

Du protocole de recherche à la publication scientifique: bonheurs et défis du statisticien en recherche biomédicale.

1/54

Conférence pour les étudiants du 1^{er} cycle - DMS

Université de Montréal

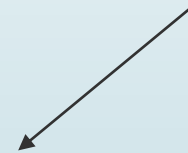
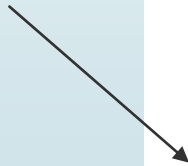
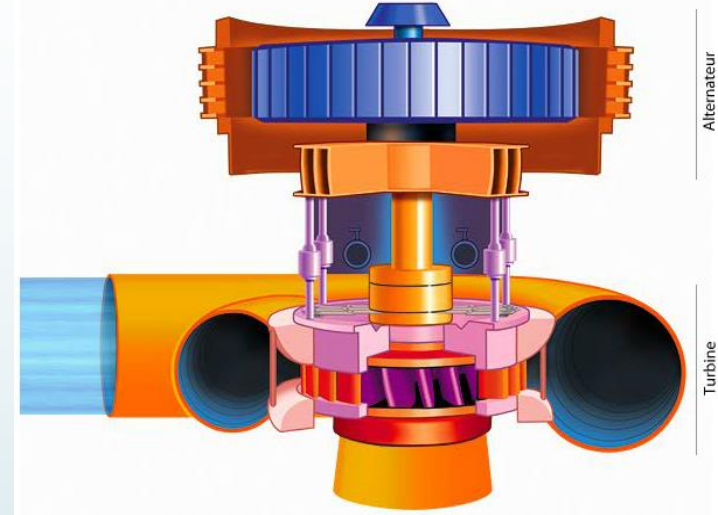
Paule-M. Bodson-Clermont

24 novembre 2024

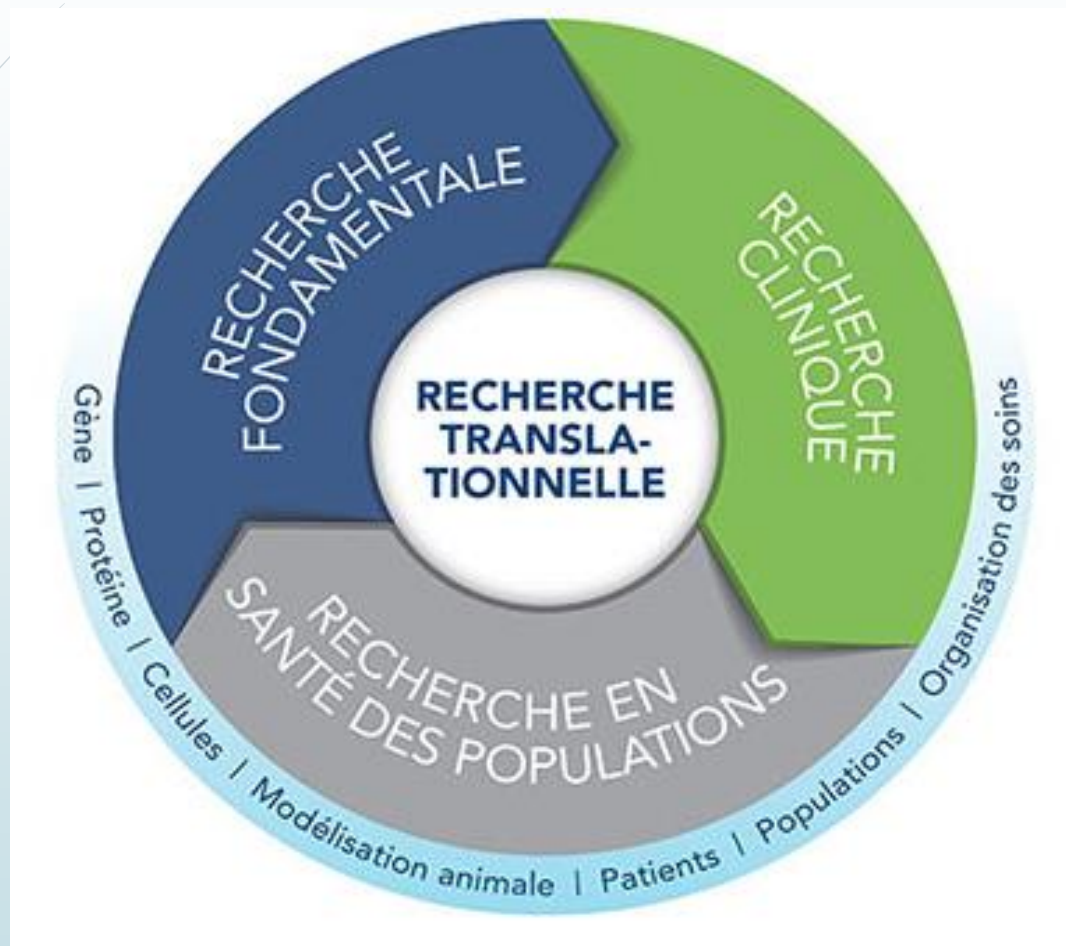
Sommaire

- Présentation
- Cycle d'une étude scientifique
- Projets recherche clinique
- Projets futurs à CITADEL
- Questions

