



Ma thèse en 5 questions clés

Dr Bastien Genet

Gériatre et Chef de Clinique en Santé Publique



If my thesis was a person



De quoi va-t-on parler ?

- Revue et Zotero
- Qualitatif et **Quantitatif**
- Réglementaire
- Collection de données
- Analyse de données
- **Rappel de stats de base**

Mais déjà ... Les questions clés

1. Quelle étude dois-je choisir et pourquoi ?
2. A quelles réglementations et démarches vais-je avoir affaire?
3. Comment collecter mes données ?
4. Comment analyser mes données ?
5. De quels outils et soutiens puis-je disposer ?

Quelle étude ?



The question is ...

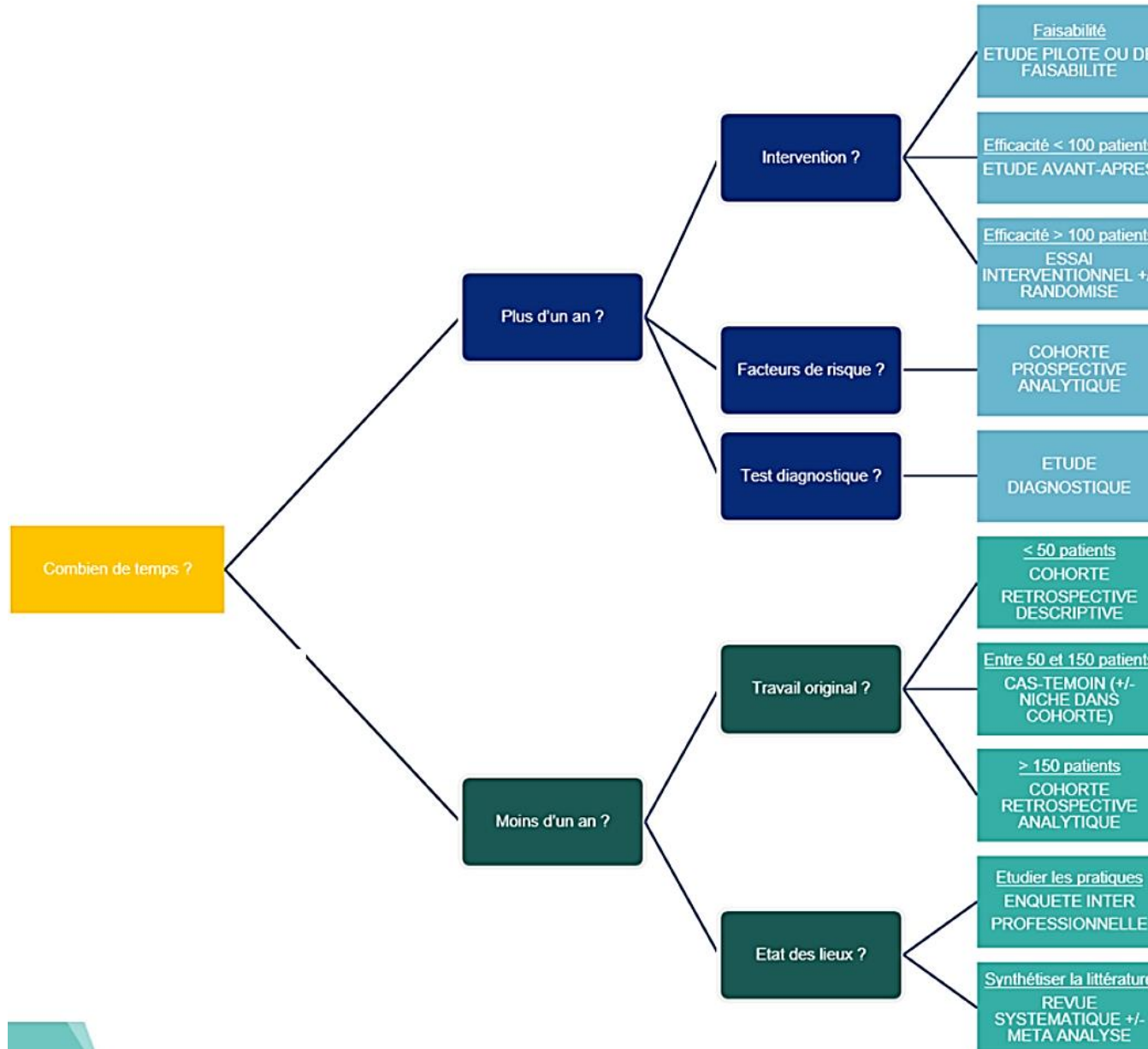
Quantitatif ou qualitatif ?

Raisonnement déductif, décrire avec des chiffres
Quantitatif

Raisonnement inductif, comprendre en profondeur
Qualitatif

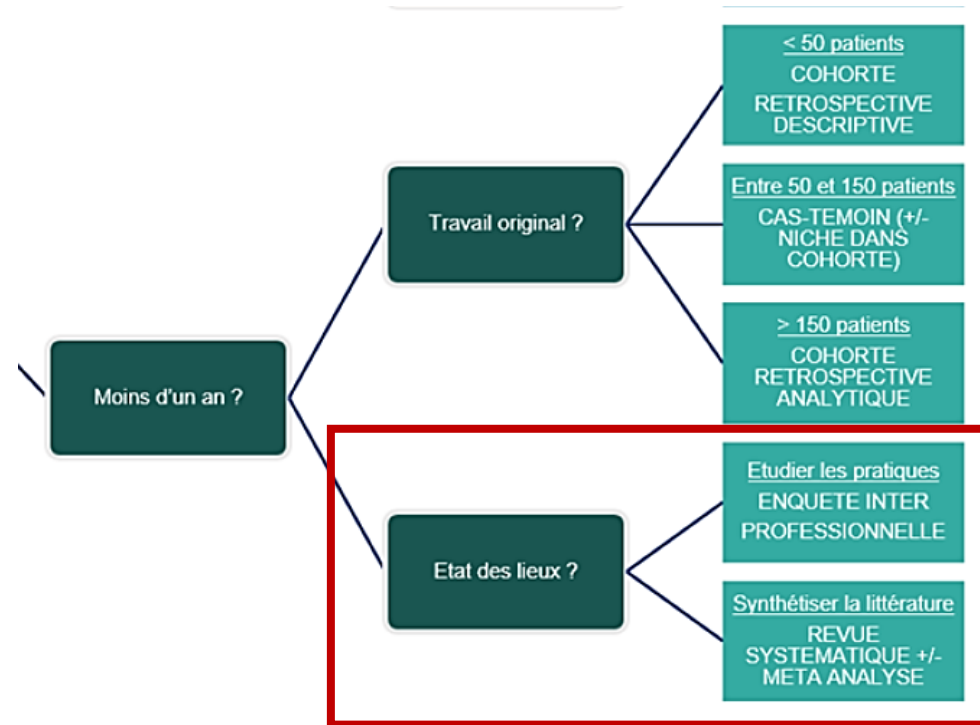
Les 2 mon cher Watson !
Méthodes Mixtes

Le Quanti – Comment ?



<https://www.figma.com/proto/sBKTkKcGoWrmAaZ5VOcvft/ALGO-THESE-V1?node-id=26-2&scaling=scale-down&page-id=0%3A1&starting-point-node-id=26%3A2&hotspot-hints=0&hide-ui=1>

Le Quanti – Comment ?



Le Quanti – Comment ?

Etudier les pratiques

On s'oriente ici plutôt vers une :

Enquête interprofessionnelle

Pourquoi et comment ?

