

Evaluation d'un facteur de risque ou pronostique dans l'étiologie d'une maladie

1

1

Plan du cours

- Classifications des enquêtes
 - Cas-témoin (+appariement), cohorte (plusieurs types), transversale
 - Définitions, design, avantages, disadvantages
- Tester l'existence de la relation
- Quantification
 - Indicateurs,
 - formules, exemples
 - quand on peut les calculer en fonction de la collection des données
 - interprétation statistique, clinique
 - interprétation des intervalles de confiance + relation avec la valeur p
 - Régression logistique simple et multiple
- Causalité
- Generalisabilité
- Evaluation de la validité interne de l'étude
- Exemples des études
- Exemples des exercices

2

2

Évaluation d'un Facteur de Risque ou Pronostique

(Les Enquêtes Etiologiques)

3

3

Les enquêtes étiologiques

But:

- ❑ **Etudier un lien** entre un facteur donné et une maladie

Objectif: évaluer:

- ❑ **Existence** du lien
 - ❑ **Quantification**
 - ❑ **Causalité**
- **Analytiques** et **observationnelles**
 - Facteur de **risque** / **protective** / **neutre**

4

4

Les enquêtes étiologiques

■ cas-témoins

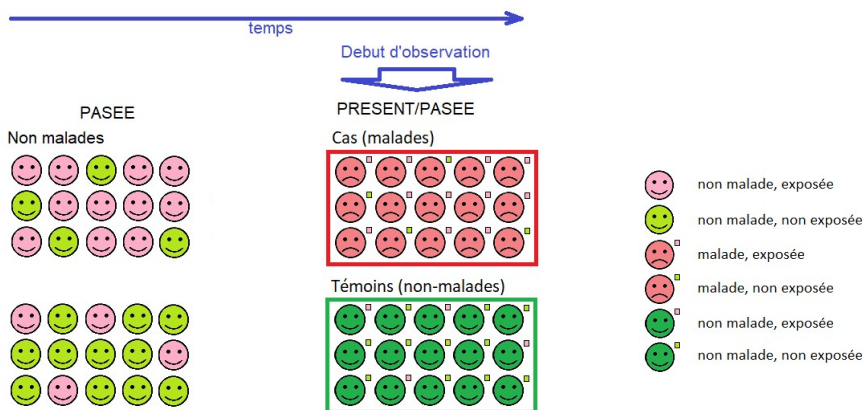
- ❑ but - évaluation **a posteriori** de la fréquence d'un facteur d'exposition (ou plusieurs) dans un **groupe de malades** par **comparaison** à un **groupe de témoins** – évaluation de la relation entre un facteur d'exposition (ou plusieurs) et une seule maladie. Donc il s'agit d'une seule maladie (pas plusieurs)
- ❑ **la maladie** constitue le **facteur d'inclusion**
- ❑ **le/les facteurs** – facteurs aléatoires (observées dans le passé – avant le commencement d'étude)
- ❑ **les cas – les malades** – en utilisant des critères clairs, mesurables, objectifs (observées dans le présent ou dans le passé)
- ❑ **les témoins – les non malades** – doit être sélectionnée de la même population que les cas. Si ils ne sont pas bien sélectionnés on peut avoir des biais de sélection

5

5

Les enquêtes étiologiques

■ cas-témoins



Graphique (Daniel Leucuta)

6

6

Les enquêtes étiologiques

■ cas-témoins

- vers le passée - rétrospective

Avantages	Désavantages
<ul style="list-style-type: none">• ils sont applicable aux maladies rares ou au long incubation• utile pour générer des hypothèses• permet identifier des prédicteurs d'un objectif<ul style="list-style-type: none">• leur existence• quantifier l'importance (OR)• (Etant rétrospective) =><ul style="list-style-type: none">• plus simples à organiser• permettent une logistique plus légère d'habitude• permettent un moindre coût d'habitude• durée plus courte	<ul style="list-style-type: none">• on peut étudier un seul objectif (outcome, maladie)• risque de<ul style="list-style-type: none">• biais de sélection (sampling bias)<ul style="list-style-type: none">• cas et témoins depuis des populations différents• biais d'observation (étant rétrospective)<ul style="list-style-type: none">• biais de mémoire (recall bias)• Les variables de confusion (confounding variables) sont une problème importante. (pour toutes les études observationnels)• (du a la collection par colonnes) =><ul style="list-style-type: none">• on ne peut pas calculer le taux d'incidence ou d'attaque (+ RIE, RIN, RR, RA, FREE – voir diapositives suivantes)

7

7

L'utilisation des techniques d'appariement (matching)

- Pour des études comparant des groupes (pas seulement pour les études cas-témoin).
- Cela signifie que pour **chaque sujet** dans le **groupe cas** / exposition on **identifie** un **sujet** présentant des **caractéristiques similaires** dans le **groupe de comparaison** (contrôle).
- Ex. appariement après: même sexe, même âge (identique, ou + / - 2 ans), les mêmes facteurs de risque (diabète) – en principe des facteurs de confusion.
- Le **but** est de **diminuer le biais de confusion** (pour les variables apariées)
- L'appariement peut être fait 1:1 ou 1:2, 1:3, 1:04 (c'est à dire, un cas et 1/2/3/4 contrôle/s)

8

8

L'utilisation des techniques d'appariement (matching)

- Les effets de l'utilisation de la technique d'appariement 1:2, 1:3, 1:4
 - ❑ augmente la puissance de l'étude entier,
 - ❑ diminue le risque beta (1-puissance),
 - ❑ diminuent la largeur de l'intervalle de confiance pour l'indicateur OR
 - ❑ diminuent la largeur de l'intervalle de confiance pour les témoins (augmente la précision), mais pas pour les cas
 - ❑ diminuent le biais de confusion (pour les variables appariées)
- ❑ Augmente le cout, l'effort

9

9

Les enquêtes étiologiques

- **Enquêtes/ études transversales** (cross sectional - en anglais)

Définition: Des études où les sujets inclus dans l'enquête sont tous les sujets présents **au moment de l'étude**

Caractéristiques:

- On ne suit pas les sujets dans le temps, on les **observe une seule fois**
- On a des informations sur **un seul moment** dans le temps
- Représente une **photographie d'un phénomène** de santé
- la question posée dans les enquêtes transversales: que se passe-t-il à un moment donné ?
- les sujets ne sont pas sélectionnés:
 - ❑ ni sur l'exposition comme dans les études de cohortes,
 - ❑ ni sur la maladie comme dans les enquêtes cas-témoins

Exemple:

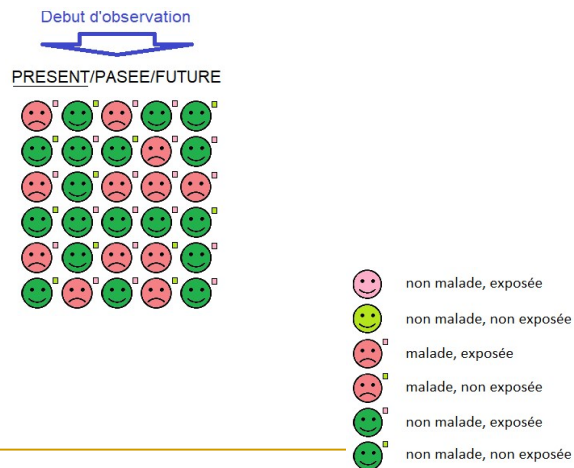
- on évalue la relation entre obésité et hypertension à un moment donné
- On évalue la fréquence des caries des enfants

10

10

Les enquêtes étiologiques

- **Enquêtes/ études transversales** (cross sectional - en anglais)



Graphique (Daniel Leucuta) 11

11

Les enquêtes étiologiques

- **Enquêtes/ études transversales** (cross sectional - en anglais)
- Caractéristiques (continuation):
 - l'accès à l'information dans une enquête transversale est rapide
 - Un jour/semaine/mois, parfois plus
 - on calcule
 - le pourcentage (d'une maladie (ex. diabète) / facteur de risque (tabagisme)),
 - moyenne, déviation standard (ex. de l'âge), médiane, quartiles (ex. du cholestérol)
 - avec l'intervalle de confiance

Avantages	Désavantages
<ul style="list-style-type: none"> • Les meilleurs pour déterminer la prévalence • On peut étudier plusieurs objectives (outcomes, maladies) • plus simples à organiser • permettent une logistique plus légère d'habitude • permettent un moindre coût d'habitude • durée plus courte 	<ul style="list-style-type: none"> • on ne peut pas <ul style="list-style-type: none"> • observer si un facteur précède un résultat d'intérêt • montrer que la cause est avant l'effet • on ne peut pas calculer le taux d'incidence ou d'attaque (+ RIE, RIN, RR, RA, FREE) • Les variables de confusion (confounding variables) <u>sont une problème importante</u> (pour toutes les études observationnels)

12

12

Les enquêtes étiologiques

■ cohorte

- La cohorte est un **group** des sujets avec une **caractéristique commune** (une maladie, aspects démographiques (âge, domicile, ...), ...
- Le **but** est l'étude d'un ou plusieurs facteurs de risque, et l'apparition d'un ou plusieurs maladies (ou complications, ou autres objectifs d'intérêt) – le critère de jugement
- La cohorte est **suivi** sur une période du temps – soit en continu, soit à des visites prédéterminées

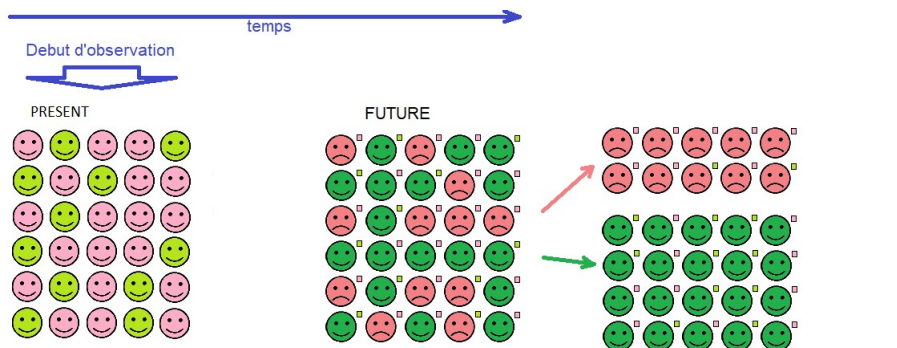
13

13

Les enquêtes étiologiques

■ Cohorte

1. cohorte prospective



Graphique (Daniel Leucuta)

