

L2 Santé Société Humanité UE10 Epidémiologie

Etudes cas-témoin

Dr Camille COUFFIGNAL

Unité de recherche clinique Paris Nord Val de Seine URC PNVS
Département d'Epidémiologie, Biostatistique et Recherche
Clinique
HUPNVS, site Bichat

24/04/2017

Histoire naturelle



Facteurs de risque

Non-malades

**Cause
Risque
Prévention primaire**



Début de la maladie

Malades

**Fréquence
Incidence
Prévention secondaire**

Diagnostic

**Tests diagnostiques
Performances diagnostiques**

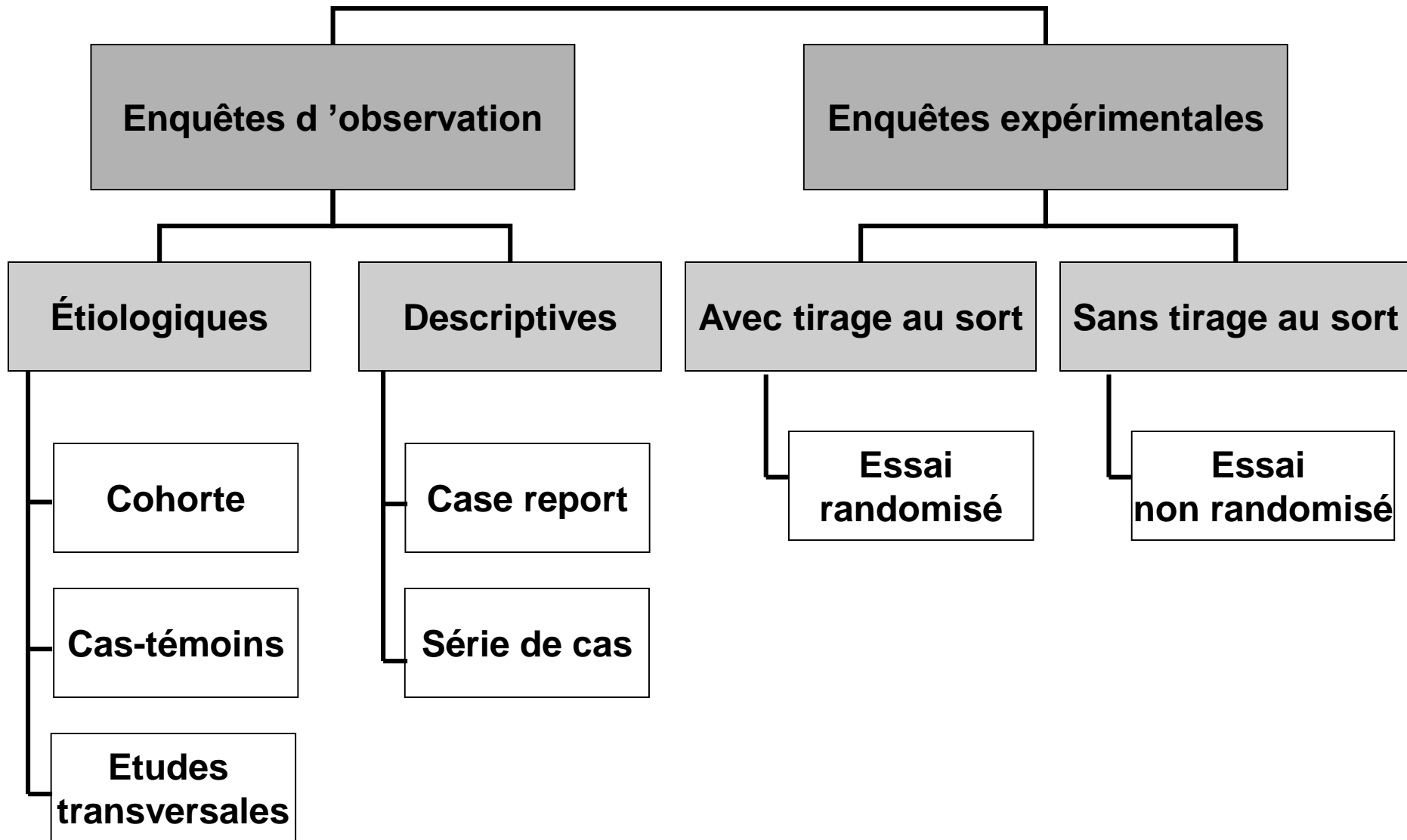
Traitement

Essais cliniques



Devenir

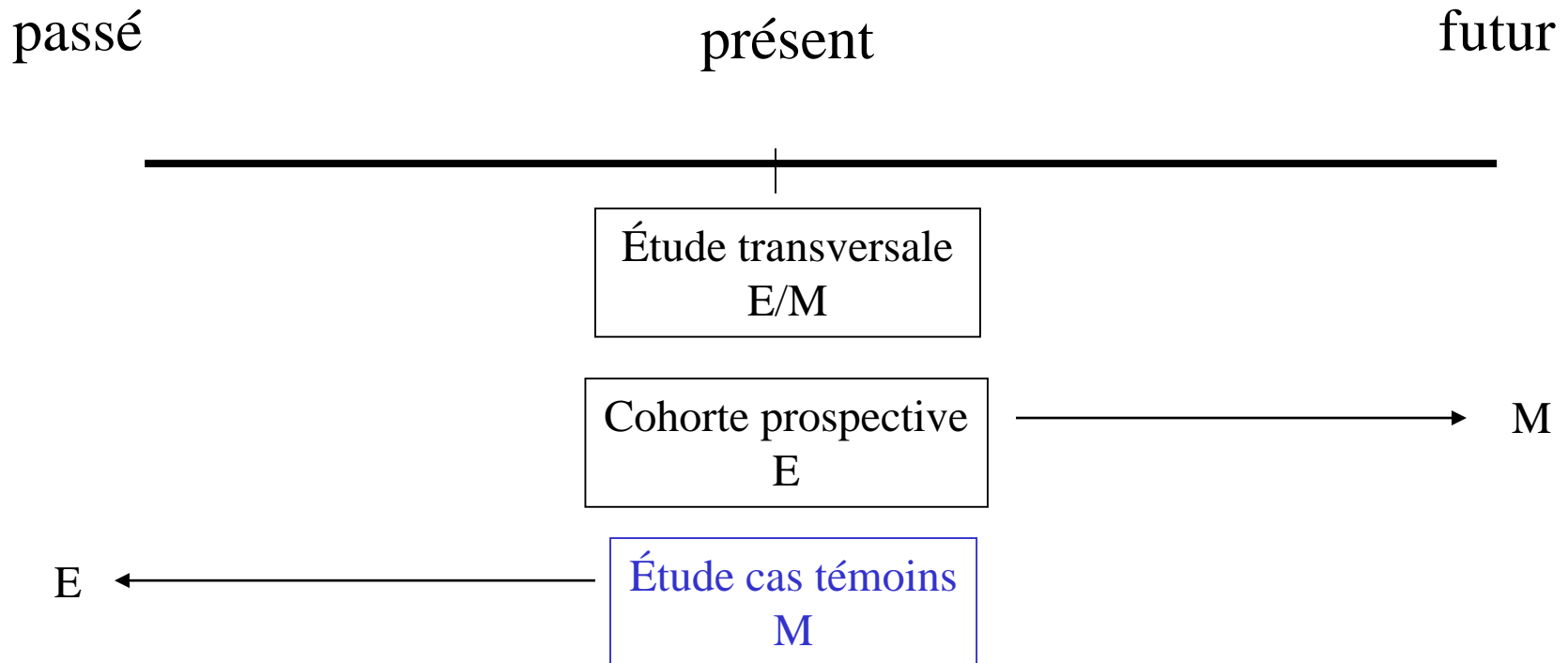
Études pronostiques



Comment classer les différentes enquêtes épidémiologiques?

E = exposition

M = maladie



Études analytiques

Étude de cohorte



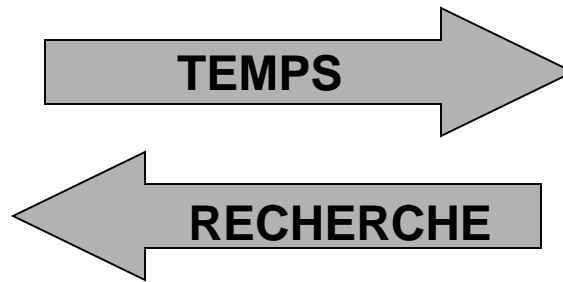
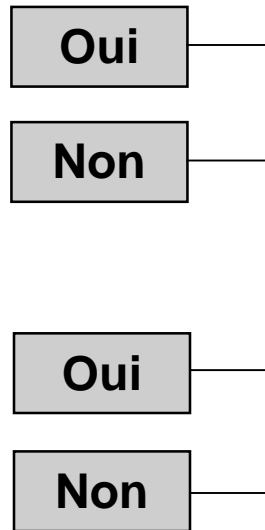
Enquête cas / témoins



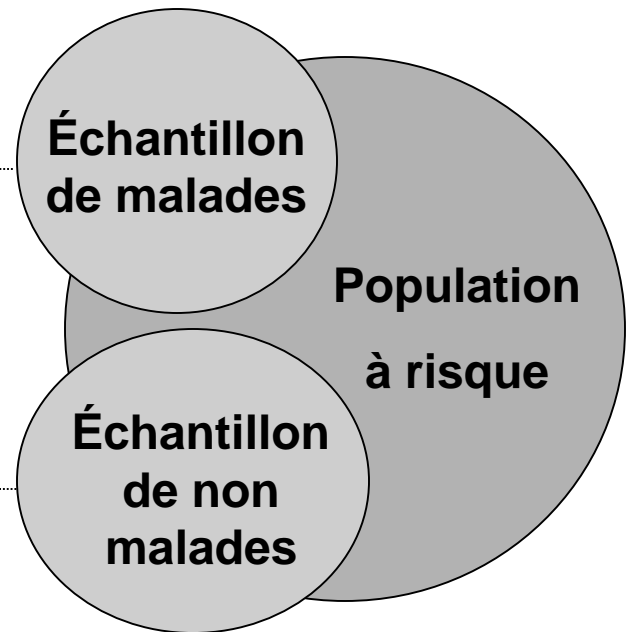
Temps



Exposition au facteur d'exposition



Cas



Témoins

Étude cas-témoins (rétrospective)

Etude cas-témoins

- Principe : comparer la fréquence d'une exposition antérieure entre un groupe de cas et un groupe de témoins
- Plutôt utilisé pour identifier des **facteurs de risque de maladie**, plutôt que des facteurs pronostiques
- Faisabilité : seule étude adaptée quand l'évènement est **très rare**

Exemples

- Association entre consommation de tabac ou de café et survenue d'un cancer de la vessie
- Association entre leucémie aiguë myéloïde et expositions professionnelles

Méthodologie

Définition et sélection des cas

- **Définition précise de la maladie impérative**
 - ➔ 1940 : Cancer de l'utérus = cancer du col ou du corps de l'utérus. En réalité FDR très différents
- Relation médicament-malformations congénitales :
1 seule substance tératogène n'augmente pas uniformément le risque de toutes les maladies congénitales

Définition et sélection des cas

- **Sélection des individus malades**
 - Basée sur une population
 - Idéalement : sélection de tous les sujets malades d'un échantillon randomisé d'une population définie
 - Evite la survenue de biais mais logistiquement et financièrement coûteux
 - Restrictions générales : âge, sexe, lieu d'habitation, origine géographique
 - Basée sur une population particulière (malades hospitalisés,...)

