments pris durant la grossesse que les témoins ayant mis au monde un enfant en bonne santé [13]. Il en est de même des patients chez qui l'on a diagnostiqué un cancer: ils sont plus fortement conscients d'une anamnèse familiale positive que les témoins en bonne santé [14].

Vous devez donc, en lisant une étude, être attentif au fait que les informations sur les participants ont été consolidées par un questionnement des proches ou un dossier médical établi ou sont de nature purement subjective.

### Biais de confusion – quels sont les facteurs de perturbation?

Le biais de confusion (*confounding*) peut affecter tous les types d'études d'observation. Il apparaît lorsque la caractéristique étudiée (par ex. un facteur de risque ou un traitement) est associée à un facteur de perturbation complémentaire et que ce dernier fait croire à l'existence d'une relation de cause à effet entre la caractéristique étudiée et la maladie. Exemple: une étude atteste un quotient intellectuel plus élevé aux propriétaires de chats qu'aux propriétaires de chiens. Loin de démontrer l'existence d'une relation causale, cette différence résulte d'un facteur de perturbation: le temps libre, resp. la possibilité de planifier sa vie privée. Les personnes ayant une meilleure formation ont fréquemment un horaire de travail plus irrégulier; par conséquent, elles ne disposent pas du temps nécessaire pour promener un chien et ont tendance à préférer un chat comme animal domestique.

Autre exemple en médecine clinique: les premières études sur les effets indésirables des contraceptifs oraux ont porté à croire à un risque fortement accru d'infarctus du myocarde. Il s'est révélé par la suite que cette association découle du taux plus élevé de tabagisme chez les femmes recourant aux contraceptifs hormonaux [15]. En lisant une étude d'observation, vous devez donc accorder une attention méticuleuse aux caractéristiques des participants. Les indications sur les maladies associées, les facteurs de risque, l'alcoolisme et le tabagisme, le poids corporel, l'anamnèse familiale, etc., peuvent révéler un risque de confusion.

A-t-on pris la peine de déterminer les facteurs de confusion éventuels avant de réaliser l'étude? A-t-on tenu compte des facteurs d'influence éventuels lors de l'interprétation des résultats? Un article de qualité devrait toujours préciser les limites de validité d'une étude et décrire clairement les possibilités de confusion. Et lorsque la pertinence d'un essai n'est pas discutée, il faut se montrer particulièrement vigilant.

### Digression sur les études de dépistage non randomisées

Les études non randomisées qui évaluent l'efficacité des mesures de dépistage présentent des biais de types très variés. Nous allons expliquer brièvement deux d'entre eux:

Si une mesure de dépistage, par exemple du cancer du sein ou du cancer du poumon, détermine la date à laquelle un diagnostic est posé et qu'aucun traitement n'influence le pronostic, l'analyse du temps écoulé entre le moment du diagnostic et celui du décès conclura néanmoins à un avantage de survie factice. Les personnes soumises au dépistage ont un temps d'avance (*lead time*) sur celles qui ne l'ont pas été, ce qui donne lieu au «biais d'avance au diagnostic» (*lead-time bias*). Un autre type de biais résulte du fait qu'il est fréquent que les tumeurs à croissance rapide, dont le pronostic est plus mauvais, se manifestent déjà sur le plan clinique entre deux examens de prévention. Le dépistage a donc tendance à sélectionner les cancers dont la croissance est plus lente et le pronostic plus favorable. On nomme «biais de sélection des cas de meilleur pronostic» (*length-time bias*) l'effet de l'exclusion des tumeurs plus agressives, consistant à entraîner un prolongement factice de l'espérance de survie chez les patients faisant l'objet d'un dépistage. Il correspond au biais de Neyman évoqué plus haut à propos des études de cas-témoins.

L'évaluation correcte d'une mesure de dépistage exige que la distribution au sein des différents groupes soit randomisée et que la période d'observation soit suffisamment longue pour permettre d'observer l'évolution des malades et de déterminer la mortalité d'ensemble en tant que critère décisif. Les études non randomisées ne se prêtent pas à l'évaluation d'un dépistage.

### Biais du sujet immortel – pourquoi les acteurs nominés à l'Oscar ne vivent pas plus longtemps que les autres acteurs

Ce type de biais représente la combinaison d'un biais de sélection et d'un biais d'évaluation. Le «biais du sujet immortel» (*immortal-time bias*) survient lorsque c'est le plan même d'une étude qui implique l'existence d'un intervalle de temps pendant lequel un événement clinique ou un décès ne peuvent pas intervenir. Prenons l'exemple d'une étude de cohorte évaluant l'impact des corticoïdes sur la réduction du taux de mortalité chez les patients asthmatiques: les participants à l'étude doivent nécessairement survivre pendant la période allant

de la pose de diagnostic au début du traitement, ce qui revient à dire qu'ils sont définis comme «immortels» pendant cet intervalle. Supposons à présent que l'éva-

#### Les participants à l'étude sont-ils soumis à des risques concurrents?

luation du groupe de contrôle se fasse sur la base d'intervalles commençant au moment du diagnostic et tienne ainsi compte des décès précoces: cette différence d'approche lors de l'évaluation donnerait l'impression que les corticoïdes diminuent le risque de mortalité même s'ils n'avaient aucun effet. On trouve d'autres exemples d'études d'observation montrant que le biais du sujet immortel peut aller jusqu'à inverser un résultat: un risque semble diminuer alors qu'en réalité il augmente [16].