- 1) Pas le meilleur niveau de preuve certes mais ...
- 2) Un but CONCRET : améliorer les pratiques et les conditions de travail
- 3) En faisant des questionnaires à destination des professionnels de santé

Synthétiser la littérature

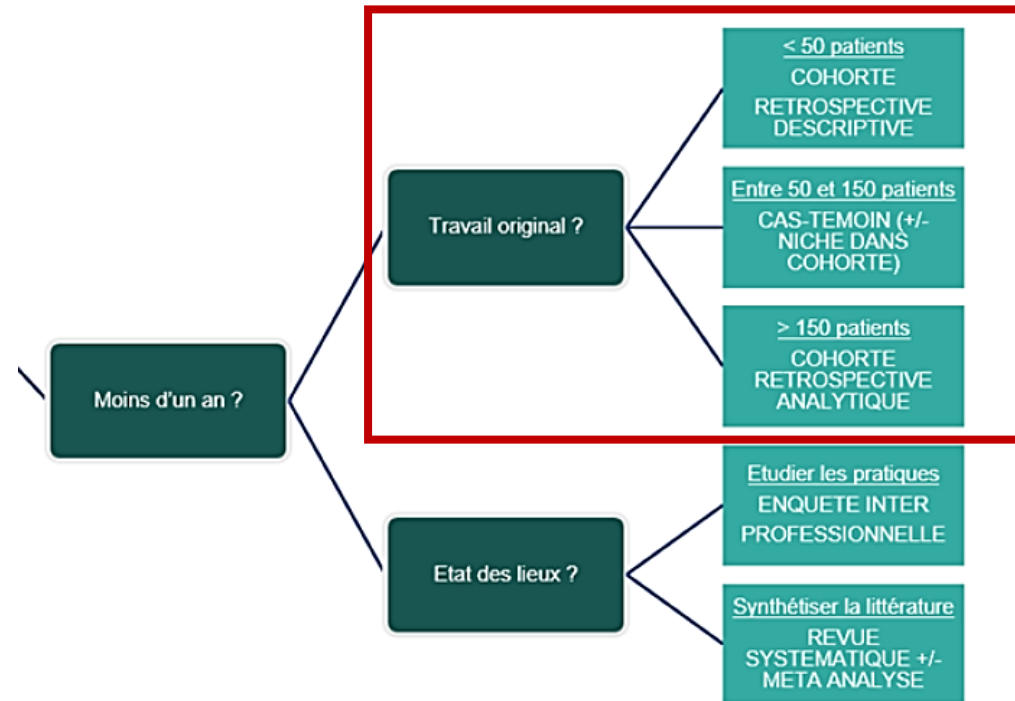
On s'oriente ici plutôt vers une :

Revue systématique +/- Méta analyse

Pourquoi et comment ?

- 1) Articles et références à lire sans collection de données
- 2) Comparaison standardisée des études sélectionnées (critères PRISMA)
- 3) Méta-analyse selon hétérogénéité des données

Le Quanti – Comment ?



Le Quanti – Comment ?

Combien de patients
sont recrutables ?

Objectif
Décrire une
population

Moins de 50

On s'oriente ici plutôt vers une :

Cohorte Rétrospective Descriptive

Pourquoi et comment ?

- 1) Données à récolter simplement (pas de protocole, certificat de non opposition des patients)
- 2) Pas assez de puissance pour faire de l'analytique (univarié mais multivarié non possible)

Combien de patients
sont recrutables ?

Objectif Rechercher
des facteurs de
risque

Entre 50 et 150

On s'oriente ici plutôt vers un :

Cas-Témoins (+/- niché dans une cohorte)

Pourquoi et comment ?

- 1) Données à récolter simplement (pas de protocole, certificat de non opposition des patients)
- 2) Pas assez de puissance pour une cohorte rétrospective analytique
- 3) Le cas-témoin niché dans une cohorte sera toujours meilleur

Combien de patients
sont recrutables ?

Objectif Rechercher
des facteurs de
risque

Plus de 150

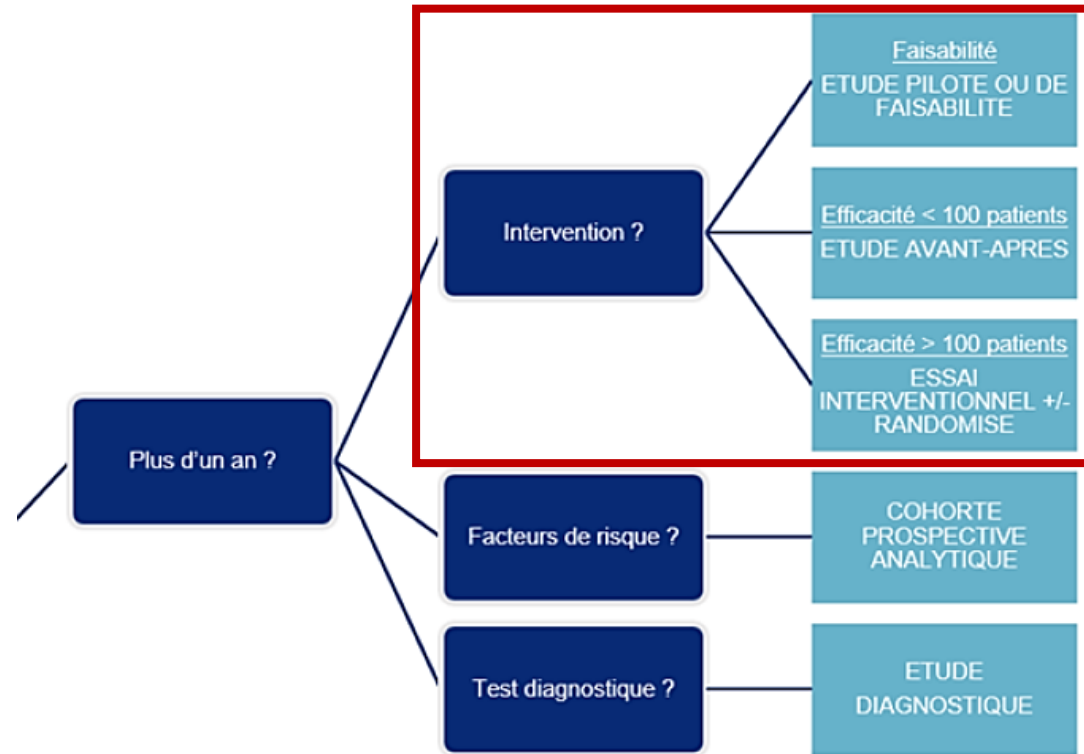
On s'oriente ici plutôt vers une :

Cohorte Rétrospective Analytique

Pourquoi et comment ?

- 1) Données à récolter simplement (pas de protocole, certificat de non opposition des patients)
- 2) Assez de puissance pour faire de l'analytique sans utiliser une étude cas-témoins

Le Quanti – Comment ?



Le Quanti – Comment ?

Combien de patients sont recrutables ?

Objectif
Tester l'efficacité d'une intervention

Plus de 100

On s'oriente ici plutôt vers un :

Essai interventionnel +/- randomisé

Pourquoi et comment ?

- 1) On a le temps et les moyens, c'est parti pour l'étude d'élite de la recherche biomédicale
- 2) Attention cependant aux coûts et à la réglementation plus exigeante

Combien de patients sont recrutables ?

Objectif
Tester l'efficacité d'une intervention

Moins de 100

On s'oriente ici plutôt vers une :

Etude de type Avant - Après

Pourquoi et comment ?

- 1) On veut tester l'efficacité de notre intervention
- 2) Mais on ne possède pas assez de patients pour trouver une éventuelle différence
- 3) On prend donc le patient comme son propre témoin

Combien de patients sont recrutables ?

