

Association, biais, confusion

Diapositives adaptées de celles des Prof. Angèle Gayet-Ageron et
Prof. Thomas PERNEGER

Prof. Christophe COMBESCURE

Unité d'Appui Méthodologique du CRC

Hôpitaux Universitaires de Genève & Faculté de Médecine

Révision du cours précédent

- **Etude cas-témoins**: plan d'étude pour examiner des associations entre un ou plusieurs facteurs de risque/protecteurs et une maladie
 - Surtout pour maladies rares
 - Efficace et économique
- Odds ratio: ratio des odds d'exposition des cas sur les témoins
- Odds ratio: mesure d'association applicable aux études cas-témoins (mesure de risque non interprétable car ratio cas/témoins arbitraire) : **ad/bc**

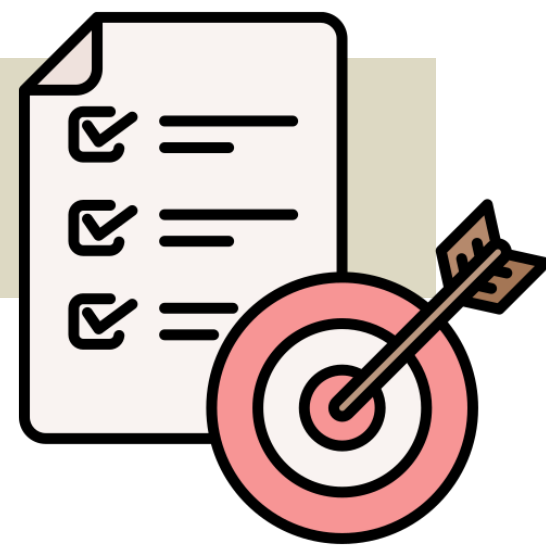
	Cas	Témoins
Exposé	a	b
Non-exposé	c	d

Révision du cours précédent

- Régression logistique
 - Méthode de modélisation pour des variables dépendantes (outcomes) binaires
 - Permet de rapporter des odds ratios pour variables indépendantes continues ou catégorielles
- Appariement des témoins aux cas
 - Méthode pour contrôler des facteurs de confusion; on choisit des témoins identiques aux cas sur les facteurs d'appariement choisis à l'avance
 - On analyse des paires [cas; témoin]
 - Odds ratio de McNemar: ratio du nombre de **paires discordantes**:
[cas exposé & témoin non-exposé] / [cas non-exposé & témoin exposé]

Objectifs de la séance

- Approfondir les notions suivantes
 - Association, et ses formes statistiques
 - Qu'est-ce qui peut expliquer une association
 - Causalité
 - Biais de sélection et biais d'information
 - Effet de confusion



Hypothèses scientifiques

- Portent souvent sur l'**association** entre 2 variables
- Est-ce que la dexaméthasone accélère la guérison de la pharyngite chez les adultes?
 - Variable 1: traitement: dexaméthasone ou placebo (exposition)
 - Variable 2: état du patient à 48 heures: guéri ou pas guéri (outcome)
- Est-ce que les personnes plus grandes ont aussi un poids plus élevé?
 - Variable 1: taille
 - Variable 2: poids
- Est-ce que le test urinaire rapide permet de diagnostiquer l'infection urinaire?
 - Variable 1: résultat du test rapide (nouveau test ou test index)
 - Variable 2: gold standard (culture des urines) (test de référence)

Question de recherche

- Les essais cliniques spécifient:

- Population (ou patients)
- Intervention (p. ex. nouveau traitement)
- Comparateur (p. ex. placebo)
- Outcome (ou variable d'évaluation, ou critère de jugement)

PICO

- Les études observationnelles spécifient:

- Population (ou patients)
- Exposition (cause, facteur de risque, ...)
- Outcome (survie, maladie, problème de santé, ...)

PEO

➤ La question de recherche porte sur **l'association** entre les 2 variables principales

Analyses statistiques

- **Décrivent** l'association dans l'échantillon de l'étude
 - Description et comparaison de 2 moyennes
 - Description et comparaison de 2 proportions
 - Nuage de points; pente de régression linéaire, corrélation
 - Description et comparaison de courbes de survie
 - Sensibilité et spécificité
- Tentent de déterminer si l'association existe aussi dans l'univers des paramètres (**inférence** statistique)
 - Estimation, intervalle de confiance
 - Test statistique, valeur p

Mesures d'association

- Différence entre 2 moyennes
- Pente de régression linéaire (différence moyenne de B pour un changement d'une unité de A)
- Différence de risque, ou « number needed to treat »
- Risque relatif, ou réduction relative du risque
- Ratio de risques instantanés (hazard ratio)
- Odds ratio
- Sensibilité et spécificité
- Aire sous la courbe ROC
- ...

Cinq explications possibles pour une association observée entre A et B

1. **Causalité**: A cause B
2. Causalité **inverse**: B cause A
3. Effet de **confusion** (« confounding »)
 - une troisième variable C produit l'association observée entre A et B
4. **Biais** (erreur systématique)
 - dû aux méthodes utilisées
5. **Hasard** (erreur aléatoire)

Toujours envisager ces 5 possibilités !!



Cinq explications possibles pour une association observée entre A et B

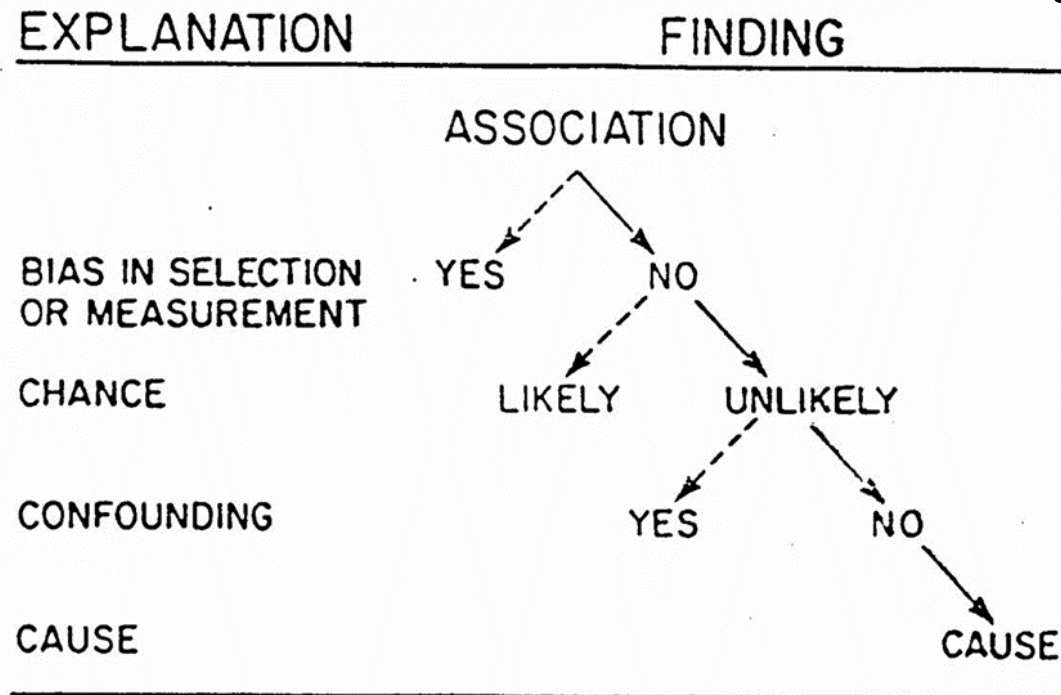


Figure 11.4. Association and cause.



