

Les biais en épidémiologie

Bernard BRANGER

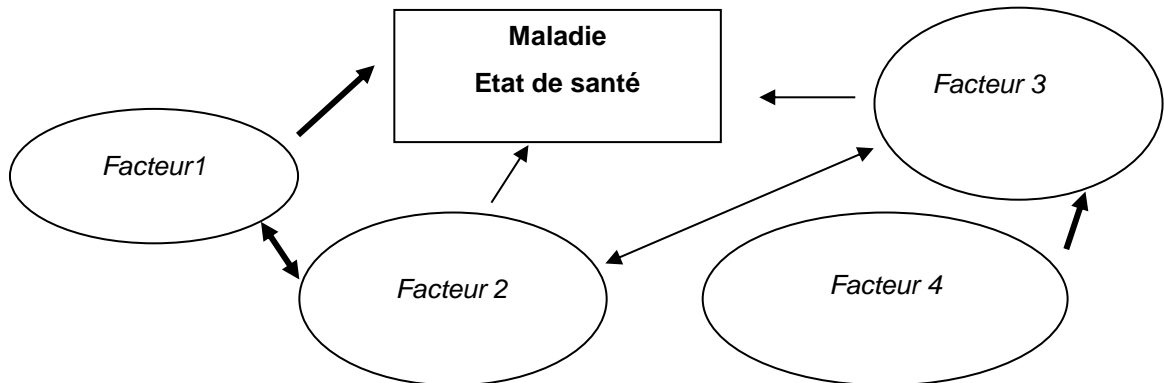
Epi-Sûre

(Association de conseil en épidémiologie et statistiques)

11 bis, rue Gabriel Luneau - 44000 NANTES

Mail : branger44@gmail.com - Tél 06 32 70 33 80

NB : Utilisation libre; merci de citer la référence sur le Web (<https://www.epi-sure.com/>)



I. Liens possibles entre une maladie M et une exposition E

Si l'analyse statistique permet de conclure qu'il y a un lien entre une exposition et une maladie, la conclusion sur la causalité de ce lien ne s'impose pas, car plusieurs possibilités de liens existent.

1. On croit $E \rightarrow M$, alors que $M \rightarrow E$!

Par exemple, on veut étudier si la présence d'un obstétricien à la naissance (ici E) donne de meilleurs résultats qu'en son absence (dans une enquête d'observation). On observe, en fait, que la présence de l'obstétricien augmente la mortalité et les séquelles de l'enfant (ici M) ! La raison en est que, lorsque la naissance se passe mal (prédisposant à M), on appelle l'obstétricien (E), alors que, lorsque tout se passe bien, on ne l'appelle pas. Donc la mortalité et les séquelles (M) ne sont pas la conséquence de la présence de l'obstétricien (E) mais la cause.

Il en est peut-être de même du lien entre « prendre froid » (E) et apparition d'un syndrome grippal (M). Il est possible que les premiers frissons, assimilés à un coup de froid, soient le signe que la température augmente et donc que E n'est pas la cause, mais la conséquence de la maladie.

2. E et M évoluent ensemble sans lien

On a observé aux Etats-Unis un lien entre la baisse de la pratique religieuse et la baisse de la tension artérielle. Il s'agit sans doute d'une évolution temporelle sans lien causal. Il en est de même du lien entre la durée du cycle lunaire (28 – 29 jours) et la durée des cycles menstruels chez la femme. On pourrait dire la même chose si on compare le taux d'incidence du SIDA en France de 1983 à 1990 avec le taux de chômage...

3. E et M sont liés mais c'est un autre facteur qui est cause de E et de M

Le taux de transfert des nouveau-nés (M) dans un service de nouveau-nés est lié avec la pratique de la césarienne (E). En fait, ce n'est pas la césarienne elle-même qui est facteur de l'état de santé du bébé (qui entraîne le transfert), mais c'est le fait que l'on pratique une césarienne quand l'enfant va mal qui conditionne le transfert. D'où, si l'état du bébé est nommé B, on peut écrire $E \rightarrow B$ et $B \rightarrow M$, mais il n'y a pas de lien direct entre E et M.

4. $E \rightarrow X \rightarrow M$

On peut mettre en évidence un lien entre alcool et cancer du poumon, mais il n'y a pas de lien causal direct. On sait en effet que les alcooliques sont plus fumeurs. En conséquence, « alcool » (E) est lié à « tabac » (X) qui est lié au cancer du poumon.

5. E n'est pas une cause, mais une condition favorable à M

Norbert Elias rapporte qu'au Moyen Age les médecins déconseillaient de prendre des bains (E) car cette pratique était liée à l'apparition de maladies (M). En fait, les bains étant publics, la promiscuité était un facteur de transmission de maladie, et donc c'est le caractère public du bain qui est facteur de maladie et non le caractère « bain » seulement.

De même, il existe un lien entre position de sommeil du nourrisson (sur le ventre) et la mort subite. Ce n'est pas une cause directe mais une prédisposition au risque d'asphyxie en cas de problème chez le nourrisson.

6. $E \rightarrow M \rightarrow E$

De manière générale, les vaccinations protègent des maladies. Il en est ainsi pour la vaccination contre la rubéole qui fait baisser l'incidence de la rubéole. Cependant, on observe que la vaccination de 60 % d'une population augmente le taux d'incidence des rubéoles des femmes enceintes (avec risque de malformations des nouveau-nés). En fait, un taux de couverture de 60 % n'empêche pas le virus de circuler, mais retarde le moment de la primo-infection rubéolique qui de l'enfance auparavant survient chez de jeunes adultes non vaccinés ou vaccinés non-répondeurs. D'où, l'exposition à la vaccination (E) entraîne une baisse de la maladie (M) chez les enfants, mais que cette baisse de la maladie augmente, chez les non-vaccinés ou les vaccinés non-répondeurs à un âge plus tardif, la fréquence du contact avec le virus.