Objectif
Tester l'acceptabilité d'une intervention

Une faisabilité

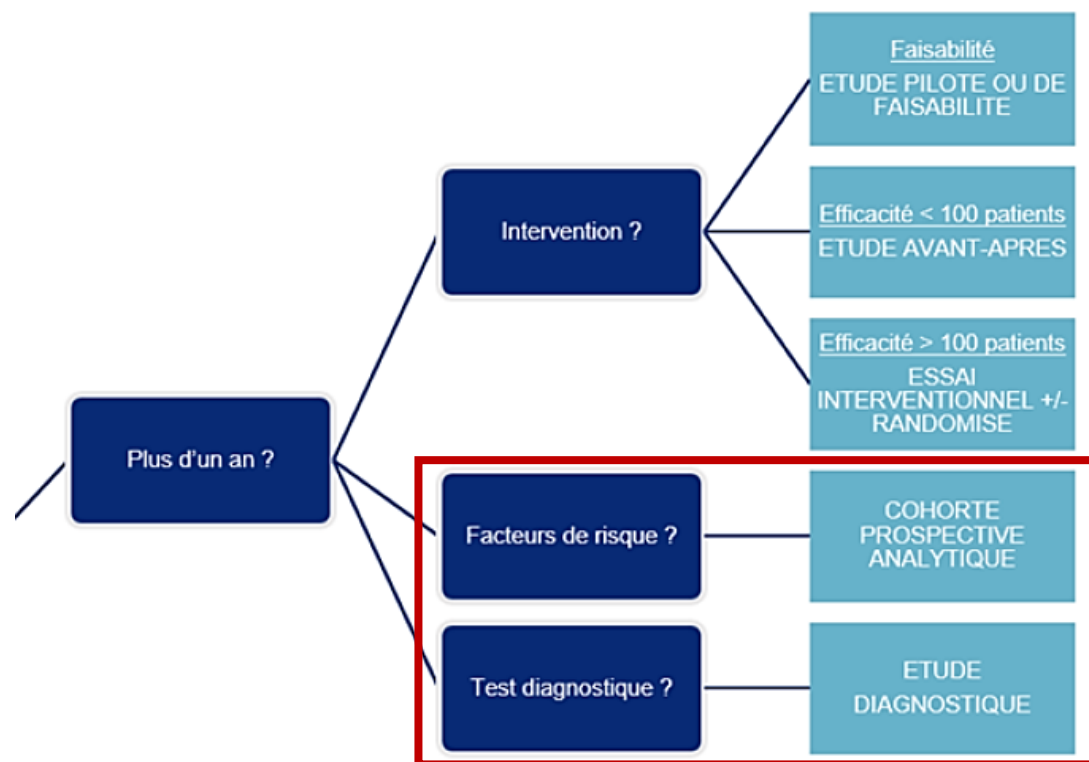
On s'oriente ici plutôt vers une :

Etude pilote ou de faisabilité

Pourquoi et comment ?

- 1) Notre objectif est de voir si notre intervention est acceptée par la population
- 2) Elle est généralement préliminaire à un essai interventionnel de plus grande envergure

Le Quanti – Comment ?



Le Quanti – Comment ?

Qu'est-ce que
j'évalue ou étudie ?

Objectif
Tester la
performance d'un
test

Un test diagnostic

On s'oriente ici plutôt vers une :

Etude Diagnostique

Pourquoi et comment ?

- 1) On cherche à évaluer la performance de notre test
- 2) On dispose d'un effectif et d'un délai correct (au moins 300 patients nécessaires)

Qu'est-ce que
j'évalue ou étudie ?

Objectif Rechercher
des facteurs de
risque

Des facteurs de risque

On s'oriente ici plutôt vers une :

Cohorte Prospective Analytique

Pourquoi et comment ?

- 1) C'est la meilleure des études observationnelles et ...
- 2) On a le temps pour rédiger le protocole et collecter + analyser les données

Quelles réglementations ?



Le Quanti – Le Réglementaire

Recherche sur la personne humaine en vue du développement des connaissances biologiques et médicales					
Recherches sur le médicament	Investigations cliniques (IC)*	Catégorie 1 Recherches interventionnelles	Catégorie 2 Recherches interventionnelles à risques et contraintes minimales	Catégorie 3 Recherches non interventionnelles (RNI)	Autres Recherches
Règlement EU 536/2014	Règlement EU 2017/745 pour les DM Règlement EU 2017/746 pour les DMDIV	Code de la Santé Publique (CSP) (Loi Jardé, partie réglementaire)			CSP (Partie réglementaire)
Essais cliniques sur des médicaments	Investigations cliniques sur les dispositifs médicaux, et les dispositifs de diagnostic in vitro	Recherches hors produits de santé Qui comporte une intervention non justifiée par sa prise en charge habituelle	Recherches à risque minime* *définies dans une liste fixée par arrêté	Recherches observationnelles* *définies dans une liste fixée par arrêté	Recherches rétrospectives Enquête de satisfaction Evaluations de pratiques de soins Recherches en sciences humaines et sociales. ...
Enregistrement (N° EuCT)	Enregistrement (N° EUDAMED)		Enregistrement (N° ID-RCB)		
Autorisation ANSM / EMA	Autorisation ANSM selon le DM	Autorisation ANSM	Information ANSM		
Avis du CPP					Comité éthique recommandé (GNEDS ou autre)
CNIL : Engagement de conformité MR001 Sinon autorisation CNIL Cf 0062-DI-100 PROM-REGL Démarches réglementaires protection des données				CNIL : Engagement de conformité MR002/MR003, sinon autorisation CNIL Cf 0062-DI-100	CNIL : MR004/MR005/MR006 si éligible sinon GESREES + CNIL Cf 0062-DI-100
Assurance (sauf pour certaines IC selon le DM utilisé)					
Consentement exprès écrit <i>Dérogation au consentement exprès en situation d'urgence sur autorisation du CPP</i>			Consentement exprès (écrit/oral) <i>Écrit : Recherches entrant dans le champ de la loi Bioéthique Dérogation au consentement exprès en situation d'urgence</i>	Recueil de la non-opposition (écrite ou orale) <i>Écrit : Recherches entrant dans le champ de la loi Bioéthique</i>	Droit d'opposition Non opposition écrite ou orale
Enregistrement clinical trial avant 1 ^{ère} inclusion					Pas d'enregistrement clinical trial

Figure 2. Résumé des différentes institutions réglementaires impliquées dans la RC en France d'après l'Inserm⁸