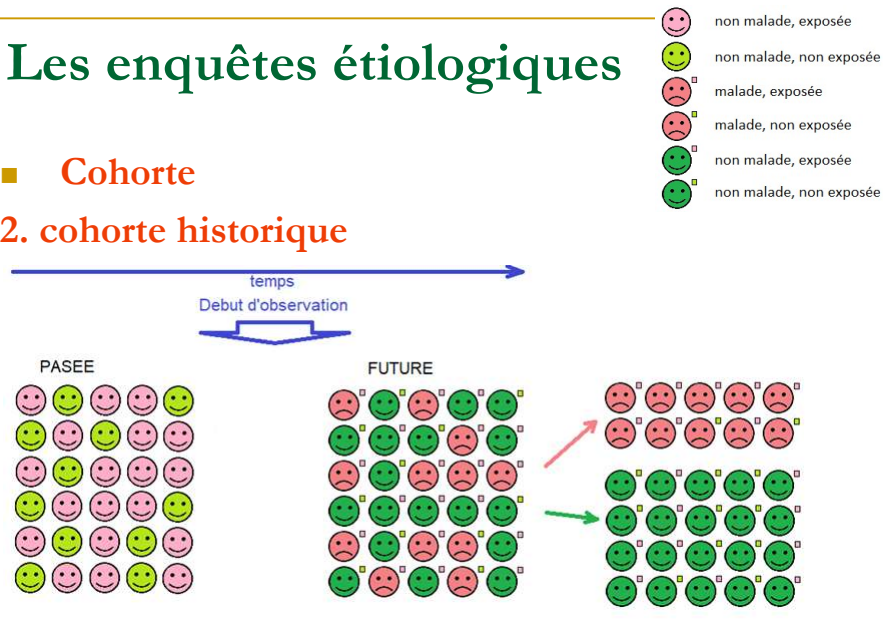
14

14

Les enquêtes étiologiques

■ Cohorte

2. cohorte historique



Graphique (Daniel Leucuta)

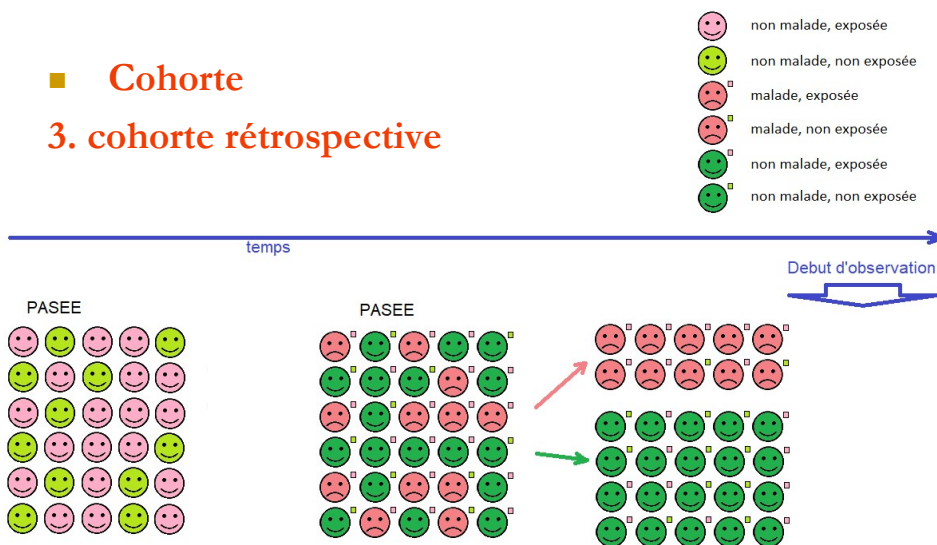
15

15

Les enquêtes étiologiques

■ Cohorte

3. cohorte rétrospective



Graphique (Daniel Leucuta)

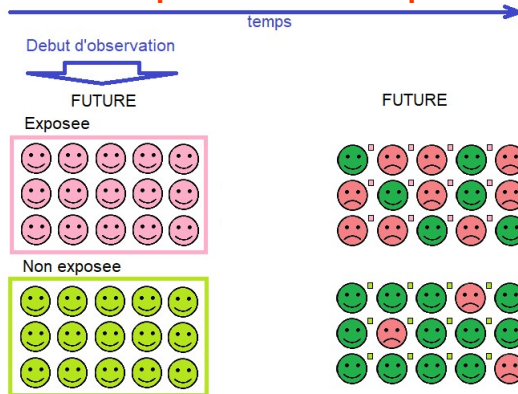
16

16

Les enquêtes étiologiques

■ Cohorte

4. cohorte exposée – non exposée



- non malade, exposée
- non malade, non exposée
- malade, exposée
- malade, non exposée
- non malade, exposée
- non malade, non exposée

Le début d'observation peut être dans le présent, et l'exposition d'être dans le passée

Graphique (Daniel Leucuta) 17

17

Les enquêtes étiologiques

■ cohorte

- Si la collecte est prospective on a les caractéristiques connues:

■ vers le futur

Avantages	Désavantages
<ul style="list-style-type: none"> Mesurent les événements dans une séquence temporelle donc on peut observer que une facteur précède l'objectif (ca aide a montrer la causalité) Peuvent mesurer plusieurs facteurs pronostiques, mais aussi plusieurs objectives/critères de judgement a la fois des meilleures informations par rapport a des études rétrospectives et transversales on peut calculer le taux d'incidence ou d'attaque (+ RIE, RIN, RR, RA, FREE – voir diapositives suivantes) 	<ul style="list-style-type: none"> durée est d'habitude longue plus difficile a organiser, plus chère, prends plus de temps les perdus de vue sont une risque potentiel pour ce type d' enquêtes plus la suivi est longue le risque du biais d'attrition (du a la perte des sujets) augmente la modification dans le temps <ul style="list-style-type: none"> des critères diagnostiques de l'influence du facteur d'exposition <ul style="list-style-type: none"> sont des risques potentiels pour ce type d' enquêtes Les variables de confusion (confounding variables) sont une problème importante. (pour toutes les études observationnels) nécessites parfois des effectifs importants

18

18

Les enquêtes étiologiques

■ Types des études de cohorte

- ❑ une enquête de cohorte est dite **pure** si les **mêmes sujets** sont examinés tout au long de l'étude
- ❑ une enquête de cohorte est dite **mixte** en cas de **départs ou d'intégration de nouveaux sujets** dans la cohorte au cours du suivi
- ❑ une enquête **permanente** est une généralisation d'une enquête longitudinale dont **le terme ne serait pas défini**
- ❑ si le **début** de l'enquête est **postérieur** au début de **l'exposition** mais **antérieur** à la déclaration de **la maladie** on parle alors de cohorte **historique**

19

19

Les enquêtes étiologiques

Types des études de cohorte

- les cohortes **"historiques"** ont l'avantage de **réduire la durée** du suivi en particulier lorsque le temps de latence entre l'exposition et la maladie est grand.
- dans les enquêtes **"exposés - non exposés"** le **début** de l'enquête **coïncide** avec le **début** de **l'exposition**.

20

20

Les enquêtes étiologiques

■ Etude pronostique

- But – analyser des facteurs prédictifs / pronostiques pour l'apparition d'un évènement (maladie, complication, épisode aiguë, récurrence, métastase, décès)
- D'habitude portée sur un seul group des patients qui sont suivi pour voir l'apparition de l'évènement.

21

21

Tests statistiques pour le tableau de contingence des études des facteurs de risque

- Pour montrer **l'existence de la relation**/ lien/ association/ dépendance entre le facteur pronostique et la maladie – analyse uni variée – simple – avec des tests statistiques (sans tenir compte d'autres facteurs). Pour tenir compte d'autre facteurs (de confusion) on doit utilise des techniques multi variée (eg. Régression logistique)
 - Etude sans appariement (matching – anglais)
 - Test **Chi deux** relation entre deux variables qualitatives
 - Si <20% des cellules du tableau théorique (attendue) ont des valeurs < 5, ou une cellule a des valeurs 1 ou 0.
 - Test **Fisher exact** relation entre deux variables qualitatives
 - Si >20% des cellules du tableau théorique (attendue) ont des valeurs < 5
 - Etude avec appariement
 - Test **Mc Nemar** – relation entre deux variables dichotomiques FR - Maladie

22

22

Indicateurs du risque individuelle:

- Sont des indicateurs qui **ne quantifient pas le lien facteur de risque maladie**
- Ils **présent la fréquence de la maladie**

$$\text{RIE} = p(M+ | FR+) = a/(a+b)$$

Risque individuel des exposées
(fréquence de la maladie
dans le groupe exposée)

	M ⁺	M ⁻	
F ⁺	a	b	ne ⁺
F ⁻	c	d	ne ⁻
	nm ⁺	nm ⁻	N

$$\text{RIN} = p(M+ | FR-) = c/(c+d)$$

Risque individuel des non-exposées
(fréquence de la maladie
dans le groupe exposée)

F – facteur pronostique (de risqué, ou protectif); M - maladie

23

23

Indicateurs du risque individuelle:

■ Interprétation statistique du RIE, RIN :

RIE = 15% (il y a 15 % risque d' avoir la maladie pour ceux qui sont expose)

RIN = 5 % (il y a 5 % risque d' avoir la maladie pour ceux qui ne sont pas exposées)

Interprétation clinique du RIE, RIN:

Est subjective, dépend de quoi on évalue, et du domaine d'application. Ex:

Un RIE de 5% des décès est cliniquement important

Un RIE de 5% des douleur a le place d'injection n'est pas cliniquement important

24

24

Indicateurs de quantification du lien facteur-maladie:

Risque attribuable:

$$RA = RIE - RIN = a/(a+b) - c/(c+d)$$

Risque relative:

$$RR = RIE/RIN = [a/(a+b)]/[c/(c+d)]$$

Odds ratio:

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

	M ⁺	M ⁻	
F ⁺	a	b	ne ⁺
F ⁻	c	d	ne ⁻
	nm ⁺	nm ⁻	N

$$FREE = (RIE - RIN) \times 100/RIE$$

F – facteur pronostique (de risqué, ou protectif); M - maladie₂₅

25

Indicateurs de quantification du lien facteur-maladie:

■ Interprétation statistique:

RA = 15% (il y a 15 % **plus** de risque d'avoir la maladie pour ceux qui sont expose compare a ceux qui ne sont pas expose)

RR = 1,4 (il y a 1,4 **fois** plus de **risque** d'avoir la maladie pour ceux qui sont expose compare a ceux qui ne sont pas expose)

OR= 1,3 (il y a 1,3 **fois** plus de **chance** d'avoir la maladie pour ceux qui sont expose compare a ceux qui ne sont pas expose)

26

26

Indicateurs de quantification du lien facteur-maladie:

■ Interprétation clinique pour le risque attribuable RA

Est **subjective**, dépend de quoi on évalue, et du domaine d'application.