7. $E \rightarrow M$

C'est le cas que l'on retient après avoir éliminé tous les autres et passé le raisonnement selon les critères de Hill. On parle de critère de causalité.

II. Biais de mesure

Mesurer, c'est donner une valeur par rapport à une valeur de référence (le plus souvent arbitraire). Un biais de mesure consiste à donner une valeur fautive par rapport à ce système.

Les biais de mesure peuvent concerner :

1. La précision

Les biais de précision dépendent de l'objet à mesurer ou de la méthode de mesure (balance non précise pour le poids, appareil à tension non contrôlé). Des données comme la durée de l'accouchement ne sont pas précises en raison de la difficulté de préciser le début. Idem en cas de diarrhée. Discuter de l'ouverture du col de l'utérus au cours de l'accouchement.

2. La validité

Une mesure censée mesurer ce qui est prévu de mesurer. L'obésité ne se mesure pas que par le poids, car il dépend de la taille. L'indice de masse corporelle (IMC ou BMI en anglais) est censé mieux représenter l'augmentation de la masse grasse (P/T^2). On peut essayer d'autres mesures à valider (P/T^3)...

3. La pertinence

Plusieurs critères permettent de déterminer la pertinence d'une mesure.

- ✓ Est-ce que la mesure correspond à l'état du patient, à l'état biologique que l'on veut mesurer ?
- ✓ Est-ce que les variations de la mesure reflètent les variations de l'état biologique ?
- ✓ Est-ce que le critère apporte une information nouvelle / à une autre variable utilisée antérieurement ?

Exemple : densité osseuse, dosage des globules blancs, **score composite** de plusieurs variables, score de gravité des patients. Pertinence statistique.

4. La reproductibilité

Les variables sont mesurées soit par la même personne, mais à des moments différents, soit par des personnes en des lieux et des moments différents. C'est le principal biais de mesure dans les grandes enquêtes. On distingue : la reproductibilité intra-observateur et la reproductibilité inter-observateur. Tests statistiques : test de concordance kappa pour les variables qualitatives, coefficient de corrélation, différence des moyennes, analyse de variance, coefficient de variation pour les variables quantitatives.

5. La connaissance de l'appartenance

Dans les enquêtes d'observation, le patient connaît son exposition (il sait qu'il fume..) ; ses réponses peuvent être biaisées avec le risque de minimiser sa consommation ou les conséquences sur sa santé. De même, un médecin, un enquêteur, s'il sait à quel type d'exposition appartient la personne, ou s'il sait si elle reçoit le traitement ou non. Ce sont des biais difficiles à lever.

Dans les enquêtes d'expérimentation (médicaments), ce type de biais est levé par le « double bind » ou « double aveugle » et le tirage au sort des ordres de traitement. Certains parlent même de triple aveugle où le statisticien ignore qui du groupe A ou B est traité.

6. L'effet « réverbère »

« On cherche là où on recueille l'information facilement. Les variables d'exposition sont recueillies dans la mesure où on peut les recueillir ou bien si elles sont faciles à recueillir. Exemple dans les infections de site opératoire : les variables recueillies (âge, sexe, type d'intervention, durée d'intervention, classe de contamination, score ASA du patient, urgence) sont peu modifiables et expliquent une petite partie des infections. Il est probable que la préparation de l'opéré, l'antisepsie, les conditions d'hygiène à la fermeture de la paroi... soient plus importants et il est certain qu'ils sont modifiables (prévention).

7. Biais de mémoire

Surtout dans les enquêtes rétrospectives. Exemple dans les enquêtes sur les malformations à la naissance et le lien avec les médicaments pris pendant la grossesse.

III. Biais de sélection des sujets

A. Généralités

Dans les enquêtes, on travaille sur des échantillons de patients tirés d'une population (dite aussi population cible). Toute la difficulté des enquêtes est de sélectionner les sujets de l'échantillon tels que l'échantillon ressemble plus possible à la population, ou, du moins, tels que, s'ils diffèrent, on sait en quoi ils diffèrent.

Exemple : envoi d'un questionnaire aux médecins d'un département sur les formations qu'ils ont suivies dans une année. Rechercher les biais de sélection possibles ; comment les contrôler.

B. Selon le type d'enquête

- Dans les enquêtes transversales : les biais assez constants sont le risque de sur-expositions chez les malades. Exemple en milieu hospitalier.
- Dans les cohortes, on risque le biais sur la maladie (les malades sont plus recherchés, ou les malades sont suivis ailleurs), ou biais sur l'exposition (les exposés sont plus retrouvés).
- Dans les enquêtes cas-témoins, le risque est que les témoins soient sur-exposés ou sous-exposés.

C. Biais plus généraux

- Biais de Berkson : sur-représentation des hospitalisés exposés.
- Biais de non-réponses (NR) : il est rare que les NR soient complètement aléatoires. Plusieurs situations : les NR sont aléatoires, les NR sont liées à l'individu, les NR sont liées à la question, les NR sont liées aux deux.
- Biais « healthy worker effect » ou biais du travailleur en bonne santé, pour lequel les travailleurs exposés sont en meilleure santé du fait de leur résistance aux tâches les plus difficiles.
- Biais de suivi dans les cohortes : enquête en entreprise dans laquelle on suit les pathologies. Risque de sortie de l'entreprise des plus malades !

D. Le choix des témoins en général

Exemple : étude d'un lien entre tabagisme des parents et bronchiolite des nourrissons. Cas = nourrissons hospitalisés âgés de 1 mois à 2 ans avec une bronchiolite. Où prendre les témoins ?