Présentation



Centre de recherche du CHUM

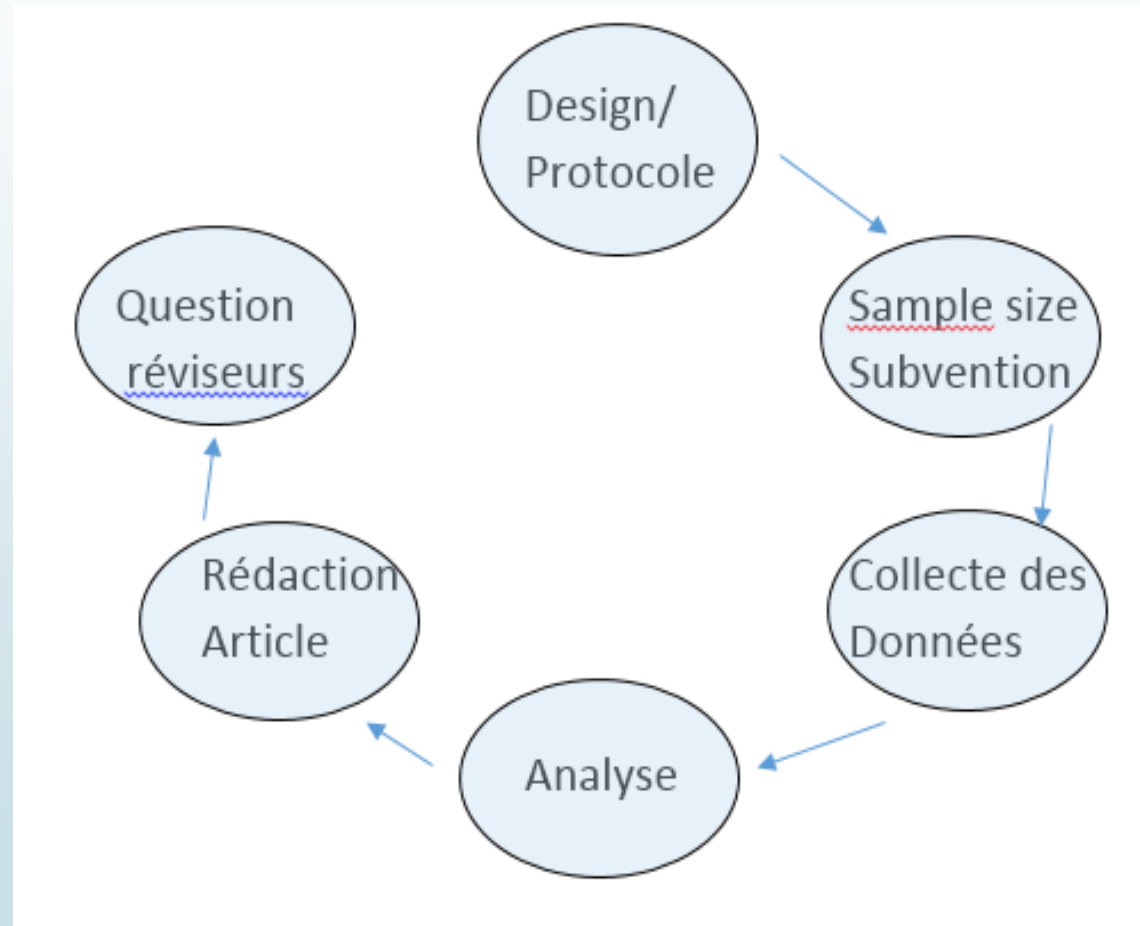


- Plus de 400 chercheurs
- Plus de 900 publications par année
- Axes de recherche:
 - Cancer
 - Cardiométabolique,
 - Imagerie et ingénierie
 - Neurosciences
 - etc.

Un biostatisticien à CITADEL, ça fait quoi?

- 1) Projets courts: consultation statistique à différents stades du cycle d'une étude, dont les demandes de subvention
- 2) Projets longs – recherche clinique: suivre une étude du début à la fin (avec gestion de données ou non)
- 3) Projets hospitaliers
- 4) Comité scientifique

Cycle d'une étude



Projets « courts »

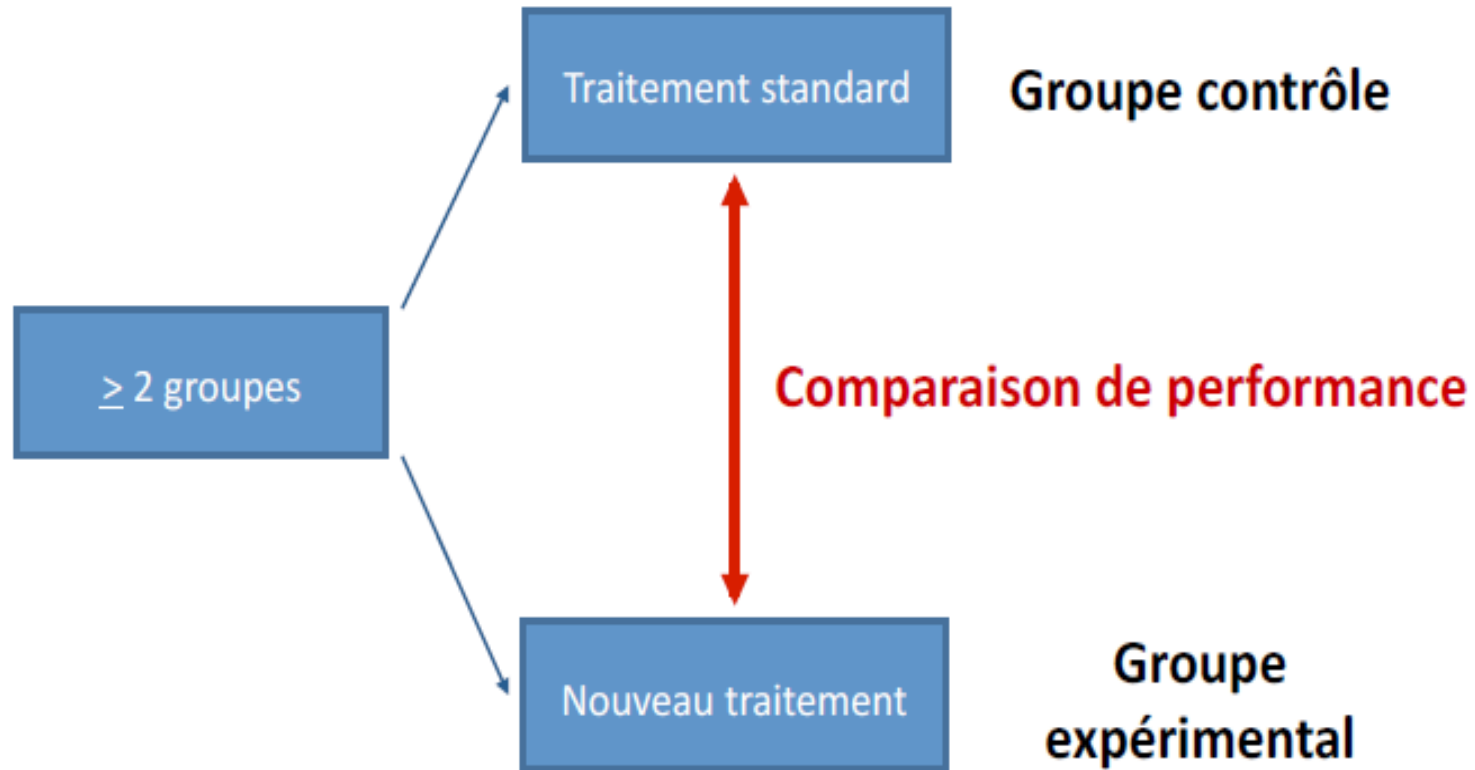
- Le chercheur peut venir consulter un statisticien à chacune de ses étapes
- Idéalement au début!

"To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of."

