

UE10 Épidémiologie

Pr Eloy

Le 20/04/2020 de 15h30 à 17h30

Ronéotypeur : Othmane Bennani/Florian Ayoul

Ronéoficheur : Othmane Bennani/Florian Ayoul

# COURS 4 : ETUDE DE COHORTES

## Plan du cours

- Généralité
- Indicateurs de sante calculables dans la cohorte
- Étude étiologique
- Principe des études de cohorte
- Méthodologie
- Interprétation des résultats
- Conclusion

## I) Généralités

### A) Objectifs de l'épidémiologie

Rappels :

- **Décrire et surveiller l'état de santé** (décrire un phénomène de santé en fonction du temps, de l'espace, des caractéristiques de la population) et **émettre des hypothèses** : c'est l'**épidémiologie descriptive**

-**Rechercher les causes d'un phénomène de santé** : c'est l'**épidémiologie analytique**

-**Décider et évaluer** (fournir des outils d'aide à la décision et évaluer un traitement, une technique diagnostique ou de dépistage, un programme de santé publique) : c'est l'**épidémiologie évaluative**

### B) Types d'étude ( voir le schéma dans la fiche)

Il y a deux types d'études : **observationnelles** et **interventionnelles**

Les études **observationnelles** comprennent :

- les études **analytiques/ étiologique** : cohorte, cas-témoins et études transversale
- les études **descriptives** : transversale et longitudinale

Les études **interventionnelles** :

-**avec tirage au sort** : essai randomisé

-**sans le tirage au sort** : essai non randomisé

Nous allons nous intéresser dans cette première partie aux études analytiques avec les études de cohorte et aux études descriptives avec les études longitudinales.

### C) Objectif de la cohorte

Les objectifs de la cohorte sont

-Descriptif :

$\text{prevALence} = \frac{\text{ALL cases}}{\text{Population @ risk}}$
$\text{iNcidence} = \frac{\text{New cases}}{\text{Population @ risk}}$

-Analytique/ étiologique : étude des cas et des facteurs d'une maladie (facteur de risque d'une maladie et facteur pronostics d'un évènement)

Proportion de malades  
dans une population donnée, à un instant donné

—————> PREVALENCE

Vitesse de production de nouveaux cas  
d'une maladie dans une population non malade

—————> TAUX d'INCIDENCE

Probabilité pour un sujet non malade  
de devenir malade au cours d'une période fixée

—————> RISQUE

## II) Incidence de santé calculables dans la cohorte

### A) Incidence- incidence cumulée- taux d'attaque

- Indicateur dynamique : prend en compte la vitesse de survenue de la maladie
- Nombre de nouveaux cas d'une maladie survenue pendant une période donnée au numérateur
- Plusieurs calculs : Incidence cumulée (taux d'incidence) et densité d'incidence

L'incidence cumulée est le nombre de nouveaux cas de la maladie dans la population pendant une période donnée (le numérateur) divisé par le nombre total de personnes à risque de développer la maladie dans cette population pendant la même période (le dénominateur).

Attention : Cela concerne une population fixe à risque pendant une période donnée ou un nombre de sujets au début et à la fin de la période connus, c'est une moyenne. C'est un taux : probabilité de développer la maladie : risque

#### Exemple :

- Le personnel soignant d'un service de soins intensifs veut savoir s'il y a eu une augmentation du nombre de nouveaux cas de pneumonie.

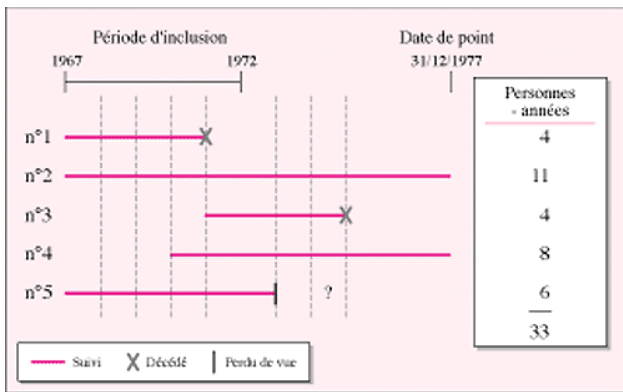
On choisit le mois de janvier comme période d'étude.