Les valeurs **proche de 0** indiquent importance clinique **absente ou réduite** (0 – absence de la relation, absence de la différence entre les exposées et les non exposées)

De **plus en plus on s'éloigne de 0**, l'importance clinique **augmente**

$5\% < 10\% < 20\% \dots$

Pour la force d'association -10% est équivalent à $+10\%$, ont la même importance clinique, la même force d'association, seulement la **direction de l'association est différente**

Ex:

Un RA de 5% des décès est cliniquement important

Un RA de 5% des douleurs à la place d'injection n'est pas cliniquement important (parce que le décès est plus grave que la douleur en général)

27

27

Indicateurs de quantification du lien facteur-maladie:

■ Interprétation clinique pour le risque relatif ou odds ratio - OR, et RR

Est **subjective**, dépend de quoi on évalue, et du domaine d'application.

les valeurs **proche de 1** indiquent importance clinique **absente ou réduite** (1 – absence de la relation, absence de la différence entre les exposées et les non exposées).

De **plus en plus on s'éloigne de 1**, l'importance clinique **augmente**

$1 < 1,01 < 1,5 < 2 < 5 < 10 \dots$

$1 < 0,99 < 0,7 < 0,5 < 0,2 < 0,1 \dots$

Les valeurs proches de 0 et celles proches de $+\infty$ – sont très grandes.

Pour la force d'association $0,1$ ($1/10=0,1$) est équivalent à 10 ($1/0,1=10$), et $0,5$ ($1/2=0,5$) est équivalent à 2 ($1/0,5=2$). La direction est clairement opposée, mais la force de l'association est de même intensité

Pour la force d'association $0,1$ (distance vers absence de la différence(1): $1-0,1=0,9$) est plus grand que $0,5$ (distance vers absence de la différence(1): $1-0,5=0,5$).

Pour la force d'association 10 est plus grand que 2 .

Pour la force d'association 2 par rapport à une force d'association de $0,1$ (2 est équivalent à $1/2=0,5$). Donc la distance vers absence de la différence(1): $1-0,5=0,5$ est plus petit que pour $0,1$ (ou la distance vers l'absence de la différence est $1-0,1=0,9$). La direction est clairement opposée

28

28

Indicateurs de quantification du lien facteur-maladie:

OR – **odds ratio** = **chance** (en anglais: odds) d'avoir la maladie chez les exposées / chance d'avoir la maladie chez les non exposées

- **ce n'est pas une probabilité!**

- **ce n'est pas un rapport des probabilités!**

La chance = probabilité d'un événement / probabilité d'absence de l'événement

La probabilité d'absence de l'événement = 1 - probabilité d'un événement

OR est une approximation du RR quand la prévalence de la maladie est réduite
(<10% ou <5%)

29

29

Indicateurs de quantification du lien facteur-maladie:

- **Explication – OR – n'est pas un rapport des probabilités – donc pas rapport des risques**

$$OR = \frac{\text{Chance}(M+|F+)}{\text{Chance}(M+|F-)} = \frac{\frac{\text{Probabilité}(M+|F+)}{1-\text{Probabilité}(M+|F+)}}{\frac{\text{Probabilité}(M+|F-)}{1-\text{Probabilité}(M+|F-)}} = \frac{\frac{\frac{a}{a+b}}{1-\frac{a}{a+b}}}{\frac{\frac{c}{c+d}}{1-\frac{c}{c+d}}} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} \cdot \frac{c+d}{a+b} = \frac{a*d}{b*c}$$

30

30

Indicateurs de quantification du lien facteur-maladie:

- Explication – OR est calculable dans une collection expose-non exposée, mais aussi cas témoin, et bien sur échantillon représentative
- Dans une collection avec regroupement en fonction d'exposition on peut calculer a l'intérieur des groupes avec et sans exposition: $a/(a+b)$, $b/(a+b)$, $c/(c+d)$, $d/(c+d)$

$$OR = \frac{\text{Chance}(M+|F+)}{\text{Chance}(M+|F-)} = \frac{\frac{\text{Probabilité}(M+|F+)}{1-\text{Probabilité}(M+|F+)}}{\frac{\text{Probabilité}(M+|F-)}{1-\text{Probabilité}(M+|F-)}} = \frac{\frac{\frac{a}{a+b}}{\frac{b}{a+b}}}{\frac{\frac{c}{c+d}}{\frac{d}{c+d}}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a*d}{b*c}$$

- Dans une collection avec regroupement en fonction de la présence de la maladie - cas- témoin on peut calculer a l'intérieur des groupes avec et sans maladie: $a/(a+c)$, $c/(a+c)$, $b/(b+d)$, $d/(b+d)$

$$OR = \frac{\text{Chance}(F+|M+)}{\text{Chance}(F+|M-)} = \frac{\frac{\text{Probabilité}(F+|M+)}{1-\text{Probabilité}(F+|M+)}}{\frac{\text{Probabilité}(F+|M-)}{1-\text{Probabilité}(F+|M-)}} = \frac{\frac{\frac{a}{a+c}}{\frac{c}{a+c}}}{\frac{\frac{b}{b+d}}{\frac{d}{b+d}}} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a*d}{b*c}$$

31

31

Indicateurs de quantification du lien facteur-maladie:

- Explication – comment l'OR est une approximation du RR pour les maladies rares (prevalence < 5% ou < 10%)
- La prevalence = nombre malades / nombre total des sujets
- La prevalence = $(a + c) / N$
- Si la prevalence est petit, donc $a+c$ est petit, donc a est petit par rapport a le b – donc $a+b \approx b$, et c est petit par rapport a d – donc $c+d \approx d$

$$RR = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{\frac{a}{a+b \approx b}}{\frac{c}{c+d \approx d}} = \frac{a*d}{b*c} \approx OR$$

	M ⁺	M ⁻	
F ⁺	a	b	ne ⁺
F ⁻	c	d	ne ⁻
	a+c	b+d	N

32

32

Indicateurs calculables dans un tableau de contingence en fonction de la collection des données pour les études des facteurs de risque/ pronostiques:

■ **Calculs possibles en fonction du type de collection des données pour les études des facteurs de risque/ pronostiques**

- un seul group / échantillon représentatif / une cohorte
 - Fréquence de la maladie
 - Fréquence du facteur de risque
 - RIE, RIN, RA, RR, FREE, OR
- en fonction du facteur de risque (avec, sans exposition) – exposée – non exposée
 - Calculs sur des lignes:
 - RIE, RIN, RA, RR, FREE, OR
 - en fonction de la maladie (avec, sans maladie) cas-témoins
 - Calculs sur des colonnes:
 - OR

33

33

Indicateurs de quantification du lien (évaluer la force/ importance de la relation) facteur-maladie, en fonction de la collection des données pour les études des facteurs de risque/ pronostiques:

■ **Calculs possibles en fonction du type de collection des données pour les études des facteurs de risque/ pronostiques**

- un seul group / échantillon représentatif / une cohorte
 - RA, RR, FREE, OR
- en fonction du facteur de risque (avec, sans exposition) – exposée non exposée
 - Calculs sur des lignes:
 - RA, RR, FREE, OR
 - en fonction de la maladie (avec, sans maladie) cas-témoins
 - Calculs sur des colonnes:
 - OR

34

34

Quantification de l'importance de la relation pour plusieurs facteurs

Problématique:

- Les **relations** entre différents variables médicales sont souvent **complexes**, et **plusieurs** facteurs/ **variables**/ **caractéristiques** sont **impliquées**.
- Avec le **AR, FREE, RR, OR** (calculée $a*b/c*d$) on peut **évaluer** la **relation** entre **seulement deux variables** – donc une analyse uni varie (entre une variable d'intérêt/ dépendante – et une variable explicative/ indépendante). On le nomme des indicateurs non ajustée, brutes, crûes, non contrôlées.
- Mais **on a besoin de tenir compte de l'effet d'autres variables** (nommée de **confusion**) qui sont connues qu'ils influence la relation qu'on étudie. On doit faire une étude bibliographique pour identifier ces variables de confusion. L'ajustement peut réduire le biais de confusion (mais pas de sélection, d'attrition, de classement)
- Pour tenir compte des plusieurs variables on **doit utiliser des techniques multi variées** (eg. Régression logistique, logistique conditionnelle, Cox, ...)

Examen LCA - ECN

35

35

Quantification de l'importance de la relation pour plusieurs facteurs

■ Types des variables dans une régression logistique:

- La **variable dépendante**/ prédite de la régression (la variable qu'on essaie de prédire) doit être **une** variable **dichotomique**!
 - La présence de la maladie (ou le décès)
 - (ex. diabète, hypertension, infarctus oui/non)
- Les **variables indépendantes** (explicatives) (les variables avec lesquelles on essaie de prédire la variable dépendante) – **une ou plusieurs, de différents types**
 - Qualitatives (facteurs) (ex. consommation d'alcool, tabagisme, obésité)
 - Quantitatives (ex. l'âge)

Examen LCA - ECN

36

36

Quantification de l'importance de la relation pour plusieurs facteurs

Régression logistique uni variée/ simple, multivariée/ multiple, odds ratio crue/ajuste...

- Une régression logistique avec une seule variable indépendante
 - permet une analyse uni variée – on parle d'un régression logistique simple
 - nous offre un odds ratio brut/crue (« crude »/ » unadjusted» en anglais - sans ajustement, pour la variable indépendante
- Une régression logistique avec une plusieurs variables indépendantes
 - permet une analyse multi variée – on parle d'un régression logistique multiple
 - nous offre un odds ratio ajusté (« adjusted » en anglais) pour la variable indépendante – avec ajustement pour les autres variables indépendants dans la régression
- L'odds ratio ajusté peut nous rapprocher plus a la vérité que le odds ratio brut (« crude »/ » unadjusted» – obtenu a l'aide d'une régression logistique simple, ou calcul $a*d/b*c$), parce que on peut tenir compte (on peut contrôler) d'autres facteurs de risque qui agit dans le même temps

Examen LCA - ECN

37

37

Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio crue/brut/non ajuste d'une variable qualitative dichotomique pour un régression logistique simple (uni variée)

- L' équation du régression logistique simple (uni variée), avec un variable indépendante qualitative dichotomique:
- Maladie (oui/non) = $OR_{\text{facteur}_1} * \text{facteur}_1$ (oui/non)
- Ex: cheilopalatoschisis (oui/non) = $1,90 * \text{alcool}$ (oui/non)
- L' interprétation du Odds ratio crue/brut, sans ajustement
- Si la variable est qualitatives dichotomiques (ex. consommation d'alcool):
 - il y a 1,90 fois plus de la chance d'avoir la maladie - cheilopalatoschisis pour ceux avec exposition – avec consommation d'alcool comparée a ceux qui n'ont pas été exposée a l'alcool.

