

Epidémiologie : cours n°3

Professeur : Isabelle BOUTRON

Lundi 10 décembre 2007 à 8h30

Ronéotypeuse : Bahia DAHMANI

Enquêtes cas-témoins

Enquêtes de cohortes

Plan du cours :

I. Facteurs de risque et facteurs pronostiques

II. Enquêtes de cohortes

- a. Principes
- b. Avantages
- c. Limites
- d. Mesure d'association

III. Enquêtes cas témoins

- a. Principes
- b. Avantages
- c. Limites
- d. Mesure d'association

IV. Biais

V. Liens de causalité

NB : le diaporama de ce cours est disponible sur le site de Clément.

I. Facteurs de risque et facteurs pronostiques

Le risque

- Le **risque** est la probabilité qu'un groupe de personnes exposé à certains facteurs (facteurs de risque) développe par la suite une maladie donnée
- Il est utile dans le domaine médical d'identifier différents facteurs de risque
Par exemple, le tabac est un facteur de risque de survenue du cancer pulmonaire, de maladies cardio-vasculaires et de décès.
- **Estimation du risque** : quand on veut mesurer un risque, on va mesurer dans une population quelle est la probabilité que ce groupe développe par la suite un cancer ou une maladie cardio-vasculaire, on va donc donner une estimation à ce risque, qui repose sur les statistiques.

Les facteurs de risque

Les **facteurs de risque** sont des déterminants associés à une augmentation du risque de devenir malade. On distingue :

- ➔ **Les facteurs intrinsèques** sur lesquels on ne va pas pouvoir agir sur le plan thérapeutique : sexe, âge, facteurs génétiques.
- ➔ **Les facteurs extrinsèques** contre lesquels on peut lutter *via* des politiques de santé, par l'éducation médicale, la prévention :
 - Facteurs comportementaux (tabac, sédentarité)
 - Facteurs environnementaux (qualité de l'air, qualité de l'eau, exposition à l'amiante, agents infectieux, précarité)

Identification des facteurs de risque

Cette identification est facile quand par exemple, une personne est exposée à un agent infectieux (plat contaminé par un staphylocoque...) puis, quelques heures plus tard une gastroentérite se déclenche. Le **temps de latence** entre l'exposition et la survenue de la maladie est très court, on peut ainsi faire le lien entre les deux très facilement.

Cependant, la reconnaissance du risque peut être difficile du fait de différents facteurs :

- Période de latence longue
 - C'est le délai entre l'exposition et la survenue de la maladie
 - Ex : cancer à l'âge adulte de la thyroïde après une radiothérapie pour une amygdalite dans l'enfance
- Exposition fréquente
 - Ex : Tabac : si tout le monde fume, on aura du mal à comparer la survenue de cancer chez ceux qui fument et ceux qui ne fument pas.

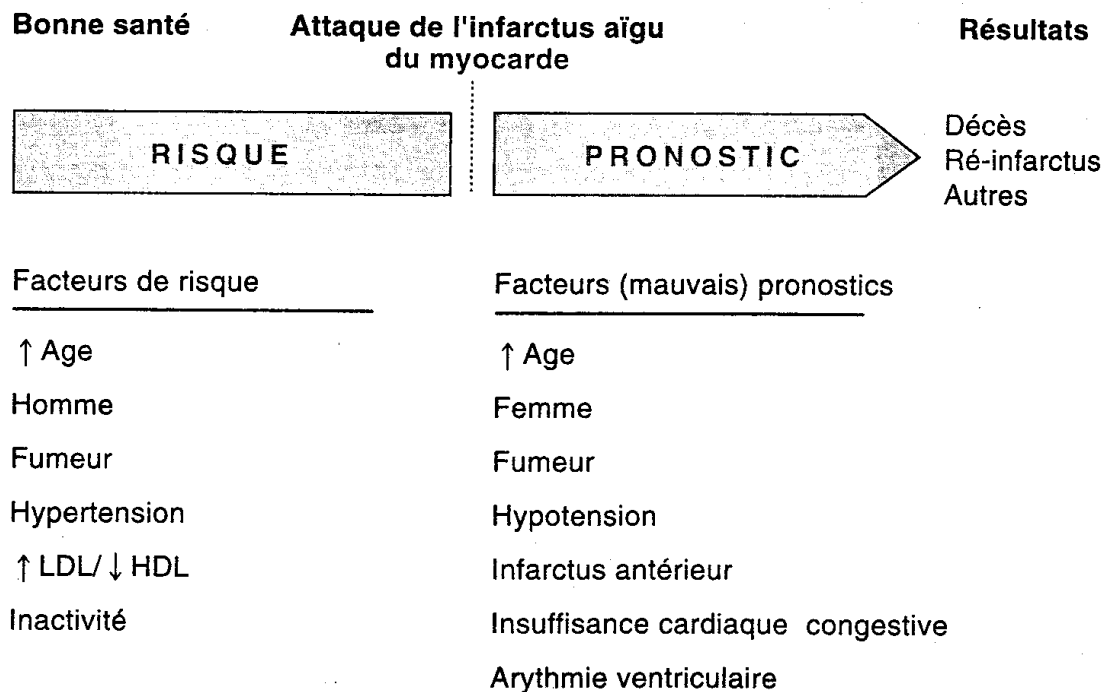
- Maladie rare
 - Cancer du poumon chez le patient tabagique
 - Incidence du cancer du poumon = 2/1000
- Risque faible
 - Exemple : la contraception orale qui augmente très peu le taux de cancer du sein