MÉTHODE DE DÉFINITION D'UN CAS

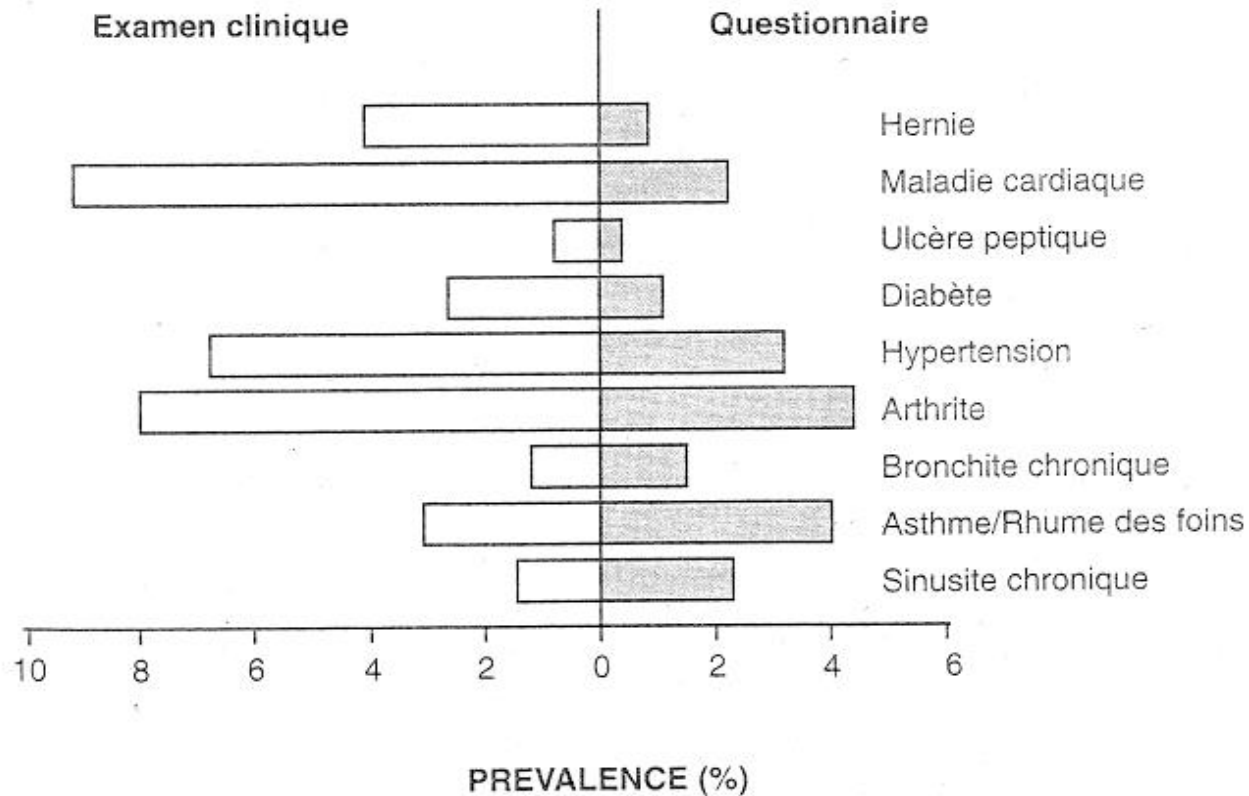


Fig. 4-3. La prévalence dépend de la définition du cas. La prévalence des maladies dans une population générale d'après les opinions des patients (étude) et de l'évaluation clinique. (D'après Sanders BS. Have morbidity surveys been oversold ? *Am J Public Health* 1962 ; 52 : 1648-1659)

Définition des cas

Homogénéité des cas

- Même localisation (athérome cœur différents des vaisseaux cérébraux)
- Même histologie (cancers du poumon : cancer épidermoïde et cancer anaplasique)
- Même groupe d'âge
- Eventuellement si problème de certitude diagnostique, classer les malades en :
 - Sûrs
 - Probables
 - Possibles

Représentativité des cas

Généralisation possible vs validité

Ex : infarctus du myocarde

- Echantillon randomisé des cas d'infarctus survenus dans la population préférable à un choix d'hommes de 45 à 75 ans hospitalisés dans un hôpital ?
- Echantillon randomisé → plus facilement généralisable à la population d'où est issu l'échantillon
- Il est nécessaire pour la validité de l'étude d'avoir une information complète et exacte sur la maladie
- La validité de l'étude ne doit pas faire l'objet de compromis
- Si pas de validité, pas de généralisation envisageable

Cas incidents / Cas prévalents

- **Cas incident** : inclus dans l'étude au moment où il devient malade
- **Cas prévalent** : malade avant d'être inclus dans l'étude. C'est donc un cas incident ayant survécu ou n'ayant pas été perdu de vue à la date d'inclusion dans l'étude. Dans le cas des maladies graves, les cas prévalents ont été sélectionnées sur la survie. Si la survie est elle même associée au facteurs de risque étudié, la mesure de l'association entre exposition et maladie sera **biaisée** (ex : malformations congénitales).

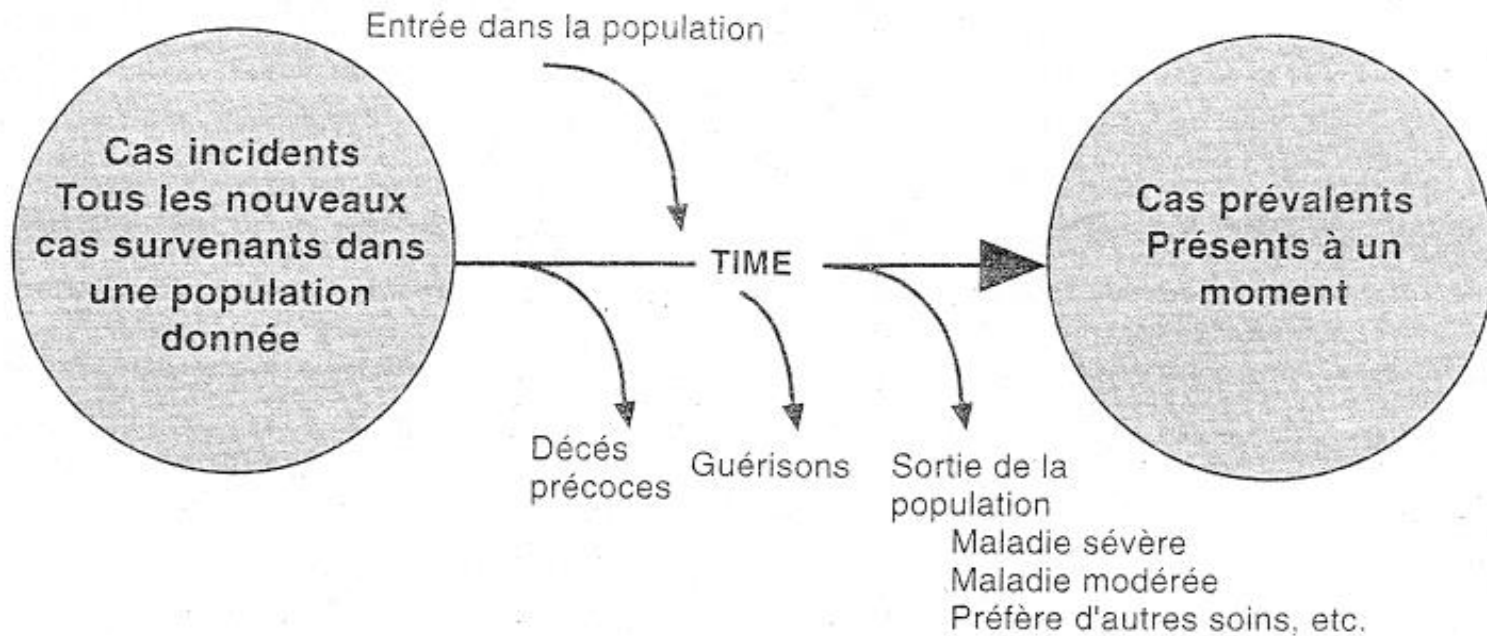


Fig. 4-5. La différence de cas entre les études de prévalence et d'incidence.