IV. Biais de classement

Le biais concerne des erreurs de classement des malades/ non-malades et les erreurs de classement de l'exposition.

Deux exemples simples

- ✓ cancer du poumon versus non-cancer → être sûr que les témoins traités pour autre chose ou non-malades n'aient pas le cancer du poumon par la réalisation d'une radiographie pulmonaire
- ✓ fumeurs et non-fumeurs : nombre de cigarettes comme seuil, anciens fumeurs, fumeurs actuels...

Il s'agit en fait de bien connaître la valeur des signes en fonction de leur sensibilité et de leur spécificité.

On distingue :

→ **Les biais différentiels** qui affectent soit l'exposition seule, soit la maladie seule. Les critères d'association comme RR ou OR peuvent être sous-estimés, soit sur-estimés.

→ **Les biais non différentiels** qui affectent et l'exposition et la maladie : les RR ou OR sont toujours sous-estimés (tendance vers 1).

V. Biais de l'étude elle-même

A. Biais de codification et d'enregistrement

Il s'agit d'erreurs d'enregistrement ou de saisie. Biais fréquent qui nécessite un temps souvent important de validation des données. Certaines données peuvent être validées sur le questionnaire lui-même (sexe des personnes et antécédents de grossesse par exemple, ou cancer de la prostate chez des enfants..). D'autres nécessitent un retour aux données des dossiers ; d'autres ne sont pas repérables. Problème de la double saisie : logiciels et coûts en temps-personnes. Déclarations trompeuses ou déclarations biaisées (la réponse « oui » est plus valorisante dans certaines circonstances.

B. Biais de déclarations des sujets

Biais de compréhension, biais de mémoire. Questions trop intimes ; consommation d'alcool, de drogues, questions sexuelles. Exemple : la taille déclarée par les personnes est en moyenne de 1.02 cm supérieure à la réalité (mesure avec une toise) et le poids de 0.65 kg inférieure à la réalité (mesure avec une balance), et donc l'IMC risque d'être sous-estimé de 0.2 à 0.5 !

C. Biais de non respect du protocole

D. Biais des perdus de vue

Attention, partir du principe que les perdus de vue ne sont pas dus au hasard, mais qu'ils ont un lien avec la maladie (des décès par exemple !) ou à l'exposition. Toujours vérifier, si on le peut, la comparaison, sur des critères d'inclusion, entre suivis et perdus de vue).

VI. Biais de confusion : le raisonnement analogique

Une tendance naturelle (culturelle) est de lier et de mettre en situation de causalité deux phénomènes apparus ou analysés en même temps. Dans le domaine de la santé, ce type de raisonnement a été longtemps utilisé en phytothérapie (le saule qui pousse près de l'eau peut servir à guérir les maladies attribuées à l'eau ; les plantes jaunes peuvent servir à l'ictère...).

Les médecines dites « parallèles » comme l'homéopathie se servent de ce raisonnement. C'est un fond culturel fort comme le lien entre le froid et les maladies, la vitamine C et les défenses, les dents et les maladies des enfants.... Dans le domaine météorologique, le raisonnement analogique est la base de raisonnement.

Ce qui est en jeu est l'existence de facteurs de confusion non pris en compte.... Le rôle du scientifique est de se prémunir de ce raisonnement et de s'en méfier a priori avant de l'étudier....

- facteur de confusion ou facteur confondant : facteur d'exposition associé à la maladie mais aussi associé à un autre facteur, connu ou non, recherché ou non, lequel est causal ou a une force d'association supérieure ou plus plausible scientifiquement.

- prévention de la « confusion » :

= appariement, stratification avant l'étude : on sépare l'analyse en fonction de catégories d'exposition.

= ajustement après l'étude : analyse bivariée, analyse multivariée avec une modélisation telle que $E(Y) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_{1k} X_k$. Les coefficients sont ajustés entre eux et lèvent les facteurs de confusion. Exemple : régression multiple, analyse discriminante, régression logistique, modèle de Cox. Modèle complet, modèle pas-à-pas...

➔ **Exemples de biais de confusion** : à New York, une enquête a été faite pour déterminer les **facteurs associés à la contamination VIH**. Le tableau suivant a été établi selon l'habitat dans un quartier (Q.)

Tableau I : Lien entre sérologie HIV et origine dans des quartiers

	HIV +	HIV -	Total
Q. défavorisés	300	150	450
Q. favorisés	150	300	450
Total	450	450	900

HIV+ dans les quartiers défavorisés est de $300/450 = 67\%$ et dans les quartiers favorisés de $150/450 = 33\%$. Il s'agissait d'une étude cas-témoins où, à un VIH +, on associait un VIH -. L'OR vaut ici $300 \cdot 300 / 150 \cdot 150 \approx 4$ (**3.03 – 5.28**). On pourrait penser que les quartiers défavorisés ont une susceptibilité plus grande au virus HIV ? En fait, il existe un tiers facteur, qui est la nature de la drogue utilisée : les Q. défavorisés utilisaient surtout l'héroïne (4 injections par jour), et les Q. favorisés la cocaïne (sans injection). On peut alors dresser le tableau suivant :

Tableau II : Lien entre sérologie HIV et origine dans des quartiers selon la nature de la drogue

<i>Héroïne</i>	HIV +	HIV -	Total
Q. défavorisés	275	25	300
Q. favorisés	125	25	150
Total	400	50	450

OR (Quartiers défavorisés de HIV +) = $275 \cdot 25 / 25 \cdot 125 = 2.19$ (1.20 – 3.99)

91.7 % HIV + versus 83.3 %

<i>Cocaïne</i>	HIV +	HIV -	Total
Q. défavorisés	25	125	150
Q. favorisés	25	275	300
Total	50	400	450

OR (Quartiers défavorisés de HIV +) = $25 \cdot 275 / 25 \cdot 125 = 2.19$ (1.20 – 3.99)

16.7 % versus 8.3 %

On voit que le risque est le même pour la cocaïne et l'héroïne vis-à-vis du quartier. Le quartier n'est pas un facteur causal, mais un facteur confondant; c'est la drogue qui est facteur causal (en première analyse).