Examen LCA - ECN

38

38

Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio crue/brut/non ajusté d'un variable quantitative pour un régression logistique simple (uni variée)

- L' équation du régression logistique simple (uni variée), avec un variable indépendante quantitative:
 - Maladie (oui/non) = $OR_{variable_1} * variable_1$
 - Ex: cheilopalatoschizis (oui/non) = $1,10 * \text{âge de la mère (années)}$
- L' interprétation du Odds ratio crue/brut, sans ajustement
- Si la variable est quantitative (ex. âge de la mère):
 - pour chaque unité de mesure de la variable (pour chaque année de la mère) en plus la chance d'avoir cheilopalatoschizis augmente 1,10 fois

Examen LCA - ECN

39

39

Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio ajusté d'un variable qualitative dichotomique, et d'un variable quantitative pour un régression logistique multiple (multivariée)

- L' équation du régression logistique multiple (multivariée):
- Maladie (oui/non) = $OR_{facteur_1} * facteur_1 (oui/non) + \dots OR_{variable_n} * variable_n$
- Ex: cheilopalatoschizis (oui/non) = $2,10 * alcool (oui/non) + 1,15 * \text{âge de la mère (années)}$
- L' interprétation du Odds ratio ajusté pour des variables
 - Qualitatives dichotomiques (ex. consommation d'alcool): il y a 2,10 fois plus de la chance d'avoir la maladie - cheilopalatoschizis pour ceux avec exposition – avec consommation d'alcool comparée a ceux qui n'ont pas été exposée a l'alcool si on tiennent les autres variables constantes (on ajuste les autres variables, on contrôle les autres variables) (ici l' âge)
 - Quantitatives (ex. âge de la mère): pour chaque unité de mesure de la variable (pour chaque année de la mère) en plus la chance d'avoir cheilopalatoschizis augmente 1,15 fois si on tiennent les autres variables constantes (on ajuste les autres variables, on contrôle les autres variables) (ici l'alcool)

40

40

Évaluer la causalité

Problématique de l'évaluation de la causalité

- Une **seul étude observationnel** n' est pas suffisante pour tirer des **conclusions fermes de cause à effet** !
- On a besoin des **plusieurs études** (voir diapositif suivant – les critères du **Hill**) **observationnels** pour essayer comprendre la nature causale d'un relation
- Il y a **beaucoup des facteurs qui ne sont pas contrôlée** dans les études observationnels qui peuvent expliquer la relation observe dans l' étude.
- Une **étude expérimental bien réalise** peut montrer la **causalité** (dans cet étude toutes les facteurs sont contrôlées, la seule différence sont les groupes comparées)

41

41

Examen LCA - ECN

Critères nécessaires pour évaluer la causalité (après A.B. Hill)

Critères intrinsèques (a l' intérieur d'un étude) (le résultat doit être significatif $P < 0,05!$)	Critères extrinsèques (dans plusieurs études) (le résultat doit être significatif $P < 0,05!$)
Séquence temporelle : l'exposition doit précéder l'apparition de la maladie, et cela avec un temps de latence compatible avec celui de la maladie. (seulement dans des études longitudinales ou l'absence de la maladie a été prouvée au début de l'étude. N'est pas possible dans des études transversales)	Constance de l'association : la relation doit être retrouvée dans plusieurs études, si possible dans des populations différentes. (eg. OR statistiquement significative et dans la même direction, dans plusieurs études)
Force de la liaison : plus la relation est forte, moins elle peut être expliquée par un tiers facteur. (plus OR , ou RR , ou HR sont importants [eg. 3; 4; 5 – comme facteur de risque, ou 0,5; 0,4 ; 0,2 – comme facteur protectif] ou RA est important [eg. 20%; 30% – comme facteur de risque, ou - 20%; - 30% – comme facteur protectif])	Cohérence temps-espace : si un facteur est cause de maladie, la distribution de la maladie dans le temps et dans l'espace doit être calquée sur celle du facteur. (pays avec prévalence élevée du tabagisme ont prévalence élevée des maladies liées a le tabagisme. L' évolution du prévalence de tabagisme est suivie dans le temps par l' évolution de la prévalence des maladies liées a le tabagisme)
Relation dose-effet : plus l'exposition est longue et/ou intense, plus la fréquence de la maladie doit être élevée, analyses qui montre une tendance croissante (eg. OR plus grand pour des su groupes avec expositions plus grandes, par rapport a de su groupes avec expositions plus réduites)	Relation dose-effet inter-études : entre études à taux d'exposition différents, la fréquence de la maladie doit être en rapport. (eg. OR plus grand dans les études avec expositions plus grandes)
Effet de la suppression de l'exposition (réversibilité de l' effet) : si, à l'inverse l'exposition est supprimée, on s'attend à une diminution de la fréquence de la maladie. (eg. OR plus petit pour le group avec suppression de l'exposition par rapport a le group sans suppression de l'exposition)	Spécificité de la relation : rarement vérifié. Le facteur doit être nécessaire à l'apparition d'une seule maladie. (eg. L'asbeste produit le mésothéliome – un cancer de la pleure qui entoure le poulmon)
Cohérence dans les sous-populations : la relation doit se retrouver , par exemple, chez les hommes et les femmes de l'étude. (eg. OR statistiquement significatif dans différents su groupes et dans la même direction)	Apport des études expérimentales : le contrôle du facteur entraîne un contrôle de la fréquence de la maladie. Par exemple dans des études sur des animaux.
Indépendance de l'association vis-à-vis des autres facteurs connus : dans une analyse multi variée le OR reste statistiquement significative – donc facteur de risque indépendant)	Plausibilité biologique et physiopathologique.

42

42

Critères pour la generalisabilite des résultats

- L'inclusion proche de l' exhaustivité (presque toute la population) des sujets dans l' étude
- Un échantillon de grande taille
- Plusieurs villes
- Plusieurs pays
- Une prévalence similaire du chose d' intérêt (maladie) a d'autres études dans le même pays
- Le tirage au sort (aléatoire) des plusieurs hôpitaux, cabinets de médecine de famille, ...
- Une étude prospective

Examen LCA - ECN

43

43

EVALUATION DE LA VALIDITÉ INTERNE DE L'ÉTUDE AVEC NEWCASTLE OTAWA SCALE +NHLBI

- Pour les études **cas- témoin**
- La **sélection de sujets** (influence le biais de sélection)
 - Les **critères d'inclusion** dans les échantillons ont-ils été **clairement définis**? et Les **sujets d'étude** et le **cadre** (type d'hôpital, cabinet de médecin de famille) ont-ils été **décrits en détail**?
 - Les **cases**
 - 1) Le **diagnostic** de maladie est **sur** (idéal: critères **objectives, standardisées**, valide, fiable, et mis en œuvre de manière cohérente chez tous les participants à l'étude, des **preuves de bonne qualité** (radiographie, dossier médical, Codification de la maladie dans une base des données, réalisée par un médecin ou commission diagnostique-idéal avec validation indépendante par plusieurs personnes)
 - 2) Les cas sont **représentatives** (série **consécutive** des cas [toutes les cas], ou échantillon **aléatoire**, qui vient dans l'hôpital ou populationnel – base des données du pays)
 - Les **contrôles**
 - 3) La **définition de control** – (voir les cases) explicite écrit que ils n'ont pas eu la maladie
 - 4) La **sélection des contrôles** - sélectionnées depuis la même population que les cas, et eux peuvent être des cas aussi, si ils ont eux la maladie (including the same timeframe)
- 5) La **comparabilité** de sujets (influence le biais de confusion)
 - 5.1. Les auteurs ont identifiée des **possibles facteurs de confusion essentiels** et ont les ajustées?
 - 5.2. Des **techniques d'aide a la comparabilité**:
 - La présence de l'**ajustement** (meilleure comparabilité) – une analyse multivariée une **analyse multivariée** (eg. **Régression** logistique, Cox, ANCOVA, **propensity score matching**) avec ajustement pour plusieurs facteurs de confusion (eg. Age, comorbidités, autre facteurs de risque qui influencent la maladie (ou critère de jugement)
 - **Design appariée** - Un design d' étude avec appariement pour plusieurs facteurs qui influence la maladie (sexe, diabète, âge, ...) et/ou **analyse stratifiée** et/ou **exclusion des sujets avec facteurs de confusion**
 - Valeurs similaires des facteurs de confusion (et différences non statistiquement significatifs) ne constituent pas une indice de comparabilité! Toujours on va avoir du biais du confusion résiduel
- L'**évaluation de l'exposition** au facteur/s de risqué (influence le biais d'observation, de classification, d'information)
 - 6) La **présence/absence du facteur d'exposition est sure**, ou l' évaluation du facteur s'exposition est sure, objective, valide, fiable (eg. Un sujet qui consomme alcool affirme qu'il ne consomme pas d'alcool – n'est pas sure) et mis en œuvre de manière cohérente (y compris sur la même période) pour tous les participants à l'étude ?
 - 7) La **manqué d'information sur le facteur d'exposition** (eg. biais non répondeur)
 - 8) Les enquêteurs ont-ils pu **confirmer** que l'**exposition/le risque** s'est produit **avant** le **développement** de la **maladie** condition ou de l'événement qui a défini un participant comme un cas ?
 - 9) Les **évaluateurs** de l'**exposition/du risque** **n'étaient-ils pas informés** du statut de **cas ou de contrôle** des participants ?

<https://www.cribi.ca/health/health-topics/otawa-otawa-measurement-tool>
<https://www.cribi.ca/otawa/otawa-measurement-tool>

44

Evaluation de la validité interne de l'étude avec Newcastle Ottawa scale +NHLBI

Pour les études de **cohorte**

- La **sélection** de sujets (influence le biais de sélection)
 - Les critères d'inclusion dans les échantillons ont-ils été clairement définis? et Les sujets d'étude et le cadre (type d'hôpital, cabinet de médecin de famille) ont-ils été décrits en détail?
- L'**exposition**
 - 1) Les sujets exposés sont **représentatives** pour la situation d'intérêt (série **consécutives** des cas [toutes les cas], ou échantillon **aléatoire**, qui vient dans l'hôpital/ou populationnel – base des données du pays)
 - 2) La **présence/absence du facteur d'exposition est sûre**, ou l'évaluation du facteur s'exposition est sûre (idéal: critères **objectives, standardisées, valides, fiables**) Un sujet qui consomme alcool affirme qu'il ne consomme pas d'alcool - influence le biais d'observation, de classification, d'information) et mis en œuvre de manière cohérente (y compris sur la même période) pour tous les participants à l'étude?
 - 3) La **sélection des non exposés** - sélectionnés depuis la même population que les exposés
- 4) La preuve que l'objectif (le critère de jugement [eg. La maladie], n'était pas présent au début de l'étude
- 5) La **comparabilité** de sujets (influence le biais de confusion)
 - 5.1. Les auteurs ont identifié des possibles facteurs de confusion essentiels et ont les ajustés?
 - 5.2. Des techniques d'aide à la comparabilité:
 - La présence de l'ajustement (meilleure comparabilité) – une **analyse multivariée** (eg. **Régression** logistique, Cox, ANCOVA, **propensity score matching**) avec ajustement pour plusieurs facteurs de confusion (eg. Age, comorbidités, autre facteurs de risque qui influencent la maladie (ou critère de jugement)
 - **Design appariée** - Un design d'étude avec appariement pour plusieurs facteurs qui influence la maladie (sexe, diabète, âge, ...) et/ou **analyse stratifiée** ou/et **exclusion des sujets avec facteurs de confusion**
 - Valeurs similaires des facteurs de confusion (et différences non statistiquement significatifs) ne constituent pas une indice de comparabilité!
- L'évaluation de l'objectif (le critère de jugement [eg. La maladie]) (influence le biais d'observation, de classification, d'information)
 - 6) La présence/absence du l'objectif (le critère de jugement [eg. La maladie]) est **sûre** (idéal: critères **objectives, standardisées, valide, fiable**, et mis en œuvre de manière cohérente chez tous les participants à l'étude; des **preuves de bonne qualité** (radiographie, dossier médical, Codification de la maladie dans une base des données, réalisée par un médecin ou commission diagnostique))
 - **Evaluation indépendante et en aveugle ou preuves de bonne qualité (voir avant)**
 - 7) Le **temps de suivi des patients a été suffisant** pour que l'objectif (le critère de jugement [eg. La maladie]) peut arriver
 - 8) la **perte de plus de 20% des sujets**, ou la **perte des sujets est inégale entre les exposés et les non exposés** (eg. Biais d'attrition, biais de sélection)