Cas incidents / Cas prévalents

Souvent pour maladies rares on prend les cas prévalents mais risque de biais car :

- Séquence chronologique exposition/maladie toujours potentiellement litigieuse dans les études cas-témoins, encore plus difficile si l'on utilise les cas prévalents
 - Critères diagnostiques peuvent changer au cours du temps (diagnostics les plus récents sont plus précis)
- ➔ Si possible n'utiliser que les cas incidents

Avantages et inconvénients d'inclusions des cas incidents et prévalents

	Cas incidents	Cas prévalents
Sources habituelles	Registres	Structure des soins
Durée nécessaire à l'inclusion	Longue	Courte
Biais de sélection par le pronostic	Peu vraisemblable	Possible : <ul style="list-style-type: none"> • Evolution plus longue → pronostic meilleur • Parfois au contraire mauvais pronostic
Biais de sélection par le recours aux soins	Peu vraisemblable	Fréquente

Méthodologie

Définition et sélection des témoins

- Non atteints par la maladie étudiée
- Mêmes restrictions générales et spécifiques que les cas
 - ➔ Issus de la même population que les cas
 - ➔ L'opportunité d'être exposé doit être la même chez les cas et les témoins.
- Sélection des témoins indépendante de l'existence ou non d'une exposition

Population générale

hospitalisés

proches du malade

Plusieurs groupes de témoins

Sélection des témoins

- Problème plus difficile
 - Pas de groupe témoin idéal en toutes circonstances
 - Sélection des témoins pour un groupe particulier de cas
-
- Il faut prendre en compte :
 - Les caractéristiques et l'origine des cas
 - La nécessité d'obtenir des informations comparables chez les cas et les témoins
 - Les considérations pratiques
 - Ils doivent être avant tout comparables à la population des sujets malades et que toute exclusion ou restriction intervenant dans la sélection des cas doit leur être également appliquée et vice-versa

Sélection des témoins

- **Une des sources les plus fréquentes** : malades hospitalisés pour une maladie autre que celle faisant l'objet de l'étude

Ex : étude cas-témoin usage de cigarettes et infarctus du myocarde

Cas : soins intensifs des hôpitaux

Témoins : services chirurgicaux et médicaux de ces hôpitaux

Sélection des témoins hospitalisés

- **Avantages :**

- Avantages pratiques et scientifiques :

- Facilité de recrutement (diminution coût)
- Du fait de leur hospitalisation, ils ont plus de chance que des sujets sains de connaître leurs antécédents ou leurs expositions antérieures.

Leur manière de fournir l'information est comparable à celle des cas → diminution de la possibilité de biais de mémoire

- Hospitalisés dans le même hôpital → biais de sélection (socio-économiques ou autres) identiques aux cas sauf si réputation de l'hôpital diffère pour certaines pathologies
- Coopération potentielle plus facile que celle de sujets en bonne santé → diminution biais liés à l'absence de réponse

Sélection des témoins hospitalisés

- **Inconvénients**

- Ils sont par définition malades et différents donc d'individus en bonne santé par un certain nombre de caractères qui peuvent avoir un rapport avec leur maladie ou avec l'hospitalisation

→ ATCD d'exposition peuvent ne pas être très représentatifs de ceux de la population dont sont issus les cas

- Patients hospitalisés ont plus tendance à fumer, utiliser des contraceptifs oraux, boire de l'alcool

Sélection des témoins hospitalisés

- **Catégorie de malades à inclure dans le groupe témoin**

Les maladies ayant motivé l'hospitalisation des témoins constituent un problème majeur car ils peuvent également être associées aux F de R que l'on veut étudier

➔ Patients atteints de maladies connues pour être associés plus ou moins à l'exposition étudiée seront exclus du groupe témoin

ex : tabac : F de R des maladies respiratoires (bronchite, pneumopathie)

Etude cas-témoins sur tabac et cancer du poumon

Utilisation comme témoins de patients atteints de bronchite

➔ sous-estimation du véritable risque lié au tabagisme (proportion de fumeurs dans les 2 groupes étant artificiellement identique)

Sélection des témoins de population

- **Groupe de témoins peut provenir de la population générale**

Recrutement à partir de listes électorales ou de l'annuaire téléphonique

Mais :

- Identification des témoins difficile
- Difficile à contacter
- Refusant fréquemment de participer
- Rapport avec le risque de développer la maladie peut différer systématiquement chez les individus qui acceptent de participer et chez ceux qui refusent de participer ou que l'on n'arrive pas à contacter

*Ex : étude F de R activité physique et maladie coronarienne
témoins issus de la population générale recrutés à domicile*

- ➔ Groupe de témoins avec une plus grande proportion de patients moins actifs car ceux-ci ont plus de chance d'être chez eux lors de l'appel des enquêteurs

Sélection d'autres témoins

- **3ème source de témoins**

- Groupe particuliers : amis, parents, conjoints, voisins
- Avantages des témoins issus de la population générale : bien portants
- Plus de chance d'être coopératifs
- Permettant également de contrôler certains facteurs de confusion importants en relation avec l'origine ethnique, le statut socio-économique ou l'environnement. Contrôle difficilement envisageable autrement

Attention si le facteur étudié est le même chez les amis, les membres de la famille,... → sous estimation de l'effet réel de l'exposition

Par exemple, si on s'intéresse aux habitudes alimentaires ou tabagiques, les individus proches auront probablement les mêmes expositions → ne sont pas des groupes témoins appropriés

Comparaison des avantages et inconvénients des témoins hospitalisés et recrutés dans la population dans les enquêtes cas-témoins