Tableau III : Lien entre la sérologie HIV et la nature de la drogue

	HIV +	HIV -	Total
Héroïne	400	50	450
Cocaïne	50	400	450
Total	450	450	900

88.9 % versus 11.1 %

OR = $(400 \cdot 400) / (50 \cdot 50) = 64.0$ (42.2 – 96.9)

Tableau IV : Lien entre quartiers et nature de la drogue

	Q. défavorisés	Q. favorisés	Total
Héroïne	300	150	450
Cocaïne	150	300	450
Total	450	450	900

66.7 % versus 33.3 %

OR = $(300 \cdot 300) / (150 \cdot 150) \approx 4$ (3.03 – 5.28)

➔ **Exemples de biais de confusion : Taux d'infection de site opératoire après césarienne**

Tableau V : Infection nosocomiale après césarienne

Variables	Infection n=12	Pas infection n=137	RR	p
Age moyen (ans)	24.9 ± 3.9	29.2 ± 5.1	--	< 0.01
- Taux ≥ 35 ans	9.2 %		∞	NS
- Taux < 35 ans	0 %			
Urgence				
- Oui	10.0 %		3.9	NS
- Non	2.6 %		0.5 - 29.2	
Antibioprophylaxie				
- Oui	13.4 %		3.7	< 0.05
- Non	3.7 %		1.1 – 13.0	
NNISS				
= 0	9.2 %		2.7	NS
> 0	3.4 %		0.4 – 19.8	
Durée intervention (mn)	54 ± 25	52 ± 14	--	NS
- Durée > 45 mn	8.3 %		1.1	NS
- Durée < 45 mn	7.5 %		0.3 – 3.5	
Rupture membranes (h)	17 ± 33	6 ± 12		NS
> 6 heures	11.8 %		1.9	NS
≤ 6 heures	6.1 %		0.6 – 5.7	
Liquide teinté				
- Oui	13.0 %		1.8	NS
- Non	7.1 %		0.5 – 6.2	
Durée du travail (h)	6.5 ± 5.1	4.9 ± 5.3	--	NS
- Durée > 6 heures	12.3 %		2.3	
- Durée ≤ 6 heures	5.4 %		0.7 – 6.7	
Fièvre pendant le travail				
- Oui	0 %		0	NS
- Non	8.8 %			
IMC	28.1 ± 2.8	27.6 ± 4.8	--	NS
> 30	9.8 %		1.3	NS
≤ 30	7.4 %		0.4 – 4.1	
Sonde de tocolyse				
- Oui	12.8 %		2.0	NS
- Non	6.4 %		0.7 – 5.9	

Le seul facteur significatif est l'antibioprophylaxie, mais comme facteur aggravant ! En fait, l'antibioprophylaxie est très liée à des facteurs de risques d'infection et est de fait un facteur de confusion probable.

Tableau VI : Facteurs liés à l'antibioprophylaxie

Variabes	Taux ATBPRO	p
Urgence = oui	54.5 %	< 10 ⁻⁴
Urgence = non	17.9 %	
Rupture ≤ 12 h	40.9 %	< 10 ⁻³
Rupture > 12 h	68.2 %	
Utérus cicatriciel = oui	27.0 %	< 0.05
Utérus cicatriciel = non	50.9 %	
Sonde toco = oui	59.0 %	< 0.05
Sonde toco = non	40.0 %	
Oxymétrie = oui	73.1 %	< 0.01
Oxymétrie = non	38.0 %	

➔ **Exemples de risque de biais de confusion : tabagisme de la femme enceinte et poids de naissance**

Le poids moyen de naissance des enfants de fumeuses est plus faible de 200 g que celui des enfants de non-fumeuses. Ce n'est pas un argument pour dire que le tabac pendant la grossesse est CAUSAL du petit poids de naissance. Car les fumeuses sont très différentes des non-fumeuses : âge plus jeune, parité plus grande, âge gestationnel plus faible, taille plus petite, poids plus élevé, CSP différente, alimentation différente, antécédents de GEU et d'IVG plus fréquents, relations sexuelles

plus fréquentes avec des partenaires plus nombreux, niveau de scolarité plus faible, habitat différent... Alors, pourquoi se fixer sur le tabac ?

Pour pouvoir affirmer le rôle "propre" du tabac (ou facteur indépendant), il faut "tenir compte" (c'est-à-dire ajuster sur) des autres facteurs, plus une hypothèse confirmée en laboratoire (le tabac agit sur la perfusion placentaire ++).

➔ **Recherche de biais de confusion : l'obésité des enfants à 5-6 ans, rôle du tabagisme pendant la grossesse** (von Kries. Am J epidemiol 2002 ;156 :954-61). Munich.

Sur 6 483 enfants de 5 et 6 ans avec 3 847 non-fumeuses (59.3 %) et 2 636 fumeuses au jour de l'enquête (638 fumeuses (9.8 % du total) pendant la grossesse et 1998 non-fumeuses (30.8 % du total) au moment de la grossesse).

« Toutes les variables ont été considérées comme confondantes »... Analyse multivariée avec régression logistique : 4 facteurs restent significatifs = **tabagisme maternel (ORa = 2.06 ; 1.31 – 3.23)**, obésité d'un des parents (ORa = 4.55 ; 3.07 – 6.74) et poids de naissance > 90^{ème} perc. (ORa = 2.23 ; 1.27 – 3.90) et télévision ou gameboy > 1 heure/jour (ORa = 1.86 ; 1.22 – 2.79). Dose-effet.