45

Evaluation de la validité interne de l'étude avec Newcastle Ottawa scale+NHLBI

Pour les études **transversales** (cross-sectional – en anglais), avec un seul échantillon

- La **sélection** de sujets (influence le biais de sélection)
 - Les critères d'inclusion dans l'échantillon ont-ils été clairement définis? et Les sujets d'étude et le cadre (type d'hôpital, cabinet de médecin de famille) ont-ils été décrits en détail?
 - 1. Les sujets sont **représentatives** pour la situation d'intérêt (série **consécutives** des cas [toutes les cas], ou échantillon **aléatoire**, qui vient dans l'hôpital)
- L'**observation/la mesure des variables** (influence le biais de mesure/ observation)
 - L'**exposition**
 - 2. La **présence/absence du facteur d'exposition est sûre**, ou l'évaluation du facteur s'exposition est sûre (idéal: critères **objectives, standardisées, valides, fiables**)
 - L'**objectif**
 - 3. La présence/absence du l'objectif (le critère de jugement [eg. La maladie]) est **sûre** (idéal: critères **objectives, standardisées, valide, fiable**, et mis en œuvre de manière cohérente chez tous les participants à l'étude; des **preuves de bonne qualité** (radiographie, dossier médical, Codification de la maladie dans une base des données, réalisée par un médecin ou commission diagnostique-idéal avec validation indépendante par plusieurs personnes))
- 4. La **comparabilité** de sujets (influence le biais de confusion)
 - 4.1. Les auteurs ont identifié des possibles facteurs de confusion essentiels et ont les ajustés?
 - 4.2. Des techniques d'aide à la comparabilité:
 - La présence de l'ajustement (meilleure comparabilité) – une **analyse multivariée** (eg. **Régression** logistique, Cox, ANCOVA, **propensity score matching**) avec ajustement pour plusieurs facteurs de confusion (eg. Age, comorbidités, autre facteurs de risque qui influencent la maladie (ou critère de jugement)
 - **Design appariée** - Un design d'étude avec appariement pour plusieurs facteurs qui influence la maladie (sexe, diabète, âge, ...) et/ou moins bien: **analyse stratifiée** ou/et **exclusion des sujets avec facteurs de confusion**
 - Valeurs similaires des facteurs de confusion (et différences non statistiquement significatifs) ne constituent pas une indice de comparabilité!

* Le checklist a été adapté didactique

46

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

- l'intervalle de confiance indique si le résultat est statistiquement significative
 - Si la valeur correspondante a l'hypothèse nulle du test statistique est en dehors de l'intervalle de confiance – l'hypothèse nulle est rejeté, et l'hypothèse alternative est acceptée. Donc $p < 0,05$. Donc résultat statistiquement significatif
 - Si la valeur correspondante a l'hypothèse nulle du test statistique est dans l'intérieur l'intervalle de confiance – l'hypothèse nulle n'est pas rejeté - mais elle n'est pas acceptée!). Donc $p > 0,05$

Examen LCA - ECN

47

47

Examen LCA - ECN

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

- Pour des indicateurs de type taux (risque relatif - RR, odds ratio - OR)
 - L'hypothèse nulle: il n'y a pas de différence entre les groupes
 - RR – risque relatif ($= RIE/RIN$): $RIE = RIN$ (risque des exposées est égale a le risque des non exposées) $\Rightarrow RR = 1$
 - OR – odds ratio ($= OE/ON$) – le taux de la chance: $OE = ON$ (la chance des exposées est égale a le chance des non exposées) $\Rightarrow OR = 1$
 - L'hypothèse alternative: il y a une différence entre les groupes $RR \neq 1$, $OR \neq 1$
 - Si 1 est en dehors de l'intervalle de confiance – l'hypothèse nulle est rejeté, et l'hypothèse alternative est acceptée. Donc $p < 0,05$. Donc résultat statistiquement significatif
 - Si 1 est dans l'intervalle de confiance – l'hypothèse nulle n'est pas rejeté - mais elle n'est pas acceptée!). Donc $p > 0,05$

48

48

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

- Pour des indicateurs de type taux (RR, OR)
 - IC 95% de RR de 2,3 a 5
 - La valeur correspondante a l' hypothèse nulle ($RR = 1$) du test statistique est en dehors de l'intervalle de confiance – l' hypothèse nulle est rejeté, et l' hypothèse alternative est acceptée. Donc $p < 0,05$
 - IC 95% de OR de 0.4 a 0.9
 - La valeur correspondante a l' hypothèse nulle ($OR = 1$) du test statistique est en dehors de l'intervalle de confiance – l' hypothèse nulle est rejeté, et l' hypothèse alternative est acceptée. Donc $p < 0,05$

49

49

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

- Pour des indicateurs de type taux (RR, OR)
 - IC 95% de RR de 0,6 a 30
 - La valeur correspondante a l' hypothèse nulle ($RR = 1$) du test statistique est dedans l'intervalle de confiance – l' hypothèse nulle n'est pas rejeté. Donc $p > 0,05$, le résultat n'est pas statistiquement significative
 - IC 95% de OR de 0,4 a 50
 - La valeur correspondante a l' hypothèse nulle ($OR = 1$) du test statistique est dedans l'intervalle de confiance – l' hypothèse nulle n'est pas rejeté. Donc $p > 0,05$, le résultat n'est pas statistiquement significative

50

50

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

■ Pour des indicateurs de type différence des pourcentages (Risque attribuable - RA)

- L' hypothèse nulle: il n'y a pas de différence entre les groups
 - RA – risque attribuable (=RIE-RIN): RIE = RIE (risque des exposées est égale a le risque des non exposées) => RA = 0
- L' hypothèse alternative: il y a une différence entre les groups $\neq 0$
- Si 0 est en dehors de l'intervalle de confiance – l' hypothèse nulle est rejeté, et l' hypothèse alternative est acceptée. Donc $p < 0,05$. Donc **résultat statistiquement significatif**
- Si 0 est dans l'intervalle de confiance – l' hypothèse nulle n'est pas rejeté - mais elle n'est pas acceptée!). Donc $p > 0,05$

Examen LCA - ECN

51

51

Interpreter si le resultat est statistiquement significative ($p < 0,05$) en regardant l'intervalle de confiance

■ Pour des indicateurs de type différence des pourcentages (RA)

- IC 95% de RA de 10% a 20%
 - La valeur correspondante a l' hypothèse nulle (RA = 0%) du test statistique est en dehors de l'intervalle de confiance – l' hypothèse nulle est rejeté, et l' hypothèse alternative est acceptée. Donc $p < 0,05$, le résultat est statistiquement significative
- IC 95% de RA de -10% a 20%
 - La valeur correspondante a l' hypothèse nulle (RA = 0%) du test statistique est dedans l'intervalle de confiance – l' hypothèse nulle n'est pas rejeté. Donc $p > 0,05$, le résultat n'est pas statistiquement significative

52

52

Exemple article - MD

BMC Public Health. 2012 Nov 16;12:989. doi: 10.1186/1471-2458-12-989. Available from: <http://www.biomedcentral.com/1471-2458/12/989>

Assessment of dental caries predictors in 6-year-old school children - results from 5-year retrospective cohort study.

Masood M¹, Yusof N, Hassan MI, Jaafar N.

BACKGROUND: This was a **retrospective cohort** study undertaken to assess the rate and pattern of dental caries development in 6-year-old school children followed-up for a period of 5 years, and to identify baseline risk factors that were associated with 5 years caries experience in Malaysian children.

METHODS: This 5-years retrospective cohort study comprised primary school children initially aged 6 years in 2004. Caries experience of each child was recorded annually using World Health Organization criteria. The rates of dental caries were recorded in prevalence and incidence density of carious lesions from baseline to final examination. Risk assessment was done to assess relative risk for caries after 5 years in children with baseline caries status. Simple and multiple logistic regression analysis were performed to identify significant independent risk factors for caries.

RESULTS: The sample consisted of 1830 school children. All components of DMFT showed significant differences between baseline and final examination. Filled teeth (FT) component of the DMFT showed the greatest increases. Results revealed the **initial baseline caries level in permanent dentition was a strong predictor for future caries after 5 years (RR=3.78, 95% CI=3.48-4.10, P<0.001). Logistic regression analysis showed significant association between caries occurrence and residence (urban/rural) (OR=1.80, P<0.001).** However, it was not significantly associated with gender and ethnicity. The incidence of 5.80 persons/100 person-year of observation after 5-years was 0.76 teeth/100 teeth.

CONCLUSION: The majority of 1 concentrated in only a small proportion of 6 years was a strong predictor permanent teeth caries at six years of almost no cost, questions the need for and cost-effectiveness of expensive technology-based commercial caries predictions kits.

Table 4 univariate and Multivariate logistic regression model for independent variables' effect on caries increment at baseline and after five years

Independent Variable			Crude			Adjusted		
			OR	95% CI ^{††}	P [†]	OR	95% CI ^{††}	P [†]
Gender	Male	Female	1.19	0.97-1.46	0.088	1.16*	0.95-1.41	0.143
Urban/Rural			1.80	1.46-2.22	<0.001	1.75***	1.42-2.1	<0.001

Odds ratio adjusted for ethnicity and urban/rural (*), gender and urban/rural (**), gender and ethnicity (***)

53

Exemple article - MD

BMC Public Health. 2012 Nov 16;12:989. doi: 10.1186/1471-2458-12-989. Available from: <http://www.biomedcentral.com/1471-2458/12/989>

Assessment of dental caries predictors in 6-year-old school children - results from 5-year retrospective cohort study.

Masood M¹, Yusof N, Hassan MI, Jaafar N.