Soumission et Acceptation d'Article

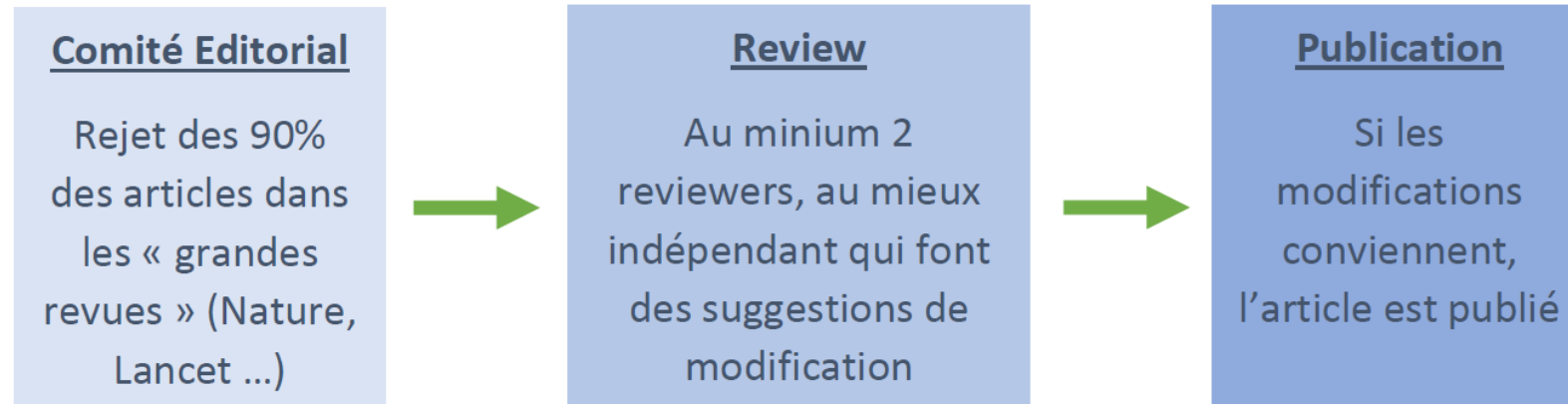


Figure 3. Schéma résumant le processus de publication d'un article de la soumission à la publication

Impact Factor (IF)

nb de citations moyen (dans cette revue et les autres) des articles publiés par ce journal dans les deux dernières années

SIGAPS

*Somme de pour chaque article
Rg revue (selon IF) * Place dans les auteurs*

Comment Collecter mes Données ?



Le Quanti – Base de donnée

1. **Pas de couleur** sur la police ou le fond de tes cellules
2. **Pas d'explication du codage** de tes variables dans les cellules *du genre « Sexe 1= H et 0 = F »* mais juste « Sexe » puis dans ton Plan de Collection des Données tu expliqueras
3. Nom de variable simple **sans espace** (dans ce cas mets des « _ »), **virgule** ou **caractère spécial**
4. **Pas de virgule** pour les décimales mais des points
5. Pour les **données manquantes = cellule vide**

Le Quanti – Pour chaque données ...

1. La **collecte** de données est-elle facilement **réalisable** ?
2. Ma variable est-elle **quantitative ou qualitative** ?
3. Ma variable (score ou marqueur) est-elle **performante, reproductible et déjà utilisée** dans la littérature ?
4. Est-ce que je risque d'avoir beaucoup de **données manquantes** ?
5. **Quand** cette variable va-t-elle avoir de **l'intérêt** dans mon étude ?

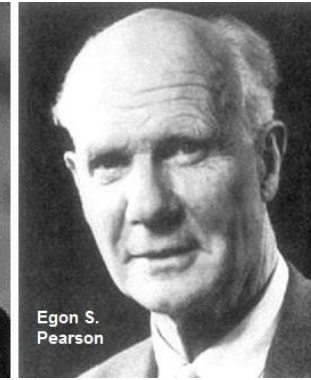
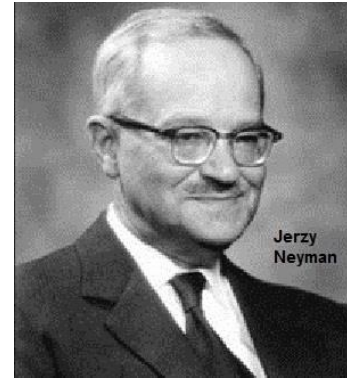
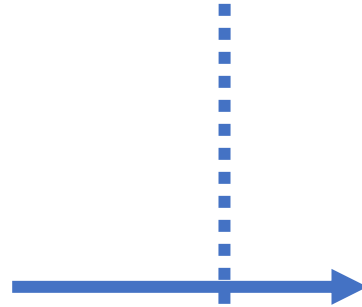
Comment Analyser mes Données ?

Error 404
file not found

Le Quanti – Plan d'Analyse statistique

1. **Objectif** principal et secondaire(s) de mon étude
2. **Critère de jugement** principal et secondaire(s) associé(s) de mon étude
3. **Hypothèse statistique** impliquée dans le critère de jugement principal et **NSN**
4. **Définition des nouvelles variables** créées à partir de celles du Plan de Collection des Données
5. **Figures et Tableaux** prévus
6. **Tests statistiques** utilisés pour chacune des descriptions de la partie
7. **Modèle utilisé** pour les analyses ajustées ou multivariées
8. Gestion des **données manquantes**

Le Quanti – L'Hypothèse (1)



H_0 (hypothèse nulle) : Médicament n'a pas d'effet $\Rightarrow \mu = 1,25$ (même avec injection)
 H_1 (hypothèse alternative) : Médicament a un effet $\rightarrow \mu \neq 1,25$ (avec injection)

On suppose que H_0 est vraie

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{100}} = \frac{5}{10} = 0,5$
 $\mu_{\bar{x}} = \mu = 1,25$

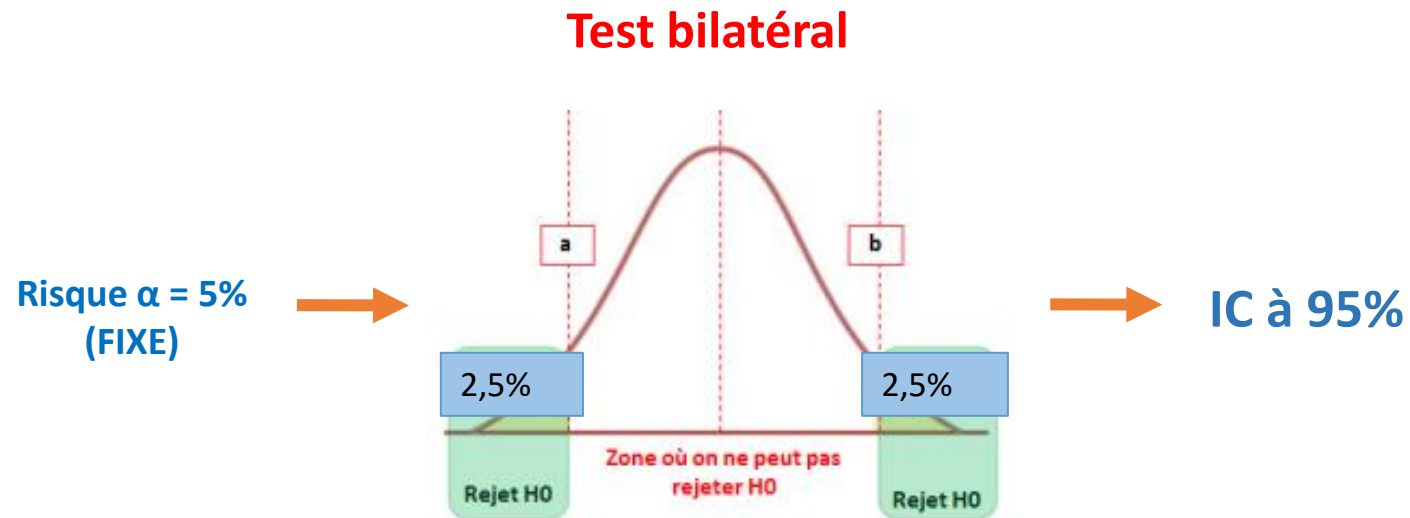
A normal distribution curve with a dashed vertical line at the mean $\mu_{\bar{x}} = \mu = 1,25$.

Le Quanti – L'Hypothèse (2)

EN RESUME : Les risques et tests d'hypothèse

Pour H_0 = absence de différence et H_1 : Traitement A > Traitement B on peut dire que :

- α = risque de conclure à tort d'une supériorité
- β = risque de conclure à tort de l'absence de supériorité
- $1 - \beta = P$ = Puissance statistique (min à 80%)
- $1 - \alpha$ = risque de conclure à raison de la supériorité



Le Quanti – P value ou IC ?