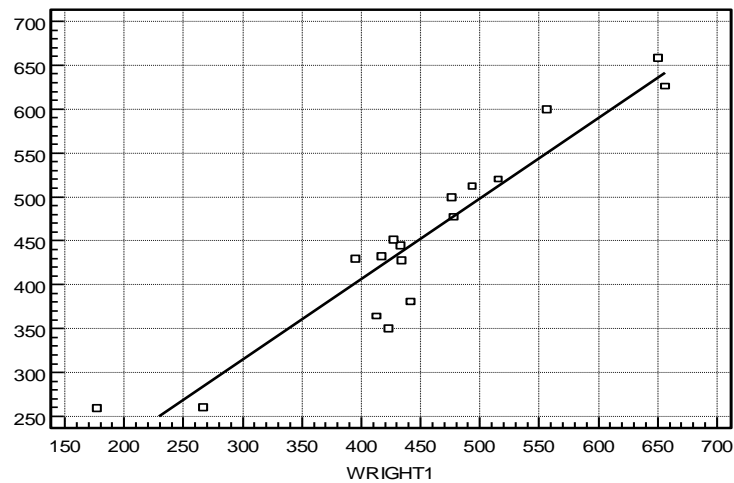
Tableau VII : Facteurs pouvant expliquer l'obésité à 5-6 ans, analyse univariée et multivariée

Variables	OR	IC 95 %	ORa	IC 95 %
Tabagisme maternel pdt G	2.96	1.97 – 4.46	2.06	1.31 - 3.23
Niveau études parents > 10 ans	0.48	0.36 – 0.65	0.68	0.45 - 1.03
IMC d'un parent > 30	4.51	3.31 – 6.15	4.55	3.07 - 6.74
Poids de naissance > 90 ^{ème} perc.	2.24	1.47 – 3.40	2.23	1.27 - 3.90
Parent seul	1.30	0.86 - 1.99		
Habitat concentré	1.16	0.77 - 1.76		
Télévision ou gameboy > 1 h/j	2.06	1.54 – 2.75	1.86	1.22 - 2.79
Sport en club	0.65	0.49 – 0.87		
Allaitement	0.55	0.41 - 0.73	0.91	0.60 - 1.38
Aliments solides introduits avant 4 mois	1.39	0.98 - 1.97		
Biberon avec sucre pour dormir	1.48	1.11 – 1.99		
Repas pris seul	1.60	1.02 – 2.49		
Repas chaud le soir	0.86	0.64 - 1.16		
Repas devant la télévision	1.58	1.15 – 2.18	1.35	0.84 - 2.15

8. Le piège du coefficient de corrélation

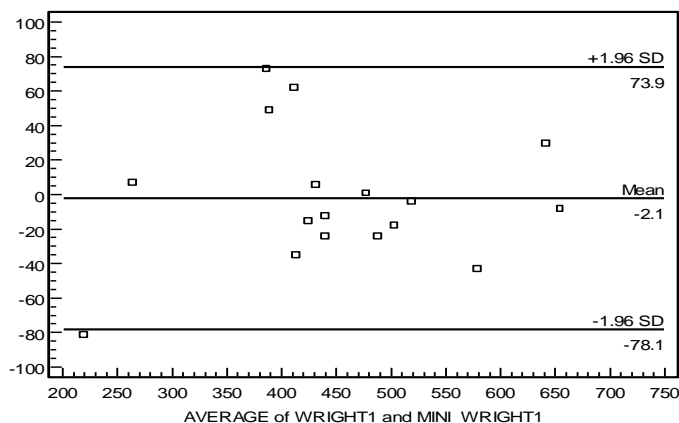
Deux méthodes de mesure du débit-expiratoire de pointe : les valeurs sont forcément corrélées avec un $r = 0.94$ ($p < 0.001$)

Figure 8 : Droite de régression entre une mesure du débit-mètre de pointe par un appareil comparé à un autre appareil.



Il faut utiliser **la méthode de Bland et Altman** telle que les écarts entre la méthode de référence et la nouvelle méthode soit mise en évidence. Voir Bland JM, Altman DG (1986) Statistical method for assessing agreement between two methods of clinical measurement. The Lancet, i, 307-310.

Figure 9 : Diagramme de Bland et Altman



- Exemples triviaux de facteurs de confusion

- ▶ Confusion de la cause et de la conséquence
 - ✓ 90 % des personnes meurent dans leur lit...
 - ✓ Les voitures rouges vont plus vite que les autres
- ▶ Deux phénomènes sans rapport entre les deux (rechercher un tiers facteur expliquant les deux indépendamment)
 - ✓ Liens entre nombre de divorces et nombre de salles de bains
 - ✓ Alcool et cancer du poumon
 - ✓ Césarienne et mortalité périnatale
- ▶ Deux phénomènes avec des rapports complexes entre eux deux
 - ✓ **Redoublement scolaire** (« élèves en retard ») et performances scolaires ultérieures.
Tiré de « Dossiers de l'enseignement scolaire » N°166 (mai 2005)