Results revealed the **initial baseline caries level in permanent dentition was a strong predictor for future caries after 5 years (RR=3.78, 95% CI=3.48-4.10, P<0.001). Logistic regression analysis showed significant association between caries occurrence and residence (urban/rural) (OR=1.80, P<0.001).**

Interpretations:

Du RR - il y a 3,78 fois plus de **risque** des futures caries (dans les suivant 5 ans) pour ceux qui ont eu des caries au première évaluation par rapport au ceux qui n'ont pas eu

Du OR **crude - brut** (voir tableau - Crude) - (brut signifie qu'il s'agit - entre deux variables seulement - une variable dépendante (la présence des caries) et une variable indépendante (le milieu urbain/rural) - d'un analyse univariée - une seule variable indépendante) il y a 1.8 fois plus de **chance** des futures caries (dans les suivant 5 ans) pour ceux qui sont dans le milieu urbain par rapport au ceux qui sont dans le milieu rural.

Du OR **ajust - adjusted** (voir tableau - Adjusted) - (ajusté signifie qu'il s'agit d'un analyse entre une variable dépendante (la présence des caries) et plusieurs variables indépendantes (le milieu urbain/rural et le sexe - gender)- analyse multivariée - plusieurs variables indépendantes) il y a 1.8 fois plus de **chance** des futures caries (dans les suivant 5 ans) pour ceux qui sont dans le milieu urbain par rapport au ceux qui sont dans le milieu rural, **on tenant constantes (on ajustant, on contrôlant) les autres variables (sexe - gender).**

Les variables qu'on ajuste/qu'on contrôle avec la régression logistique sont des facteurs de confusion (en anglais - **confounding factors**)

Table 4 univariate and Multivariate logistic regression model for independent variables' effect on caries increment at baseline and after five years

Independent Variable			Crude			Adjusted		
			OR	95% CI ^{††}	P [†]	OR	95% CI ^{††}	P [†]
Gender	Male	Female	1.19	0.97-1.46	0.088	1.16*	0.95-1.41	0.143
Urban/Rural			1.80	1.46-2.22	<0.001	1.75***	1.42-2.1	<0.001

34

54

Exemple article - MG

Environ Health. 2015 Jan 7;14(1):1. Vallières E, Pintos J, Parent ME, Siemiatycki J. Available from: <http://www.chjournal.net/content/14/1/1>

Occupational exposure to wood dust and risk of lung cancer in two population-based case-control studies in Montreal, Canada.

BACKGROUND: Wood dust is one of the oldest and one of the most common occupational exposures in the world. The present analyses examine the effect of lifetime exposure to wood dust in diverse occupational settings on lung cancer risk.

METHODS: We conducted two population-based case-control studies in Montreal: Study I (1979-1986) included 857 cases and two sets of controls (533 population and 1349 cancer controls), and Study II (1996-2001) comprised 736 cases and 894 population controls. Detailed job histories were obtained by interview and each job was evaluated by expert chemist-hygienists to estimate the likelihood and level of exposure to many substances, one of which was wood dust. Odds ratios (ORs) were computed in relation to different indices of exposure to wood dust, adjusting for several covariates including smoking. Three datasets were analysed: Study I with population controls, Study I with cancer controls, and Study II.

RESULTS: The most frequently exposed occupations in our study population were in construction, timber and furniture making industries. We found **increased risks of lung cancer for substantial cumulative exposure to wood dust in Study I with cancer controls, (OR = 1.4: 95% confidence interval 1.0;-2.0) and in Study II (OR = 1.7: 95% confidence interval 1.1-2.7).** There were no excess risks of lung cancer in any of the three datasets among workers whose cumulative exposure was not substantial. These tendencies held equally within strata of low smokers and heavy smokers.

CONCLUSION: There was evidence of increased risk of lung cancer among workers with substantial cumulative exposure to wood dust.

55

55

Exemple article

Table 3 Odds ratio for lung cancer associated with occupational exposure wood dust in two case-control studies

	Controls	Cases	OR ₁ *	OR ₂ **	95% CI (OR ₂)	
Study I						
	Population Controls					
No exposure	389	630	1.0	1.0	(ref)	
Any level of exposure	144	227	0.8	0.7	0.5	1.0
Any level ≤ 20 years	88	141	0.8	0.7	0.5	1.0
Any level > 20 years	56	86	0.7	0.7	0.4	1.1
Non-substantial level	74	113	0.8	0.7	0.5	1.0
Substantial level	70	114	0.7	0.7	0.5	1.0
Study I						
	Cancer Controls					
No exposure	1072	630	1.0	1.0	(ref)	
Any level of exposure	277	227	1.2	1.1	0.9	1.5
Any level ≤ 20 years	179	141	1.2	1.1	0.8	1.5
Any level > 20 years	98	86	1.2	1.2	0.8	1.7
Non-substantial level	161	113	1.0	1.0	0.7	1.3
Substantial level	116	114	1.5	1.4	1.0	2.0
Study II						
	Population Controls					
No exposure	640	501	1.0	1.0	(ref)	
Any level of exposure	254	235	1.2	1.1	0.9	1.5
Any level ≤ 20 years	165	139	1.1	1.0	0.8	1.4
Any level > 20 years	89	96	1.4	1.3	0.9	1.9
Non-substantial level	201	167	1.1	1.0	0.7	1.3
Substantial level	53	68	1.9	1.7	1.1	2.7

*adjusted for age, ethnolinguistic group, years of education, median family income, respondent status and cigarette index.

**adjusted for the same covariates as above, as well as IARC Group 1 occupational carcinogens (asbestos, diesel exhaust, formaldehyde, cadmium, chromium VI, nickel and silica).

56

Exemple article - MG

Environ Health. 2015 Jan 7;14(1):1. Vallières E, Pintos J, Parent ME, Siemiatycki J. Available from: <http://www.chjournal.net/content/14/1/1>

Occupational exposure to wood dust and risk of lung cancer in two population-based **case-control** studies in Montreal, Canada.

RESULTS: The most frequently exposed occupations in our study population were in construction, timber and furniture making industries. We found **increased risks of lung cancer for substantial cumulative exposure to wood dust in Study I with cancer controls, (OR = 1.4: 95% confidence interval 1.0;-2.0) and in Study II (OR = 1.7: 95% confidence interval 1.1-2.7).**

Study	OR	95% CI	OR	95% CI	OR	95% CI
Substantial level	53	68	1.9	1.7	1.1	2.7

*adjusted for age, ethnolinguistic group, years of education, median family income, respondent status and cigarette index.

**adjusted for the same covariates as above, as well as IARC Group 1 occupational carcinogens (asbestos, diesel exhaust, formaldehyde, cadmium, chromium VI, nickel and silica).

Interpretation

Du OR **ajust - adjusted** (voir tableau) – (ajusté signifie qu'il s'agit d'une analyse entre une variable dépendante (la présence du cancer du poumon) et plusieurs variables indépendantes/explicatives (la poussière de bois, et l'âge, groupe ethnolinguistique, ..., autres carcinogènes, asbeste, formaldéhyde, cadmium, ...) – analyse multivariée) il y a 1.7 fois plus de **chance** de cancer du poumon pour ceux qui ont été exposés au travail à la poussière de bois par rapport aux ceux qui n'ont pas été exposés au travail à la poussière, on tenant constantes (on ajustant, on contrôlant) les autres variables (l'âge, groupe ethnolinguistique, ..., autres carcinogènes, asbeste, formaldéhyde, cadmium, ...).

Les variables qu'on ajuste/qu'on contrôle avec la régression logistique sont des facteurs de confusion (en anglais – confounding factors)

57

57

Exemples des questions pour l'examen

1) Une étude a été réalisée pour évaluer la relation entre la consommation exagérée du the noir comme facteur de risque pour l'apparition de l'hypertension. Ils ont obtenu : RR=1,5 ; OR=1,9 ; RIE (avec exposition exagérée) =0,3 ; RIN (avec exposition normale au the) =0,20. Les résultats sont statistiquement significatifs. L'hypertension est une maladie fréquente. Lesquelles des suivantes affirmations sont correctes ?

- a) il y a 10% moins de risque d'avoir hypertension pour ceux avec consommation normale du the par rapport aux ceux avec la consommation exagérée du the
- b) le risque d'avoir hypertension pour ceux avec consommation exagérée du the est plus grand par rapport aux ceux avec consommation normale du the
- c) il y a 1,5 fois plus de chance d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the
- d) il y a 1,9 fois plus de risque d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the
- e) il y a 1,9 fois plus de chance d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the

Réponse: a, b, e

2) Une étude a été réalisée pour évaluer la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'hypertension. Ils ont obtenu : RR=0,5 ; OR=0,4 ; RIE (avec exposition au the)=0,15 ; RIN (sans exposition au the)=0,30. Les résultats sont statistiquement significatifs. L'hypertension est une maladie fréquente. Lesquelles des suivantes affirmations sont correctes ?

- a) il y a 15% moins de risque d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- b) il y a 0,5 fois plus de risque d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- c) il y a 0,5 fois plus de chance d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- d) il y a 0,4 fois plus de risque d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- e) il y a 0,4 fois plus de chance d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the

Réponse: a, b, e

58

Exemples des questions pour l'examen

5) * Dans une étude d'évaluation de thé vert comme un facteur de protectif pour l'apparition du cancer, en utilisant une collection de données de type expose-non expose, selon les 4000 buveurs de thé vert 20 ont développé un cancer et selon les 2990 sujets qui n'ont pas bu du thé vert 10 ont développé un cancer. Calculez le RIE et le RIN.

- a) 0.5% et 0.3%
- b) ne peut pas être calculé pour ce type d'étude
- c) 3% et 5%
- d) 99.7% et 0.3%
- e) 66.67% et 42.8%

Réponse: a

6) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 200 buveurs de café 60 ont eu un changement important du couleur des dents et selon les 100 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Odds Ratio.

- a) 0,1
- b) 1,5
- c) ne peut pas être calculé pour ce type d'étude
- d) 1,71
- e) 1,7%

Réponse: d

59

Exemples des questions pour l'examen

7) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 240 buveurs de café 80 ont eu un changement important du couleur des dents et selon les 140 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Risque relatif et le Risque attribuable.