Facteurs de risque / Facteurs pronostiques

- **Facteurs de risque** : Caractéristiques de **sujets sains** associées au **développement de la maladie**. (ex : tabac)
- **Facteurs pronostiques** : Conditions associées à **l'évolution de la maladie**. Concernent des **sujets malades**. (ex : si un patient a un cancer du poumon et qu'il continue de fumer, le tabac constituera un facteur de mauvais pronostic qui entraînera des complications)

Différence entre facteurs de risque et facteurs pronostiques

Exemple de l'infarctus du myocarde



Les facteurs de mauvais pronostic augmentent la probabilité de décès du patient.

Identification du risque et du pronostic, pourquoi ?

- Facteurs de risque
 - Prévenir la survenue de la maladie (politiques de santé, consultations médicales)

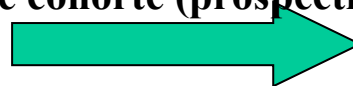
- Facteurs pronostiques
 - Informer patient/famille/correspondant etc... à propos de l'espérance de vie, des complications, de l'avenir en terme de qualité de vie...
 - Adapter le traitement à la gravité de la maladie
 - Orienter les examens complémentaires
 - Aide à la prise de décision (ex : greffe de moelle ou chimiothérapie ?)
- Attention la prédiction au niveau d'un groupe n'est pas identique à la prédiction pour un individu
- Exemple :
 - Un consommateur de tabac a 20 fois plus de risque de développer un cancer du poumon par rapport à un non fumeur
 - Il a 1 chance sur 100 de développer un cancer du poumon dans les 10 prochaines années

Identification du risque, comment ?

Études observationnelles, études analytiques

Enquête de cohorte (prospective)

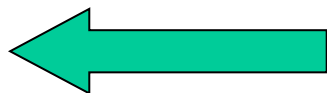
Exposition
Oui / non



Cas / maladie

Enquête cas / témoins (rétrospective)