	Témoins hospitalisés	Témoins de population
Accès aux personnes	Facile (regroupés)	Long et coûteux (dispersés)
Motivation à participer	Forte (malades)	Faible (en bonne santé)
Disponibilité (temps)	Souvent grande	Souvent limitée (occupations)
Taux de réponse	Souvent élevé	Parfois moyen ou faible
Origine sociogéographique	Même zone d'attraction que les cas	Souvent plus large que les cas (si hospitaliers)
Biais de sélection	En principe important	En principe limité
Biais de mémoire (différentielle)	En principe réduit	Possible
Surexposition au facteur	Possible	Peu probable
Biais de mesure de l'exposition	Possible (si état de santé connu)	Possible (lieu d'interrogatoire différent)
Généralisation des résultats	Limitée	large

Nombre de sujets témoins

Rapport optimal 1/1

1 cas, 1 témoin (en particulier si nombre de cas important)

Si nombre de cas faible

la puissance de l'étude augmente avec le nombre de sujets témoins, jusqu'à 4 témoins pour 1 cas

L'appariement

- Chaque cas peut être apparié avec un ou plusieurs témoins ayant des caractéristiques communes
 - en général sexe, âge, lieu de résidence
 - autres facteurs de risque connus considérés comme importants
- **But** : Réduit les différences entre cas et témoins et ainsi augmente la puissance de l'étude
- **Inconvénient** : L'appariement excessif : on apparie sur un facteur lui-même lié à l'exposition étudiée → les cas et les témoins auront vraisemblablement une même histoire d'exposition → OR biaisé vers 1
- *Exemple* : lien prise d'anti-inflammatoires non stéroïdiens et hémorragie digestive en appariant sur les symptômes d'arthrose.

Méthodologie

Mesure de l'exposition

- Rétrospective
 - Sujets eux-mêmes (interview ou questionnaire postal)
 - D'un proche (conjoint)
 - Mère des enfants
 - Informations dans l'observation médicale

→ *Erreur liée à la mémoire*
- Recueil d'informations aussi semblable que possible chez les malades et les témoins
- **Aveugle des enquêteurs** sur la catégorie d'individus (malades ou témoins)
 - Limiter le biais d'observation

- **Information sur la maladie**

- Sources nombreuses :

- Registres de décès
 - Registre de cas d'une maladie
 - Registre des malades sortis d'un hôpital (malades sortis avec le diagnostic d'infarctus du myocarde par exemple)

- **Mesure de l'exposition**

- erreurs liées à la mémoire surtout

- Si faits anciens

- Quantification demandée

- Biais de déclaration

- la déclaration de l'exposition à certains facteurs peut faire espérer une réparation (maladie professionnelle) ou faire craindre l'opprobre (exposition à l'alcool,...)

Mesure d'association

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

$$RR = \frac{a/a+b}{c/c+d}$$

Le rapport du nombre de cas au nombre de témoin est fixé par le chercheur



Les proportions $a/a+b$ et $c/c+d$ sont arbitraires et ne représentent pas une mesure du risque

Mesure d'association

Exemples

1 ^{ère} étude	Cas	Témoins
Exposés	80	50
Non exposés	120	150
	200	200

2 ^{nde} étude	Cas	Témoins
Exposés	80	100
Non exposés	120	300
	200	400 x 2

Mesure d'association

Exemples

1 ^{ère} étude	Cas	Témoins
Exposés	80	50
Non exposés	120	150
	200	200

$$RR = (80 / 130) / (120 / 270)$$
$$RR = 0.6 / 0.4 = 1.5$$

2 ^{nde} étude	Cas	Témoins
Exposés	80	100
Non exposés	120	300
	200	400 X 2

$$RR = (80 / 180) / (120 / 420)$$
$$= 0.4 / 0.3$$
$$= 1.3$$



Les enquêtes cas-témoins ne permettent pas de calcul direct du RR

Odds Ratio et Risque Relatif

Mesures association entre exposition et maladie

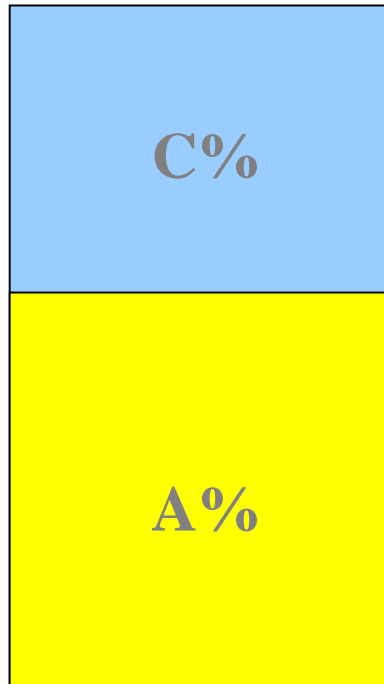
L'odds ratio a la même interprétation qu'un risque relatif

- étude de cohorte : RR ou OR
- étude cas témoin : OR

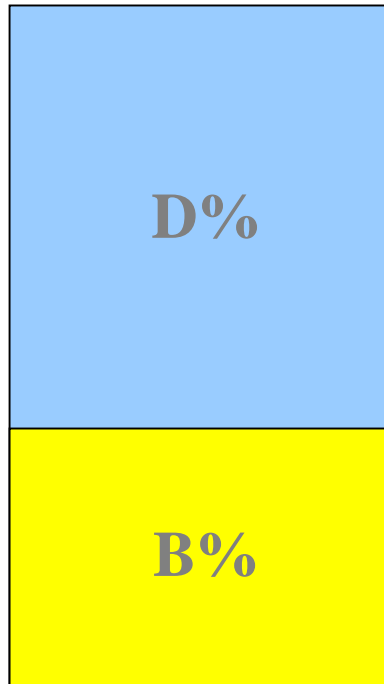
Concept de l'Odds

- Odds (« côte ») = $P(\text{succès})/P(\text{échec})$
= $P/(1-P)$
- Ex : turfistes
 - Epi Beauty $P = 60\%$
 - Odds = $60/40 = 1,5/1$
= 1,5 contre 1