EN RESUME : p value *versus* IC (⚠ avis personnel)

2 situations sont possibles en RC :

- L'**observationnel** (recherche de facteurs de risque, protecteurs ou pronostiques) où le principal est la **généralisation** des résultats ⇒ INTERVALLE DE CONFIANCE
- L'**interventionnel** où le principal est de **ne pas se tromper** à tort sur l'efficacité d'une intervention ⇒ p VALUE

Le Quanti – Le NSN (1)

Le NSN dépend de plusieurs facteurs qui se retrouvent de manière plutôt logique : le risque α , le risque β et la différence que l'on souhaite mettre en évidence (appelée Δ).

Il existe deux formules principales pour le calculer selon que la différence est une différence d'effectif (1 = critère qualitatif) ou de moyenne (2 = critère quantitatif). A noter que le 10 est un arrondi pour $(z_{1-\alpha} - z_{1-\beta})^2$ pris à $\alpha = 5\%$ et puissance $(1-\beta) = 90\%$ (d'autres valeurs sont sur la **Figure 5**). A noter également l'existence de deux groupes A et B.

- 1) $10 \times 2 \times \sigma^2 / \Delta^2$ avec σ = différence attendue (littérature, différence connue) et Δ = différence attendue des auteurs
- 2) $10 \times 2 \times (P_A(1 - P_A) + P_B(1 - P_B)) / (P_A - P_B)^2$ avec P_A = patients présentant le critère 1a dans le groupe A et P_B = patients présentant le critère 1a dans le groupe B.

Le Quanti – Le NSN (2)

TABLE 2

VALEUR DE

$$H = 2 (z_{1-\alpha} + z_{1-\beta})^2 \text{ en fonction de } \alpha \text{ et } \beta$$

$\beta \backslash \alpha$	0,01	0,025	0,05	0,10
0,05	31,5	26,0	21,6	17,1
0,10	26,0	21,0	17,1	13,1
0,20	20,1	15,7	12,4	9,0
0,50	10,8	7,7	5,4	3,3

Figure 5. Différentes valeurs de $2 \times (z_{1-\alpha} + z_{1-\beta})^2$ selon A. Laplanche et al¹⁷

Le Quanti – Décrire les variables qualitatives

EN RESUME : Les tests des variables qualitatives

Grands effectifs \Rightarrow Test du χ^2 (H_0 : l'effectif attendu est le même que l'observé)

Petits effectifs \Rightarrow Test exact de Fisher (H_0 : les variables sont indépendantes)

Données appariées \Rightarrow Test de Mc Nemar (H_0 : les paires discordantes sont comparables dans les groupes)

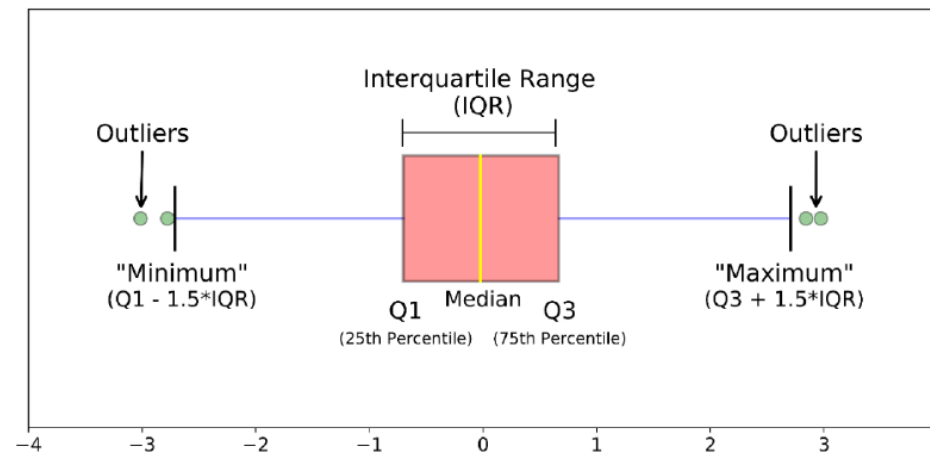
Le Quanti – Décrire les variables quantitatives (1)

- Trait jaune = Médiane

- $Q1 = 1^{\text{er}}$ quartile, $Q3 = 3^{\text{ème}}$ quartile et $Q1-Q3 =$ intervalle interquartile (= IQR) qui contient 50% des valeurs

- Entre $Q1$ et Min et $Q3$ et Max on a de chaque côté ENVIRON les 50% restants (plus précisément 1.5 fois l'IQR + le quartile concerné)

- Les sujets hors des clous définis au-dessus seront considérés comme des « outliers » ou sujets « aberrants »



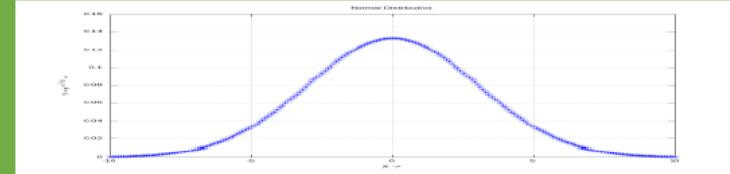
Le Quanti – Décrire les variables quantitatives (2)

Nota Bene : Si on cherche à comparer une variable quantitative entre plus de 2 groupes on utilisera une ANOVA qui est une comparaison de variance pouvant être grossièrement assimilée à une régression linéaire.

EN RESUME : Les tests des variables quantitatives

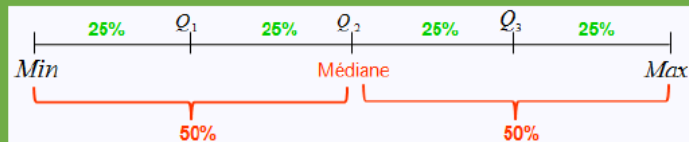
Cas 1 : La variable est continue et normale

- ⇒ On décrit avec moyenne (\pm SD)
- ⇒ On teste avec t.test (\pm apparié)
- ⇒ H0 : « la normalité est comparable »



Cas 2 : La variable est discrète ou non normale

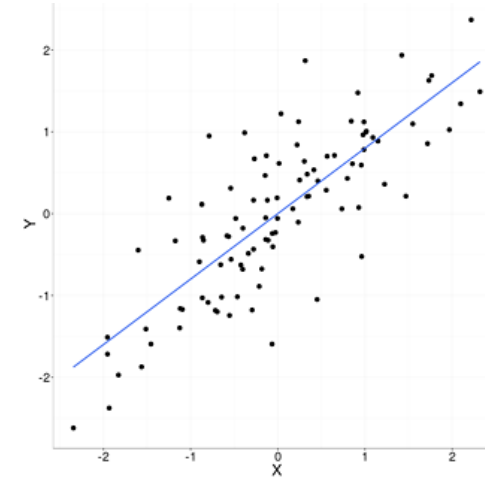
- ⇒ On décrit avec médiane (m) [Q1-Q3]
- ⇒ On teste avec Wilcoxon-Mann-Whitney (\pm apparié)
- ⇒ H0 : « la distribution autour de m est comparable »



Le Quanti – Analyse multivariée

Le Quanti – Régression linéaire

$$Y = a + bx_1 + cx_2 + dx_3 \dots$$



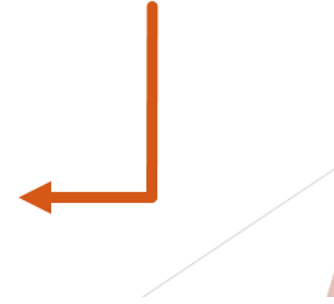
« Si x_1 vaut 1 ou 1 unité, alors
Y augmente de b »

