

Biais et Facteurs de Confusion

Delphine Magnin, Philippe Vanhems

Département d'Hygiène, Epidémiologie et Prévention, Hôpital Edouard Herriot, Lyon

Introduction

Les études épidémiologiques, qu'elles soient descriptives, étiologiques ou évaluatives effectuées en hygiène hospitalière produisent des résultats concernant entre autres des facteurs de risques ou des résultats d'investigation dans un but d'intervention sur le terrain. Il est nécessaire d'acquiescer un réflexe de prudence et même de critique avant d'envisager ces applications. Tout professionnel doit évoquer les facteurs qui peuvent avoir altéré la qualité des résultats des investigations avant de considérer ceux-ci comme acquis. Au moins trois points doivent être discutés à cet effet, il s'agit de l'impact des biais, la prise en compte des facteurs de confusion et la puissance statistique de l'étude. Ce dernier point ne sera pas traité dans cette fiche méthodologique.

Biais

Il s'agit d'erreurs systématiques survenues au cours de l'étude liées à l'investigateur, aux participants, ou au questionnaire par exemple, qui va induire une estimation d'un paramètre (moyenne, risque) qui va différer systématiquement de la valeur réelle, par excès ou par défaut. Un biais diffère d'une erreur aléatoire qui est un concept probabiliste et qui est distribuée de manière identique dans les différents groupes étudiés. Une erreur aléatoire entraîne seulement un manque de précision.

Biais de sélection ou d'échantillonnage

Ce biais survient lorsque l'échantillon diffère systématiquement de la population dont il est supposé provenir. Les données recueillies peuvent différer de celles de la population cible et les résultats peuvent être erronés. Pour apprécier la qualité de la représentativité de l'échantillon, il est possible de comparer la répartition en son sein de certaines variables (âge, sexe, type de structure...), à celle de la population. Ceci impose de connaître les caractéristiques de la population cible au même moment. Ce biais se retrouve à la fois dans des études descriptives et des études étiologiques. Dans ce dernier cas, ils affectent

plus facilement les études cas-témoins.

Ce biais peut provenir **de la méthode de constitution des échantillons** qui ne sont alors pas représentatifs de la population. Par exemple, une étude sur la prévalence des accidents d'exposition au sang chez le personnel soignant effectuée en été peut ne pas être représentatif de la population des soignants de l'hôpital en question en raison des congés annuels et du nombre important de remplaçants engagés à cette période comme des étudiants par exemple.

Biais de recrutement

Ce biais existe à chaque fois que la probabilité d'inclusion des sujets dans l'étude est liée à un des facteurs étudiés. Ce biais s'observe surtout lorsque le recrutement s'effectue dans les institutions de soins. En effet, lors d'une étude cas-témoins, les cas ne sont pas toujours représentatifs de l'ensemble des cas atteints de la maladie, puisque les patients hospitalisés présentent souvent un stade plus sévère de la maladie. De plus, par commodité, on inclut souvent des témoins issus du même établissement que les cas. La représentativité par rapport à la population générale n'est alors pas forcément assurée et ce point doit être discuté au moment de la définition des témoins. Il est donc nécessaire de s'assurer que les critères de sélection utilisés ne vont pas définir un groupe particulier qui présenterait une caractéristique commune qui les différencierait de l'ensemble de la population.

Biais d'autosélection

Ce biais apparaît lorsque les sujets entrent dans l'étude en fonction d'une décision qu'ils prennent eux même et qui peut être liée aux phénomènes étudiés (mauvais état de santé, exposition au facteur étudié...).

Biais lié aux perdus de vue

C'est un des biais principal des études longitudinales où les patients sortent de la cohorte durant le suivi avant la survenue de l'évènement étudié. Il faut s'efforcer d'obtenir des informations sur ces individus afin de les retrouver et de les inclure si possible à nou-

veau dans l'étude. Il sera nécessaire de comparer leurs caractéristiques à celles des sujets restés dans l'étude afin de corriger ou d'estimer un biais potentiel. Des évènements extrêmes peuvent se rencontrer. Par exemple, ces patients peuvent se sentir en pleine forme et refuser de se soumettre aux examens répétés afin d'évaluer leur santé. Ces patients perdus de vue peuvent aussi être décédés sans que l'investigateur le sache. Cette éventualité est importante car le décès peut-être lié à la maladie étudiée.

