

# Sources de variabilité et d'erreurs dans les études épidémiologiques

## Notions de précision & validité

---

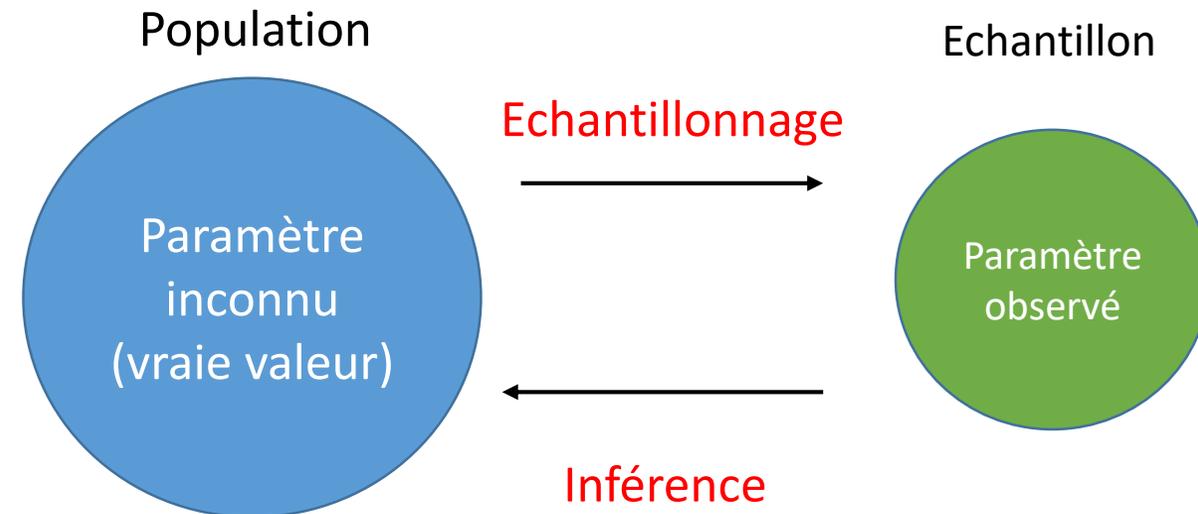
Ter Tiero Elias DAH MD, Msc, PhD  
AHU Santé Publique – Option Epidémiologie  
Université de Ouahigouya

# Position du problème

## Etude épidémiologique

- Exercice de mesure servant à décrire et à comparer
- Mais toute mesure est sujette à l'erreur
- Le résultat obtenu est une combinaison de la vraie valeur (celle que l'on observerait si l'on pouvait étudier la population entière des sujets auxquels on souhaite appliquer les résultats) **et d'erreurs** commises à différentes étapes de l'enquête.
- Tout le travail de l'épidémiologiste consistera à reconnaître, prévenir et maîtriser ces erreurs. A toutes les étapes d'une étude (conception, mise en place, déroulement et analyse), il est donc utile de réfléchir aux sources potentielles d'erreur.

## Estimation d'un paramètre



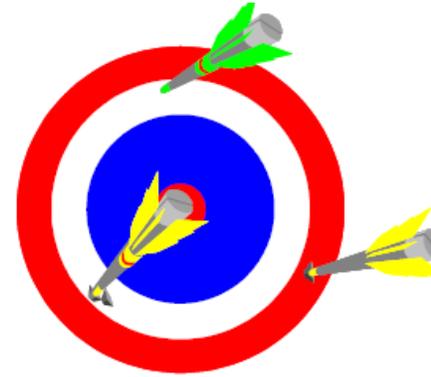
# Précision & validité

# Deux types d'erreurs

## 1. Erreur aléatoire

- Liée aux fluctuations d'échantillonnage (hasard)
- Ne modifie pas la valeur du paramètre estimé de façon systématique
- Augmente la variance, Et donc altère la précision
- Dépend de la taille de l'échantillon et de la variabilité de l'évènement étudié