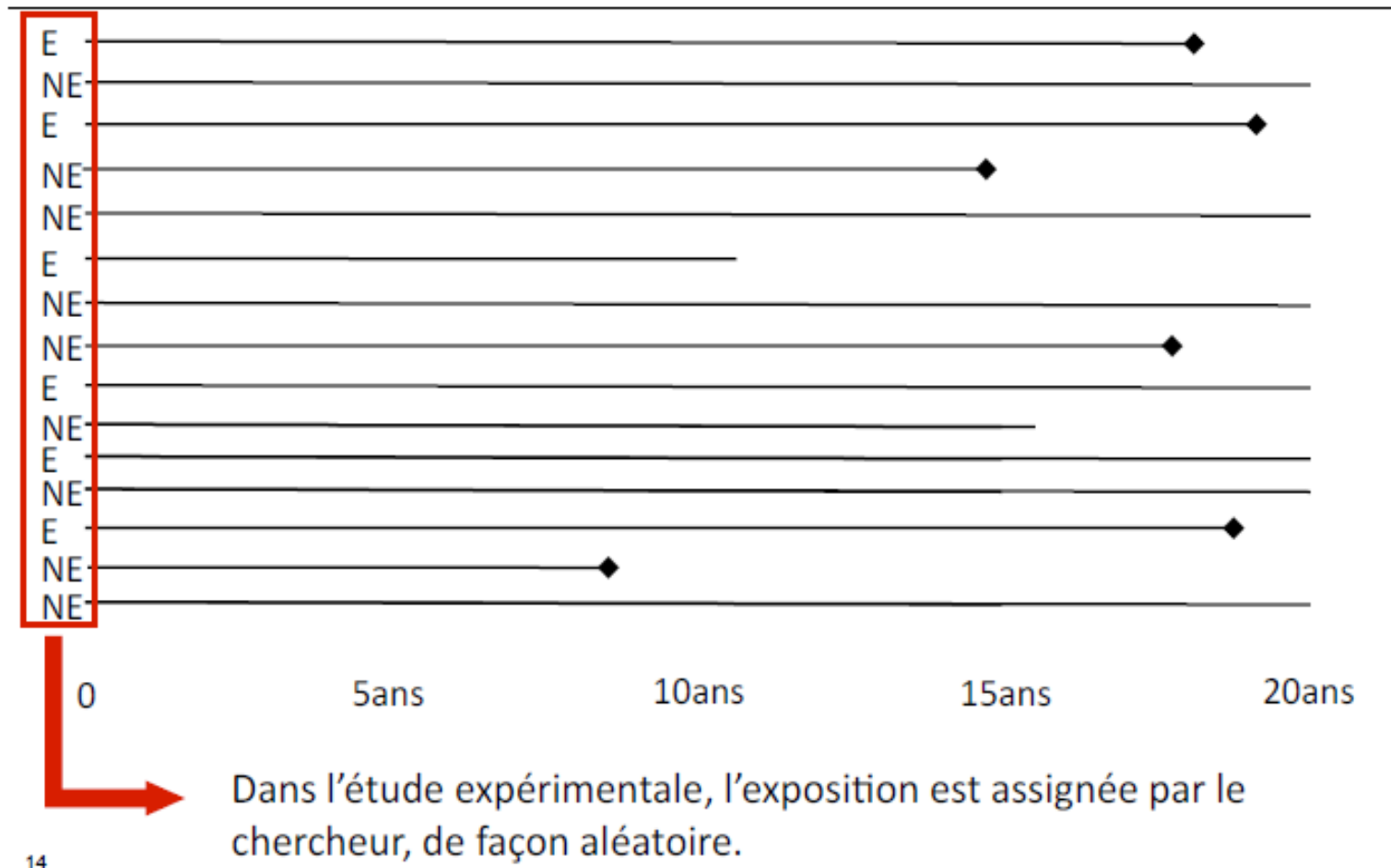
Ronald Fisher

« Appeler un statisticien après que l'expérience soit finie, c'est un peu comme lui demander de faire une autopsie: il ne pourra que constater la cause de l'échec de l'expérience »

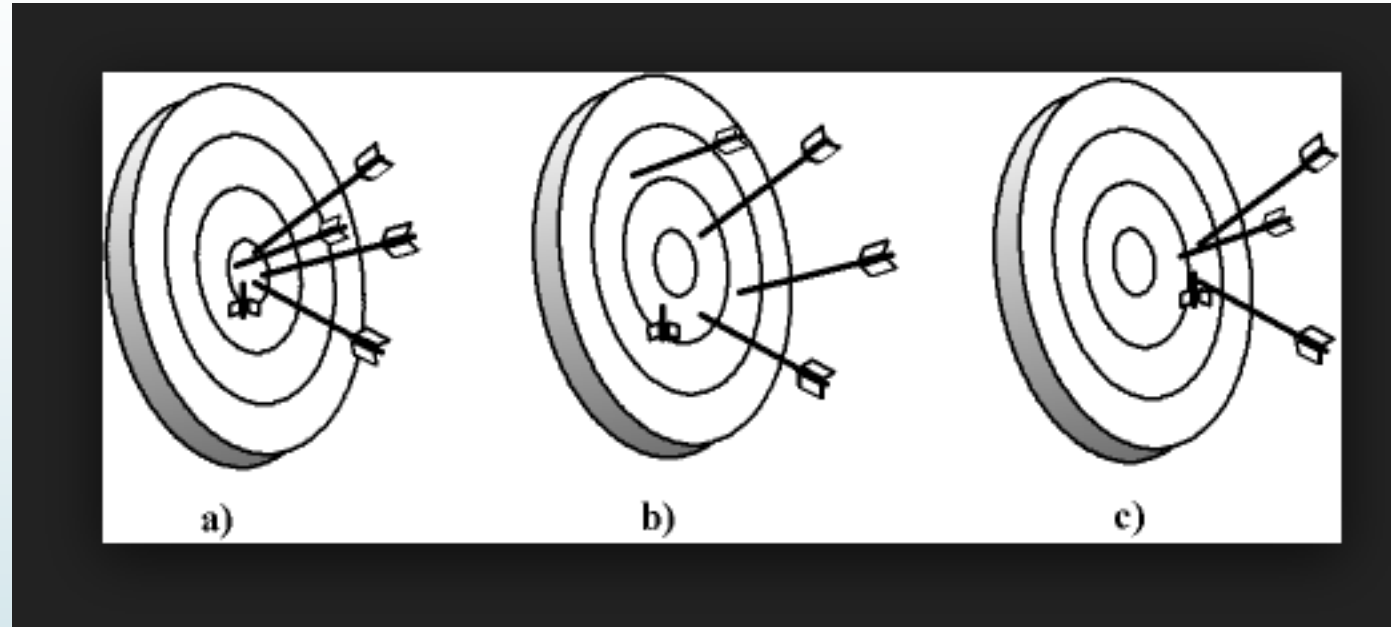
Étude expérimentale



Études de cohorte



Biais et Variabilité



On voudrait avoir une réponse:

- Valide
- Non biaisée
- Précise

Qui pourra éventuellement convaincre la communauté scientifique (jusqu'à changer la pratique clinique)

Qualité d'une bonne consultation

- Écoute
- Respect mutuel: reconnaître les compétences de chacun
- Clarifier les attentes
- Être transparent
- Ne pas craindre de:
 - Répéter
 - Dire qu'on ne comprend pas
 - Dire qu'on ne sait pas

Défis d'une consultation statistique

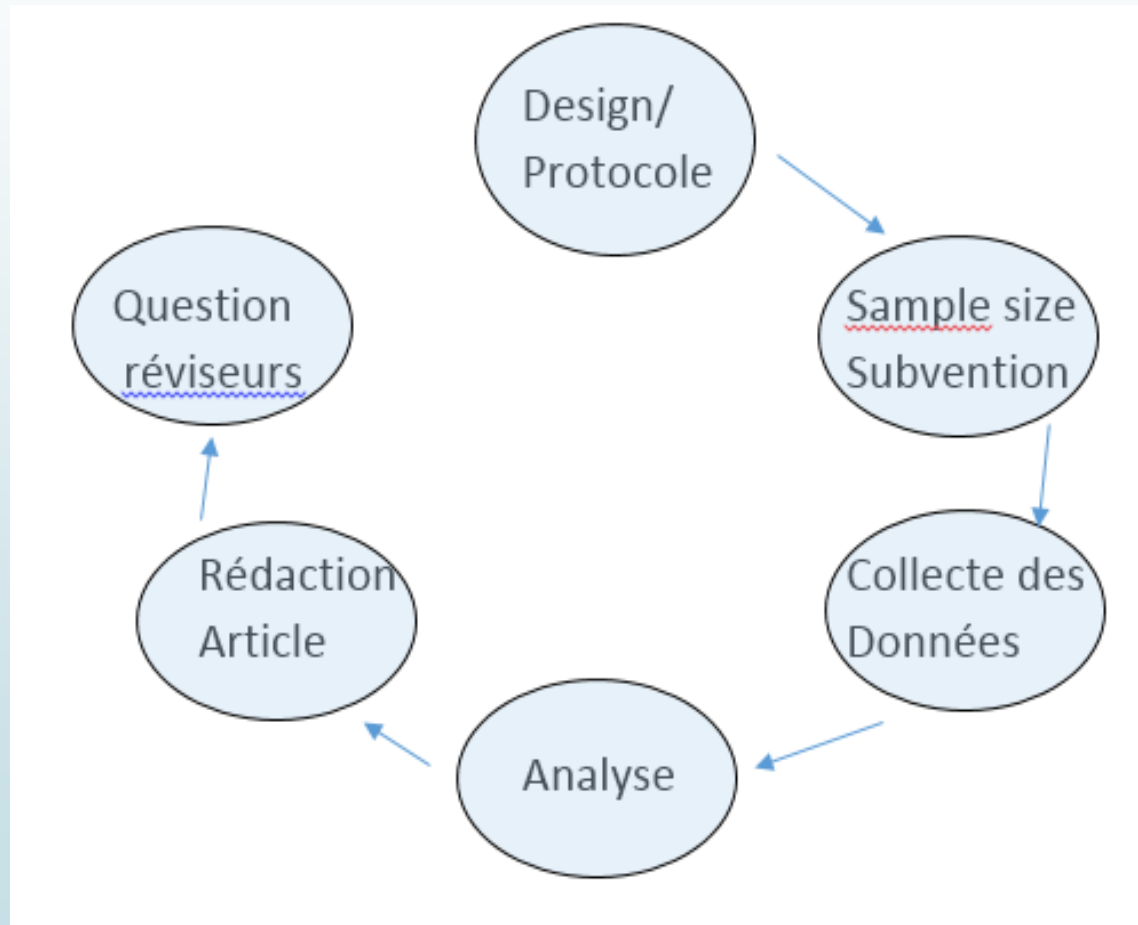
Statistics - A subject which most statisticians find difficult but in which nearly all physicians are expert. Stephen Senn, Statistical Issues in Drug Development, page 4.