Etude cas-témoins



Cas



Témoins

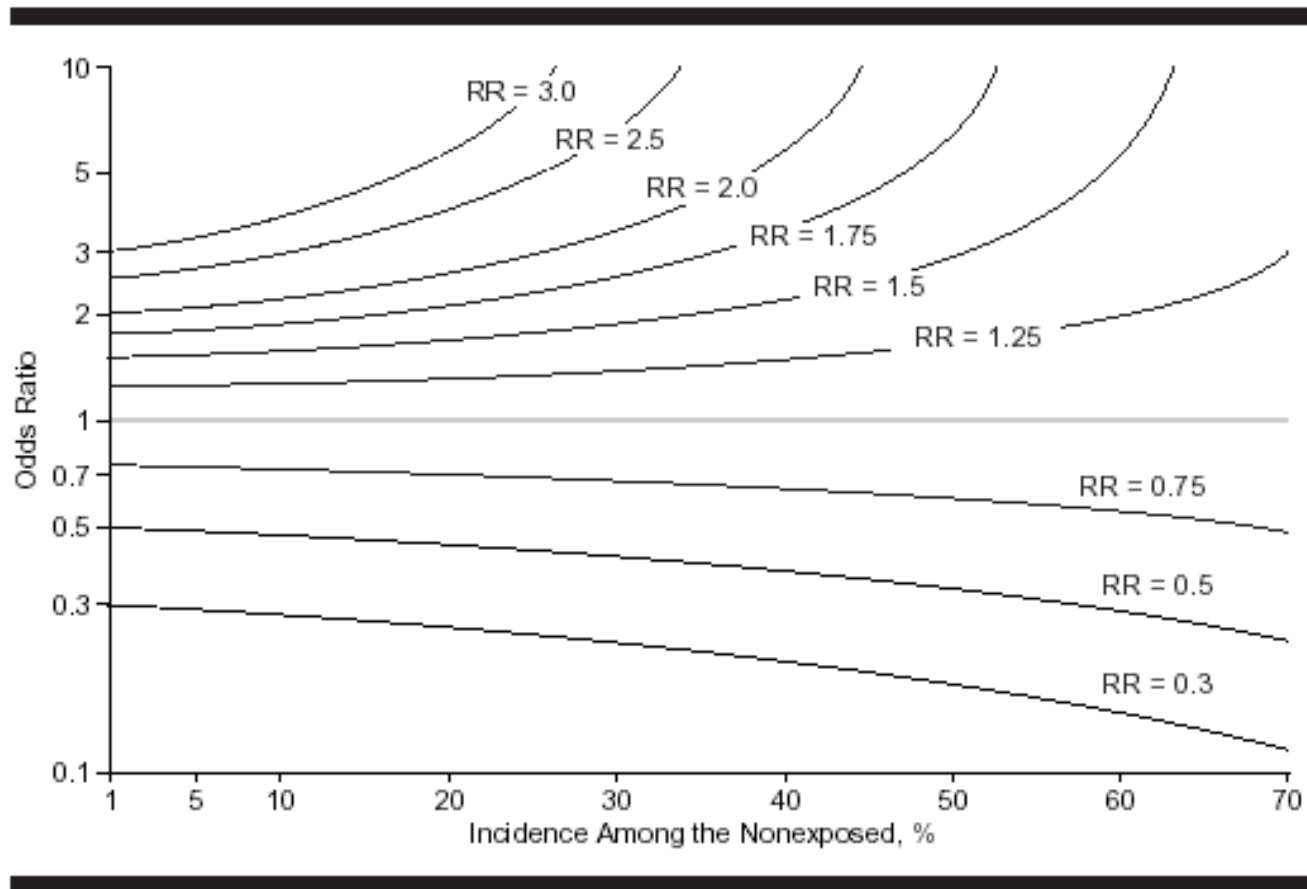
 Non exposés
 Exposés

Cote (odds) d'exposition parmi les cas = A/C

Cote (odds) d'exposition parmi les témoins = B/D

Rapport de cotes - **Odds Ratio**
 $= (A/C)/(B/D) = AD/BC$

Relation RR et OR en fonction de l'incidence de l'événement



The relationship between risk ratio (RR) and odds ratio by incidence of the outcome.

Mesure d'association

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

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

$(a / M_1) = \text{prob d'être exposé chez les cas}$

$(c / M_1) = \text{prob de ne pas être exposé chez les cas}$

Cote = $P/(1-P) = a/c$

$(b / M_0) = \text{prob d'être exposé chez les témoins}$

$(d / M_0) = \text{prob de ne pas être exposé chez les témoins}$

Cote = $P/(1-P) = b/d$

OR = ad/bc

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

Mesure d'association

- Si maladie rare (prévalence $< 5\%$) OR sensiblement équivalent au RR
- Les sujets exposés ont une probabilité d'avoir la maladie OR fois celle des non-exposés

OR = 1 signifie qu'il n'y a pas d'association entre l'exposition et la maladie

- la cote d'exposition des cas est la même que la cote d'exposition des témoins

OR = 2 (la cote d'exposition des cas est 2 X la cote d'exposition des témoins)

- Facteur de risque

OR = 0.5 (la cote d'exposition des cas est $/2$ par rapport la cote d'exposition des témoins)

- Facteur protecteur

Mesure d'association

Exemples

Exemple 1	Cas	Témoins
Exposés	80	50
Non exposés	120	150
	200	200

$$RR = (80/130) / (120/270)$$

$$RR = 0.6 / 0.4 = 1.5$$

$$OR = 80 \times 150 / 50 \times 120 = 2$$

Exemple 2	Cas	Témoins
Exposés	80	100
Non exposés	120	300
	200	400

$$RR = (80/180) / (120/420) = 0.4 / 0.3 = 1.3$$

$$OR = 80 \times 300 / 100 \times 120 = 2$$

Estimation et précision du risque

- OR = 3,3 IC_{95%} [1,8 – 5,9] \Leftrightarrow p < 0.05

Facteur de risque

- OR = 0,3 IC_{95%} [0,2 – 0,6] \Leftrightarrow p < 0.05

Facteur protecteur

- OR = 1,8 IC_{95%} [0,9 – 3,6] \Leftrightarrow p \geq 0.05

Pas de lien avec la maladie

Plus le nombre de sujets inclus dans l'étude est grand,
plus la précision de l'estimation est grande.