- – 220 patients sont passés par les soins intensifs pendant le mois.
- – 20 ont été hospitalisés en raison d'une pneumonie
- – 10 ont contracté une pneumonie pendant leur hospitalisation.  
→ ce qui laisse 200 personnes à risque.  
→ L'IC de la pneumonie est donc  $10/200 = 0,05$  cas pour 100 patients

**Le taux d'attaque** est **taux d'incidence cumulée** lorsque la population n'est exposée que pendant un temps limité (épidémies).

**Le taux d'incidence** (soit la densité d'incidence) est le nombre de nouveaux cas (le numérateur) divisé par le total des périodes où chaque personne de la population était à risque, exprimé en temps-personne (en jours-personne par exemple). Il reflète les variations de la durée d'observation des personnes à risque. Le dénominateur représente la « personne temps » :

-1 personne temps = 1 personne x temps de suivi (jusqu'à l'apparition de la maladie, la fin de l'étude, le décès, ou perdu de vue)



-1 personne année correspond à la durée de suivi de :

- – 1 personne non-malade pendant 1 an. Si cette personne n'est pas malade pendant l'année de suivi, elle aura été susceptible de produire un nouveau cas de maladie pendant 1 an.
- – 2 personnes non-malades pendant 6 mois. Il pourrait s'agir de personnes décédées au bout de 6 mois de suivi. Elles ne peuvent plus produire de nouveaux cas de maladie après leur décès.

Exemple :

- 200 patients à risque ont passé un total de 1 000 personnes jours aux soins intensifs en janvier.
- Puisque 10 des 200 patients ont contracté une pneumonie
- le TI de la pneumonie est de 10 cas par 1 000 personnes jours (soit un cas par 100 personnes jours).
- = il y a un risque qu'un nouveau cas se développe sur 100 personnes jours.

### Incidence : synthèse

- L'incidence cumulée mesure le risque de la maladie quand on assume que la durée d'observation est la même pour tous les patients inclus dans le dénominateur
- Le taux d'incidence mesure le risque quand la durée d'observation varie entre les différentes personnes incluses dans le dénominateur.

### Exercice

- On a recensé 100 800 sujets pour étudier l'incidence d'une maladie dans la population. On constate que 800 individus sont déjà atteints.
- 1) Que faire de ces sujets ?
- 2) Quel indicateur de morbidité pouvez-vous calculer ?
- On a recensé 100 800 sujets pour étudier l'incidence d'une maladie dans la population. On constate que 800 individus sont déjà atteints.
- 1) Que faire de ces sujets ? **On les élimine**
- 2) Quel indicateur de morbidité pouvez-vous calculer ? **Prévalence = 800/100 800**
- On suit les sujets conservés pour l'enquête sur une période de 1 an. Les résultats sont les suivants : 100 cas déclarent la maladie à 3 mois, 200 cas à 6 mois, 100 cas à 9 mois et 100 perdus de vue à 6 mois.
- 3) **Calculez le taux d'incidence de la maladie**
- On suit les sujets conservés pour l'enquête sur une période de 1 an. **100 000 pers sur 12 mois = 1 200 000 personnes mois** Les résultats sont les suivants : 100 cas déclarent la maladie à 3 mois = **300 PM**, 200 cas à 6 mois = **1200 PM**, 100 cas à 9 mois = **900 PM** et 100 perdus de vue à 6 mois = **600 PM**.
- **$1\ 200\ 000 - (300+1200+900+600) = 1\ 197\ 000\ PM$**
- **$1\ 197\ 000 / 12 = 99\ 750\ Personnes\ années$**
- **TI = nombre de nouveaux cas / population était à risque, exprimé en personne-temps =  $400 / 99\ 750 = 0,00401\ personnes\ années = 4,01\ nouveaux\ cas\ pour\ 1000\ personnes\ par\ an.$**

### III) Étude étiologique (= cohorte)

#### A) General

Les principes de cette étude sont le suivi longitudinal + échelle individuelle + groupe de sujets. Il y a différentes durées d'observation des sujets et différentes données recueillies selon les objectifs scientifiques.