- Comprendre ce que le chercheur veut
- Se familiariser avec le vocabulaire
- Parler par personne interposée : résident vs chercheur
- Échéancier et temps de travail

Qualités d'un bon consultant en statistique

- Bonne communication:
 - Écoute active
 - Poser des questions
- Solides connaissances statistiques (design, survie, régression, Modèles mixtes, GEE, diagnostic, bayésien, etc.)
- Rigueur
- Imagination
- Débrouillardise

1) Design et rédaction du protocole



Rédaction du protocole

- Bien définir la question de recherche
- Choisir le design approprié
- Répondre à la bonne question

RÔLE DU STATISTICIEN: poser des questions, soulever des problèmes, proposer des solutions...

- Ex. upgrading

Biopsie ----- chirurgie – mesures de variables (scan)
3-4 mois

Question de recherche PICO

► Bien définir la question de recherche

Critère PICO	Signification	Explication
P	Population ou Patient ou problème médical	Caractéristiques du patient et ou le problème qu'il pose
I	Intervention évaluée (ou exposition)	Nouveau traitement, test diagnostic
C	Comparateur	Placebo, traitement, ou test de référence
O	Outcome en anglais, critère d'évaluation	Taux de mortalité à 1 an, taux de complications à 30 jours.

Intérêt pour la science: est-ce que ça va changer la pratique?

Exemple

- P: patients cancéreux pré-chirurgie
- I: méditation
- C: séance d'exercice légère
- O: anxiété

- Question de recherche: Chez les patients cancéreux pré-chirurgie , est-ce que la méditation comparativement à des exercices légers réduit l'anxiété?
 - D (Domaine de la question: prévention, diagnostique, thérapeutique...)
 - E (Type de devis: expérimental ou non, ECR, observationnel, cas-contrôle, etc.)

Rédaction du protocole

- ▶ Mais encore....
 - Tous les patients cancéreux? Inclusion exclusion
 - Stade 4? Cancer spécifique?
 - Importance des définitions
 - Pendant combien de temps?
 - Réduit l'anxiété à quel moment?
 - Quelle est la mesure utilisée? Adéquate à la population?

- ▶ Question de recherche: Chez les patients cancéreux pré-chirurgie , est-ce que la méditation comparativement à des exercices légers réduit l'anxiété?

Choisir le bon design

- Erreurs fréquentes
 - Étude pilote/Étude de faisabilité
 - Thabane 2010: [A tutorial on pilot studies: the what, why and how](#)
 - Être trop ambitieux (ex. RCT avec intervention numérique gratuite multicentrique)
 - Taux de recrutement/population trop restreinte
 - Non-infériorité vs supériorité
 - Schumi 2011: [Through the looking glass: understanding non-inferiority - NCBI - NIH](#)

Non- inferiorité

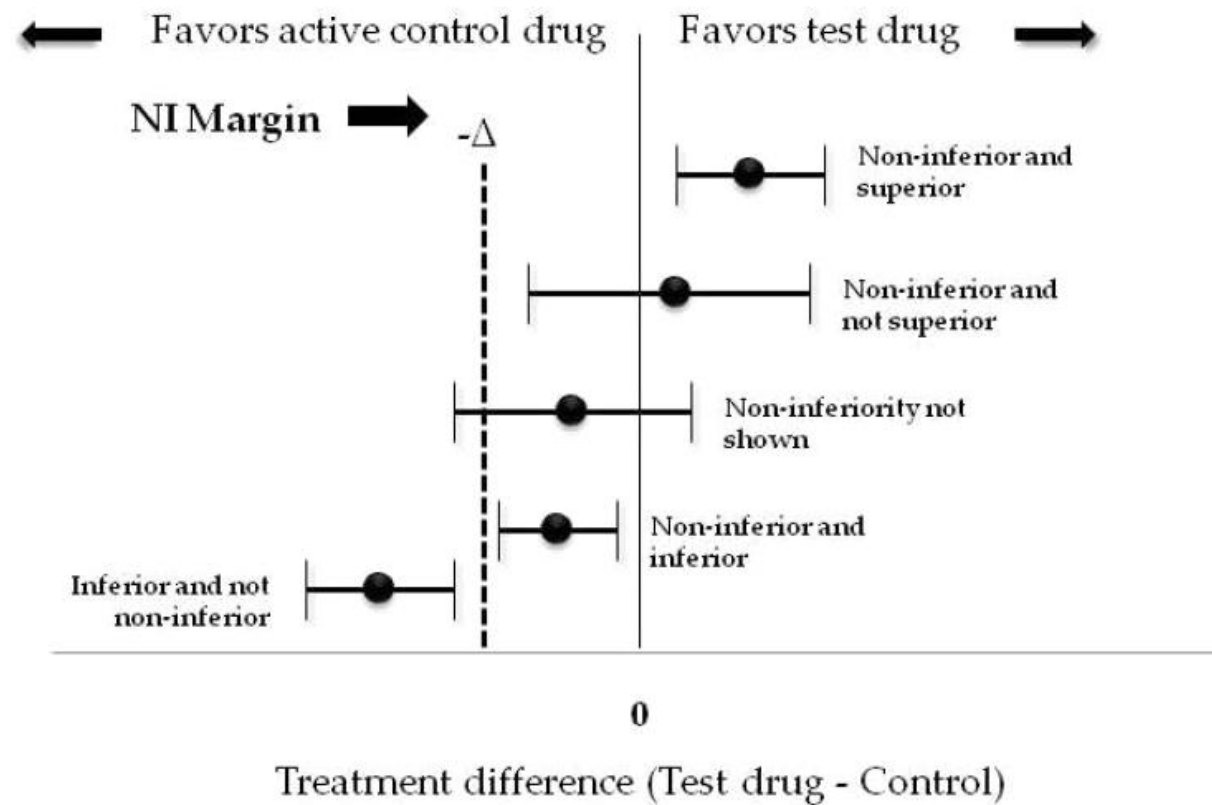
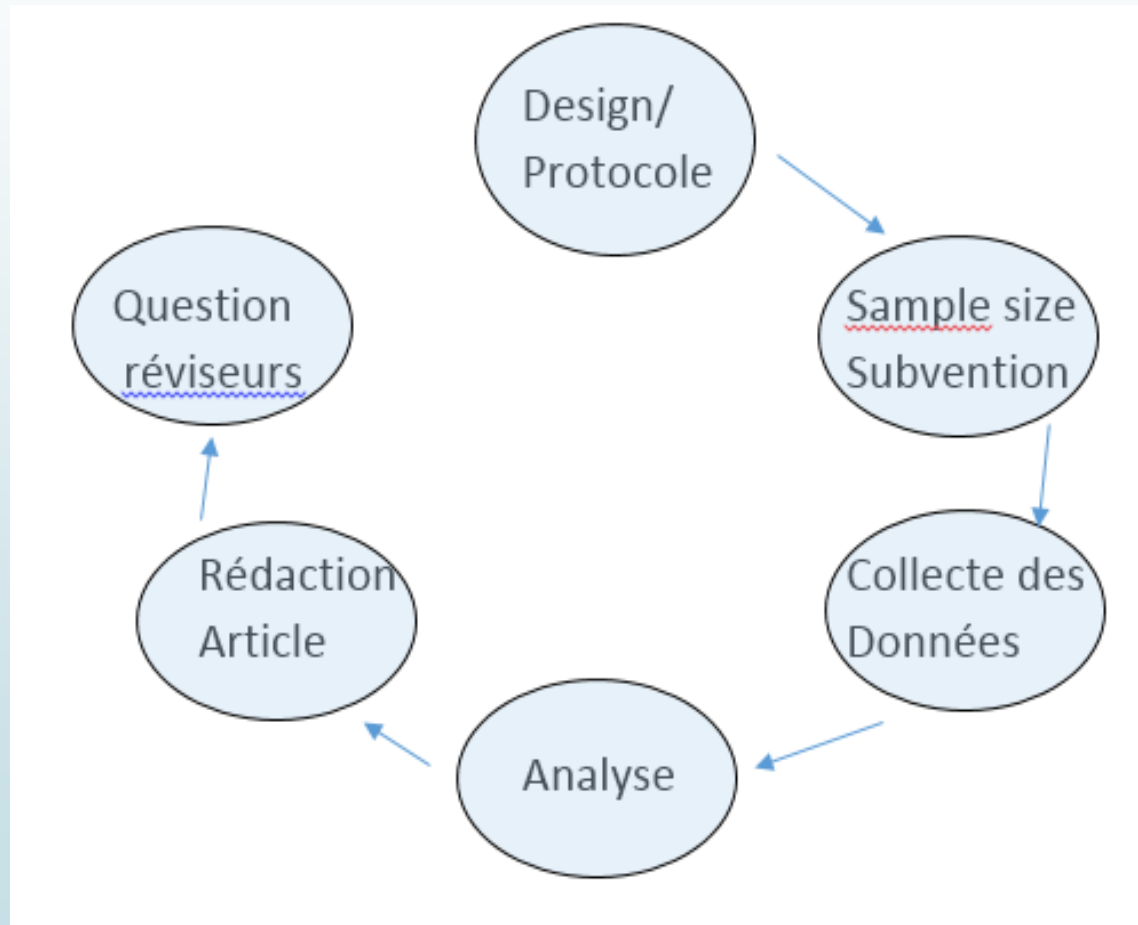