Toujours envisager ces 5 possibilités !!

Une bonne étude scientifique....

- Apporte des arguments en faveur ou en défaveur de la causalité
- Limite et/ou décrit les effets de confusion résiduels
- Limite et/ou décrit le risque de biais liés au choix des méthodes
- Limite et/ou décrit le rôle du hasard

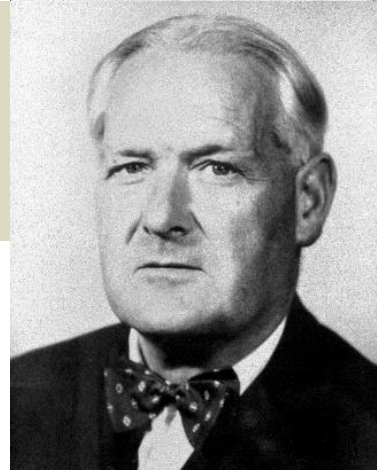


Causalité (entre l'exposition et l'*outcome*)

- Démontrée par une étude **expérimentale**: si on modifie la cause, l'effet change
 - expérimentation peut être éthiquement inacceptable
- Séquence **temporelle**: la cause doit précéder l'effet
- Autres arguments:
 - association forte
 - relation dose-réponse graduelle
- Dépend de la compréhension qu'on a des mécanismes sous-jacents (biologiques, moléculaires)
- N'est jamais prouvée par l'observation d'une association statistique, sans intervention expérimentale

9 Arguments en faveur de la causalité

selon Austin Bradford Hill, 1965



1. Force de l'association
2. Consistance (reproductibilité dans des situations différentes)
3. Spécificité (une cause, un effet)
4. **Temporalité** (cause précède l'effet)
5. Gradient (plus de cause implique plus d'effet)
6. Plausibilité (mécanisme causal décrit)
7. Cohérence (entre épidémiologie et biologie)
8. **Expérimentation** (manipulation de la cause induit changement de l'effet)
9. Analogie (avec des mécanismes semblables)

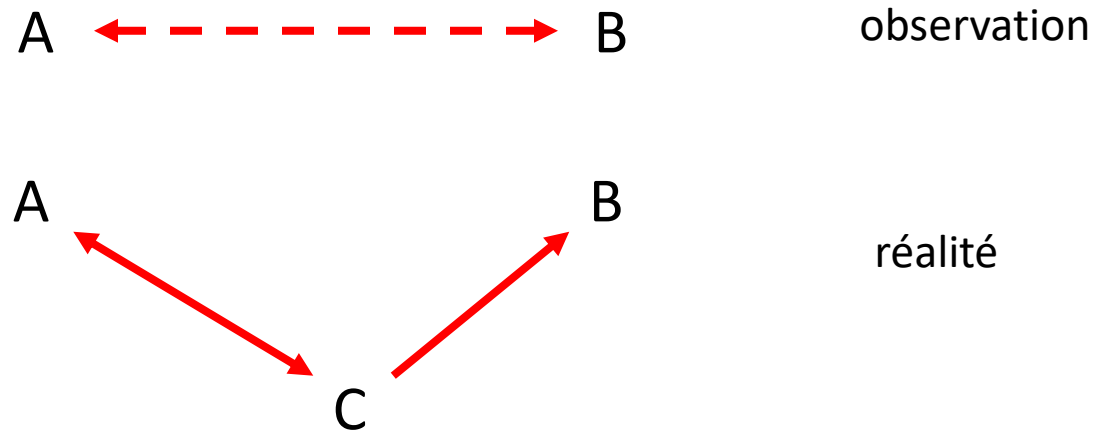


Effet de confusion



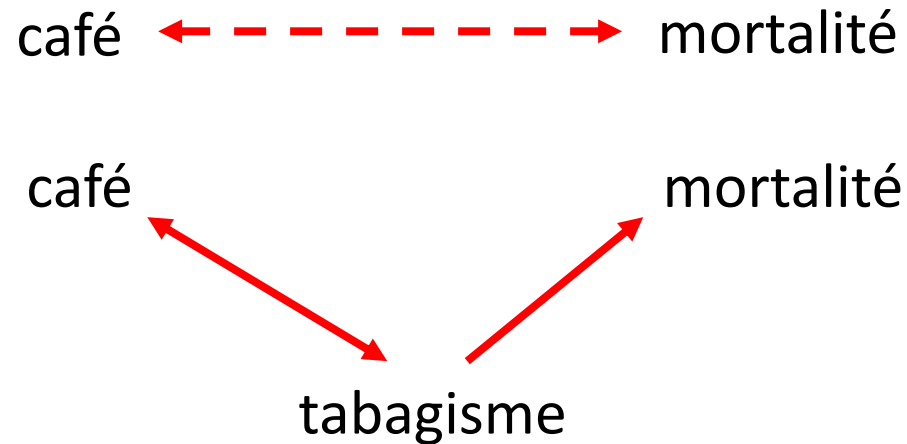
Effet de confusion

- Association apparente entre A et B est en fait due à une autre variable (facteur de confusion C)



Exemple

- Est-ce que la consommation de café augmente la mortalité?



- Si les buveur-ses de café sont plus susceptibles de fumer, on peut voir une association entre consommation de café et mortalité même si le café n'a aucun effet
- Le tabagisme est dans ce cas un facteur de confusion

Conditions à vérifier

- Le facteur de confusion C
 - doit être **une cause de B**
et
 - doit être **associé avec A**
mais
 - ne doit **pas être une conséquence de A**

Dans notre exemple:

- Le tabac cause le décès
et
- Il est associé au café
mais
- Le tabac n'est pas une
conséquence du café

Association of Coffee Drinking with Total and Cause-Specific Mortality

N Engl J Med 2012;366:1891-904.

Coffee is one of the most widely consumed beverages, but the association between coffee consumption and the risk of death remains unclear.

We examined the association of coffee drinking with subsequent total and cause-specific mortality among 229,119 men and 173,141 women in the National Institutes of Health–AARP Diet and Health Study who were 50 to 71 years of age at baseline. Participants with cancer, heart disease, and stroke were excluded. Coffee consumption was assessed once at baseline.