**Précision = absence d'erreur aléatoire**



a. Erreur aléatoire

# Deux types d'erreurs

## 1. Erreur aléatoire

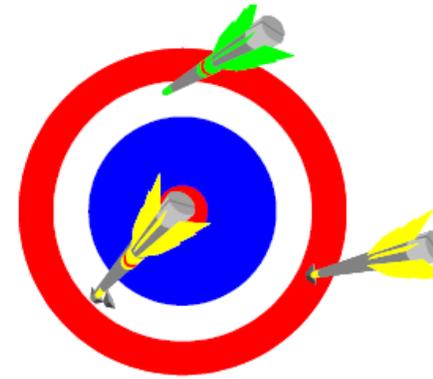
- Liée aux fluctuations d'échantillonnage (hasard)
- Ne modifie pas la valeur du paramètre estimé de façon systématique
- Augmente la variance, Et donc altère la précision
- Dépend de la taille de l'échantillon et de la variabilité de l'évènement étudié

**Précision = absence d'erreur aléatoire**

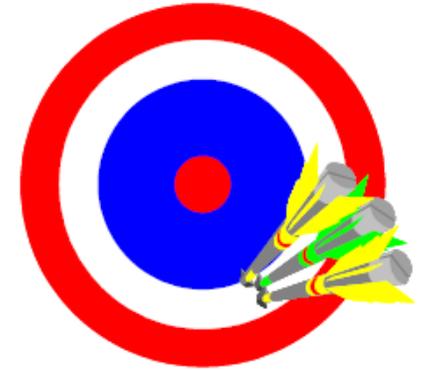
## 2. Erreur systématique ou biais

- Pas due au hasard
- Liée à la conception de l'étude : sélection des groupes, mesure de l'évènement
- Altère la validité

**Validité = absence d'erreur systématique**



a. Erreur aléatoire



b. Erreur systématique

# Précision d'une estimation

## Précision

- Absence d'erreur aléatoire

## Comment assurer au mieux une bonne précision ?

- Limiter l'erreur aléatoire
  - Réaliser l'étude sur un échantillon le plus représentatif possible
  - Calculer les intervalles de confiances des différentes mesures (fréquence, association, impact)
  - Avoir une taille d'échantillon suffisamment grande
    - Se calcule en fonction de la puissance, de la précision souhaitées, du schéma d'étude, et de la population source