Figure 2 Possible outcomes of a non-inferiority trial.

Choisir le bon design (suite)

- Cohorte rétrospective mal sélectionnée
 - * Temps de suivi différent selon les groupes de comparaison
 - * Ne pas mettre de groupe comparatif (mais vouloir évaluer l'effet)
 - * Inclure seulement les données complètes dans la collecte des données dans le cas d'étude rétrospective
- Ne pas mesurer des variables confondantes potentielles
- Ne pas tenir compte des données hiérarchiques/corrélation entre les données
 - Genders 2012: [Methods for calculating sensitivity and specificity of clustered data - NCBI](#)
- Anticiper les biais: Biais de contamination (design par cluster), biais de recrutement, etc.

2) Calcul de taille d'échantillon



2. Calcul de taille d'échantillon

- ▶ Souvent dans le cadre des demandes de subvention
 - ▶ IRSC
 - ▶ Fondations
 - ▶ Etc.
- ▶ « un calcul très rapide »
 - ▶ Le chercheur pense que ça se fait très rapidement
 - ▶ Questions de recherches encore embryonnaires → design
 - ▶ Nécessité de se procurer les « intrants »

Calcul de la taille d'échantillon



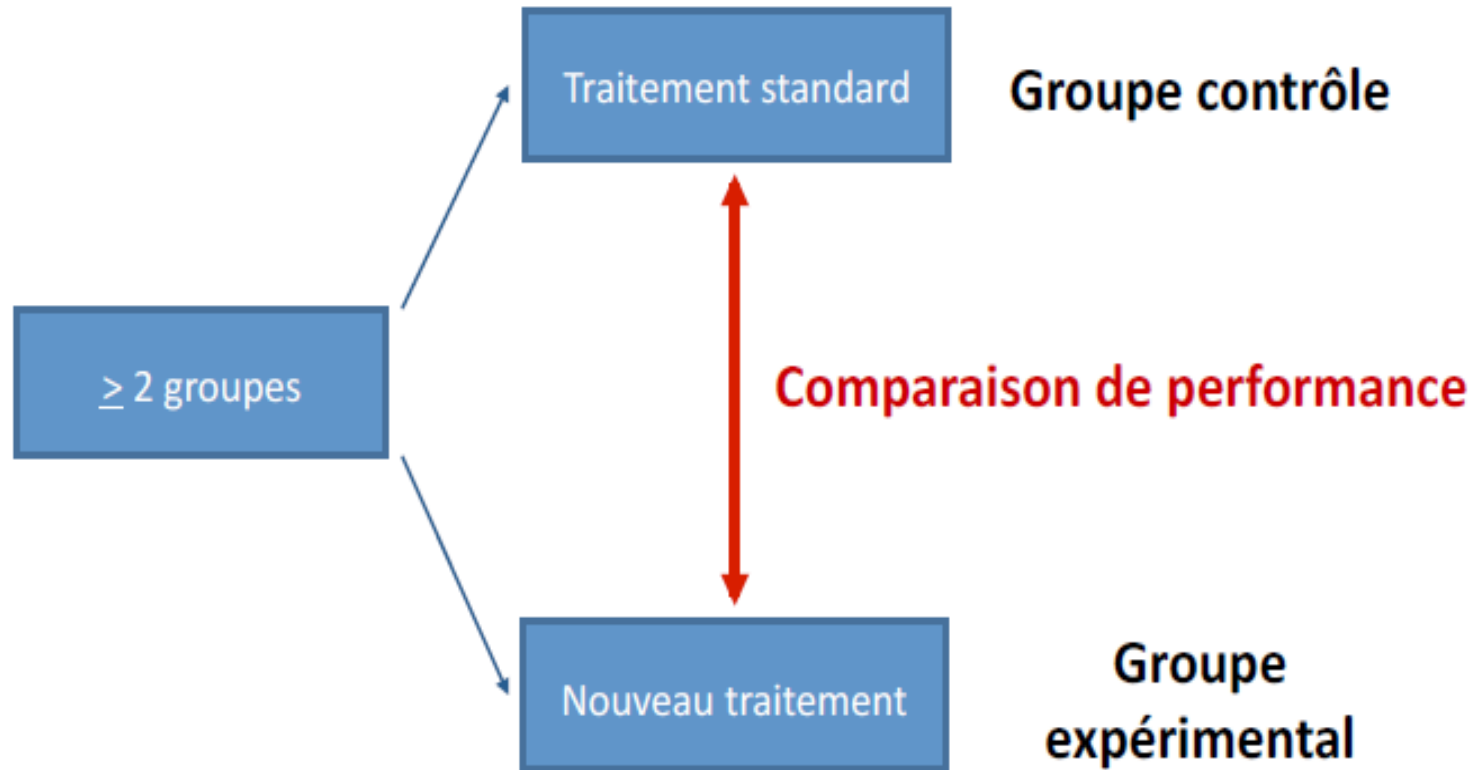
$$n = \frac{Z^2 p(1 - p)}{d^2}$$



Taille d'échantillon: pas une science exacte...