Exposition



Cas / maladie
Oui / non

Temps

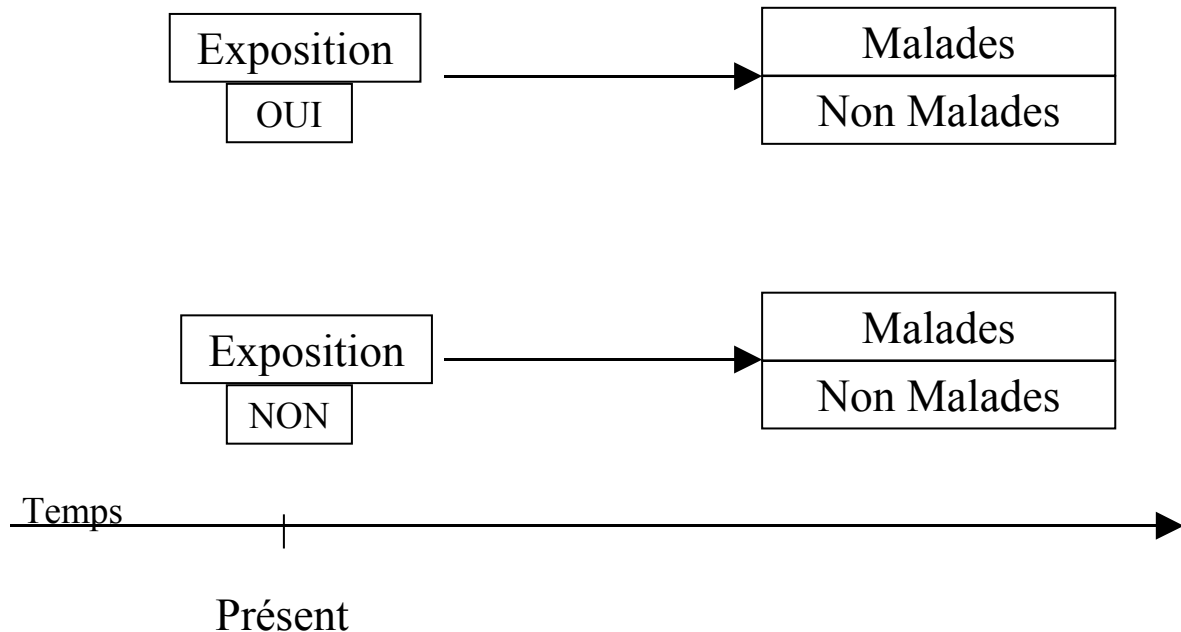


Présent

II. Enquêtes de cohortes

a. Principes

- Une **cohorte** = groupe de sujets suivi dans le temps



b. Avantages

- Étude du risque de plusieurs maladies
 - Ex: tabac → AVC, emphysème, cancer pulmonaire, cancer oro-pharyngé, pathologie cardiovasculaire etc.
- Expositions rares : on aura suffisamment de patients exposés pour démontrer quelque chose sur le plan statistique
- Séquence chronologique exposition/maladie : les patients sont suivis dans le temps donc on a une bonne information sur la chronologie entre l'exposition et la survenue de la maladie (au bout de combien de temps la maladie survient-elle ? quels sont les premiers signes ? etc.)
- Limite les biais de sélection et de mémorisation car suivi prospectif (données collectées au fur et à mesure)
- Données d'incidence (taux de nouveaux cas de maladie dans le temps) et donc une meilleure estimation du risque (**risque relatif** RR et **risque absolu** RA) car l'étude est prospective.

c. Limites

- Souvent coûteuse et longue
- **Biais de sélection** : dans l'idéal, les groupes des exposés et des non exposés diffèrent uniquement par l'exposition, mais en pratique, c'est impossible (différence d'âge...)
- **Perdus de vue** : patients qui déménagent, qui ont d'autres maladies...
- Peu performant pour les maladies rares, ou à temps de latence long

Exemple

Étude **prospective** d'une cohorte de 17 800 sujets de l'union des travailleurs exposés à l'amiante de 1967 à 1975 puis comparaison à une population non exposée (population générale, ou population travaillant dans un autre domaine). Ceci a permis de démontrer que l'amiante est facteur de risque de cancer.

d. Mesure d'association

	Malade	Non malade	
Exposés	a	b	$E_1 = a + b$
Non exposés	c	d	$E_0 = c + d$
	$M_1 = a + c$	$M_0 = b + d$	

Exposés : taux d'incidence = Risque = $Re = a/E_1$

Non exposés: taux d'incidence = Risque = $Rne = c/E_0$

Risque attribuable à l'exposition: $Re - Rne$

Risque relatif = $RR = Re / Rne$
--

Exemple

	Malade	Non malade	
Exposés	10	90	$E_1 = 100$
Non exposés	5	95	$E_0 = 100$
	$M_1 = 15$	$M_0 = 185$	