- a) 2,33 et 0,19
- b) ne peut pas être calculé pour ce type d'étude
- c) 2,33 et 19%
- d) 3 et 0,19
- e) 3 et 0,33

Réponse: b

8) Quelles affirmations, en concernant la régression logistique pour plusieurs facteurs avec l'équation : $\text{infarctus myocardique (oui/non)} = 1,5 * \text{hypertension (oui/non)} + 1,10 * \text{âge (années)}$, sont correctes:

- a) la variable dépendante de la régression est l'hypertension
- b) la variable dépendante de la régression est l'infarctus myocardique
- c) une variable indépendante (explicative) est l'hypertension
- d) il y a 1,5 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont exposés à l'hypertension comparé à ceux qui ne sont pas exposés en ajustant pour la variable âge
- e) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir infarctus myocardique 1,10 fois en ajustant pour la variable hypertension

Réponse: b, c, d, e

60

Exemples des questions pour l'examen

9) Quelles affirmations, en concernant la régression logistique pour plusieurs facteurs avec l'équation : $\text{infarctus myocardique (oui/non)} = 1,5 * \text{hypertension (oui/non)} + 1,10 * \text{âge (années)}$, sont correctes:

- a) l'odds ratio ajustée pour la variable hypertension est 1,5
- b) l'odds ratio ajustée pour la variable âge est 1,1
- c) l'odds ratio crude/brut pour la variable hypertension est 1,5
- d) il y a 1,5 fois plus de risque d'avoir l'infarctus myocardique pour ceux qui sont exposés à l'hypertension comparés à ceux qui ne sont pas exposés sans ajustement pour la variable âge
- e) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir un infarctus myocardique 1,10 fois sans ajustement pour la variable hypertension

Réponse: a, b

10) Quelles affirmations, en concernant la régression logistique pour plusieurs facteurs avec l'équation : $\text{infarctus myocardique (oui/non)} = 1,5 * \text{hypertension (oui/non)} + 1,10 * \text{âge (années)}$, sont correctes:

- a) cette régression est une analyse multivariée
- b) la variable l'hypertension est une variable de confusion (confounder en anglais) pour l'âge
- c) la variable âge est une variable de confusion (confounder en anglais) pour l'hypertension
- d) il y a 1,5 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont exposés à l'hypertension comparés à ceux qui ne sont pas exposés en contrôlant pour la variable âge
- e) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir un infarctus myocardique 1,10 fois en tenant constante la variable hypertension

Réponse: a, b, c, d, e

61

Exemples des questions pour l'examen

11) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 200 buveurs de café 60 ont eu un changement important du couleur des dents et selon les 100 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Risque relatif.

- a) 0,1
- b) 1,5
- c) ne peut pas être calculé pour ce type d'étude
- d) 1,71
- e) 1,7%

Réponse: c

12) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 200 buveurs de café 60 ont eu un changement important du couleur des dents et selon les 100 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Risque attribuable.

- a) 0,1
- b) 1,5
- c) ne peut pas être calculé pour ce type d'étude
- d) 1,71
- e) 1,7%

Réponse: c

62

Exemples des questions pour l'examen

12) Une étude a été réalisée pour évaluer la relation entre la consommation exagérée du the noir comme facteur de risque pour l'apparition de l'hypertension. Ils ont obtenu : $RR=1,5$; $OR=1,9$; RIE (avec exposition au the) $=0,3$; RIN (sans exposition au the) $=0,20$. Les résultats sont statistiquement significatifs. L'hypertension est une maladie fréquente. Lesquelles des suivantes affirmations sont correctes ?

- a) il y a 10% plus de risque d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the
- b) il y a 1,5 fois plus de risque d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the
- c) il y a 10% plus de chance d'avoir hypertension pour ceux avec consommation exagérée du the par rapport aux ceux avec consommation normale du the
- d) il y a 30% probabilité d'avoir hypertension pour ceux avec consommation exagérée du the
- e) il y a 20% probabilité d'avoir hypertension pour ceux avec consommation normale du the

Réponse: a, b, d, e

13) Une étude a été réalisée pour évaluer la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'hypertension. Ils ont obtenu : $RR=0,5$; $OR=0,4$; RIE (avec exposition au the) $=0,15$; RIN (sans exposition au the) $=0,30$. Les résultats sont statistiquement significatifs. L'hypertension est une maladie fréquente. Lesquelles des suivantes affirmations sont correctes ?

- a) il y a 15% plus de risque d'avoir hypertension pour ceux qui ne consomment pas du the par rapport aux ceux avec consommation modérée du the
- b) il y a moins de risque d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- c) il y a plus de risque d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the
- d) la consommation modérée du the est un facteur protectif pour l'hypertension
- e) il y a moins de chance d'avoir hypertension pour ceux avec consommation modérée du the par rapport aux ceux qui ne consomment pas du the

Réponse: a, b, d, e

63

Exemples des questions pour l'examen

13) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 200 buveurs de café 60 ont eu un changement important du couleur des dents et selon les 100 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Risque individuel des non exposées (RIN).

- a) 0,1
- b) 1,5
- c) ne peut pas être calculé pour ce type d'étude
- d) 1,71
- e) 1,7%

Réponse: c

14) * Dans une étude d'évaluation de café comme un facteur de risque pour le changement du couleur des dents, en utilisant une collection de données de type cas témoin, selon les 200 buveurs de café 60 ont eu un changement important du couleur des dents et selon les 100 sujets qui ne boivent pas du café 20 ont eu un changement important du couleur des dents. Calculez le Risque individuel des exposées (RIE).

- a) 0,1
- b) 1,5
- c) ne peut pas être calculé pour ce type d'étude
- d) 1,71
- e) 1,7%

Réponse: c

64

Exemples des questions pour l'examen

15) * Dans études des facteurs de risque avec une collection des données cas témoin on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) Aucune de ces indicateurs

Réponse: c

16) Dans études des facteurs de risque avec une collection des données exposée-non exposée on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) Aucune de ces indicateurs

Réponse: a, b, c, d

17) * Dans études des facteurs de risque avec une collection des données échantillon représentatif pour quantifier la force de la relation entre un facteur de risque et une maladie on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) RIN

Réponse: b, c, d

65

Exemples des questions pour l'examen

18) * Dans études des facteurs de risque avec une collection des données cas témoin pour quantifier la force de la relation entre un facteur de risque et une maladie on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) Aucune de ces indicateurs

Réponse: c

19) * Dans études des facteurs de risque avec une collection des données exposée-non exposée pour quantifier la force de la relation entre un facteur de risque et une maladie on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) RIN

Réponse: b, c, d

66

Exemples des questions pour l'examen

20) Dans études des facteurs de risque avec une collection des données échantillon représentatif on peut calculer les suivantes indicateurs médicales

- a) RIE
- b) RA
- c) OR
- d) RR
- e) Aucune de ces indicateurs

Réponse: a, b, c, d

21) * Dans études des facteurs de risque avec une collection des données échantillon représentatif pour quantifier l'importance de la relation entre un facteur de risque et une maladie on peut calculer les suivantes indicateurs médicales

- a) la prévalence de la maladie
- b) la fréquence du facteur de risque
- c) la fréquence de la maladie
- d) RR
- e) Aucune de ces indicateurs

Réponse: d

22) Dans études des facteurs de risque avec une collection des données échantillon représentatif pour évaluer la relation entre un facteur de risque et une maladie on peut calculer les suivantes indicateurs médicales

- a) la prévalence de la maladie
- b) la fréquence du facteur de risque
- c) la fréquence de la maladie
- d) RR
- e) Aucune de ces indicateurs

Réponse: a, b, c, d

67

Exemples des questions pour l'examen

	OR unadjusted	95% CI	P-value	OR adjusted	95% CI	P-value
Obesite	1,9	(1,3-4,9)	<0,001	1,7	(1,2-4,5)	<0,001
Age (annees)	1,15	(1,07-3,9)	<0,001	1,09	(1,04-3,1)	<0,001

23) Quelles affirmations, en concernant la régression logistique uni variée et multiple (multivariée) dans le tableau pour prédire l'infarctus myocardique en fonction de l'obésité et l'âge, sont correctes:

- a) l'odds ratio ajustée pour la variable obésité est 1,7
- b) l'odds ratio ajustée pour la variable âge est 1,09
- c) l'odds ratio crude/brut pour la variable obésité est 1,9
- d) l'odds ratio unadjusted pour la variable âge est 1,15
- e) il y a 1,7 fois plus de risque d'avoir l'infarctus myocardique pour ceux qui sont obèse compare à ceux qui ne sont pas obèse avec ajustement pour la variable âge
- f) il y a 1,7 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont obèse compare à ceux qui ne sont pas obèse avec ajustement pour la variable âge
- g) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir infarctus myocardique 1,15 fois sans ajustement pour la variable obésité
- h) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir infarctus myocardique 1,09 fois avec ajustement pour la variable obésité
- i) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir infarctus myocardique 1,09 fois en contrôlant pour la variable obésité
- j) chaque unité de mesure de la variable âge (chaque année) en plus augmente la chance d'avoir infarctus myocardique 1,09 fois en tenant constante la variable obésité
- k) il y a 1,7 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont obèse compare à ceux qui ne sont pas obèse sans contrôler pour la variable âge

Réponse: a, b, c, d, f, g, h, i, j

68

Exemples des questions pour l'examen

24) Vous avez trouvé deux études sur Pubmed qui ont été réalisées pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu le suivant OR avec l'intervalle de confiance : étude A : 2 (95% CI 1,15-41), et étude B : 2,9 (95% CI, 2,4 – 15). On considère que les valeurs supérieures à 2 sont cliniquement importants, et les valeurs inférieures à 1,3 – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'étude B a un résultat cliniquement plus important que le l'étude A
- b) l'étude A est un nombre plus grand des patients que l'étude B
- c) l'étude A n'a pas un résultat statistiquement significatif
- d) l'estimateur ponctuel d'étude A est 2
- e) l'intervalle de confiance d'étude B est cliniquement important

Réponse: a, d, e

25) Vous avez trouvé deux études sur Pubmed qui ont été réalisées pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu le suivant RR avec l'intervalle de confiance : étude A : 2 (95% CI 1,15-41), et étude B : 2,9 (95% CI, 0,7 – 65). On considère que les valeurs supérieures à 2 sont cliniquement importants, et les valeurs inférieures à 1,3 – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'étude A a un résultat statistiquement significatif
- b) l'étude A est un nombre plus grand des patients que l'étude B
- c) l'étude B n'a pas un résultat statistiquement significatif
- d) l'estimateur ponctuel d'étude A est de 1,15 à 41
- e) l'intervalle de confiance d'étude B cliniquement n'est pas claire

Réponse: a, b, c, e

69

Exemples des questions pour l'examen

26) Vous avez trouvé deux études sur Pubmed qui ont été réalisées pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu le suivant risque attribuable (RA) avec l'intervalle de confiance : étude A : 5% (95% CI 4% - 6%), et étude B : 6% (95% CI -2% – 14%). On considère que les valeurs supérieures à 7% sont cliniquement importants, et les valeurs inférieures – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'étude B a un résultat statistiquement significatif
- b) l'intervalle de confiance d'étude A cliniquement n'est pas important
- c) l'étude A n'a pas un résultat statistiquement significatif
- d) l'estimateur ponctuel d'étude A est 5%
- e) l'intervalle de confiance d'étude B n'est pas claire cliniquement

Réponse: b, d, e

27) Vous avez trouvé deux études sur Pubmed qui ont été réalisées pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu le suivant risque attribuable (RA) avec l'intervalle de confiance : étude A : 5% (95% CI 0,5% - 9,5%), et étude B : 6% (95% CI -2% – 14%). On considère que les valeurs supérieures à 7% sont cliniquement importants, et les valeurs inférieures – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'intervalle de confiance d'étude A représente se que se passent dans la population cible
- b) l'intervalle de confiance d'étude A cliniquement n'est pas important
- c) l'étude A a un résultat statistiquement significatif
- d) la limite inférieure de l'intervalle de confiance d'étude B est -2%
- e) l'intervalle de confiance d'étude B n'est pas une généralisation dans la population cible, mais la situation particulière dans l'étude

Réponse: a, c, d

70

Exemples des questions pour l'examen

28) Une étude a été réalisée pour évaluer la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu : $RR=0,5$; $OR=0,4$; RIE (consommation modérée du the vert) $=0,15$; $RIN=0,30$. Lesquelles des suivantes affirmations sont correctes ?