- ▶ Dépend des hypothèses
- ▶ La réalité reflètera-t-elle les hypothèses?
- ▶ Le endpoint (la mesure) est-il bien connu? Dans notre population?
- ▶ Que veut-on montrer?

Étude expérimentale



Calcul de taille d'échantillon

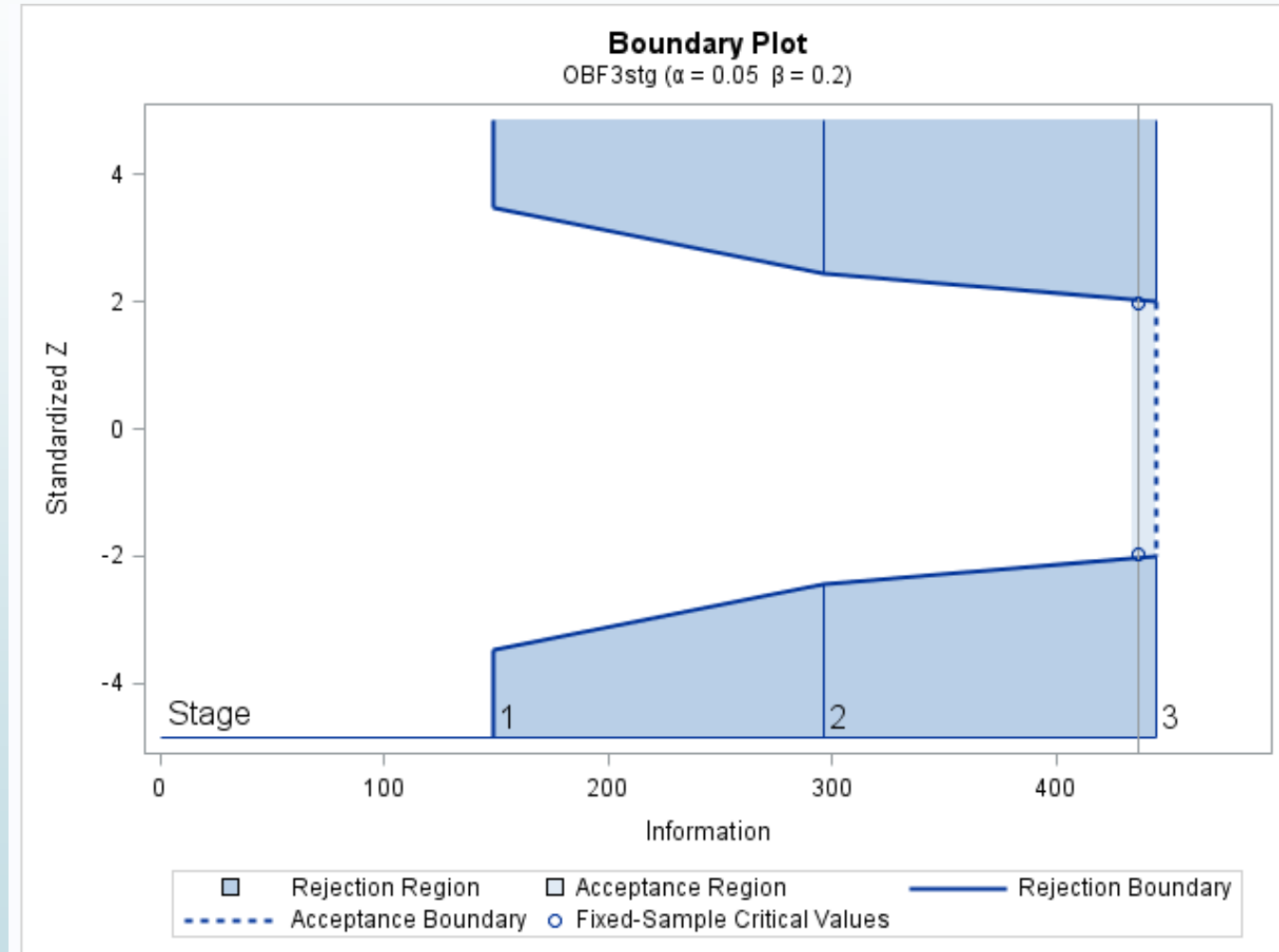
- ▶ Parfois sur des sujets stats qu'on maîtrise peu
- ▶ Importance de la collaboration entre statisticiens
- ▶ Travailler sous pression (demande de subvention)

- ▶ Pourquoi c'est important?
 - Bonne taille: \$\$ vs trouver ce qui va être possible de trouver
 - Éthique: faire des procédures non nécessaires
 - Importance de la différence clinique
 - « Est-ce que ça va changer la pratique? »
 - Différence statistique: ex. avec l'âge vs efficacité médicament



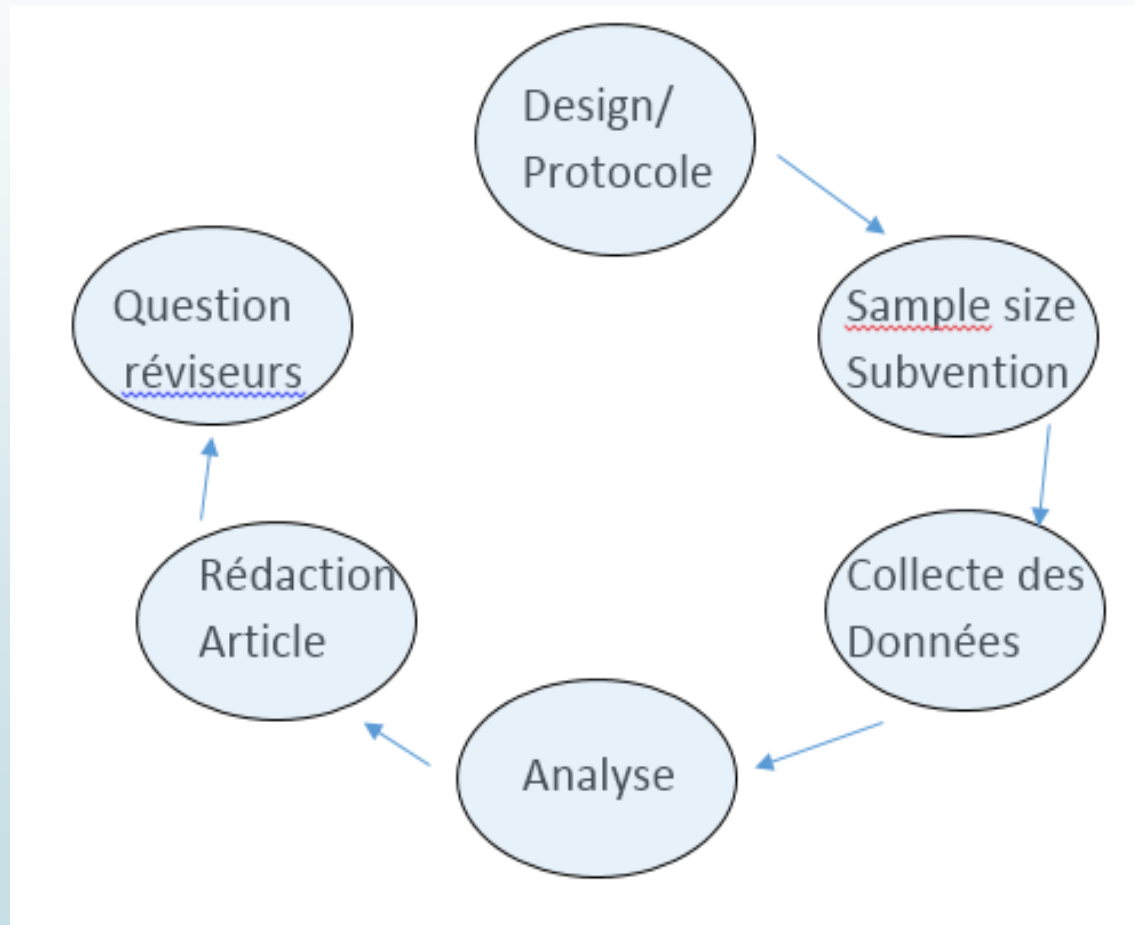
Group sequential design

- Les tests statistiques à chaque étape (look) ne sont pas indépendants



Cycle d'une étude:

3) Collecte des données



3. Collecte des données (et son *evil twin*... le ménage des données)

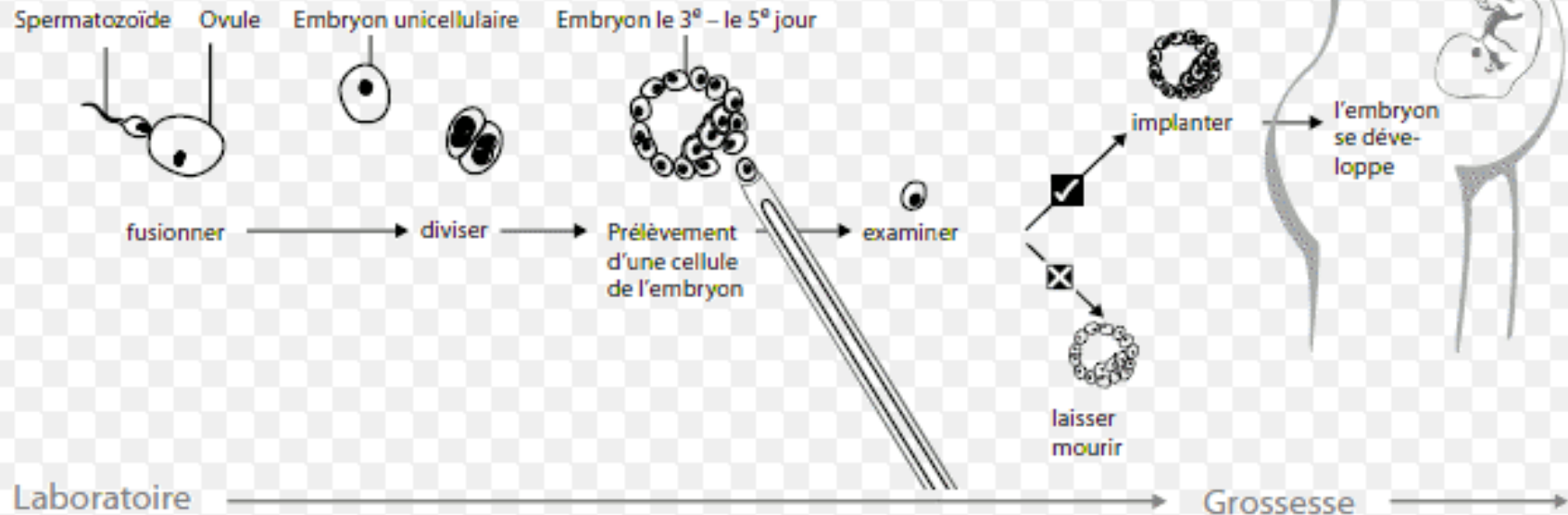
- ▶ Aucune méthode statistique ne peut améliorer les données
- ▶ On le fait: pas par goût mais par nécessité
 - Ne pas sous-estimer le temps que ça va prendre
 - Responsabilité du chercheur vs du statisticien
 - Solution: Gestion des données avec des outils Spécifiques à la recherche (RedCAP, Datafax)



Exemple en fertilité et Excel...

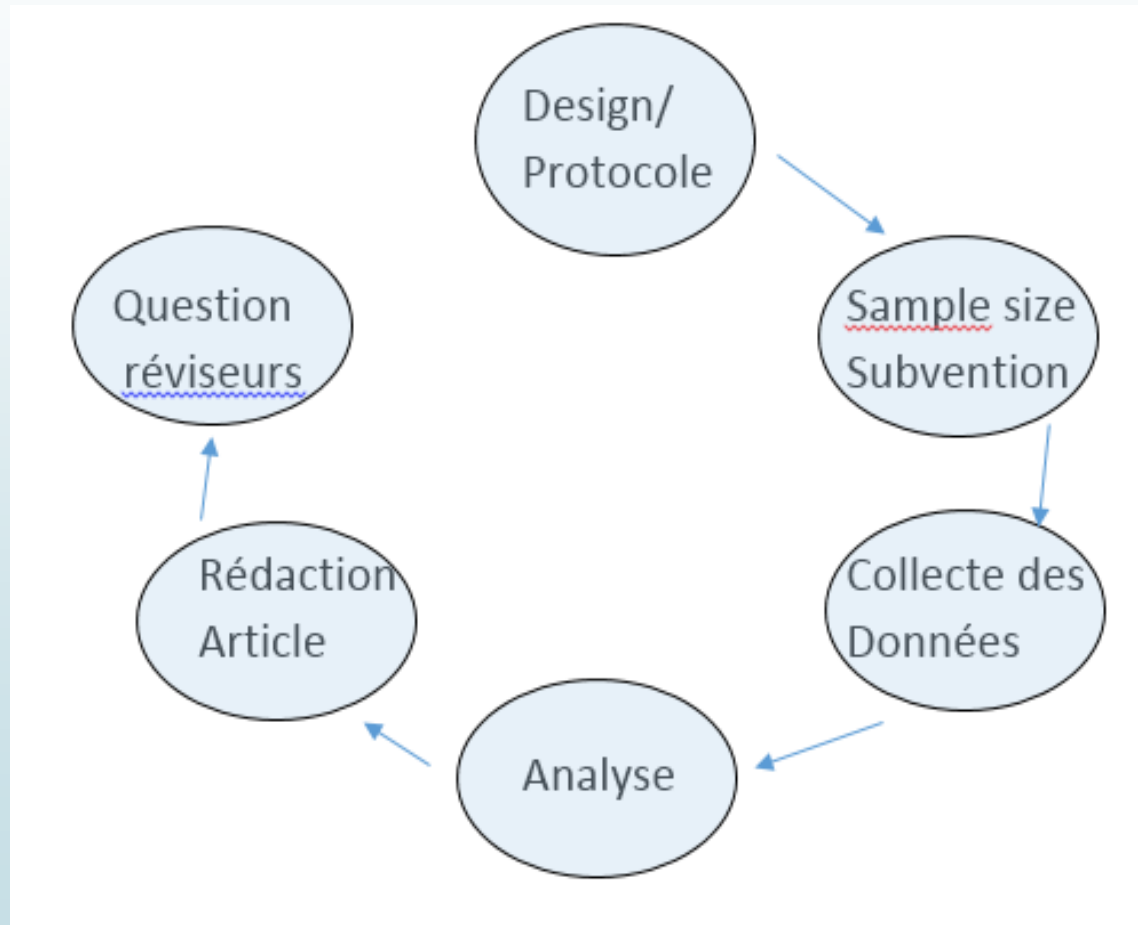
OVUL.PREM	PO.BLANCHE	EGGS	NO.VIABLE	NO.MATURE	X2PN	X2PN.CLEAV	PCT.CLEAVED	TE	PREG	CLINICAL
0	0	1	1	1	0	0	0	0	0	0
0		1	1	2	1	1	100	1	1	1
0	0	1	1	1	1	0	0	0	0	0
0	pas de ponction	1	1	1	1	1	100	1	0	0
0	0	1	1	1	1	1	100	1		1