Etude de cohorte, suivi de 1995 et 2008

Association of Coffee Drinking with Total and Cause-Specific Mortality

Les décès sont plus fréquents chez les consommateurs de café (résultat descriptif)

Table 2. Association of Daily Coffee Consumption with Total and Cause-Specific Mortality among 229,119 Men.*

Cause of Death	All Participants	No Coffee (N=21,080)	≥6 Cups (N=10,139)
All causes			
No. of deaths (%)	33,731	2766 (13.1)	1904 (18.8)
Age-adjusted hazard ratio (95% CI)		1.00	1.60 (1.51–1.69)

HR ajusté estimé à l'aide d'un modèle de régression de Cox

Le HR ajusté pour l'âge est > 1:
A âge identique, la mortalité est plus élevée chez les exposés (consommateurs de café) que chez les non-exposés (non-consommateurs de café)

Characteristic	Men	
	No Coffee (N=21,080)	≥6 Cups (N=10,139)
Age (yr)		
Median	61.1	60.3
Interquartile range	56.4–65.7	55.9–64.8
Non-Hispanic white (%) [†]	91.1	97.3
Family history of cancer (%) [‡]	46.3	48.5
Currently married (%)	83.0	83.4
College graduate (%)	53.0	37.2
Median body-mass index [§]	26.4	26.6
Diabetes (%)	8.4	8.0
Current smoker (%)	4.8	34.7
>3 Alcoholic drinks/day (%)	6.3	11.6
Vigorous physical activity ≥5 times/wk (%)	24.8	20.6
Poor or fair self-reported health (%)	7.5	9.3
Median total energy intake (kcal/day)	1869	2175
Median servings of food		
Fruits (servings/day)	1.5	0.8
Vegetables (servings/day)	1.9	1.7
Red meat (g/day)	33.1	40.5
White meat (g/day)	25.8	21.5
Use of any vitamin supplement (%)	59.6	55.5

Les consommateurs de café sont:

- plus fréquemment des fumeurs
- plus fréquemment des consommateurs d'alcool
- font globalement moins d'activités physiques
- consomment en moyenne moins de fruits
- consomment en moyenne plus de viande rouge

Association of Coffee Drinking with Total and Cause-Specific Mortality

Cause of Death	All Participants	No Coffee (N = 21,080)	≥6 Cups (N = 10,139)
All causes			
No. of deaths (%)	33,731	2766 (13.1)	1904 (18.8)
Age-adjusted hazard ratio (95% CI)		1.00	1.60 (1.51–1.69)
Multivariate-adjusted hazard ratio (95% CI)		1.00	0.90 (0.85–0.96)

Multivariate analyses were adjusted for the following factors at baseline: age; body-mass index; race or ethnic group; level of education; alcohol consumption; the number of cigarettes smoked per day, use or nonuse of pipes or cigars, and time of smoking cessation (<1 year, 1 to <5 years, 5 to <10 years, or ≥10 years before baseline); health status; diabetes (yes vs. no); marital status; physical activity; total energy intake; consumption of fruits, vegetables, red meat, white meat, and saturated fat; and use or nonuse of vitamin supplements. In addition, risk estimates for death from cancer were adjusted for history of cancer (other than nonmelanoma skin cancer) in a first-degree relative (yes vs. no).

Le HR pour la consommation de café ajusté sur tous ces facteurs est < 1:

A niveau identique des facteurs d'ajustement (listés sous la table), la mortalité est plus **basse** chez les exposés (consommateurs de café) que chez les non-exposés (non-consommateurs de café)

Interprétation

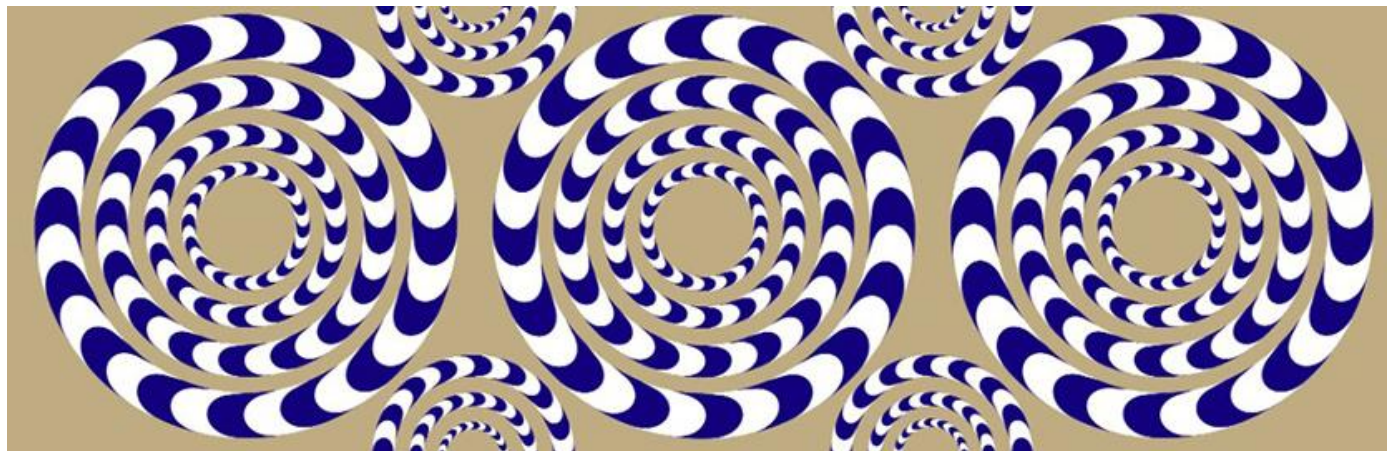
- Les consommateurs de café (≥ 6 tasses/j) ont une **mortalité observée plus élevée** que ceux qui ne boivent jamais de café: HR = 1.6
- **Si on ajuste** pour des facteurs de confusion potentiels (tabagisme, alcool, sport, alimentation), l'effet s'inverse, et **les consommateurs de café ont une mortalité réduite de 10%** par rapport aux non-consommateurs: HR = 0.9
- Le **changement** du « hazard ratio » de 1.6 à 0.9 est la conséquence de l'effet des **facteurs de confusion**
- Ceci est arrivé parce qu'en moyenne les buveurs de café ont des habitudes de vie moins saines que ceux qui ne boivent pas de café

Manifestation d'un effet de confusion

- 1) Examiner l'association de A et B sans tenir compte de C
- 2) Refaire l'analyse dans chaque niveau de C (p. ex. pour C=1, C=2, C=3...)
C étant constant au sein de chaque strate, cette variable ne peut pas expliquer l'association entre A et B
- 2_{bis}) Alternativement, ajuster l'association entre A et B pour C dans un modèle de régression
- 3) Associations sous 1) et 2)
Si elles sont **semblables** → C n'est pas un facteur de confusion
Si elles sont **différentes** → C est un facteur de confusion

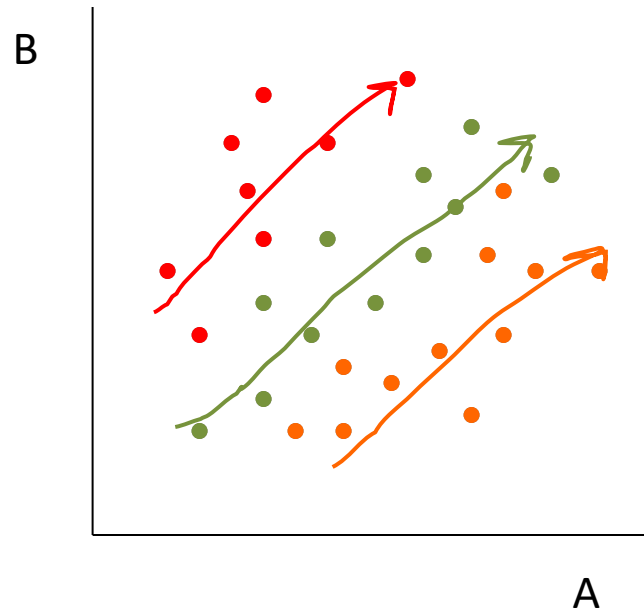
Types d'effets de confusion

- Un effet de confusion peut:
 - faire apparaître une association entre A et B là où aucune association causale n'existe
 - faire disparaître une association causale existant entre A et B
 - inverser le sens d'une association
 - renforcer ou affaiblir une association