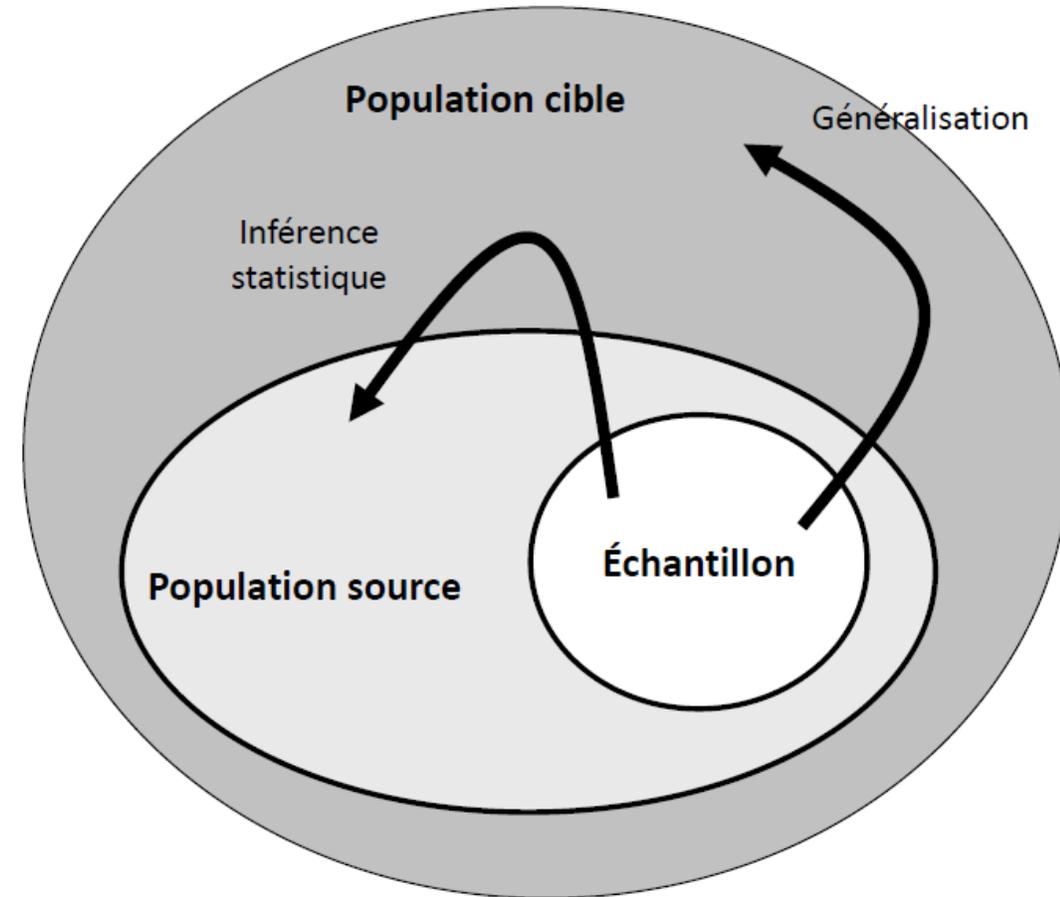
# Validité d'une étude

## Validité externe

- Concerne la population cible
- L'étude est-elle généralisable?

## Validité interne

- Concerne l'étude elle-même
  - Cherche une association entre deux facteurs
  - Le résultat reflète bien l'association entre ces deux facteurs
    - Pas de distorsion de la mesure
    - Pas d'erreur dans le sens de l'association
- = absence de biais**



# Validité externe

**Capacité de généraliser les résultats à une population autre que celle réellement étudiée dans l'enquête**

## **Aspects statistiques**

- Notion d'échantillon représentatif
- Importance du tirage au sort
- Importance de la taille d'échantillon

## **Aspects scientifiques**

- Importance des connaissances de la littérature scientifique sur la question

Attention : L'extrapolation des résultats de l'étude est possible si la validité interne est correcte

# Biais

## Définition d'un biais

Distorsion dans l'estimation de la mesure, liée à une erreur systématique

## Type de mesures concernées

- Mesures de fréquence
- Mesures d'association
- Mesures d'impact

## Etapas de survenue

- Sélection des sujets : **biais de sélection**
- Collecte de l'information : **biais d'information**
- Groupes non comparables : **biais de confusion**

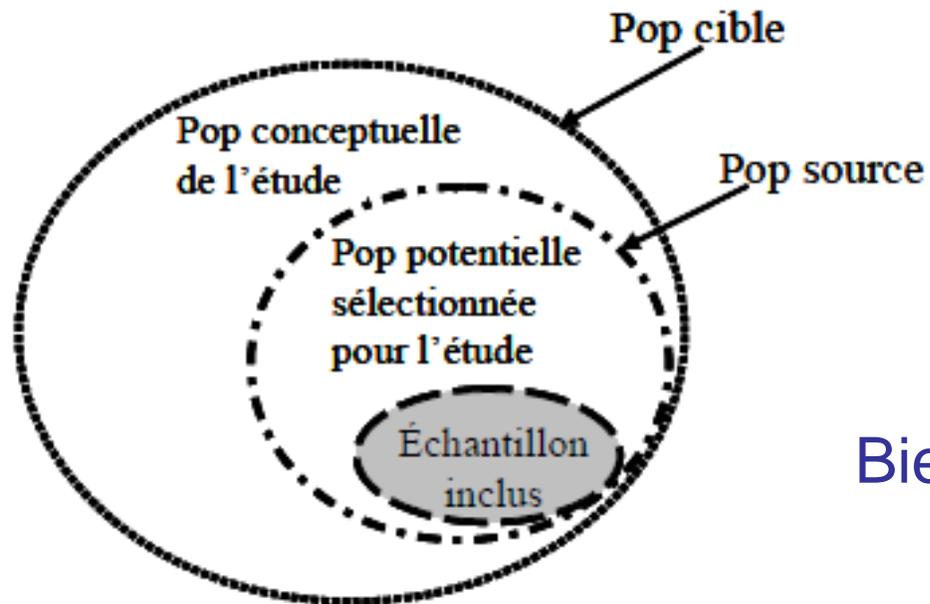
# Biais

# Biais de sélection (1)

## Définition d'un biais

Distorsion dans l'estimation de la mesure, liée à des défauts dans la sélection des sujets