- a) pour évaluer l'importance clinique de la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose on a besoin du résultat d'un test statistique
- b) pour évaluer l'importance statistique de la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose on peut utiliser le résultat d'un test statistique
- c) pour évaluer l'importance statistique de la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose on peut utiliser la valeur p d'un test statistique
- d) pour évaluer l'importance clinique de la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose on peut utiliser le RR ou l' OR
- e) pour évaluer l'importance clinique de la relation entre la consommation modérée du the vert comme facteur de risque pour l'apparition de l'ostéoporose on peut utiliser le RIE

Réponse: b, c, d

29) Une étude de cohorte prospectif a été réalisée pour évaluer la relation entre l'exposition au froid et l'apparition des infections respiratoires. L'échantillon de 10 000 sujets a été réalisé d'une manière aléatoire depuis toute la population, stratifiée sur chaque ville de la pays. Ils ont évalué si plus les températures ont été petits, plus de risque d'infection a été observée. Au début aucun sujet n'a pas eu des infections respiratoires. Le odds ratio a été de 5, $p=0,001$. Un autre étude a trouvé un odds ratio de 4,7, $p=0,008$. Lesquelles des suivantes affirmations sont correctes ?

- a) Critère de causalité observé dans l'étude: la relation dose effet
- b) Critère de causalité observé dans l'étude: Séquence temporelle
- c) Critère de causalité observé: Constance de l'association
- d) Critère qui aide à la généralisabilité des résultats: grand échantillon
- e) Critère qui aide à la généralisabilité des résultats: échantillon aléatoire

Réponse: b, c, d, e

71

Exemples des questions pour l'examen

30) Une étude de cohorte prospectif a été réalisée pour évaluer la relation entre l'exposition au froid et l'apparition des infections respiratoires. Lesquelles des suivantes affirmations sont correctes pour le risque relatif?

- a) les valeurs proche de 0 sont petites comme force de l'association
- b) les valeurs proche de (moins) - infini sont grandes.
- c) Pour la force d'association 0 est équivalent à + infini comme importance de l'association
- d) Pour la force d'association 0,8 est plus grand que 0,9 comme importance de la relation
- e) Pour la force d'association 5 est plus petit que 0,33
- f) Une valeur de 0,4 indique un facteur protectif
- g) Une valeur de 1,3 indique un facteur de risque
- h) les valeurs proche de 1 sont petites comme force de l'association

Réponse: c, d, e, f, g, h

32) Une étude de cohorte prospectif a été réalisée pour évaluer la relation entre l'exposition au froid et l'apparition des infections respiratoires. Lesquelles des suivantes affirmations sont correctes pour le risque attribuable?

- a) les valeurs proche de 0 sont petites comme force de l'association
- b) les valeurs proche de (moins) - 80% sont grandes comme force de l'association
- c) Pour la force d'association -5% est équivalent à + 5% comme force de l'association
- d) Pour la force d'association 8% est plus grand que 9% comme importance de la relation
- e) Pour la force d'association 8% est moins grand que (moins) - 9% comme importance de la relation
- f) Une valeur de (moins) - 4% indique un facteur protectif
- g) Une valeur de 5% indique un facteur de risque
- h) les valeurs proche de 1 (100%) sont très importantes comme force de l'association

Réponse: a, b, c, e, f, g, h

72

Exemples des questions pour l'examen

	OR unadjusted	95% CI	P-value	OR adjusted	95% CI	P-value
Obesité	1,9	(1,3-4,9)	<0,001	1,7	(1,2-4,5)	<0,001
Age (annees)	1,15	(1,07-3,9)	<0,001	1,09	(1,04-3,1)	<0,001
Sexe masculin	1,001	(0,9-1,01)	0,78	1,002	(0,91-1,013)	0,67

33) Quelles affirmations, en concernant la régression logistique uni variée et multiple (multivariée) dans le tableau pour prédire l'infarctus myocardique en fonction de l'obésité, l'âge, et le sexe (résultats partiellement ajustés; dans l'étude ils ont fait aussi une autre analyse ou ils ont ajusté complètement les résultats avec encore 9 variables), sont correctes:

- a) la variable obésité pourrait être facteur de confusion
- b) la variable obésité pourrait être facteur de confusion
- c) la variable sexe pourrait être un facteur de confusion
- d) Un argument en faveur d'un risque plus élevé d'obésité associée à l'infarctus myocardique est le fait que l'OR reste significatif lorsque on prend en compte des facteurs de confusion supplémentaires dans les modèles statistiques
- e) On peut rejeter l'hypothèse que le sexe est un facteur de confusion
- f) Les résultats bruts (unadjusted) sont données pour information comme est fait habituellement dans beaucoup des études
- g) Les résultats ajustés sont les seules qui présentent un intérêt réel, parce que ils sont moins sensibles à des effets de confusion
- h) Les résultats ajustés complètement sont les résultats qu'il faut retenir comme résultats « réels » dans cette étude

Réponse: a, b, d, f, g, h

73

Exemples des questions pour l'examen

	OR unadjusted	95% CI	P-value	OR adjusted	95% CI	P-value
normal	1			ref		
surpoids	1,15	(1,07-3,9)	<0,001	1,9	(1,04-3,1)	<0,001
obese	1,9	(1,07-5,9)	<0,001	2,5	(1,04-6,1)	<0,001
Age (annees)	1,05	(1,01-3,9)	0,04	1,06	(0,9-3,1)	0,06

34) Quelles affirmations, en concernant la régression logistique uni variée et multiple (multivariée) dans le tableau pour prédire l'infarctus myocardique en fonction des niveaux du poids, sont correctes:

- a) Le niveau du poids normal est le group de référence pour la variable âge
- b) Le niveau du poids normal est le group de référence pour la variable niveau du poids
- c) il y a 2,5 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont obese compare à ceux normales avec ajustement pour la variable âge
- d) il y a 1,9 fois plus de chance d'avoir l'infarctus myocardique pour ceux qui sont avec surpoids compare à ceux normales avec ajustement pour la variable âge
- e) Le résultat pour l'âge ajustée – n'est pas statistiquement significative

Réponse: b, c, d, e

35) Un étude de type cohorte prospective a évalué la relation entre la obstruction nasale et le syndrome de compression de maxillaire et a trouve une association statistiquement significative. Ils ont trouve un RR de 5 fois plus de risque d'avoir le syndrome pour ceux avec grande obstruction nasale par rapport a ceux sans obstruction, et un RR de 5 fois plus de risque d'avoir le syndrome pour ceux avec moyenne obstruction nasale par rapport a ceux sans obstruction. Lesquels sont des arguments pour la causalité de la relation?

- a) Force de l'association

74

Exemples des questions pour l'examen

35) Un étude A de type cohorte prospective a évalué la relation entre la obstruction nasale et le syndrome de compression de maxillaire et a trouve une association statistiquement significative. Ils ont trouve un OR de 5 fois plus de risque d'avoir le syndrome pour ceux avec grande obstruction nasale par rapport a ceux sans obstruction, et un OR de 3 fois plus de risque d'avoir le syndrome pour ceux avec moyenne obstruction nasale par rapport a ceux sans obstruction. Les résultats ont été les mêmes dans différentes sou populations de l' étude. 5 autre études ont observe les mêmes valeurs de OR dans différents populations. Une étude B a réalise une régression logistique ou ils ont prédit le syndrome en fonction de l'obstruction nasale, mais aussi en tenant compte des autres facteurs de confusion connues, et ils ont trouve des valeurs similaires pour le OR de l'obstruction que celui du OR de l' étude A. Une étude C expérimental sur des singes a montre que en créant une obstruction nasale, les singes ont développé un syndrome de compression de maxillaire. Différentes études ont montre que dans des pays ou il y a une prévalence plus importante de l' obstruction nasale, il y a plus du syndrome, et dans des pays ou il y a une prévalence réduite de l'obstruction nasale, il y a moins de syndrome. Autres études ont montre que dans les pays ou il y a une réduction de la prévalence de l' obstruction nasale, il y a aussi une réduction de la prévalence du syndrome Lesquels sont des arguments pour la causalité de la relation?

- a) Force de l'association
- b) Séquence temporelle
- c) Relation dose effet
- d) Cohérence dans les sous-populations
- e) Constance de la relation
- f) Effet de la suppression de l'exposition (réversibilité de l' effet)
- g) Cohérence temps-espace
- h) Apport des études expérimentales
- i) Cohérence dans les sous-populations
- j) Indépendence de l'association vis-à-vis des autres facteurs connus

Réponse: a, b, c, d, e, g, h, i, j

75

Exemples des questions pour l'examen

36) Vous avez trouvé deux études transversales, sur Pubmed qui ont été réalisée pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. L'étude A a été réalisé sur un tirage au sort des échantillons de 1000 sujets de 4 hôpitaux de 4 villes du Portugal et étude B réalisé sur un tirage au sort des échantillons de 500 sujets de 10 hôpitaux de dix pays européens. Lesquelles des suivantes affirmations soutiennent la generalizabilite des résultats ?

- a) Un échantillon de grande taille pour étude A
- b) Plusieurs pays pour étude B
- c) Plusieurs pays pour étude B
- d) Le tirage au sort (aléatoire) des plusieurs hôpitaux
- e) Une étude prospective

Réponse: a, b, c, d

76

Notions étudiées dans le cours

- Classifications des enquêtes
 - Cas-témoin (+appariement), cohorte (plusieurs types), transversale
 - Définitions, design, avantages, disadvantages
- Tester l'existence de la relation
- Quantification
 - Indicateurs,
 - formules, exemples
 - quand on peut les calculer en fonction de la collection des données
 - interprétation statistique, clinique
 - interprétation des intervalles de confiance + relation avec la valeur p
 - Régression logistique simple et multiple
- Causalité
- Generalisabilité
- Evaluation de la validité interne de l'étude
- Exemples des études
- Exemples des exercices

77

77

Merci beaucoup!

78

78