Coefficients

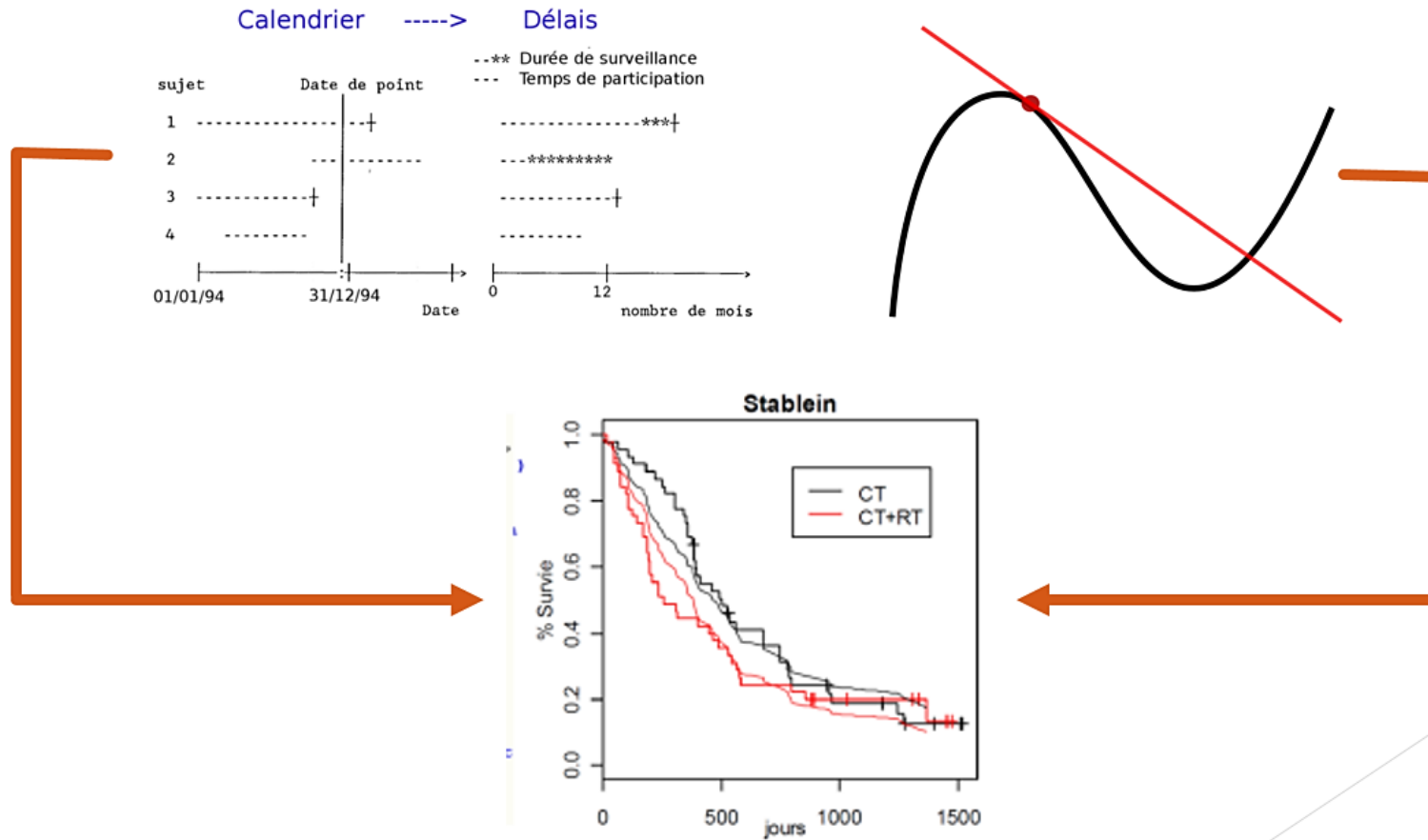
Le Quanti – Régression logistique

$\text{Log}(Y/1-Y) = a + bx_1 + cx_2 + dx_3 \dots$  OR ajustés

« Si x_1 vaut 1 ou 1 unité, alors
Y sera multiplié par $\exp(b)$ »



Le Quanti – La Survie



Le Quanti – Régression de Cox

$Y(t, Z) = h_0(t) * \text{Somme}(\exp(b * Z)) \longrightarrow$ HR ajustés

« Si Z vaut 1 ou 1 unité, alors Y sera multiplié par $\exp(b)$ »

Quels supports et quels soutiens ?



Ressources types d'étude

LEPCAM. « Identifier la question et le type d'étude adaptée », s. d.

<https://lepcam.fr/index.php/les-etapes/type-etude/>.

Saillour-Glénisson, F., et L.R. Salmi. « Évaluation des effets d'une intervention complexe ». *Revue d'Épidémiologie et de Santé Publique* 71, n° 2 (avril 2023): 101377.

<https://doi.org/10.1016/j.respe.2022.06.002>.

Ressources Zotero et codage sur R

Tutoriel Zotero

Topo vidéo Zotero de critique_ton_article

<https://www.instagram.com/p/CpmTV86rSS6/>

Tuto You Tube de Psychonotes

<https://www.youtube.com/watch?v=WKd4xploVYc>

Ressource pour codage sur R (réalisation de tes analyses statistiques)

MOOC du Pr Bruno Falissard (gratos, juste à te créer un compte FUNMOOC)

<https://www.fun-mooc.fr/fr/cours/introduction-a-la-statistique-avec-r/>

Calcul du NSN

<https://biostatgv.sentiweb.fr/?module=etudes/sujets#>

Ressources Méta-Analyses et Revues Systématiques

**Topo vidéo Méta Analyses et Revues Systématiques de
critique ton article**

<https://www.instagram.com/p/CdxUbw3IB7V/>

Ressources du site Cochrane (la référence en MA et RS)

<https://france.cochrane.org/ressources>

<https://training.cochrane.org/handbooks>

Ressources Réglementaires

Guide de la CNIL sur la protection des données

https://www.cnil.fr/sites/cnil/files/2023-04/cnil_guide_securite_des_donnees_personnelles-2023.pdf

Formulaire MR-003

<https://www.cnil.fr/fr/declaration/mr-003-recherches-dans-le-domaine-de-la-sante-sans-recueil-du-consentement>

Formulaire MR-004

<https://www.cnil.fr/fr/declaration/methodologie-de-reference-04-recherches-nimpliquant-pas-la-personne-humaine-etudes-et-evaluations-dans-le-domaine-de-la-sante>

Rôles, Fonctionnement, Composition des CPP (IDF)

<https://www.iledefrance.ars.sante.fr/comites-de-protection-des-personnes-cpp>

ANY QUESTIONS ?



MERCI !!!!

PLUS JAMAIS PEUR DE LA THESE !

En 5 questions clés



Dr Bastien GENET

Docteur en Gériatrie (DES Gériatrie Paris)

Chef de Clinique en Santé Publique (Paris)

Master 1 de Santé Publique, Paris-Saclay

*Master 2 de Méthodologie et Statistiques en Recherche Biomédicale,
Faculté de Paris-Saclay*

Doctorant en Epidémiologie Clinique

Version 1 – Janvier 2024

“Plus jamais peur de la thèse en 5 questions clés”, Bastien Genet, Janvier 2024

tipelle.com

<https://fr.tipeee.com/critiquetonarticle>



critique_ton_article



bastien.genet@aphp.fr

MERCI !!!!