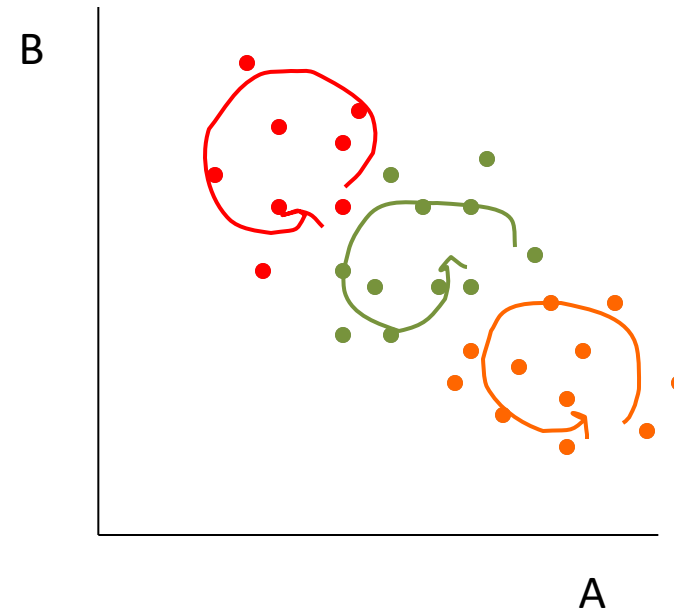
Graphiquement

3 strates de C



A et B sont associés dans chaque strate de C

3 strates de C



A et B ne sont associés dans aucune strate de C

Comment neutraliser les facteurs de confusion ?

- **Randomisation** (essais cliniques randomisés)
 - l'allocation aléatoire des patients équilibre les caractéristiques des 2 groupes de traitement, en moyenne
 - ceci casse le lien entre les facteurs de confusion C et le traitement A
 - cela marche aussi pour les facteurs de confusion inconnus!
 - attention: pas de garantie, surtout si les groupes sont petits
- **Appariement** (études cas-témoins appariées)
 - on choisit les témoins de sorte qu'ils soient identiques aux cas sur certains facteurs de confusion présumés
 - ceci casse le lien entre ces facteurs de confusion C et la maladie B
 - attention: on ne peut appairer que sur peu de variables, d'autres effets de confusion peuvent persister

Comment neutraliser les facteurs de confusion ? (suite)

- **Stratification**

- on analyse l'association entre A et B dans chaque niveau de C
- l'association entre A et B dans chaque niveau de C ne peut pas être confondue par C

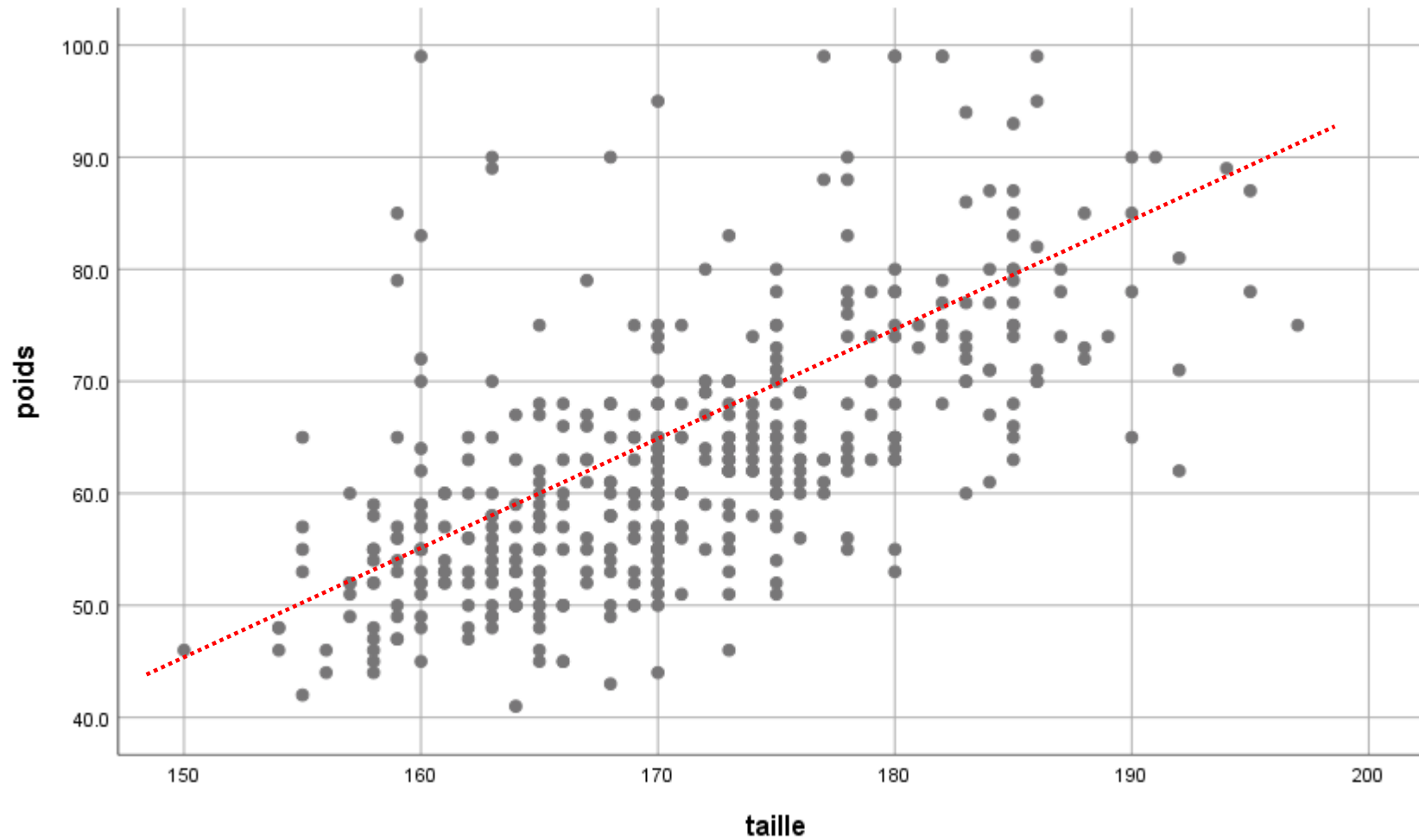
- **Ajustement statistique**

- on crée un modèle statistique: B prédit par A
- on ajuste ce modèle pour C: B prédit par A et C
- on regarde si l'association entre A et B **change** après cet ajustement pour C
- cette démarche explique une grande partie des modèles multivariés utilisés dans les articles scientifiques médicaux d'études observationnelles