Autre exemple: une étude de cohorte rétrospective, publiée en 2001, a avancé l'hypothèse que les acteurs nominés à l'Oscar vivent plus longtemps que ceux qui ne le sont pas [17]. Mais une contre-étude publiée en 2006 a pu montrer que

cet avantage était entièrement dû au biais du sujet immortel [18], car les acteurs nominés

étaient considérés comme «immortels» jusqu'à la remise du prix (c'est-à-dire jusqu'à l'âge de 79 ans pour deux d'entre eux), alors qu'aucun âge minimal de survie n'était défini pour le groupe de contrôle.

Il est facile d'éviter l'apparition du biais du sujet immortel par une approche d'évaluation appropriée (voir [19]). Dans les études de cohortes, il faut s'assurer que les périodes d'observation des différents groupes soient effectivement similaires, et déterminer si les plans d'étude impliquent l'existence d'intervalles pendant lesquels aucun événement clinique ne peut survenir.

#### Correspondance:

Dr Peter Kleist  
GlaxoSmithKline AG  
Talstrasse 3-5  
CH-3053 Münchenbuchsee  
[peter.m.kleist@gsk.com](mailto:peter.m.kleist@gsk.com)

#### Comment peut-on expliquer ces écarts?

## Conclusions

Les études randomisées contrôlées continueront à représenter le *gold standard* dans la plupart des questionnements.

La fréquente difficulté à évaluer l'ampleur d'un biais diminue fortement la validité interne et par conséquent l'importance des études d'observation. Pour neutraliser les facteurs de confusion au moyen de méthodes statistiques appropriées (par ex. la restriction ou la stratification), il faut les identifier avant le début de l'étude et relever les données correspondantes [20]. Les facteurs de confusion inconnus représentent un problème inhérent aux études d'observation, et les biais de sélection et d'information réduisent la valeur d'une étude de façon définitive.

Par ces explications, j'espère avoir réussi à vous sensibiliser et que vous serez désormais en mesure de poser les bonnes questions en rapport avec la réalisation et les résultats des études d'observation, et d'en évaluer la qualité.

Les études d'observation peuvent servir à émettre un signal ou une hypothèse et à accroître les connaissances sans exiger d'intervention, mais elles ne peuvent en aucun cas remplacer les études randomisées contrôlées lorsque ces dernières sont réalisables.

#### Références recommandées

- Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*. 2002;359:248-52.
- Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health*. 2004;58:635-41.
- Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. *Nephron Clin Pract*. 2010;115:c94-9.
- Earle C, Hébert PC. A reader's guide to the evaluation of screening studies. *Postgrad Med J*. 1996;72:77-83.
- Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167:492-9.

Vous trouverez la liste des références en ligne sous [www.medicalforum.ch](http://www.medicalforum.ch) en annexe à l'article.

# Bias in Beobachtungsstudien / Les biais dans les études d'observation

## Weiterführende Literatur (Online-Version) / Références complémentaires (online version)

- 1 Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000; 342: 1878–86.
- 2 Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000; 342: 1887–92.
- 3 Pocock SJ, Elbourne DR. Randomized trials or observational tribulations? Editorial. *N Engl J Med* 2000; 342: 1907–9.
- 4 Ioannidis JPA, Haidich A-B, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001; 286: 821–30.
- 5 Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. *Nephron Clin Pract* 2010; 115: c94–c99.
- 6 Ganguli M, Lytle ME, Reynolds MD, Dodge HH. Random versus volunteer selection for a community-based study. *J Gerontol A Biol Sci Med Sci* 1998; 53: M39–M46.
- 7 Seltzer CC, Bosse R, Garvey AJ. Mail survey response by smoking status. *Am J Epidemiol* 1974; 100: 453–7.
- 8 Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002; 359: 431–4.
- 9 Moss AR, Osmond D, Bacchetti P, et al. Risk factors for AIDS and HIV seropositivity in homosexual men. *Am J Epidemiol* 1987; 125: 1035–47.
- 10 Bender MM, Levy AS, Schlucker RE, et al. Trends in prevalence and magnitude of vitamin and mineral supplement usage and correlation with health status. *J Am Diet Ass* 1992; 92: 1096–101.
- 11 Nelson HD, Humphrey LL, Nygren P, et al. Postmenopausal hormone replacement therapy: scientific review. *JAMA* 2002; 288: 872–81.
- 12 Neyman J. Statistics: servant of all sciences. *Science* 1955; 122: 401–6.
- 13 Werler MM, Pober BR, Nelson K, Holmes LB. Reporting accuracy among mothers of malformed or nonmalformed infants. *Am J Epidemiol* 1989; 129: 415–21.
- 14 Floderus B, Barlow L, Mack TM. Recall bias in subjective reports of familial cancer. *Epidemiol* 1990; 1: 318–21.
- 15 Jain AK. Cigarette smoking, use of oral contraceptives, and myocardial infarction. *Am J Obstet Gynecol* 1976; 126: 301–7.
- 16 Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010; 340: 907–11.
- 17 Redelmeier DA, Singh SM. Survival in academy award-winning actors and actresses. *Ann Intern Med* 2001; 134: 955–62.
- 18 Sylvestre M-P, Huszti E, Hanley JA. Do Oscar winners live longer than less successful peers? A reanalysis of the evidence. *Ann Intern Med* 2006; 145: 361–3.
- 19 Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2008; 167: 492–9.
- 20 Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; 359: 248–52.