Les biais

Biais de sélection

- Inclusion
- Biais qui survient quand le critère de sélection des cas et des témoins n'est pas **indépendant de l'exposition** (la fréquence de l'exposition chez les témoins est différente entre l'échantillon et la population cible).
- Attention au biais de sélection si taux de participation peu élevé ou très différent entre cas et témoins
 - Relation exposition maladie observée chez les participants à l'étude est différente de celle des individus qui auraient pu être sélectionnés mais ont refusé de participer

Exemples de biais de sélection des témoins

Cas	Témoins	Exposition	Biais
SIDA vivant à San Francisco	Personnes consultant pour MST	Nombre de partenaires (forte prévalence d'exposition)	OR=2.9
SIDA vivant à San Francisco	Les voisins	Nombre de partenaires (faible prévalence d'exposition)	OR=52.0

Biais d'observation

- Erreur systématique lors du recueil d'information (exposition) chez les sujets retenus pour l'étude
 - Connaissance de la maladie peut modifier le rendu de l'information par le sujet, l'interprétation par l'enquêteur
 - Subjectivité
- Seuil à partir duquel un individu est considéré comme exposé

Difficile car nécessite des connaissances sur la maladie

Ex : tabac et cancer du poumon

durée de la période de tabac plus importante que quantité moyenne de tabac fumée quotidiennement

tabac et infarctus

quantité de tabac fumée est plus importante

- ▶ Intérêt +++ des informations recueillies avant la survenue de la maladie car exactitude des données ne dépend pas dans ce cas de l'apparition ultérieure de la maladie

Biais de mémoire

- Les individus qui ont été malades, ont tendance à rechercher les causes possibles de la maladie
- Leur relation de l'exposition différera en fonction de leur statut malade ou non malade

Pour limiter ces biais

- Questionnaire et interviews standardisées
- Conditions de recueil comparables
- En aveugle du statut cas ou témoin
- En aveugle de l'hypothèse de l'étude

Biais de classement

- Erreur dans le classement de la maladie ou de l'exposition
 - Erreur inévitable
 - Conséquences variables selon que l'erreur de classement d'un paramètre est indépendante de l'autre paramètre c'est à dire si la proportion des erreurs de classement portant sur l'exposition ou sur la maladie est la même dans les groupes comparés

Ex : étude cas-témoins sur le risque d'AVC

1) 80 % des cas et des témoins rapportent avec exactitude leurs antécédents hypertensifs

– Erreur de classement sur l'exposition mais cette erreur est indépendante de la maladie (erreur non différentielle)

2) Si parmi les sujets avec antécédents d'hypertension 90 % des sujets et 60 % des témoins les signalent

→ erreur de classement de l'exposition non indépendante de la maladie (erreur différentielle)

– Erreur de classement due au hasard → sous-estimation de l'OR (une association observée ne peut donc être fausse)

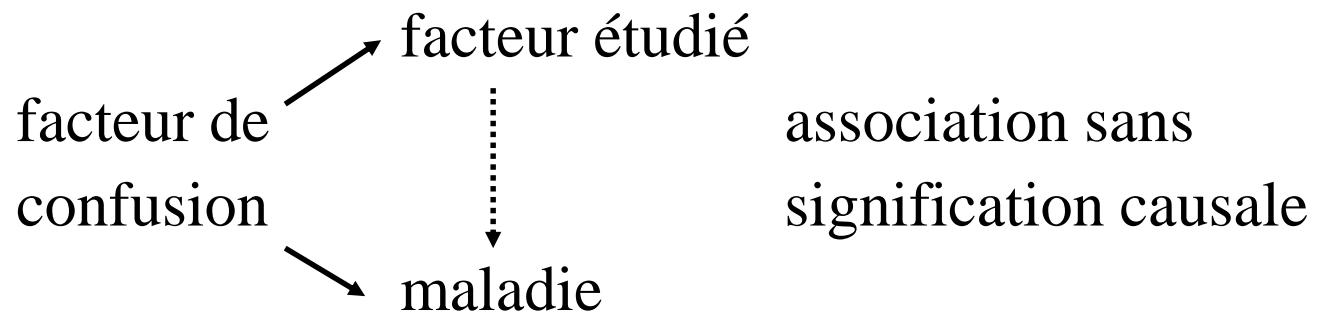
– Erreur de classement non due au hasard → sur ou sous-estimation de l'OR selon la situation (difficultés d'interprétation)

Facteurs de confusion

Tiers facteur qui est associé à la fois au facteur d'exposition et à la maladie. Sa présence déforme la réalité de l'association entre les facteurs d'exposition et la maladie. La déformation de la force d'association est due à une distribution différente du tiers facteur entre les cas et les témoins.

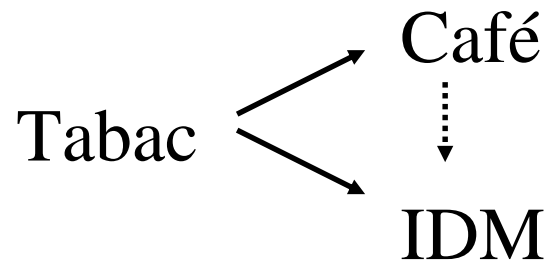
Un facteur de confusion est une variable qui doit être :

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



Facteurs de confusion

Association consommation de café et risque d'infarctus du myocarde



fume en buvant son café

si une variable est associée à la maladie et non à l'exposition ou vice versa, elle ne peut être un facteur de confusion

Prise en compte des facteurs de confusion

- Avant le début de l'étude
 - Restriction lors de la sélection des cas et des témoins
 - Appariement
 - **But** : Réduit les différences entre cas et témoins et ainsi augmente la puissance de l'étude
 - **Inconvénient** : L'appariement excessif : on apparie sur un facteur lui-même lié à l'exposition étudiée → les cas et les témoins auront vraisemblablement une même histoire d'exposition → OR biaisé vers 1
- Après de début de l'étude
 - Stratification ou analyse en sous groupes
 - Ajustement
 - OR brut vs OR ajusté dans une analyse multivariée

Interprétation des résultats

- **Relation statistiquement significative**

- \neq lien causal entre exposition et maladie

Il faut rechercher des explications alternatives

- Différents biais

- Biais de sélection
- Biais d'observation
- Biais de mémoire
- Erreurs de classement
- Non prise en compte de l'ensemble des facteurs de confusion

- Chance

- Si tests d'hypothèses multiples, le risque de conclure à tort (risque α) augmente

• **Différence non significative entre cas et témoins**

Différentes explications doivent être envisagées avant de conclure à l'absence d'effet à l'exposition étudiée :

- Puissance n'est peut être pas suffisante, taille d'échantillon trop petite
- Erreurs de classement imputables au hasard

Diminution du risque relatif

- Non prise en compte de facteurs de confusion plus fréquents chez les témoins que chez les cas
- Absence d'association entre exposition et maladie liée à un manque de spécificité

Il est possible que seul un sous-groupe de malades ou seul un sous-groupe d'exposés présente un risque élevé (risque dilué si l'ensemble des malades ou l'ensemble des exposés est étudié)