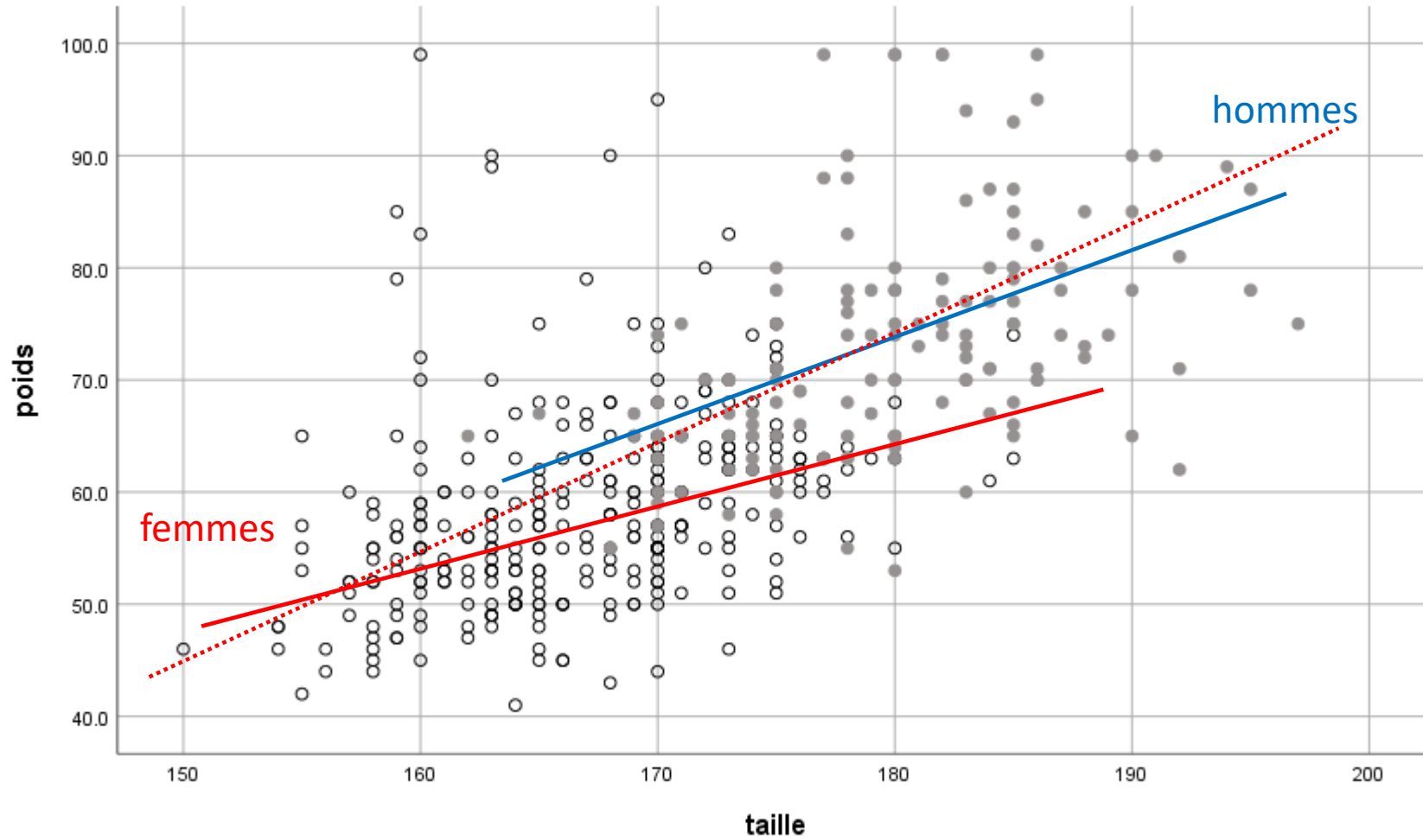
Exemple

- Est-ce que le sexe biologique confond l'association entre le poids et la taille de jeunes adultes?
- Population: étudiant-es en médecine
- Variable indépendante: taille en cm
- Variable dépendante: poids en kg
- Variable de confusion: **sexe biologique (assigné à la naissance)**

Association brute



Association par sexe



Interprétation

- La taille est associée au poids chez les étudiant-es; les plus grand-es sont aussi plus lourd-es
- La taille est associée au poids parmi les femmes et parmi les hommes, mais pour chaque sexe la pente est **moins forte** que dans l'échantillon total
- Ceci indique que le sexe est un facteur de confusion de cette association
- Ceci arrive parce que le sexe est
 - Associé à la taille: les femmes sont plus petites que les hommes
 - Associé au poids: à taille égale, les femmes sont plus légères que les hommes

Régression linéaire simple

- Poids = -79.105 + 0.832*taille

Model		Unstandardize
		B
1	(Constant)	-79.105
	taille	.832

a. Dependent Variable: poids

Sig.	95.0% Confidence Interval for B	
	Lower Bound	Upper Bound
.000	-95.558	-62.653
.000	.736	.928

-79.105 (intercept ou constante): poids en kg attendu si la taille vaut 0 cm

Cette interprétation est absurde dans cet ex, et illustre le danger de l'extrapolation au-delà des données

0.832 (pente): gain de poids en kg pour un centimètre supplémentaire de taille

Régression linéaire multiple

- $\text{Poids} = -43.595 + 0.569 * \text{taille} + 7.132 * \text{sexe}$

Model		Unstandardized B
1	(Constant)	-43.595
	taille	.569
	sexe	7.132

a. Dependent Variable: poids

Sig.	95.0% Confidence Interval for B	
	Lower Bound	Upper Bound
.000	-63.713	-23.476
.000	.439	.699
.000	4.654	9.610

- - 43.595 : poids en kg attendu si la taille vaut 0 cm et sexe=0
 - Là aussi c'est absurde, d'autant que sexe=0 n'existe pas
- 0.569: gain de poids en kg pour un centimètre supplémentaire de taille que l'on soit femme ou homme
- 7.132: supplément de poids si on est un homme, quelle que soit la taille
 - Si femme sexe=1, si homme sexe=2

Effet de confusion

- Association taille-poids brute: **+0.832** kg/cm
- Association ajustée pour le sexe/genre: **+0.569** kg/cm
- La pente est réduite d'un tiers environ après ajustement sur le sexe biologique
- Le sexe est un facteur de confusion parce que
 - les femmes sont plus petites que les hommes
 - à taille égale, les femmes sont plus légères que les hommes

Dans l'autre sens aussi...