Exposés : taux d'incidence = Risque = $Re = a/E_1 = 10/100 = 10\%$

Non exposés: taux d'incidence = Risque = $Rne = c/E_0 = 5/100 = 5\%$

Risque attribuable à l'exposition: $Re - Rne = 10\% - 5\% = 5\%$

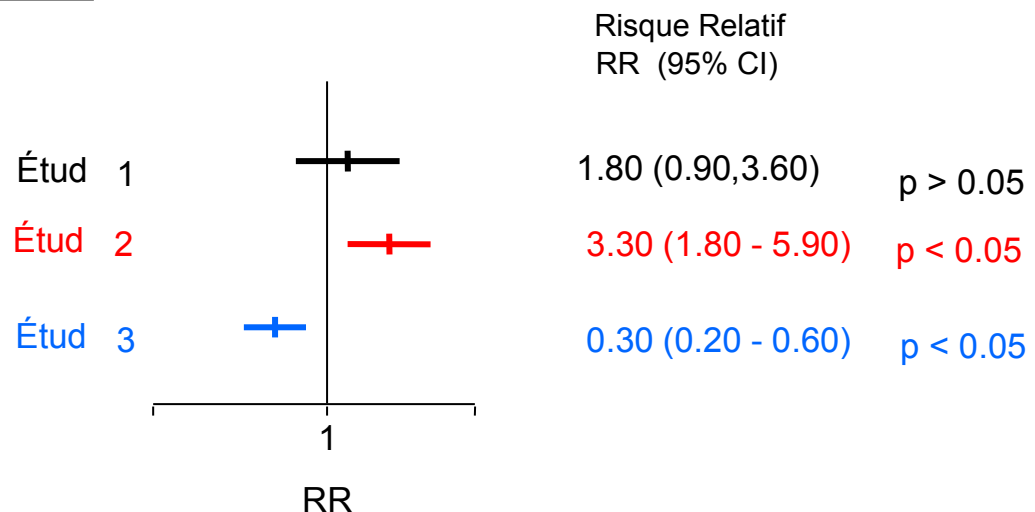
Risque relatif = $RR = Re / Rne = 10\%/5\% = 2$

Mesures d'association

Intervalle de confiance du risque relatif :

- Une enquête de cohorte est rarement réalisée sur l'ensemble de la population, mais sur un **échantillon représentatif** de cette population.
- Le RR est donc une **variable aléatoire** qui subit une fluctuation d'échantillonnage
- On calcule donc un **intervalle de confiance à 95% du RR** (IC 95%) : on a 5% de chances que le RR soit en dehors de cet intervalle

Interprétation du RR



Étude 1: absence de risque

Étude 2: facteur de risque

Étude 3: facteur protecteur

Étude 3: facteur protecteur

Etude 1 :

Si $RR = 1$, risque des exposés = risque des non exposés. Il ne s'agit donc pas d'un facteur de risque.

Ici, l'IC contient la valeur 1, il n'y a donc pas de risque sur le plan statistique, aucun lien ne peut être démontré.

Etude 2 :

L'IC ne contient que des valeurs supérieures à 1, le facteur étudié est donc un facteur de risque.

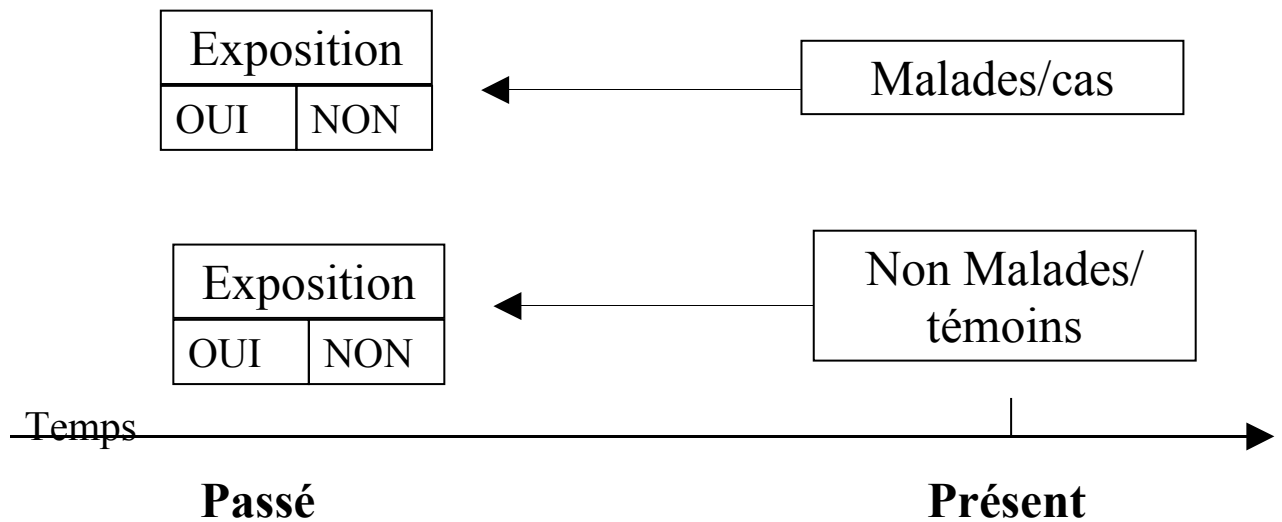
Etude 3 :