Biais de détection ou d'information

Ce biais est lié à la subjectivité de l'enquêteur lorsqu'il existe un lien éventuel entre l'exposition au facteur et l'inclusion des cas. Ainsi, par exemple on risque d'inclure plus volontiers dans le groupe des cas un sujet de réanimation porteur d'une pneumopathie si il était intubé. Ou alors, l'investigateur peut-être tenté de mieux documenter les expositions aux risques chez les cas que chez les témoins.

Une solution réside dans le travail en l'aveugle d'une part, c'est-à-dire que les enquêteurs ignorent si le patient est exposé ou non (ou malades ou non malade), et d'autre part la réalisation d'un questionnaire rigoureusement standardisé. Dans de nombreux cas, l'exposition au risque est évidente et l'enquêteur ne pourra pas collecter pas les données en « aveugle ».

Biais de survie sélective

Ce biais se retrouve dans les études de cohortes. Il s'agit des perdus de vue qui sont plus nombreux chez les exposés, du fait des conséquences de l'affection survenue.

Biais des non-réponses

Les patients qui ne répondent pas à une sollicitation de participation à un étude diffèrent souvent des répondants. Ce refus de participation peut traduire une moins bonne motivation pour le problème. Par exemple, les personnes non vaccinées contre la grippe participeront moins volontiers à une enquête définissant la couverture vaccinale anti-grippale. Toutefois d'autres raisons peuvent être avancées telles

que l'incapacité de répondre, le changement de domicile, les personnes décédées, etc.

On tâchera si possible de réunir quelques informations concernant ces non-répondants afin de pouvoir décrire leur caractéristiques essentielles et de les comparer à celles des répondants. Ceci peut permettre d'estimer l'ampleur du biais.

Biais de mesure

Dans une enquête descriptive, le constat d'une répartition spatiale particulière d'un indicateur comme la prévalence d'une infection par exemple, doit faire l'objet d'une analyse critique à la recherche d'erreurs susceptibles d'expliquer à elles seules une telle répartition avant de lui rechercher une cause précise.

Dans une enquête étiologique, on mesure l'association entre l'exposition à un facteur (par exemple l'intubation ou le cathéter veineux central) et la présence ou la survenue d'une infection nosocomiale (pneumopathie ou infection sur cathéter). Le biais réside dans le recueil de ces données, conduisant à classer une personne comme exposée à un facteur ou atteinte du problème alors qu'elle ne l'est pas (ou inversement). Ces erreurs peuvent être distribuées au hasard et on les retrouve alors dans les différents groupes. Par exemple, un instrument de mesure de mauvaise qualité donnera une valeur erronée dans les deux groupes. Il s'agit d'erreurs non différentielles. Dans ce cas, l'impact de l'erreur peut être évalué et peut ne présenter qu'une altération marginale de la validité de l'étude. Lorsque ces erreurs sont dites différentielles, c'est-à-dire qu'elles altèrent d'une manière différente les groupes d'individus, la validité de l'étude peut-être entièrement remise en cause car les estimations de risque sont biaisées.

Les différentes origines des erreurs différentielles sont les suivantes:

Biais de mémorisation plutôt retrouvés dans les enquêtes cas-témoins. Un cas se souviendra plus facilement d'une exposition ancienne, car il se sent plus concerné qu'un témoin. Inversement, il aura tendance à minimiser une exposition s'il se sent « coupable » de son état actuel. Il n'y a pas de franche solution à ce problème si ce n'est d'essayer de minimiser les conséquences grâce à des questions habilement posées. Par exemple, un deuxième groupe de témoins pourra être inclus. Ce groupe sera constitué de patients atteints d'une pathologie particulière, non informative pour l'étude, mais il est probable que ces patients feront un effort identique au cas pour se souvenir d'une exposition éventuelle.

Biais de classement où l'existence d'un problème de santé rend plus efficace la recherche d'une étiologie possible (étude cas-témoins), ou inversement, l'existence d'une exposition particulière rend le sujet plus attentif à la survenue de problèmes de santé (étude de cohorte). Ainsi, on recherchera plus volontiers une pneumopathie nosocomiale chez un patient intubé depuis quelques semaines plutôt que chez un patient en ventilation spontanée. Cependant, ce biais est facile à éviter, il suffit de faire examiner les patients ou les dossiers médicaux à

l'aveugle par des médecins ignorant le statut du patient vis-à-vis de l'exposition.

En général, les biais sont plus nombreux dans les études cas-témoins et c'est la raison pour laquelle une étude de cohorte sera à privilégier si le contexte s'y prête.

Au total, il est parfois difficile de contrôler un biais, aussi celui doit être anticipé et prévenu au moment de la conceptualisation de l'étude.