Formation quali/MM présentiel ou distanciel aussi !!!!



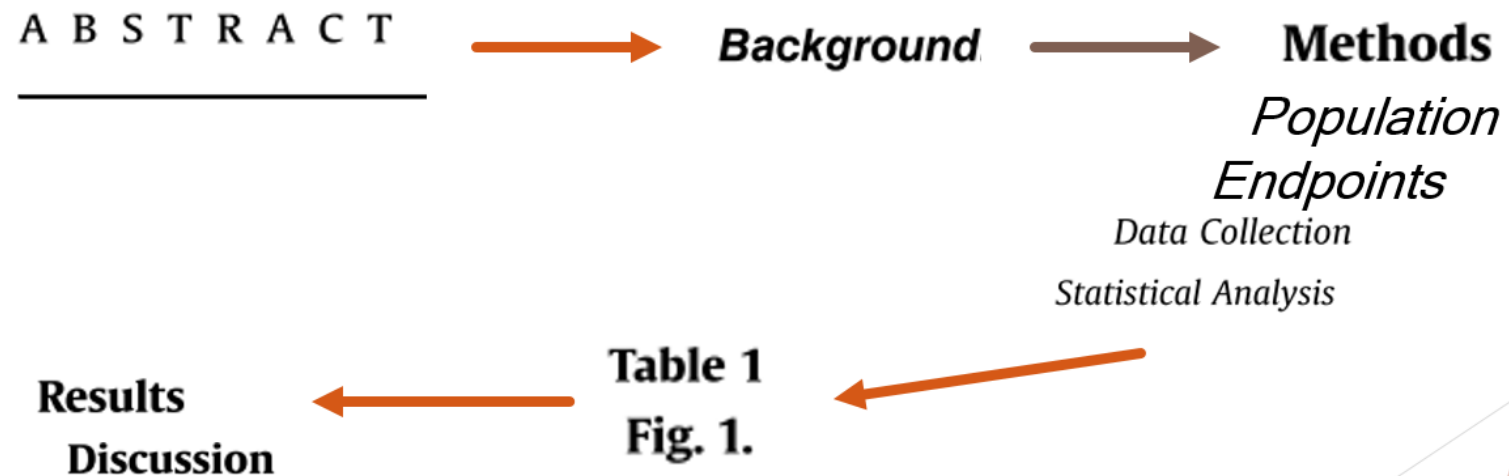
critique_ton_article



bastien.genet@aphp.fr

Bonus - Comment lire un article ?

► Quel ordre ?



Bonus - Grille de lecture d'un article

1) Introduction

Contexte ?

Définition de la Maladie ? Définition de l'Exposition ?

Population cible ?

Objectifs et Hypothèse de recherche ?

2) Méthodes (Biais +++++)

Population : Source ? Echantillon ?

Critères d'inclusion et exclusion ? (Rapport validité interne / externe +++)

Flow chart et PDV ?

Maladie : Méthode de mesure et définition ?

Exposition : Méthode de mesure ? Définition ? Contrôlée / Randomisée ?

3) Stats

Hypothèse statistique ? (Risques α et β , puissance, NSN)

Critères de jugements ? (Unique ? Pertinent ? Adéquation avec l'hypothèse ?)

Quels tests univariés ?

Quels modèles utilisés pour le multivarié ? Quelles variables d'ajustement (biais ++++++) ?

4) Tables & Figures

Biais à Baseline et populations comparables ? (**Table 1**)

Résultats univariés ? (Variables potentiellement de confusion ou interaction ou médiation ?)

Résultats multivariés ? (OR ou HR avec force d'association et IC95%)

5) A vous !

VOTRE avis sur l'article :

- Validité Externe ? (Echantillon bien compatible avec la population cible ?)
- Validité Interne ? (Méthodologie et biais ?)
- Limites ?
- Forces ?
- Réponse aux objectifs et hypothèse de base ?
- EN PRATIQUE utile et pertinent ?

6) Résultats

En cohérence avec les *Tables et Figures* ?

Présentation de TOUS les résultats ? (Certains plutôt cachés, d'autres mis en avant ? +++)

7) Discussion

Validité externe et interne de l'étude avancée par les auteurs ?

Limites et forces avancées par les auteurs ?

Cohérence par rapport à l'hypothèse et les objectifs ?

Bonus - La Causalité

Critères de Causalité d'après B. Hill

Critères internes à l'étude:

1. Existence d'une association statistique entre l'exposition et la maladie
2. Forte intensité de l'association
3. Existence d'une relation de type "dose-effet" entre l'exposition et la maladie
4. Spécificité de relation exposition <-> maladie (+/-)
5. Prise en compte correcte de l'ensemble des facteurs de confusion connus de manière à innocenter d'autres facteurs. Minimisation des biais de sélection et de classement
6. Absence d'ambiguïté temporelle

Critères externes à l'étude (bibliographie)

1. Concordance entre les résultats d'études (**Consistency**)
2. Plausibilité biologique
3. Concordance avec les expérimentations menées *in vitro* ou chez l'animal
4. Gradients géographiques parallèles de l'exposition et de la maladie
5. **Diminution de l'incidence de la maladie lorsque l'exposition est supprimée ou réduite**