Le risque chez les exposés est moins important que le risque chez les non exposés, il s'agit donc d'un facteur protecteur. (Ex : consommation de fruits et légumes et infarctus du myocarde)

III. Enquêtes cas-témoins

a. Principes

- Étude rétrospective
- Sélection de la population en fonction de la présence ou non du critère de la maladie
- Recherche rétrospective de l'exposition au facteur de risque



b. Avantages

- Intérêts
 - Résultats rapides (patients sélectionnés puis interrogés ou consultation des dossiers médicaux)
 - Coût faible
- Indications
 - Maladies dont les périodes de latence sont longues
 - Maladies rares
 - Étude de plusieurs facteurs de risque pour une seule maladie

c. Limites

- Pas adapté pour l'étude
 - De plusieurs maladies
 - Des expositions rares
 - De la relation chronologique entre exposition et maladie
- Ne permettent pas un **calcul direct des taux d'incidence** de la maladie (nombre de nouveaux cas de maladie qui apparaissent dans la population) chez les patients exposés et non exposés (car le nombre de cas est choisi par le statisticien lui même)

- Estimation du risque par **Odds Ratio OR** (et non pas par risque relatif)
- Les risques de **biais** sont important particulièrement ceux de **sélection** et de **mémoire** (recueil rétrospectif)

d. Mesure d'association

	Cas	Témoins	
Exposés	a	b	$E1 = a+b$
Non exposés	c	d	$E0 = c+d$
	$a + c = M1$ (<i>fixé</i>)	$b + d = M0$ (<i>fixé</i>)	

Estimation du RR par le calcul d'OR (Odds ratio = rapport de cotes)

Cote = probabilité d'être exposé / probabilité de ne pas être exposé
(calculée chez les cas et les témoins)

$(a / M1) = \text{prob d'être exposé chez les cas}$ $(c / M1) = \text{prob de ne pas être exposé chez les cas}$ Cote = $P/(1-P) = a/c$	$(b / M0) = \text{prob d'être exposé chez les témoins}$ $(d / M0) = \text{prob de ne pas être exposé chez les témoins}$ Cote = $P/(1-P) = b/d$
--	--

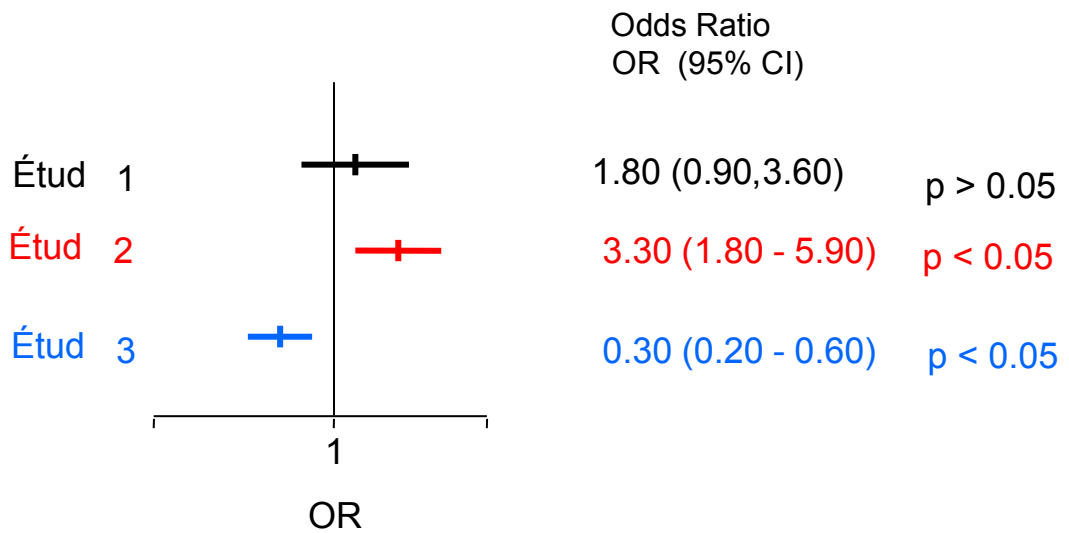
OR = ad/bc (rapport des cotes)

OR = rapport d'exposition d'un cas sur celui d'un témoin

$$OR = \frac{\text{nombre d'exposés chez les malades} \times \text{nombre de non exposés chez les témoins}}{\text{nombre non exposés chez les malades} \times \text{nombre d'exposés chez les témoins}}$$

- Si maladie rare (prévalence < 5 %) OR sensiblement équivalent au RR (on aura donc une bonne estimation du RR pas l'OR)
- Les sujets exposés ont une probabilité d'avoir la maladie OR fois celle des non exposés
- Intervalle de confiance de l'OR

Interprétation de l'OR



Étude 1: absence de risque

Étude 2: facteur de risque

Étude 3: facteur protecteur

L'interprétation est la même que pour les risques relatifs.