- Différence brute de poids entre hommes et femmes

sexe	Mean	N	Std. Deviation
1	58.365	293	9.0243
2	73.092	141	10.3551
Total	63.150	434	11.7162

- En régression linéaire:

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	43.638	1.365		31.974	.000	40.956	46.321
	sexe	14.727	.971	.589	15.163	.000	12.818	16.636

a. Dependent Variable: poids

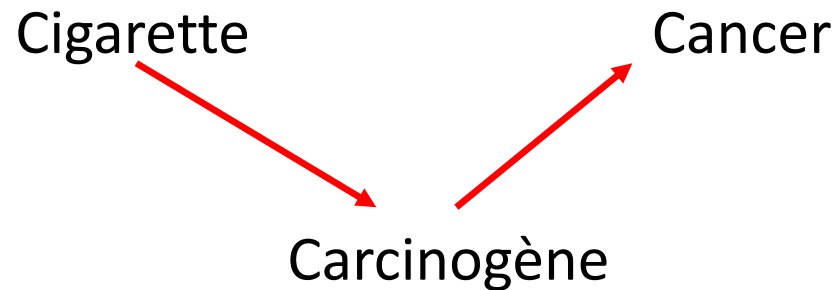
- Sans ajuster pour la taille, les hommes pèsent 14.727 kg de plus que les femmes
- A taille égale, les hommes pèsent 7.132 kg de plus que les femmes (modèle précédent)

Effet de confusion bis

- Différence de poids brute: **14.7** kg de plus pour les hommes
- Différence ajustée pour la taille: **7.1** kg de plus pour les hommes
- La différence est réduite de moitié environ après ajustement
- L'effet de confusion dû à la taille renforce la différence de poids entre hommes et femmes
- A noter
 - **Les 2 analyses sont justes!**
 - Les hommes pèsent vraiment 14.7 kg de plus que les femmes en moyenne
 - A taille égale, la différence vaut vraiment seulement 7.1 kg

Chaîne causale ≠ effet de confusion

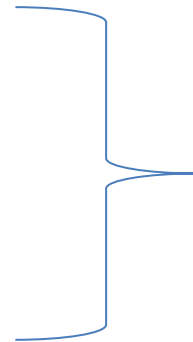
- Dans certains cas, des variables forment une chaîne causale
- Ex:
 - Fumer des cigarettes (A) entraîne
 - La présence de carcinogènes dans les alvéoles (C), qui cause
 - Le cancer du poumon (B)



- Dans ce cas on ne parle pas d'effet de confusion, mais de **cause distale** (tabagisme) et de **cause proximale** (carcinogène)

Suite du cours...

1. Causalité
2. Effet de confusion
3. Biais (erreur systématique)
4. Hasard (erreur aléatoire)



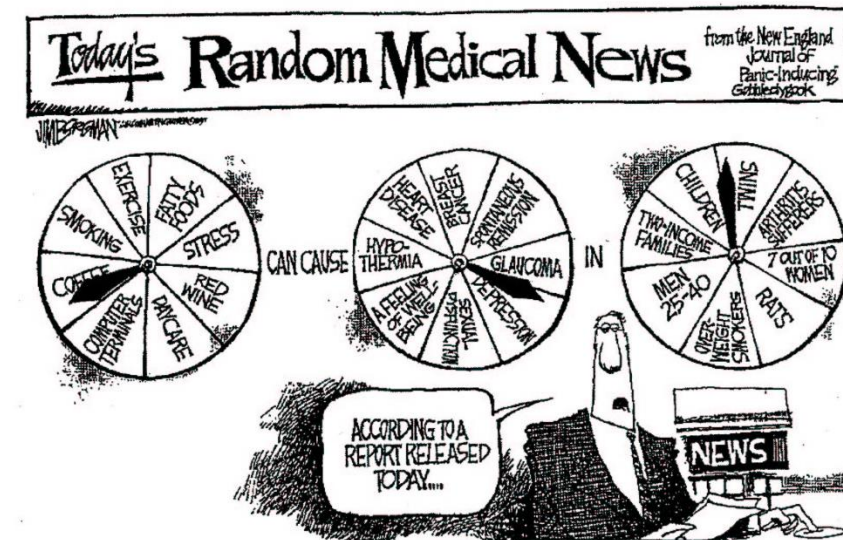
Deux types d'erreurs

Deux sortes d'erreurs

- Différences entre résultat de l'étude (ce qu'on a mesuré) et la réalité ou vérité (ce qu'on souhaite connaître)
- **Aléatoires**
 - Dues au seul hasard qui intervient lors de la sélection de l'échantillon de l'étude
 - Indépendantes : l'erreur qui affecte un échantillon ne dit rien de ce qui arrivera dans le prochain échantillon
- **Systematiques**
 - Dues à des procédures de recherche imparfaites (manière de choisir l'échantillon, de mesurer les variables, d'analyser les données,...)
 - Prévisibles: si les méthodes restent les mêmes, les erreurs systematiques iront dans le même sens

Erreurs aléatoires

- Évaluées par
 - Largeur de l'intervalle de confiance
 - Probabilités d'erreurs de type 1 et type 2
- Il faut toujours **envisager la possibilité** qu'un résultat d'étude soit une erreur de type 1 ou 2, ou que l'intervalle de confiance ne contienne pas la vraie valeur du paramètre
- **Comment les limiter ?**
 1. Échantillon suffisamment grand
 2. Mesures précises des variables (bons instruments)
 3. Usage de procédures statistiques adaptées



Erreur systématique: biais

- Caractérise une méthode (plutôt qu'un résultat)
- Si on refait l'étude de multiples fois, la moyenne des observations n'est pas la bonne valeur
 $E(\text{estimateur}) \neq \text{paramètre}$
- Un biais survient à cause de procédures non-optimales de l'étude, qui orientent le résultat dans une certaine direction
- L'erreur est dite «systématique» parce que si on refait l'étude de la même manière on aura le même type d'erreur






Types de biais

- **Biais de sélection**
 - Échantillon non-représentatif de la population visée
- **Biais d'information**
 - Mesures et recueil des données incorrects
- **Biais d'atténuation** (*misclassification bias*)
 - Affecte les mesures d'association
 - Conséquence des erreurs de mesure aléatoires
- **(Biais de publication/diffusion)**
 - Publication sélective de résultats « intéressants »



Exemples de biais de sélection

- **Biais de recrutement**  jargon
 - Les patient-es référé-es à un hôpital ont souvent des formes de maladies plus graves que ceux/celles qui sont vu-es par les médecins en ville
- **Biais de volontaire sain**  jargon
 - Les volontaires sont habituellement en meilleure santé, et ont des habitudes plus saines (abstention du tabac, sport, etc.) que dans la population générale
- **Biais de survie**  jargon
 - Les patient-es qui survivent x mois/années après le diagnostic n'ont pas le même profil que ceux/celles qui sont décédé-es plus tôt; p.ex. cas « agressifs » de cancer

Dépistage du cancer du sein

- Chez les femmes de plus de 50 ans, on recommande un dépistage du cancer du sein par mammographie
- But: diagnostiquer et traiter la maladie plus tôt, pour limiter les séquelles du traitement et améliorer la survie

Exemple fictif

- *Après l'introduction du dépistage du cancer du sein, on observe la survie des femmes dont le diagnostic a été fait par dépistage, et on la compare aux données historiques du Registre des tumeurs*
- *On observe une meilleure survie dans le groupe dépisté (90%) que dans le groupe historique (80%)*
 - **Peut-on conclure que le dépistage réduit la mortalité?**

Dépistage: biais de sélection

Exemple fictif

- Les femmes qui décident de se faire dépister pourraient être différentes des femmes qui ne le font pas:
 - Meilleure santé
 - Plus jeunes
 - Habitudes de vie plus favorables
 - Plus motivées à suivre les traitements médicaux
- Mais aussi... (effet opposé!)
 - Plus susceptibles d'avoir une parente proche atteinte d'un cancer, donc à risque plus élevé
- Si ces variables sont aussi liées à la survie, elles pourraient expliquer la différence observée en termes de survie