Il y a une **distinction** entre :

- les **cohortes de malades** souffrant d'une pathologie dont le but est d'étudier **l'évolution d'une maladie** (naturelle ou sous traitement), le nombre de sujets 100- 1000; recrutés en milieu médical, données recueillies très détaillées (investigations biocliniques ...) : **étude pronostic**
- les **cohortes en population générale** dont le but est d'étudier les **causes des maladies** (particulièrement celles aux déterminants environnementaux et génétiques multiples), le **nombre de sujets** 2000-200000, suivi long, données personnelles, de mode de vie, sociales, professionnelles et environnementales +/- bio banques : **étude du risque**

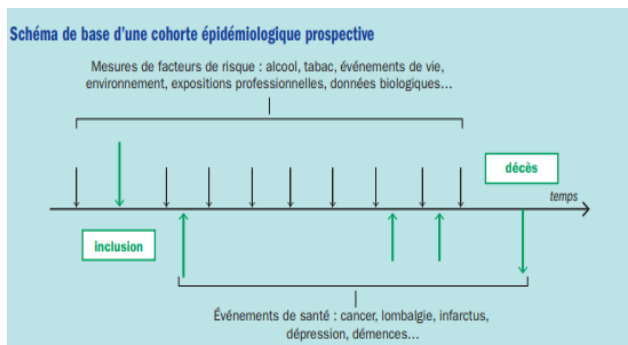
#### B) Le pronostic

Le pronostic est la prédiction de l'évolution future de la maladie après sa déclaration (Décès / survie, Rechute, métastases).

La mesure du devenir se fait par un taux = P (événement) à un temps donné Les facteurs pronostiques sont suffisamment fortement associé avec le devenir pour en prédire l'évolution. (à **différencier de la causalité**) . Ils sont **démographiques, spécifiques de la maladie** et **liés aux comorbidités**. Les facteurs pronostiques sont caractéristiques de sujets malades associées à l'évolution de la maladie.

#### C) Étude de facteurs pronostics

Cela se fait sur un **Groupe de sujets atteints d'une même maladie** / à un **même « temps »** de la maladie au **début de la période d'observation**, au **long du suivi**, on recueille pour chaque sujet **l'exposition** à des facteurs pronostics et devenir .A la fin : on calcul les **risques associés aux expositions**.



#### D) Le risque

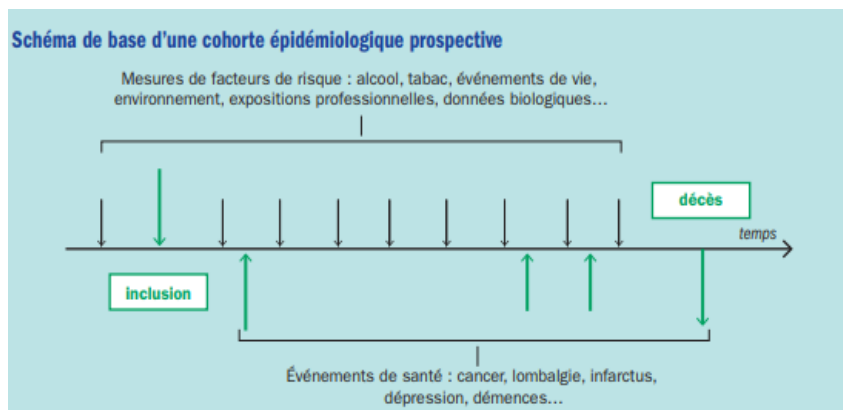
Le risque est la **probabilité qu'un groupe de personnes exposé à certains facteurs** (facteurs de risque) **développe par la suite une maladie donnée**. Les **facteurs de risque** sont des **déterminants** associés à une augmentation du risque de devenir malade :

- **Facteurs génétiques**
- **Facteurs comportementaux**
- **Facteurs environnementaux**

Les facteurs de risques sont des **caractéristiques de sujets sains** associées **au développement de la maladie**

## E) Étude des facteurs de risque

Cela concerne un groupe de sujets a priori indemne de la maladie étudiée au début de la période d'observation. Cette étude se réalise au long du suivi, on recueille pour chaque sujet l'exposition à des facteurs de risque et l'incidence des maladies. A la fin : on calcule les risques associés aux expositions.