Bonus - Les différents types d'études

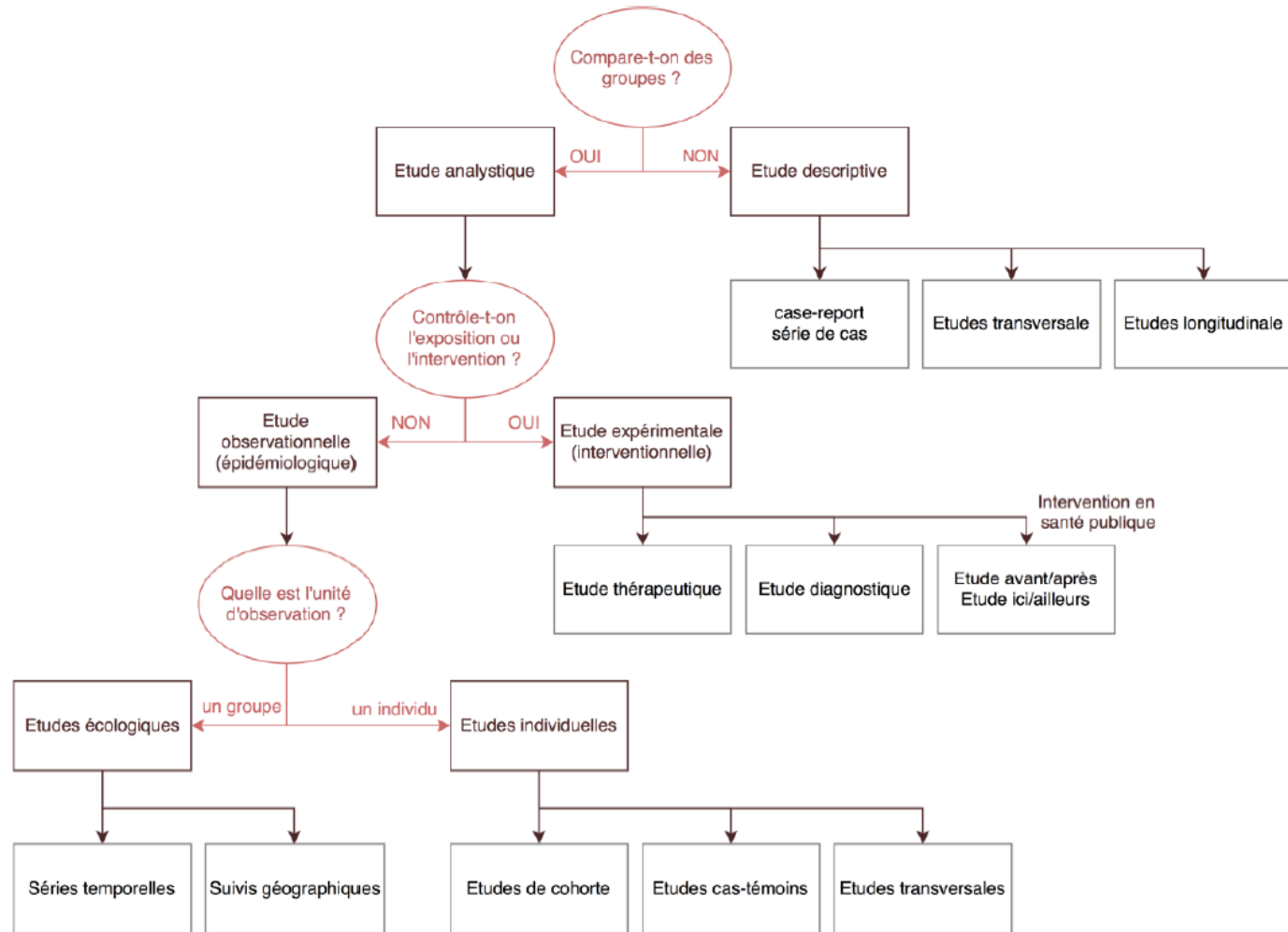


Figure 18. Proposition d'algorithme des études quantitatives par le LEPCAM²⁴

Bonus - Les rang et grades

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

Figure 19. Grade de recommandations et niveau de preuve scientifique des études selon la HAS²⁵

Bonus - Les études diagnostiques

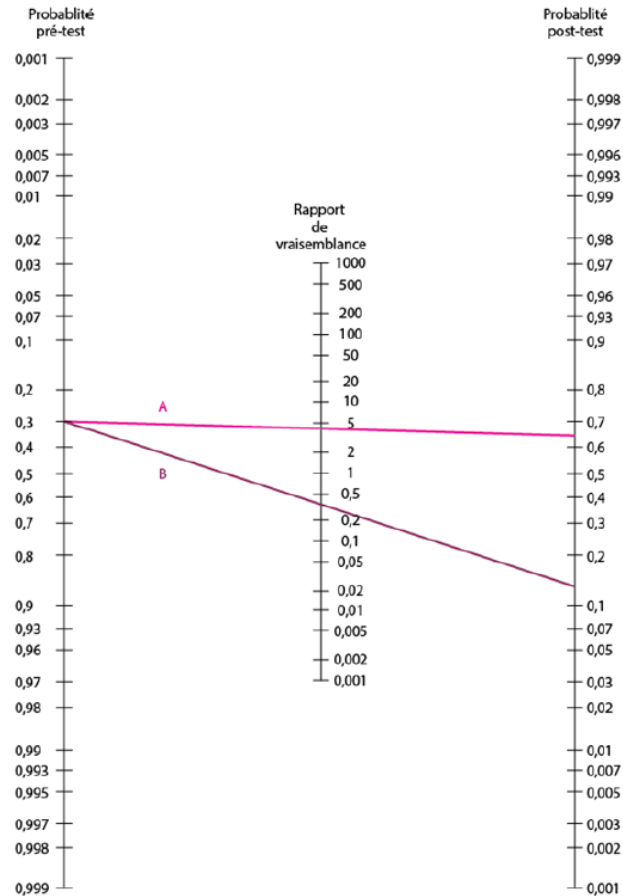


Figure 23. Exemple de Nomogramme de Fagan

EN RESUME : Les paramètres d'un test diagnostique

Capacités intrinsèques (indépendantes de P) :

$$Se = P(T+ \cap M+ / M+)$$

$$Sp = P(T- \cap M- / M-)$$

$$RVP = Se / (1 - Sp)$$

$$RVN = (1 - Se) / Sp$$

Capacités extrinsèques (dépendantes de P) :

$$VPP = P(T+ \cap M+ / T+)$$

$$VPN = P(T- \cap M- / T-)$$

Nomogramme de Fagan :

P pré-test \Leftrightarrow Capacités

intrinsèques \Leftrightarrow P post-test

Bonus - Méta-analyse et Forest Plot

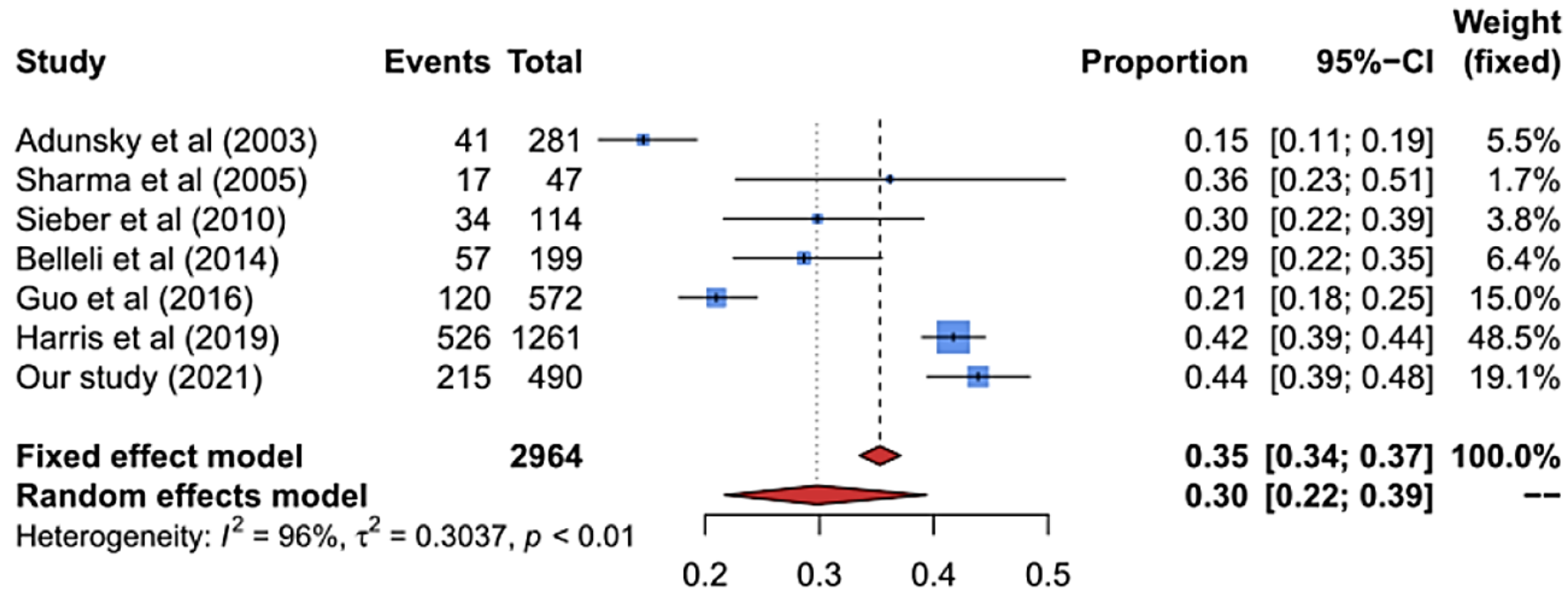


Figure 25. Exemple d'une méta-analyse faite lors d'un travail de mémoire sur confusion post-opératoire dans un contexte de fracture de hanche chez la personne âgée