Dépistage: biais de prévalence

Exemple fictif

- Le dépistage est possible lorsqu'il existe une phase pré-clinique pendant laquelle la maladie est détectable
- La durée de cette phase peut être variable
- Supposons qu'il y a autant de cas à évolution **rapide** qu'à évolution **lente**



- On dépiste plus aisément les cas à évolution lente, qui pourraient avoir une meilleure survie

Biais d'information (exemples)

- « **Recall bias** » dans une étude cas-témoins



jargon

- Les cas, malades, rapportent plus volontiers un facteur de risque que les témoins, non-malades.
- Ex: la mère d'un bébé porteur de malformation risque de se rappeler plus d'expositions à des toxiques que la mère d'un bébé qui va bien

- **Biais d'investigation ou de suivi médical**



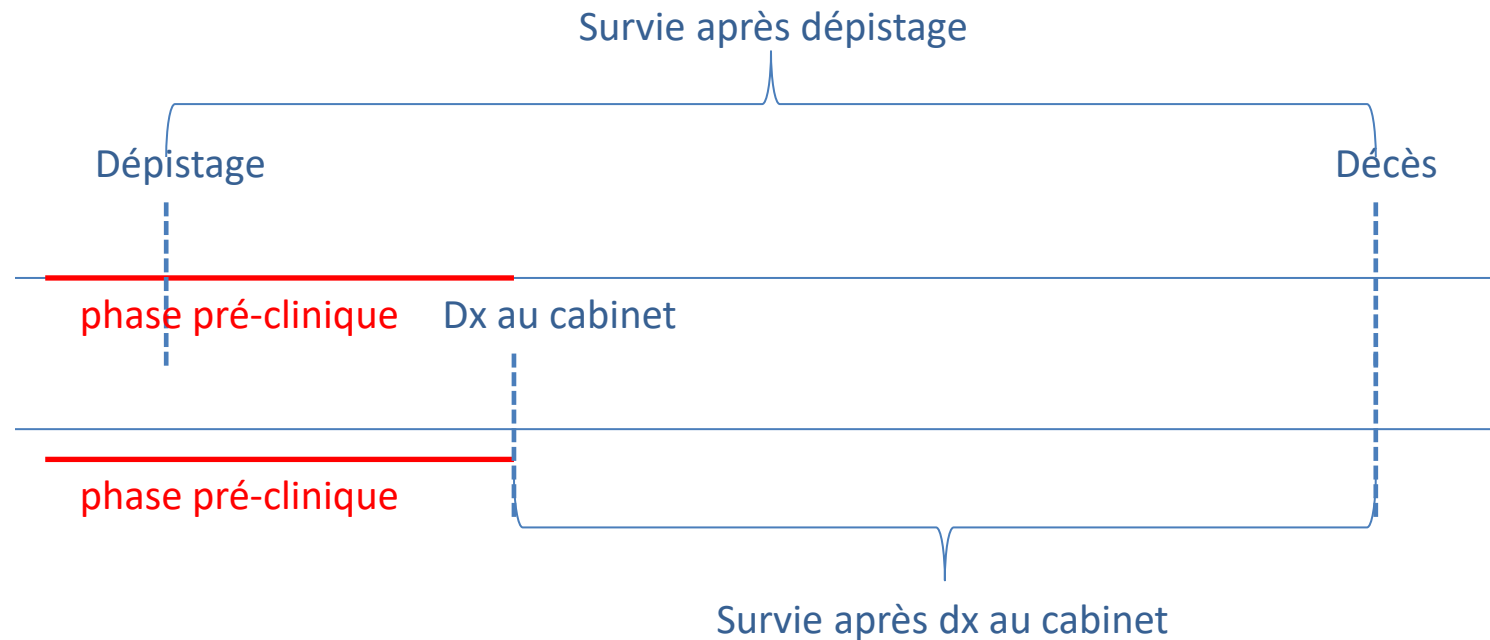
jargon

- Lorsqu'on investigate un-e patient-e, on risque de découvrir des anomalies qui seraient non-détectés dans la vie courante
- Un suivi serré (p. ex. trimestriel) va découvrir plus de complications ou diagnostics qu'un suivi plus lointain (p. ex. annuel) ou passif

Dépistage: biais de temps zéro

Exemple fictif

- On compare la survie de personnes selon le mode de diagnostic:
 - Dépistage précoce
 - Cabinet du médecin, suite à des symptômes



- La survie semble plus longue après dépistage, par avancement du temps=0

« *Misclassification bias* »

- Affecte les mesures d'**association**
- Erreurs aléatoires dans les mesures → **biais conservateur** de la mesure d'association, en direction de H_0 (p ex H_0 : OR = 1)
- Supposons que les valeurs de l'exposition sont mesurées correctement dans un sous-groupe de cas et témoins et assignées à pile ou face dans un autre