## F) Facteurs pronostics / facteurs de risque

Comment les étudier ?

Exemple :

- L'usage de drogues est-il un facteur de risque d'infection au VIH ?
- Le type de chirurgie est-il un facteur pronostique d'une moins bonne survie chez les patients atteints d'ostéosarcomes métastatiques ?
- L'hypertension artérielle est-elle un facteur pronostique de complications chez les patients diabétiques ?

## G) Niveau de preuve

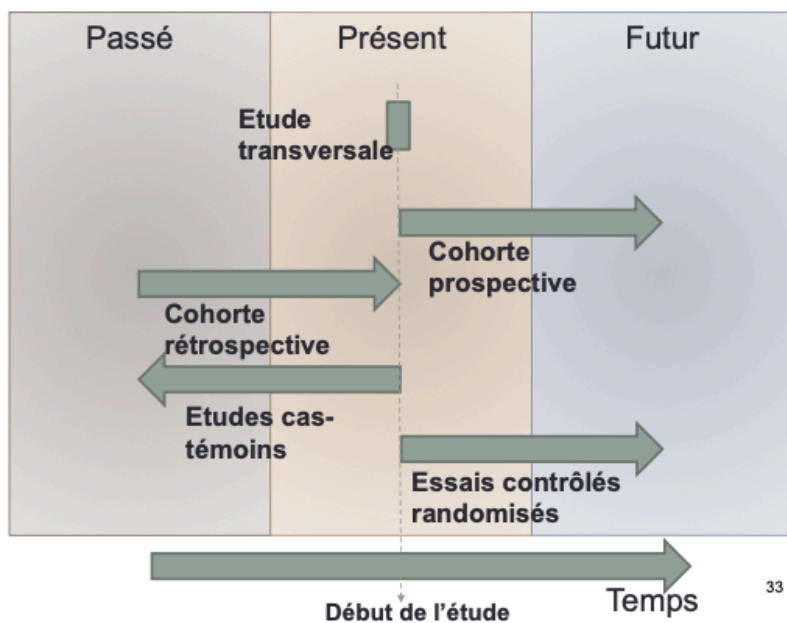
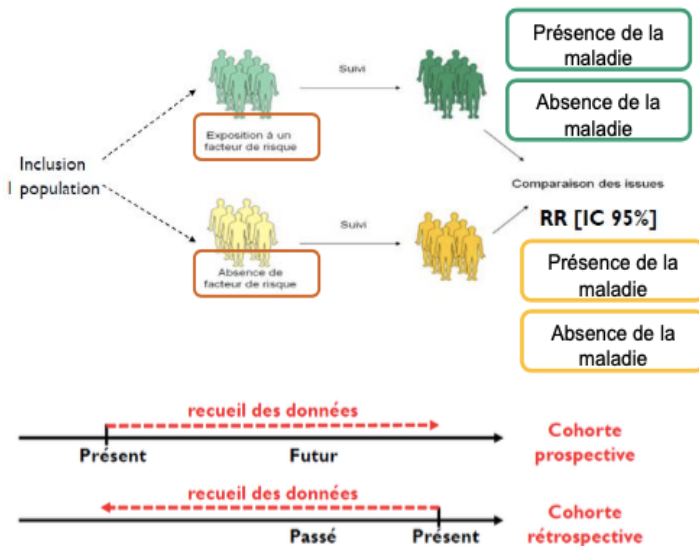
Étude analytique : cohorte → quel niveau de preuve ?