Synonyme: biais de recrutement



Bien préciser les différentes populations de l'étude

# Biais de sélection (2)

## Mécanisme

- Si la répartition exposés/non exposés ou malades/non malades change entre la population cible, la population source et la population d'étude

## Exemples courants de biais de sélection

- **Etude de cohorte**
  - Biais d'admission (fréquent dans les services hospitalier de spécialité)
    - Inclusion plus importante des patients exposés
  - Biais de survie sélective
    - Inclusion sélective des survivants d'une maladie à forte létalité, Tendance à surestimer les effets d'une intervention
  - Biais de non réponse
    - Refus de participer, de réponse ou perdus de vue
- **Etude cas-témoins**
  - Recrutement hospitalier des témoins ou population spécifique
  - Biais de survie sélective

# Biais de sélection (3)

## Détection des biais de sélection

- Comparaison avec plusieurs groupes témoins
- Comparaison répondants / échantillon de non-répondants
- Comparaison perdus de vue / personnes restant dans l'étude

## Contrôle des biais de sélection

- Au niveau de l'échantillonnage et au cours du suivi dans le cadre d'étude longitudinales
  - Bonne qualité du protocole (réflexion et anticipation)
  - Ressources pour suivi de l'étude
  - standardisation des modalités de sélection

# Biais d'information (1)

## Définition

- Distorsion dans l'estimation d'une mesure, liée à des défauts de classification des sujets
- Peut affecter la classification de l'estimation
  - De l'exposition (cas-témoins)
  - De la maladie (cohorte)
- Synonymes : misclassification, biais de mesure

## Paramètre qualitatif ou catégoriel

- L'erreur de classement concerne la survenue ou non de l'évènement mesuré

## Paramètre quantitatif

- Erreur de mesure du critère de jugement

# Biais d'information (2)

## Sources des biais d'information

- Outil de mesure imparfait
  - Questionnaire +++ s'il fait appel a la mémoire
  - Test diagnostique imparfait
- Cadre d'observation non neutre
  - Investigateur ou sujets influences par le statut (E/NE ou C/T)
- Nature des questions (questions sensibles telles que la sexualité)

## Conséquences

- Test diagnostic imparfait
  - Défaut de sensibilité : Evènements non détectés
    - Classés a tort comme « non évènement »
  - Défaut de spécificité : Faux négatifs
    - Classés a tort comme « n'ayant pas l'évènement »
- Cadre d'observation non neutre
  - Recherche plus active de la maladie chez les exposés
  - Recherche plus active de l'exposition chez les malades

# Biais d'information (3)

## Exemples de biais d'information

- Mémorisation
  - Les cas et les témoins se souviennent avec une acuité différente de leur exposition au facteur de risque
    - Une mère d'un enfant malformé se souvient mieux de ses antécédents pendant la grossesse (elle se sent coupable)
- Biais lié à l'enquêteur ou biais de subjectivité
  - L'enquêteur connaît l'hypothèse de l'étude. Et comme résultat interroge les cas de façon plus approfondie que les témoins bien portants
- Biais lié à l'enquêté
  - Prévarication: omission volontaire ou mensongère sur une question (toxicomanie)
  - Désirabilité sociale: répond faussement sur des questions sensibles (normes sociales)