Notions de causalité

- Devant une association statistiquement significative entre une exposition et une maladie, quelles questions faut-il se poser?
 - Y a-t-il un biais?
 - Y a-t-il un ou des facteurs de confusion?
 - Est-ce lié à la chance?
- Un lien n'implique pas une notion de causalité

Lien de causalité

(critères de Hill)

- Association forte (OR élevé)
- Existence d'un gradient effet-dose
- Relation temporelle logique (la cause précède l'effet)
- Spécificité de l'association entre le FDR et la maladie
- Convergence des études épidémiologiques (reproductibilité)
- Association causale plausible car cohérente avec les données de la littérature et l'expérience clinique
- Réduction de l'incidence de la maladie si diminution du FDR

Étude cas-témoins

Intérêt

- Résultats rapides
- Coût faible

Indications

- Maladies dont les périodes de latence sont longues
- Maladies rares
- Étude de plusieurs facteurs étiologiques pour une seule maladie

Principal problème

Maladie et exposition existent déjà au moment où les participants entrent dans l'étude

→ Risque de biais +++

si sélection des cas et des témoins exposés ou non s'opère différemment ou

si recueil ou souvenir de l'exposition diffère chez les sujets malades ou non malades

Étude cas-témoins

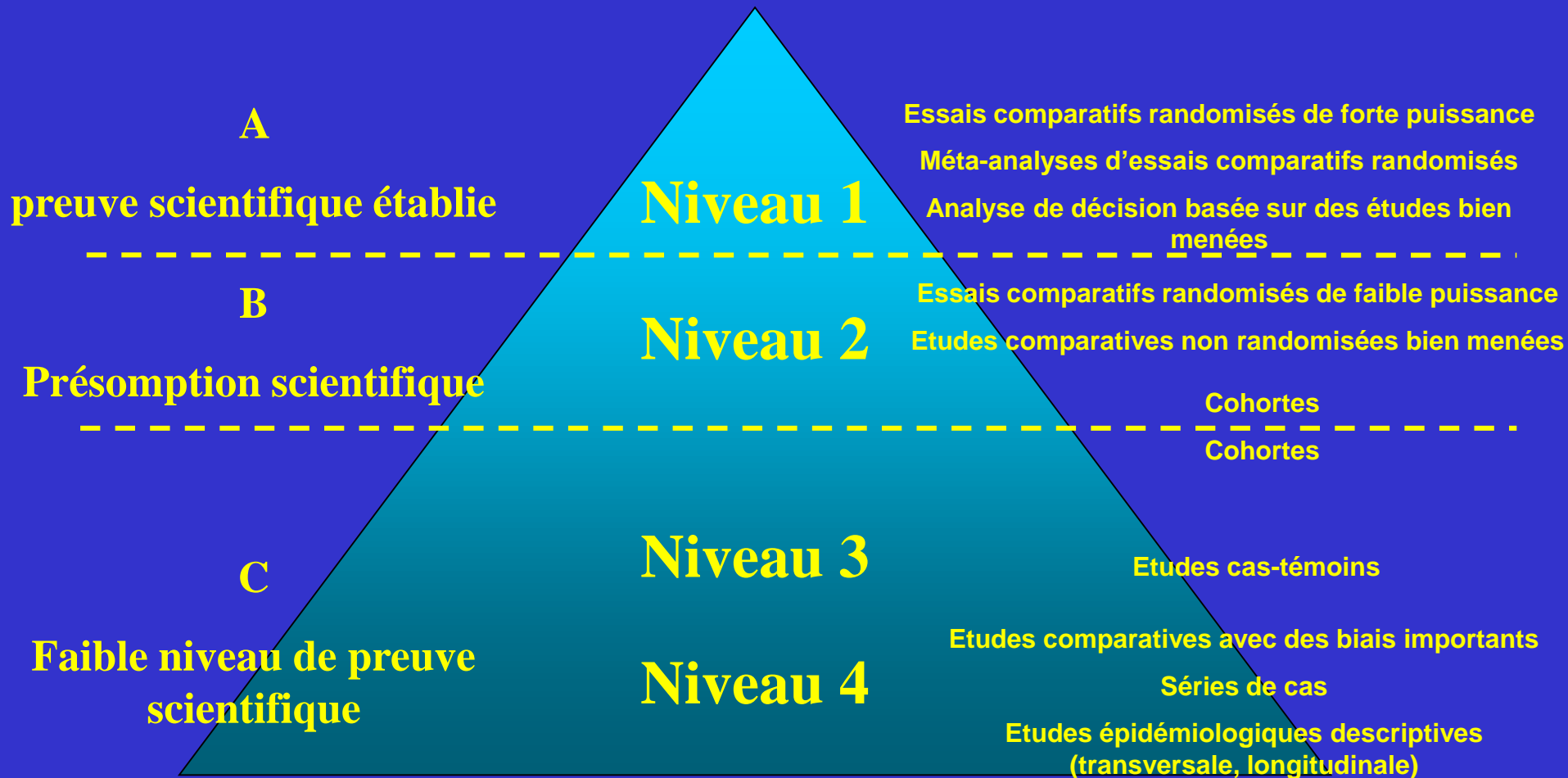
Limites

- Ne permet pas l'étude de plus **d'une maladie** à la fois
- Ne permettent pas d'évaluer les **expositions rares**, sauf si le pourcentage de risque attribuable est élevé
- Ne permettent pas un **calcul direct des taux d'incidence** de la maladie chez les patients exposés et non exposés
 - Estimation du risque par OR (bonne estimation du risque relatif si la prévalence de la maladie chez les exposés et non exposés est faible <5%)
- Les risques de **biais** sont important particulièrement ceux de sélection et de mémoire.
- La **relation chronologique** entre exposition et maladie peut être difficile à établir
- Moins bonne **qualité des données** d'exposition (recueil rétrospectif)
- Difficultés du **choix de témoins**

Force des
recommandations
HAS (grade)

Niveau de preuve
scientifique des études

Type d'études



Cohorte	Cas-témoins
Débute avec une population à risque définie	Débute avec une population à risque souvent mal définie
Inclusion de tous les cas	Sélection des cas à partir de la population disponible
Témoins non sélectionnés	Témoins sélectionnés
Exposition mesuré avant la survenue de la maladie	Exposition mesurée ou reconstruite après la survenue de la maladie
Incidence Risque relatif	Odd ratio