- Niveau de preuve = confiance
- Fonction : Du schéma des études + De leur adéquation à la question posée + De leur qualité méthodologique
- Différents systèmes d'évaluation. Ex : HAS

Grade des recommandations	Niveau de preuve scientifique fourni par la littérature
A Preuve scientifique établie	Niveau 1 - essais comparatifs randomisés de forte puissance ; - méta-analyse d'essais comparatifs randomisés ; - analyse de décision fondée sur des études bien menées.
B Présomption scientifique	Niveau 2 - essais comparatifs randomisés de faible puissance ; - études comparatives non randomisées bien menées ; - études de cohortes.
C Faible niveau de preuve scientifique	Niveau 3 - études cas-témoins.
	Niveau 4 - études comparatives comportant des biais importants ; - études rétrospectives ; - séries de cas ; - études épidémiologiques descriptives (transversale, longitudinale).

#### IV) Principe des études de cohorte

Cette étude se réalise sur **deux groupes** de sujets (au moins) : l'un **exposé**, l'autre **non exposé** au facteur étudié = **sélection en fonction de l'exposition**. Elles ne présentent pas initialement l'évènement étudié. Elle nécessite un **suivi dans le temps**, une étude de l'apparition de la maladie puis une **comparaison** pour **l'incidence de cet évènement**. L'exposition n'est pas contrôlée par l'équipe réalisant la recherche (contrairement à un schéma interventionnel). Les **mesures de fréquence** concerne **l'incidence de la maladie** et les **mesures d'association** représente le **risque relatif**.



##### A) Cohorte prospectives, rétrospectives

Études **prospectives** : la survenue d'évènements de santé dont on étudie le lien avec les facteurs d'exposition est enregistrée uniquement postérieurement à l'inclusion grâce à un suivi prospectif.

A l'inverse, dans les cohortes **rétrospectives** : tout ou partie de la période de suivi des sujets de la cohorte se situe dans le passé. (+ : faisabilité, adapté aux maladies apparaissant après une longue période d'exposition ou de latence, - : recueil plus difficile des expositions individuelles passées)



## VRAI OU FAUX

- A - Dans une enquête de cohorte historique, les individus sont classés au sein de la cohorte selon le fait d'être exposé ou non au facteur étudié
- B - Dans une enquête de cohorte prospective, les individus exposés au facteur de risque étudié sont comparés aux individus non exposés
- C - Dans une enquête de cohorte historique, les sujets ne sont pas malades au moment du recrutement
- D - Dans une enquête de cohorte prospective, les sujets ne sont pas malades au moment du recrutement

Reponse : V-V-F-V

Le suivi des sujets peut être réalisé de manière "**passive**", c'est à-dire sans contact direct avec les participants de la cohorte, par consultation de **bases de données**.

Dans d'autres cohortes, le recueil de données est effectué de manière **active** :

- questionnaires administrés par des enquêteurs en face à face ou par téléphone ou des auto-questionnaires envoyés par la Poste ou remplis sur internet
- prélèvements biologiques ou environnementaux
- examens médicaux, mesures anthropométriques ou tests psychologiques.

### V) Méthodologie

L'échantillon sélectionné doit être représentatif de la population d'intérêt

- Sélection de 2 populations distinctes : exposée et non exposée

Attention biais de sélection!

- **Etude f. de risques**: les non-exposés doivent avoir un risque égal de développer la maladie que la population cible : même âge, même sexe, même CSP...
- **Etude f. pronostics** : Les populations doivent être homogènes au regard du pronostic (Stade d'évolution de la maladie, Facteurs de comorbidité), sauf facteur étudié.

A) Evaluation de l'exposition d'intérêt :

- Définition du facteur pronostique ou du facteur de risque
  - Dose
  - Nature de la consommation
  - Durée d'exposition

Attention biais de classement !

- Méthode d'évaluation valide, reproductible
- Évaluer la qualité des sources de données utilisées pour juger de l'exposition ou non des participants
- Standardisation de l'évaluation du facteur pronostique ou FDR
- Evaluation similaire pour tous les sujets

Éviter ++ les biais différentiels

B) Biais de classement

- **Biais de classement différentiel**

la probabilité de l'erreur de classement n'est pas la même selon les groupes.