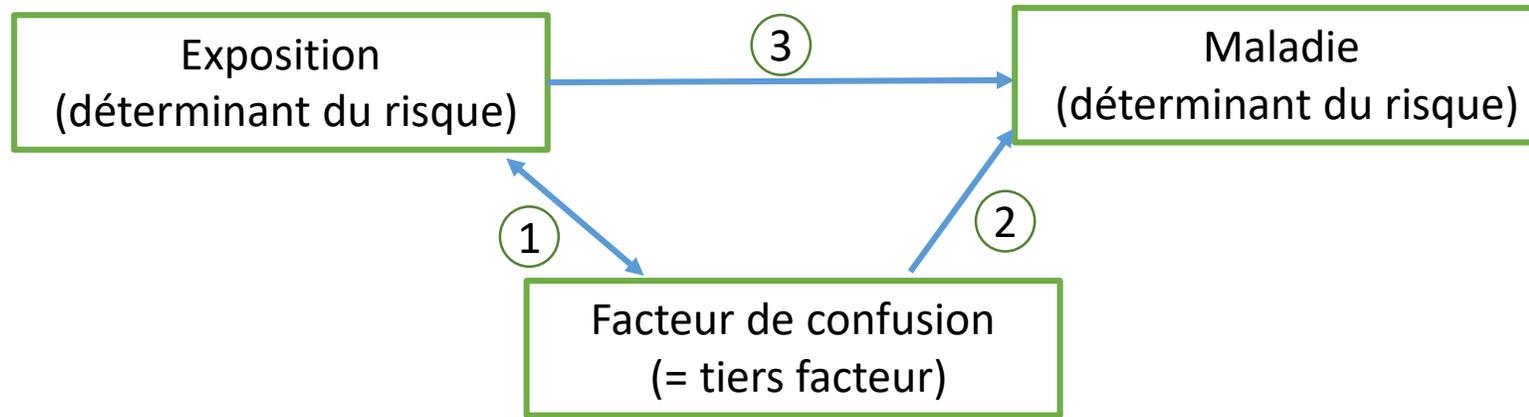
## Contrôle des biais d'information

- Surtout lors de la planification de l'étude
  - Standardisation de l'outil et du mode de recueil des données
  - Formation des enquêteurs

# Biais de confusion (1)

## Définition

Lorsqu'un facteur externe, "tiers facteur", entraîne une distorsion de l'association entre exposition et maladie



### 3 conditions

1. Association entre exposition et tiers facteur
2. Association entre tiers facteur et maladie
3. Tiers facteur n'est pas sur le chemin causal de la maladie



Relation entre exposition et événement différente si mesurée dans l'ensemble de l'échantillon ou dans les différentes catégories du facteur de confusion

# Biais de confusion (2)

## Contrôle du biais de confusion

- lors de la constitution des groupes de comparaison
  - Randomisation
    - Tirage au sort de l'affectation dans les groupes
  - Appariement
    - Recrutement de chaque personne B en fonction de la personne A
  - Assortiment de fréquence : garantit une égale répartition des facteurs de confusion dans les 2 groupes
    - Identifier le facteur de confusion, décrire dans le groupe A, choisir un groupe B ayant des caractéristiques similaires
  - Restriction
    - Le facteur de confusion (variable confondante) est restreinte à certaines catégories
- lors de l'analyse des données:
  - Ajustement (si les données concernant le tiers facteur ont été recueillies au cours de l'étude)

# Documents ressources

1. François Dabis, Jean Claude Desenclos. Epidémiologie de terrain. Méthodes et applications 2017.
2. Jean Bouyer. Méthodes statistiques. Médecine-Biologie. 2017.
3. Thierry Ancelle. Statistique-Epidémiologie, 4<sup>ème</sup> édition. 2017.
4. Paul-Marie Bernard, Claude Lapointe. Mesures statistiques en épidémiologie. 1998
5. R Bonita, R Beaglehole, T Kjellström. Basic epidemiology 1rd edition, 1993.
6. R Bonita, R Beaglehole, T Kjellström. Basic epidemiology 2nd edition, 2006.
7. Ahrens Wolfgang, Pigeot Iris. Handbook of epidemiology, 2006.

**Merci pour votre attention**