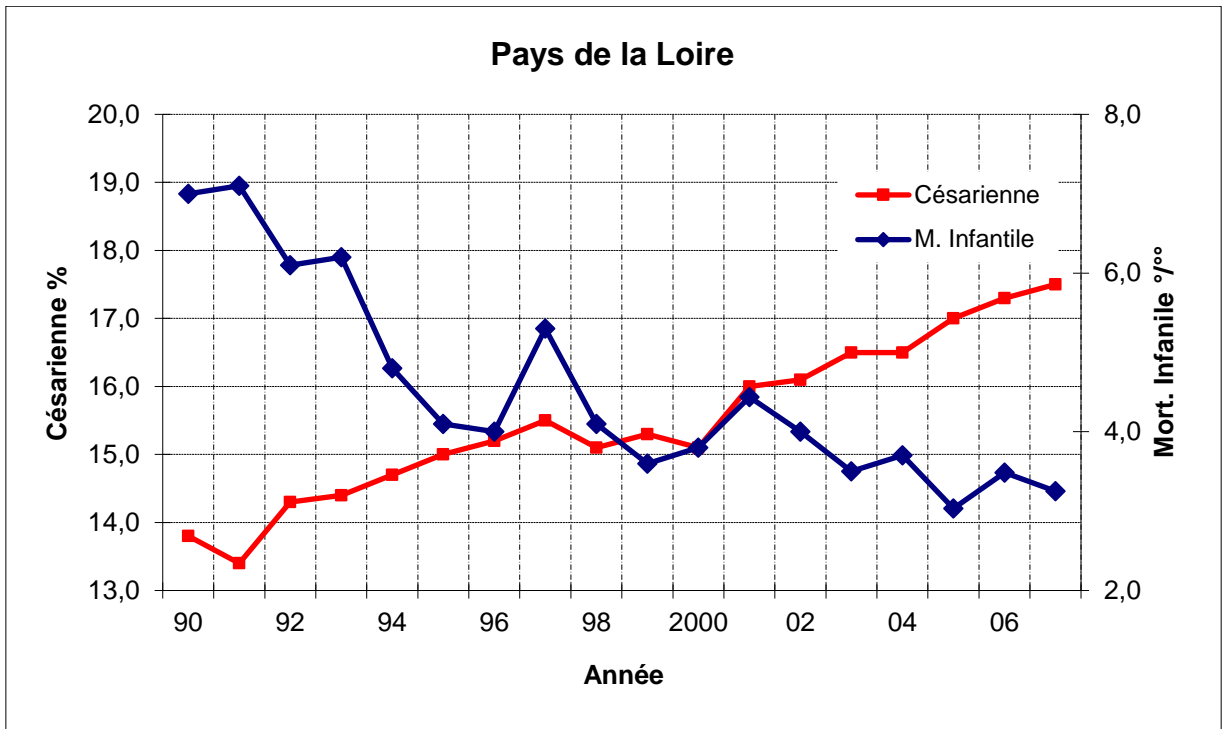
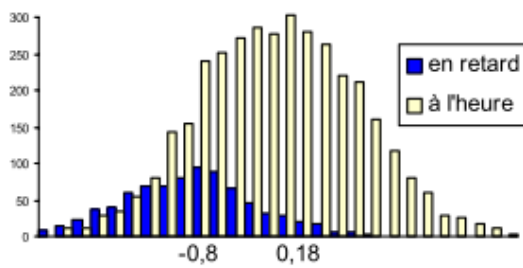
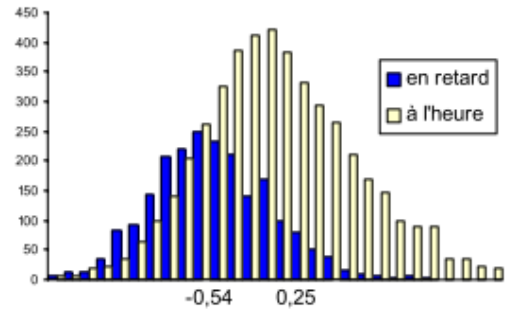


Figure 13 : Répartition des élèves selon le score à l'évaluation en fin de CM2



Source : évaluation bilan école (Gibert et al., 2004)

Figure 14 : Répartition des élèves selon le score à l'évaluation en fin de 3^{ème}



Source : évaluation bilan collège (Dauphin & Trosseille, 2004)

Conduite accompagnée : succès au permis de conduire, risque d'accidents

- Réussite au permis : 69 % versus 59 %
- Accidents : - 25 % entre 18 et 19 ans, et - 20 % entre 20 et 21 ans

✓ **Récidive d'infractions après un emprisonnement**

Tableau 2
Taux de récidive des libérés de 2002
selon des caractéristiques sociodémographiques
et pénales dans les 5 ans qui suivent la libération

	Taux de recondam- nation	Taux de prison ferme
Ensemble des libérés	59	46
Sexe		
Hommes	60	47
Femmes	34	24
Âge à l'écrou		
Mineurs	75	66
Maieurs	58	45
Âge à la libération		
Mineurs	78	68
18-29 ans	63	49
30-49 ans	55	44
50 et plus	29	19
Situation matrimoniale		
Mariés	38	30
Autres situations	61	48
Situation au regard de l'emploi au moment de l'incarcération		
Sans emploi	61	49
Avec emploi	55	39
Nationalité		
Français	64	49
Étrangers	44	37

▪ 59 % récidivent dans les 5 ans : plus pour les condamnations antérieures pour de petits délits (vols, recels), moins si il y a une libération conditionnelle ou un aménagement de peine (Réf « Cahier d'études pénitentiaires - Risque de récidive - Mai 2011 - n°36)

Condamnations antérieures		
1 condamnation antérieure	34	22
2 et plus	70	57
Mode d'exécution de la peine		
Condamnation sans peine privative de liberté	45	32
Peine couverte par la détention provisoire	59	40
Fin de peine sans aménagement	63	56
Bénéficiaires d'aménagement hors libération conditionnelle	55	47
Libération conditionnelle	39	30
Autres	62	55
Quantum de la peine prononcée		
moins de 6 mois	62	57
6 à moins de 12 mois	62	52
1 an à moins de 2 ans	64	57
2 à moins de 5 ans	53	50
5 ans et plus	37	29