L'exposition / la maladie n'est pas mesurée de la même façon dans les 2 groupes

- **Biais de classement non-différentiel**

◇ La probabilité d'erreurs sur la mesure de l'exposition ou la maladie est la même selon les groupes

◇ très souvent de nature aléatoire, et peuvent être dues à la mauvaise qualité des instruments de mesure 41

C) Evaluation du critère de jugement (évènement / maladie)

• **Choix du critère de jugement**

- Pertinence clinique
- Fiable et reproductible
- Si possible objectif (décès)

• **Évaluation du critère de jugement**

- Prospective
- Standardisée
- Si possible en insu de l'exposition ou non au facteur de risque ou facteur pronostic
- Importance dépend du niveau de subjectivité du critère de jugement
- Identique pour tous les individus.

D) Evaluation des données

• **Sources de données multiples**

- Interrogatoire
- Examens cliniques et para cliniques
- Registres de décès
- Dossier médical

**Attention biais de classement !**

- Biais de mesure (d'information)
- Même information recueillie, mêmes mesures, même suivi
- Interrogatoire à l'aveugle
- Questionnaires standardisés
- Etude de variabilité inter enquêteur

E) Suivi des participants

**Patients perdus de vue (PDV)**

- Prévoir les modalités de suivi et la stratégie pour limiter le nombre de PDV
- Description de la fréquence, des raisons et des caractéristiques des PDV
- PDV ont généralement un pronostic différent (moins bon ou meilleur)

**Risque de biais si le nombre de PDV est important**

- Pas de seuil → mais si plus 1/3 : doute sur la validité des résultats
- Dépend de la proportion de PDV par rapport au taux d'évènement
- Dépend de la % de PDV dans chaque bras / causes ≠
- Dépend du taux d'évènement chez les perdus de vue (DM)

## F) Gestion des données manquantes

Long-term outcomes of morbidly obese patients treated with gastrogastrostomy.  
Thompson KS et al. *J Gen InternMed* 1986;1:85-99.

---

Cohorte de 123 patients obèses suivis 19 à 47 mois après la chirurgie

Succès : perte de plus de 30% de l'excès de poids

103 patients (84%) évalués donc 20 perdus de vue

Taux de succès  $60/103 = 58\%$

→ Analyse pire cas / meilleur cas

	<b>Meilleur cas</b>	<b>Pire cas</b>
Nombre de succès	$60 + 20 = 80$	60
Taux de succès	$80 / 123 = 65 \%$	$60 / 123 = 49 \%$ <sup>45</sup>

## G) Changement d'exposition

- « contamination » entre bras
- Migration d'un groupe à l'autre ex : si l'on étudie comme facteurs pronostiques les habitudes tabagiques et l'activité physique
  - Certains vont conserver un mode de vie à haut risque
  - Certains vont conserver un mode de vie à bas risque
  - Certains vont évoluer d'un mode de vie à haut risque vers un mode de vie à bas risque
  - Certains vont évoluer d'un mode de vie à bas risque vers un mode de vie à haut risque

## H) Facteurs de confusion

- Identification + Définition **AVANT** début de l'étude
  - Mesure systématique standardisée valide pendant l'étude
  - Avant le début de l'étude
    - Restriction lors de la sélection des patients
  - Après le début de l'étude
    - Stratification ou analyse en sous groupes
    - Ajustement
- Peut résulter en une association, alors qu'elle n'existe pas  
→ Peut résulter en une absence d'association, alors qu'elle existe

## Appariement / stratification / ajustement

- **Appariement** : Technique qui consiste à **contrôler** a priori un (ou plusieurs) **facteur de confusion en formant des paires homogènes** quant à ce facteur de confusion, dans une étude épidémiologique.
- **Stratification** : Technique qui consiste à répartir a priori un échantillon en sous-groupes de patients (les strates) homogènes pour un ou plusieurs facteurs de confusion.
- **Ajustement** : Technique **d'analyse statistique** visant à contrôler a posteriori l'effet d'un facteur de confusion dans la relation entre un facteur de risque et une maladie.