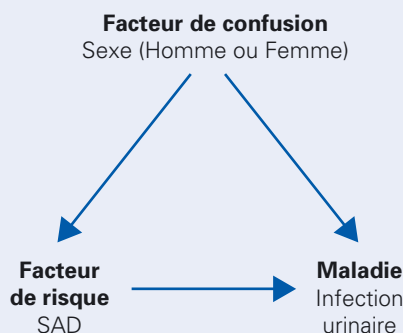
Facteurs de confusion

L'étude de la relation statistique entre un facteur et une affection ne se limite pas à tester et quantifier l'association, et il faut essentiel de tenir compte d'autres facteurs.

On parle de **facteur de confusion** pour tout facteur lié à la fois à l'affection (tous ses facteurs de risques sont donc concernés) et au facteur étudié. Un facteur de confusion sera lié à un facteur de risque indépendamment de la maladie et ce même facteur de confusion sera lié à une maladie indépendamment du facteur de risque.

Cette situation tend à « déformer » la relation réelle entre l'affection et le facteur étudié. Ces facteurs de confusion doivent être maîtrisés, d'autant plus que meilleures sont les connaissances sur l'étiologies et les cause des maladies, et des infections en particulier, plus les potentiels facteurs de confusion lors d'une étude sont nombreux.

Un exemple simple de facteur de confusion dans le domaine de l'hygiène hospitalière réside dans le l'association entre le risque d'infection urinaire (la maladie) et la sonde urinaire à demeure [SAD] (le facteur de risque) en tenant compte du sexe du patient (facteur de confusion). Le lien qui existe entre infection urinaire et SAD est bien connu ; par ailleurs il existe une association entre le risque d'infection urinaire et le sexe (les femmes sont plus souvent exposées au risque d'infection urinaire que les hommes) indépendamment du sondage, et par ailleurs les indications et les conditions de pose de SAD sont différentes chez les hommes et chez les femmes.



Leur contrôle des facteurs de confusion s'effectue le plus souvent par la réalisation d'un appariement lors de la préparation de l'étude et d'un ajustement lors de l'analyse des données.

L'appariement est un procédé de constitution des échantillons, destiné à neutraliser l'influence potentielle d'un ou plusieurs facteurs de confusion, en les répartissant de

manière équilibré entre les groupes. Il existe **l'appariement individuel** où les personnes sont recrutées par paires (une personne de chaque groupe ayant la même caractéristique pour le facteur d'appariement) ou par triplet s'il existe un cas pour deux témoins, et **l'appariement de fréquence ou par groupe** où le même pourcentage de sujets dans chaque groupe présente la caractéristique du facteur de confusion. Il s'agit dans ce cas d'effectuer une stratification sur les variables de confusion. Mais il est essentiel de retenir qu'une fois l'appariement effectué sur un facteur, on ne peut plus étudier son rôle sur la survenue du problème de santé car par définition sa distribution sera identique chez les cas et les témoins, il faut donc se limiter aux seuls facteurs de confusion connus.

L'ajustement repose sur l'analyse de la relation entre le facteur de risque et la maladie après stratification sur le facteur de confusion. On détermine un risque relatif (RR) ou un Odd-ratio (OR) pour chaque strate du facteur de confusion. On obtient ainsi une estimation non biaisée de la mesure d'association entre l'affection et son facteur de risque potentiel pour chaque strate. On utilise le test du chi-deux ajusté de Mantel-Haenszel afin de déterminer l'association entre le facteur de risque et la maladie après ajustement sur le facteur de confusion.

Pour mettre évidence l'ampleur de cette association, on détermine alors un RR ou un OR ajusté afin de rendre compte de la taille de l'effet du facteur de risque, toutes strates comprises.

Lorsqu'il existe plusieurs facteurs de confusion à prendre en compte, on utilise des méthodes statistiques permettant de construire des modèles multivariés qui ont la capacité de tenir compte de manière simultanée de l'effet de plusieurs facteurs de confusion. Les plus couramment employés sont la régression logistique ou le modèle des risques proportionnels de Cox.

Bibliographie

Ouvrages de référence

CZERNICHOV P, CHAPERON J, LE COUTOUR X. Epidémiologie, connaissances et pratiques. Ed Masson, 2001: 206-209/ 217-229.

GOLBERG M, L'épidémiologie sans peine. Ed Frison-Roche, 1990: 152-156.

RUMEAU-ROUQUETTE C, BLONDEL B, KAMINSKI M, BREART G. Épidémiologie, méthodes et pratiques. Ed Flammarion, 1997: 189-199.

BRÜCKER G, Infections nosocomiales et environnement hospitalier. Ed Flammarion, 1998.