IV. Biais

- Un **biais** est une erreur systématique lors de la planification ou de la mise en œuvre de l'enquête dont la conséquence est une estimation erronée du RR (cohorte) ou de l'OR (cas-témoins)
 - Sur estimation
 - Sous estimation
- Différents biais
 - Biais de sélection
 - Biais d'évaluation (mémoire)
 - Biais liés aux perdus de vue
 - Biais de confusion

NB : il existe beaucoup d'intitulés différents pour les biais. A l'examen, il faut utiliser les termes du cours mais il faut surtout expliquer les notions.

Biais de sélection

Il intervient lors de la sélection des sujets dans l'enquête

- Cohorte : la sélection des exposés et non exposés dépend de la survenue de la maladie
- Cas témoins : la sélection des malades et non malades dépend du facteur d'exposition (les différences associées au facteur d'exposition doivent être proscrites)
- Exemples
 - Cohorte
 - Étude de pronostic : stade de la maladie plus grave chez les exposés que chez les non exposés qui acceptent de participer (les décès seront donc peut-être plus rapides mais non liés forcément à l'exposition). Les sujets les plus malades seront plus à même de participer que les sujets sains.
 - Cas témoins
 - Inclusion de cas survivants moins exposés. Ex : enquête aux urgences sur les facteurs de risque associés à une rupture d'anévrisme. Les patients ayant eu une rupture sont comparés à ceux n'en ayant pas eu. Or, les patients les plus graves sont décédés avant d'arriver jusqu'aux urgences et ne sont donc pas inclus dans l'enquête. Il est possible que le facteur de risque soit associé à ces mêmes patients. On a alors clairement un biais de sélection car on n'inclut que les cas survivants les moins exposés.

Sélection de la population

L'appariement : permet de limiter ce biais, surtout pour les cas-témoins

- Chaque cas peut être apparié avec un ou plusieurs témoins ayant des caractéristiques communes
 - en général sexe, âge, lieu de résidence
 - autres facteurs de risque connus considérés comme importants
- **But** : Réduit les différences entre cas et témoins et ainsi augmente la puissance de l'étude
- **Inconvénient** : L'appariement excessif. Ex : si on apparie sur le sexe, on ne saura pas si le sexe masculin est un facteur de risque.

Biais d'évaluation

= **Biais d'information, biais de mémoire**

- Cohorte
 - Informations concernant la maladie sont recueillies ou mémorisées de façon différente entre les exposés et les non exposés (par exemple, les exposés ne doivent pas être consultés plus fréquemment que les non exposés)
- Cas cas-témoins
 - Informations concernant l'exposition au facteur de risque sont recueillies ou mémorisées de façon différente entre les cas et les témoins (souvent, les malades ont déjà fait un travail de mémorisation (à quoi j'ai été exposé ? etc.) donc les malades se souviennent mieux que les non malades => biais)

Biais perdus de vue

- Concerne les gens qui déménagent, ceux qui ne veulent plus venir en consultation etc.
- Principale source de biais des **études de cohortes**
- Si taux important : 30 à 40 % ➔ doutes sur la validité des résultats
- Estimation biaisée si l'importance des « **perdus de vue** » dépend de l'exposition ou de la maladie (ex : on s'intéresse au lien entre tabac et cancer pulmonaire. Le déménagement n'est probablement pas lié à l'exposition au tabac. Par contre, si un patient est perdu de vue parce qu'il a eu un événement cardio-vasculaire, ceci est certainement lié à sa consommation tabagique)
- Éventuellement on peut calculer l'**association exposition-maladie** en prenant les situations les plus extrêmes d'exposition et de maladie
 - Hypothèse 1 : tous les perdus de vue ont développé la maladie
 - Hypothèse 2 : aucun des perdus de vue n'a développé la maladie

Si les résultats ne sont pas concordants, on estime que le biais est trop important pour conclure.