## Combien de participants dans une cohorte ?

• Plus les **phénomènes d'intérêt** (maladies, expositions) sont **rares**, plus **l'effectif doit être important** pour une précision ou une puissance satisfaisante.

• Ex : Association entre exposition aux pesticides et risque de myélome multiple (l'incidence annuelle 9/100 000)

• Effectif minimum nécessaire :

– si étude = 6 ans de suivi : ~ 1 100 000 sujets

– Si étude = 10 ans de suivi : ~ 520 000 sujets.

→ Faisabilité?!

Les cohortes prospectives ne peuvent pas répondre à toutes les questions : **études type cas témoins parfois indispensables.**

## Vrai ou Faux ?

• Vous voulez réaliser une étude de cohorte ( $q^{\circ}$  = L'alcool pendant la grossesse augmente t'il le risque de TDAH chez l'enfant?). Vous voulez diminuer autant que possible le biais de classement sur la maladie. Vous allez...

- A - Recueillir des données en aveugle de l'évaluateur du statut vis-à-vis de l'exposition
- B – Faire un recueil standardisé des données d'exposition
- C – Mettre en aveugle de l'hypothèse de l'étude les participants au moment du recueil des données
- D - Recueillir rétrospectivement les données d'exposition
- E – Conseiller aux médecins de se baser sur leurs pratiques courantes pour poser le diagnostic de TDAH

Réponse : VVVFF

## VI) Interprétation des résultats

### A) Mesure d'association

– Mesure d'un lien entre la maladie et l'exposition

– Attention une association forte et positive ne signifie pas une causalité

• **Risque relatif** = facteur par lequel le risque de maladie est multiplié en présence de l'exposition (**estimable uniquement dans les cohortes**)

• Odds Ratio = lorsque la maladie est **rare**,  $OR \approx RR$  (**estimable dans tous types d'enquête**)

	Cas	Témoins	
Exposés	a	b	a+b
Non exposés	c	d	c+d
	$a + c = M_1$	$b + d = M_0$	

$$OR = (a/b) / (c/d)$$

$$RR = (a/a+b) / (c/c+d)$$

Risque de maladie : R = probabilité d'être malade

- Risque relatif :  $RR = R_e/R_0$ 
  - $R_e$  = risque chez les exposés =  $a / (a+b)$
  - $R_0$  = risque chez les non exposés =  $c/(c+d)$
- Excès de risque  $ER = R_e - R_0$  (risque attribuable)

	Malades	Non malades	Total
Exposés	a	b	a+b
Non exposés	c	d	c+d
Total	a+c	b+d	

- Exemple: une infection à l'admission en USI est –elle un facteur de décès?

	Décès = oui	Décès = non	Total
Infection = oui	37	71	108
Infection = non	31	117	148
Total	68	188	256

Risque relatif :  $RR = Re/R0$

–  $Re$  = risque chez les exposés =  $a / (a+b)$

–  $R0$  = risque chez les non exposés =  $c/(c+d)$

• Excès de risque  $ER = Re-R0$

$Re$  = risque chez les exposés =  $a / (a+b)$

$37 / (37+71) = 34\%$

•  $R0$  = risque chez les non exposés =  $c/(c+d)$

$31 / (31+117) = 21\%$

Risque relatif :  $RR = Re/R0 : 34\% / 21\% = 1,6$

• Excès de risque  $ER = Re-R0 = 34\% - 21\% = 13\%$  (attribuable à l'E)