VII. Biais d'interaction

Le chercheur a toujours tendance à considérer que les maladies sont monofactorielles (voir livre « Idées folles, idées fausses en médecine ») comme le tabac et cancers, les maladies professionnelles, les maladies infectieuses : bactériennes, parasitaires, virales, l'alimentation et maladies cardio-vasculaires et les maladies génétiques : géniques ou chromosomiques

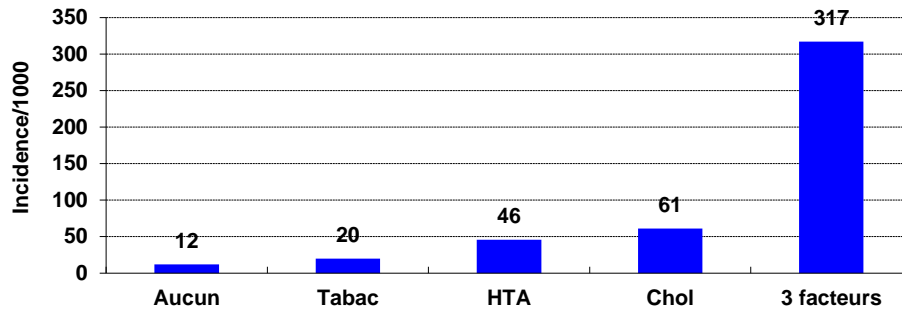
C'est souvent une idée simpliste, opérationnelle qui fait fi d'autres causes toujours présentes. Le modèle infectieux qui fait le triomphe réel des antibiotiques et autres anti-infectieux (soutenu par Koch et Pasteur) ne colle pas complètement à la réalité : tous les porteurs d'un germe ne font pas la maladie (même si tous ceux qui ne l'ont pas ne feront jamais la maladie) car le facteur « terrain » (en l'absence d'autre terme) est important.

Il peut être difficile de préciser les facteurs en cause quand il y a interaction entre les facteurs. Exemple : fréquence de mortalité cardio-vasculaire selon 3 facteurs (Fletcher)

La plupart des modèles sont plurifactorielles tels que la mort subite du nourrisson, les maladies cardio-vasculaires : « terrains » (cholestérol, HTA..) + alimentation + poids, les cancers digestifs : « terrain » + alimentation, le diabète sucré : gène + virus ou l'ulcère gastro-duodéal

C'est le modèle le plus fréquent mais qui n'est pas très opérationnel. Il pose beaucoup de questions du traitement et de la prévention. D'autre part, ce modèle, s'il est généralisé, entraîne le risque de passer à côté d'un facteur en cause.

Figure 1 : Nombre de cas de maladies cardio-vasculaires/1000 hommes en 8 ans



9. Recherche de biais en cas d'absence de différence significative : le cas de la vaccination anti-hépatite B et les complications suspectées

Exemple d'enquête en France parmi 8 études disponibles :

- Population : 17 centres de neurologie. Cas et témoins (2 par cas) idem étude de la Salpêtrière
- Recueil des données : dossiers pour la maladie, entretien téléphonique pour l'exposition, copie de certificat de vaccination ==> 64 % des cas et 71 % des témoins
- Résultats : 402 cas et 722 témoins ==> 236 (242) cas et 355 (407) témoins, 13 cas vaccinés dans les deux mois soit 4 % ? Témoins (autour de 2 %)

- OR ajusté = 1.8 (0.7 - 4.6)
- Copie du certificat : 9 cas vaccinés/152 (5.9 %) et 12 témoins/253 (4.7 %)
 - OR brut (B. Branger) = 1.26 (0.48 - 3.31)
 - OR ajusté = 1.4 (0.4 - 4.5) dans les deux mois
 - OR = 1 pour le vaccin entre 2 et 12 mois

- Analyse des biais possibles

- Biais de sélection des cas : SEP ayant consulté à l'hôpital, problème de l'absence de registre de SEP, avance au diagnostic des cas vaccinés, problème du diagnostic de première atteinte..
- Biais de sélection des témoins : malades, ils sont moins à risques d'être vaccinés
- Biais de l'exposition : perte d'information téléphonique sur l'existence de vaccins, perte d'information pour consulter le certificat
- Taille d'échantillon et puissance de l'étude : risque de deuxième espèce quand on ne trouve pas de différences
- Signification sociale de l'intervalle de confiance à 95 % de l'OR
- Politique vaccinale : protéger des individus qui ne sauront jamais s'ils auraient pu être atteints d'hépatite B, versus des individus qui font un lien chronologique entre vaccin et complications, la chronologie faisant office de causalité +++ ! ?

E. Les critères de la causalité

1. La causalité unique de Koch

Selon les postulats de Koch en 1882, l'agent causal doit : (1) être présent dans chaque cas de maladie ; (2) être isolé et cultivé en culture pure ; (3) entraîner une maladie spécifique lorsqu'elle est inoculée à l'animal ; (4) être isolé chez l'animal et identifié.

Manque le traitement apparu en 1933 avec les sulfamides et en 1940 avec la pénicilline, mais reste valable dans les modèles infectieux : → 1977 : BG- et la maladie des légionnaires ; → 1980 : VIH et SIDA avec très vite des causalités secondaires. Causalité émise un temps : les 3 H (Haïti, homosexuel, héroïnomanie)....

2. Les critères de Bradford et Hill

1. Force de l'association

C'est le risque relatif ou l'odds ratio : le RR du tabac dans le cancer du poumon est 23.7 entre les fumeurs de plus de 25 g de tabac par jour par rapport aux non-fumeurs. Le risque de choléra était de 14 entre les clients des deux compagnies (étude de Snow). Le risque de mort subite du nourrisson est de 3 à 4 pour les enfants dormant sur le ventre.