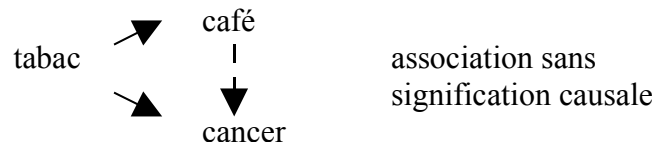
Biais de confusion

Facteurs de confusion

Un **facteur de confusion** est une variable qui doit être :

- associée à l'exposition mais ne pas en être la conséquence
- associée à la maladie indépendamment de l'exposition

Sa présence déforme la réalité de l'association entre les facteurs d'exposition et la maladie.



Le tabac est un facteur de confusion. En effet, le tabac est associé au café sans en être la conséquence, et il est aussi associé de manière indépendante à la maladie.

Si on fait l'analyse sans le tabac, on conclut que le café est un facteur de risque pour le cancer.

Mais en fait, les gens qui consomment du café sont plus fréquemment des fumeurs.

Dans le groupe exposé, on aura aussi plus d'exposés au tabac ; et dans le groupe non exposé, on aura aussi moins d'exposés au tabac.

Prise en compte des facteurs de confusion

- Il faut lister les facteurs de confusion auxquels on pense
- Lors de la planification de l'étude
 - Appariement sur le facteur de confusion (ex : autant de fumeurs chez les malades et non malades)
- Lors de l'analyse de l'étude
 - **Stratification**
 - **Ajustement**

Exemple de stratification : Complications chirurgicales parmi 1600 patients opérés

Exposition	Pas de complication	Complications	Risque	RR	IC 95%
Chirurgien A	600	200	200/800 25%	2	1.6-2.5
Autres chirurgiens	700	100	100/800 12.5%		

Ici, on conclut que le fait d'être opéré par le chirurgien A apporte 2 fois plus de risque de complications.

Or, facteur de confusion possible = utilisation de la technique X, on fait donc une stratification :

Analyse stratifiée = analyse en sous groupe

Technique X Exposition	Pas de complication	Complications	Risque	RR	IC 95%
Chirurgien A	170	230	170/400 42.5%	1.02	0.8 -1.3
Autres chirurgiens	50	70	50/120 41.7%		

Autres techniques Exposition	Pas de complication	Complications	Risque	RR	IC 95%
Chirurgien A	30	370	30/400 7.5%	1.02	0.7 -1.6
Autres chirurgiens	50	630	50/680 7.4%		

On conclut que ce n'était pas le chirurgien le facteur de risque, mais plutôt la technique X utilisée (ce qui est très différent).

- **Ajustement** : on calcule un RR qui tient compte des facteurs de confusions, c'est-à-dire des différences qui existent.
- **Analyse multivariée**

Dans l'étude de cohorte, on calcule au départ un risque relatif, mais ensuite on va tenir compte des facteurs de confusion dans une analyse multivariée. Notre résultat sera alors un OR ajusté, on peut alors estimer le risque par l'OR.

- Régression logistique
- Les résultats de la régression logistique fournissent pour chaque facteur étudié un risque (sous la forme d'un OR et d'un intervalle de confiance).
- Lorsque ce risque est significativement différent de 1, on peut affirmer que le facteur étudié est un facteur de risque **indépendamment** des autres facteurs

V. Liens de causalité

Dans toutes les études où l'on cherche à identifier les facteurs associés à la maladie, on ne démontre qu'une association et pas de lien de causalité. Cette relation de causalité peut être démontrée par l'association d'autres arguments :

- Association forte (RR et OR élevés)
- Stabilité de l'association
 - Résultats d'autres études concordants
- Spécificité de l'association
 - Un seul facteur parmi de nombreux facteurs testés
- Relation temporelle logique
- Existence d'un gradient effet-dose (ex : ceux qui fument beaucoup ont un plus grand risque de cancer)
- Association causale plausible car cohérente avec les données de la littérature et l'expérience clinique
- Réduction de l'incidence de la maladie si diminution du FDR

Conclusion

	Cas-témoins	Cohorte
Risque de biais de sélection	+++	+ si prospective
Risque de biais d'observation	+++	+ si prospective
Risque de biais liés aux perdus de vue	-	+++
Facteurs de confusion	+++	+++