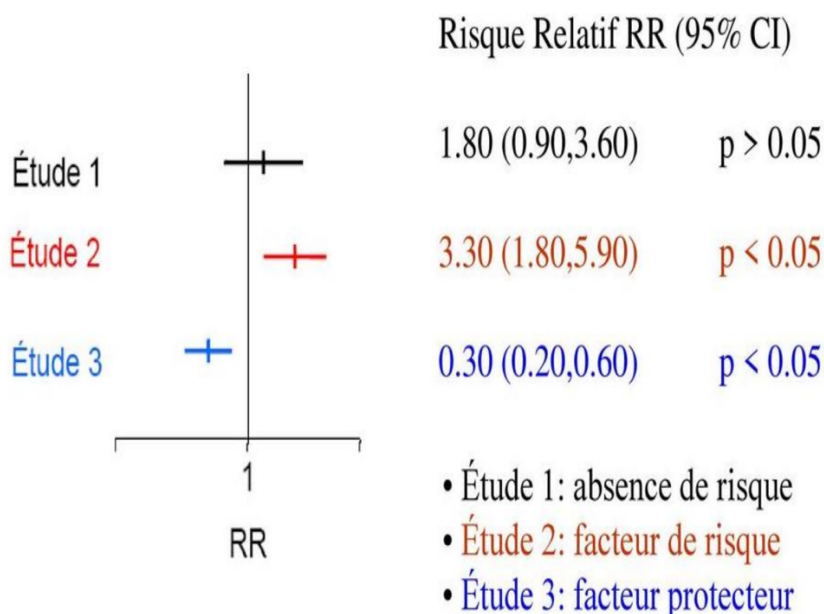
• les sujets exposés ont 1.6 fois plus de risque de développer la maladie étudiée que les sujets non exposés  
→ significatif?

### B) 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%) → aide à **conclure sur la significativité**



### C) Hazard ratio

- Rapport de risque / Hazard ratio (HR)
- Globalement équivalent au risque relatif (RR); utile lorsque le risque n'est pas constant en ce qui concerne le temps. **Il utilise les informations collectées à différents moments.**
- HR généralement utilisé dans le contexte de la survie dans le temps.
- **S'interprète comme RR**

### Vrai ou Faux ?

	Pas de mélanome oculaire	Mélanome oculaire
Emploi comme soudeur	15	7
Autre emploi	300	14

	Cas	Témoins	
Exposés	a	b	a+b
Non exposés	c	d	c+d
	a + c	b + d	

$$OR = (a/b) / (c/d)$$

$$RR = (a/a+b) / (c/c+d)$$

- A - L'odds ratio associé au type d'emploi est égal à 10.0  
B - L'odds ratio associé au type d'emploi est égal à 0.10  
C - Le risque relatif associé au type d'emploi est égal à 10.0.  
D - Le risque relatif associé au type d'emploi n'est pas calculable  
E - L'odds-ratio permet d'obtenir une approximation du risque relatif

Réponse :VFFVV

## VII) Conclusion

### Avantages

#### o Étude du risque de plusieurs maladies

– Ex: tabac, AVC, emphysème, cancer pulmonaire, cancer oropharyngé, pathologie cardiovasculaire etc.

o Expositions rares

o Séquence chronologique exposition / maladie plus facile à établir

o Limite les biais de sélection et de mémorisation car suivi

o prospectif

o Données d'incidence et donc **une meilleure estimation du risque** (RR, RA)

### • Inconvénients

o Nécessite des **moyens importants**

o Durée longue

o Risque important de perdus de vue o Ne permet pas l'étude de maladies à faible incidence (plutôt cas-témoins)

o **Faisabilité** (NSN pour une puissance et une précision satisfaisante, coûts)

o Biais de sélection

### La reconnaissance du risque peut être difficile :

– Période de **latence longue**

• Délai entre l'exposition et la survenue de la maladie

• Ex: cancer de la thyroïde après une radiothérapie pour une amygdalite dans l'enfance

– **Exposition fréquente**

• Ex: Tabac

– Maladie rare

• Cancer du poumon chez **le patient tabagique**

• Incidence du cancer du poumon = 2/1000

– Risque faible

• Exemple cancer du sein et contraception orale

« Tableau de contingence / à 4 cases »

	<b>Malades</b>	<b>Non malades</b>	
<b>Exposés</b>	Malades Exposés	Non malades Exposés	COHORTE
<b>Non exposés</b>	Malades Non exposés	Non malades Non exposés	

TRANSVERSALE

CAS-TEMOIN

21