2. Relation dose-effet

Le risque relatif passe de 6.7 pour les fumeurs de 1-14 g, à 12.3 de 15-24 g et 23.7 pour plus de 25g.

3. Chronologie et temporalité

Il faut que le tabac ait précédé la maladie cancéreuse...

4. Constance des résultats dans différentes études (et reproductibilité)

En 1963, 29 études rétrospectives et 7 études prospectives retrouvaient un lien entre tabac et cancer. En 1996, une dizaine d'études retrouvaient un lien (toujours dans le même sens) entre position de sommeil du nourrisson et mort subite. Une seule étude a fait état du lien entre le cancer du pancréas et la consommation de café.

5. Spécificité de l'association

Le facteur d'exposition ne doit entraîner que la maladie étudiée et pas une autre. Critère retrouvé pour l'amiante et le mésothéliome par exemple, mais pas pour le tabac responsable d'autres maladies. Critère non indispensable ?

6. Cohérence des résultats

Il faut que les résultats soient en accord avec l'histoire naturelle et la biologie de la maladie. Le lien entre le début de la puberté et l'existence de grossesses chez la femme avec le cancer du sein est cohérent avec l'idée de la participation hormonale dans ce cancer.

7. Plausibilité biologique

Critère pas toujours possible car il dépend des avancées de la recherche : les épidémiologistes sont souvent en avance sur la recherche fondamentale. La vitamine C dans le scorbut a été découverte un siècle ½ plus tard et le vibron du choléra 50 ans plus tard... Le rôle des extraits de fumée de tabac dans la carcinogénèse a demandé moins de temps... Le rôle du cholestérol dans les maladies cardio-vasculaires également, mais le mécanisme intime n'est pas encore connu (rôle de l'inflammation-infection).

Aucun critère n'emporte la parfaite conviction ; seule l'expérimentation permet de conclure, mais elle est rarement éthique et donc impossible ?

Raisonnement par analogie : Percival Pott avait observé un excès de cancer du scrotum chez les ramoneurs → on peut penser que les produits de fumée sont cancérigènes...

Critique possible (D. Schwartz) dans le lien causal entre le chant du coq et le lever du soleil (extraits de Chantecler, le coq, conte d'E. Rostand) : le coq se demande si son chant est la cause du lever du soleil. On y trouve au moins 4 critères de Hill : force de l'association (observation quotidienne), relation dose-effet (plus le coq chante fort, plus le soleil monte), la chronologie (le coq chante bien avant le lever « quand la nuit est encore bonne et fraîche »), spécificité (les canards ne chantent pas au lever du soleil)... Seule l'expérimentation, réalisée par la faisane dorée qui empêche Chantecler de chanter, permet de lever le doute.... Encore qu'il existe des cancers chez les non-fumeurs.....

VIII. Conclusion

A. Peut-on rattraper les biais au moment de l'analyse ?

- Non pour les biais de mesure en général. Sauf à connaître les défauts de précision..
- Non pour les biais de sélection
- Non pour les biais de classement en particulier différentiels
- Oui pour les biais de confusion à condition d'y penser avant et de les prévoir.

B. La causalité doit toujours être remise en cause

1. La causalité est l'objectif ultime de toute recherche épidémiologique dans le but de guérir ou d'empêcher les gens d'être malades.
2. La situation expérimentale est la référence pour mettre en évidence une causalité en constituant deux groupes parallèles différents uniquement par le facteur causal.
3. En situation d'observation, il faut s'entourer de précautions pour prévenir les biais.
4. En biologie et en médecine, il est probable qu'il y ait plusieurs causes de maladie. On ne cherche que celles que l'on pressent, que la société (dont fait partie le chercheur) pressent, que celles que l'on cherche, et pour lesquelles on établit des hypothèses. Il faut toujours imaginer une cause supplémentaire, inconnue d'importance plus grande (RR ou OR plus grand et se rapprochant de valeurs « causale » telles que > 20 par exemple) que celles mises en évidence. Le rôle des techniques statistiques sans hypothèse (analyses factorielles) doit être précisé.
5. La science ne progresse qu'avec des faisceaux d'arguments qui se complètent et se répondent. Le chercheur s'inscrit toujours dans un courant culturel social et scientifique. Il ne travaille jamais seul et ses travaux doivent être publiés pour que ces échanges puissent se faire.

C. Conclusion philosophique ou anthropologique

L'analyse de la notion de causalité peut être envisagée de plusieurs points de vue différents :

1. Notre raison accède directement à la réalité en soi et alors le mot de causalité désigne une tentative pour connaître en vérité l'être et la raison d'être des choses. Dans ce cas, on s'appuie sur un postulat métaphysique : la raison humaine peut saisir la complexité des choses.
2. Le terme de causalité ne dénote pas une propriété des choses mais un mode d'explication, particulièrement important d'ailleurs, des faits dont nous avons l'expérience. Dans ce cas, on se montre plus modeste et plus critique ; considérant que la connaissance début avec l'expérience, on définit la causalité comme une manière d'organiser les faits empiriques et de leur donner sens. On ne fait pas de la causalité une catégorie ontologique mais un concept épistémologique, c'est-à-dire une notion dont le sens ne s'analyse pas à la lumière de ses usages métaphysiques mais de son acceptation dans les sciences.
3. L'utilisation abusive de causalité (voire erronée) peut aliéner la liberté. En fait, dans le domaine biologique, persiste toujours une part d'incertitude.
4. La recherche de causes s'inscrit toujours dans un contexte socio-politique donné qui donne à voir ou qui autorise à voir (c'est-à-dire à publier).

=====