Mesures correctes

a	b
c	d

$$OR = ad/bc$$

Pile ou face



0.5	0.5
0.5	0.5

$$OR = 1$$

OR observé sera «dilué», entre ad/bc et 1



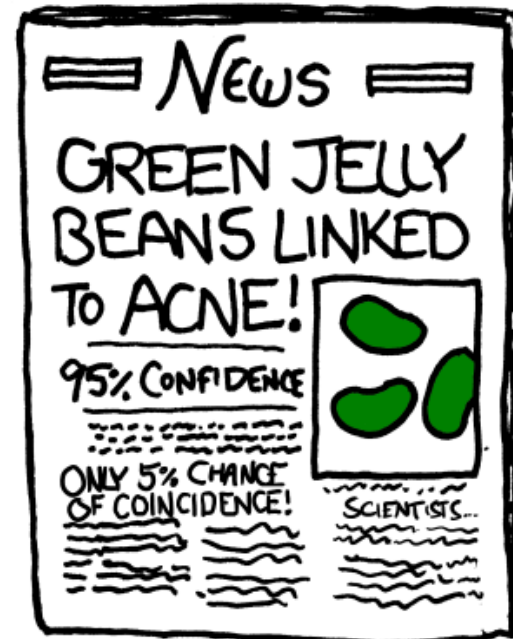
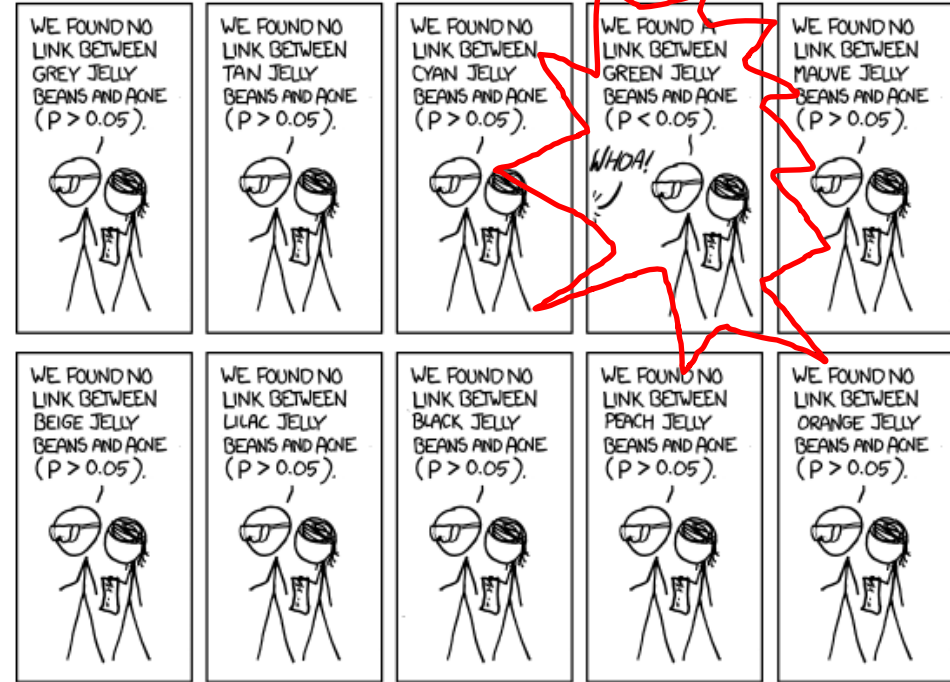
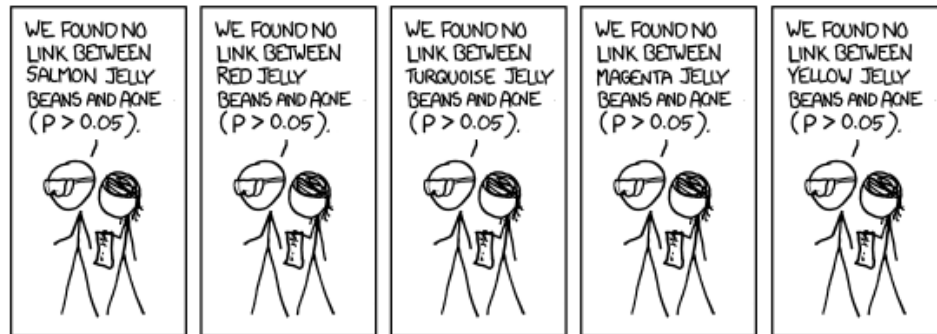
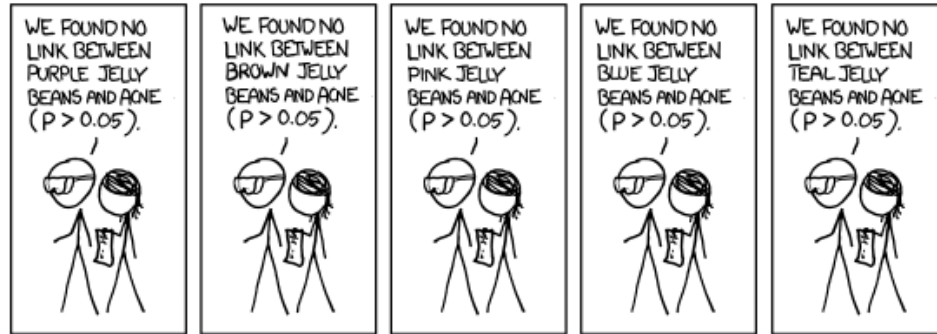
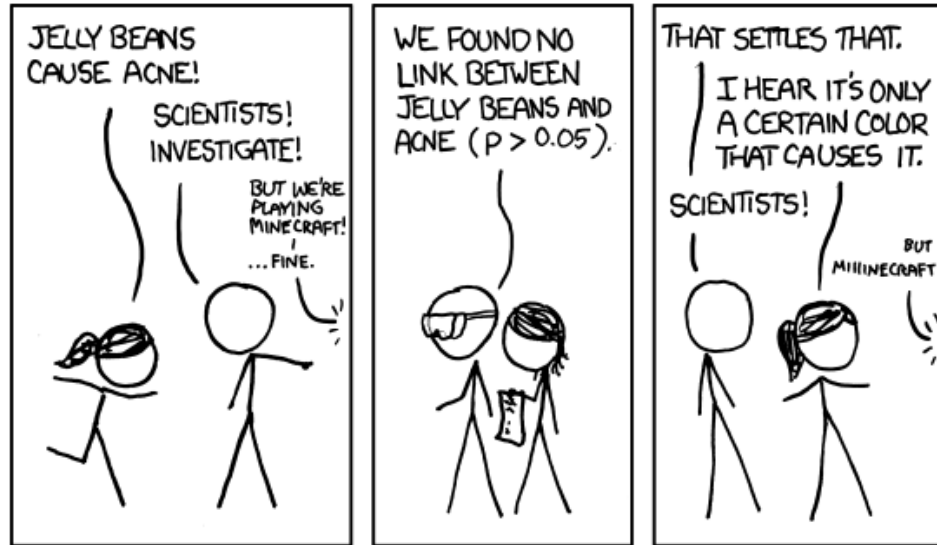
Exemples de biais de publication

- « *Fishing expedition* »  jargon
 - Une investigatrice mesure k variables chez les participant-es, et explore toutes les paires possibles ($k*(k-1)/2$ tests), dont 5% vont donner un résultat significatif même si toutes les hypothèses nulles sont vraies
- **Biais éditorial**  jargon
 - Acceptation plus aisée par les journaux professionnels d'articles qui rapportent des résultats « excitants », « significatifs », « provocateurs » ou socialement désirables
 - Remis en question très largement mais c'est un fait



SIGNIFICANT

< < PREV RANDOM NEXT > >



Comment prévenir les biais ?

- Echantillon de l'étude représentatif de la population d'intérêt (p. ex. échantillon aléatoire)
- Variables mesurées de manière précise et standardisée (avec les mêmes méthodes dans tous les groupes de l'étude)
- Questions de recherche et analyses pré-spécifiées dans un protocole
- Tous les résultats doivent être publiés

Conclusions (1)

- La plupart des questions scientifiques s'intéressent à des **associations** entre variables
- Plusieurs types (présentations) d'association:
 - Différence de moyenne, de risque
 - Risque relatif, odds ratio, etc.
- Cinq explications possibles pour une association:
 - **Causalité, causalité inverse, effet de confusion, biais, hasard**
- Causalité est démontrée par l'expérimentation
- Biais (erreur systématique) vs. hasard (erreur aléatoire)
- Deux sortes principales de biais:
 - Biais de sélection, biais d'information

Conclusions (2)

- Effet de **confusion** est dû à une variable tierce qui est associée tant au facteur de risque (A) qu'à la maladie (B)
- Pour neutraliser les effets de confusion:
 - Randomisation (essais cliniques)
 - Appariement (certaines études cas-témoins)
 - Stratification
 - Ajustement statistique (modèle multivarié)

Objectifs prochaine séance

- Comprendre les notions suivantes:
 - Revue systématique d'articles sur un thème donné
 - Méta-analyse d'essais cliniques randomisés
 - Effet commun (odds ratio, différence de moyenne,...)
 - Hétérogénéité vs. homogénéité
 - Biais de publication



Chapitres Petrie/Sabin
43: Méta-analyse