Déroulement d'un diagnostic préimplantatoire (DPI)



Cycle d'une étude:

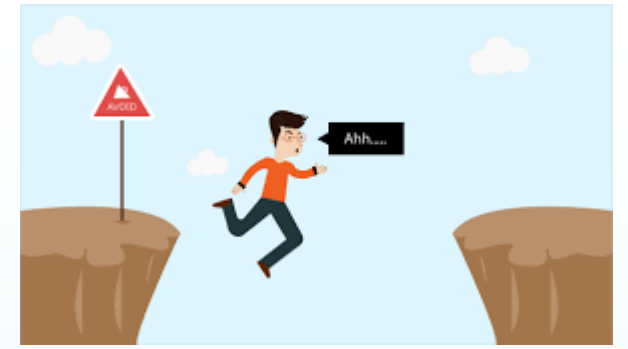
4) Analyse



4. Analyse

- ▶ Normalement, si tout le reste a été pensé, tout va bien!
- ▶ Recherche: nécessité de développer des aptitudes avec des méthodes moins usuelles
- ▶ Souvent, différents défis... qui impacteront sur le temps passé
- ▶ Importance d'avoir d'autres statisticiens à portée de main

Analyses: écueils possibles



- Données manquantes: solution? Imputation?
- Moins de données que voulues (par exemple données de survie avec dépend du nombre d'événements)
- Hypothèses des modèles
- Modèles qui ne convergent pas
- Importance des analyses de sensibilité
- Importance du dialogue avec le chercheur (modèle multivarié + facteurs confondants)
- Penser au temps que ça prend!

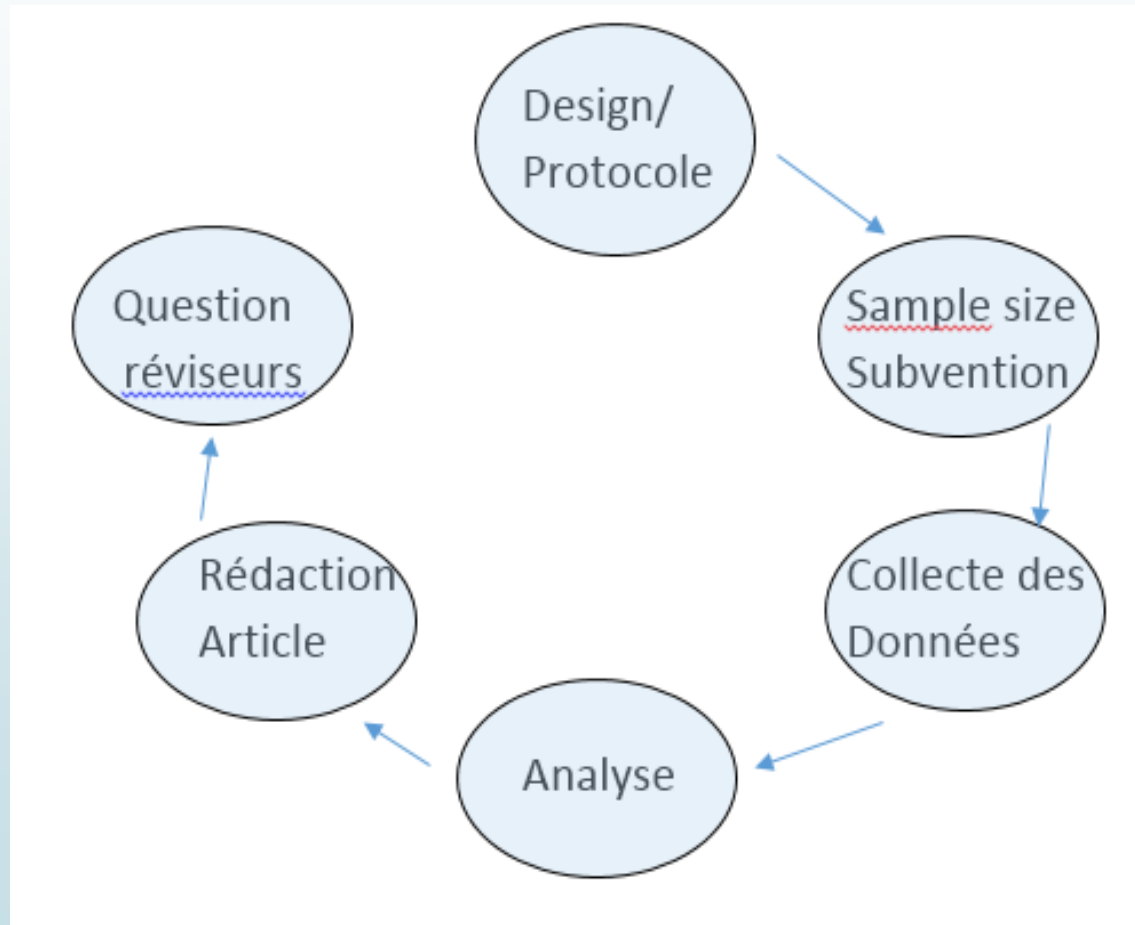
4. Analyse

Exemples de méthodes statistiques utilisées

- Survie
- Régression logistique
- Analyses en composantes principales
- ANOVA à mesurées répétées
- GEE
- Courbes ROC
- Régressions pénalisées (LASSO, Ridge, Elastic-Net)
- CART, Random Forest

Cycle d'une étude:

5) Rédaction de l'article

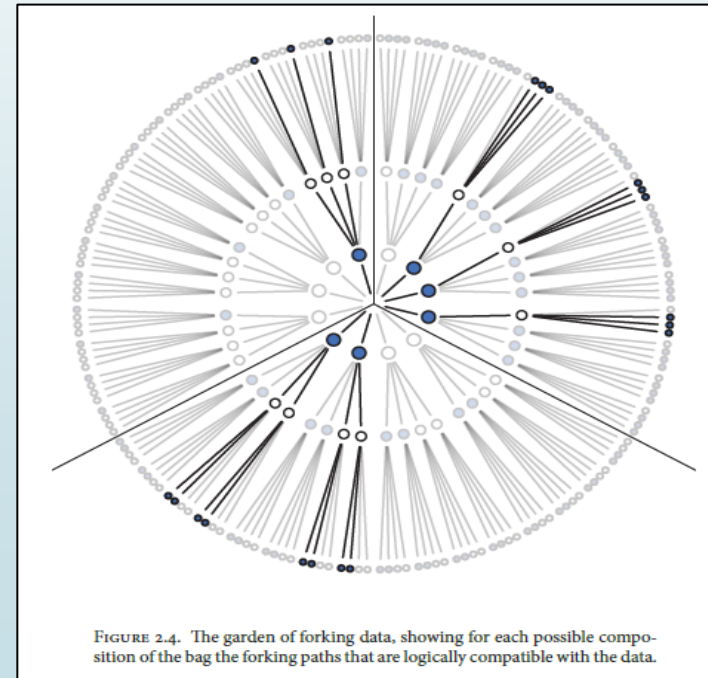


5. Rédaction de l'article

- ▶ Rôle du biostatisticien
 - Écriture du paragraphe de méthodologie
 - Révision de la section résultats
- ▶ Réalignement de l'interprétation
- ▶ Doigté et tact: faire travailler le chercheur, clarifier les responsabilités, etc.
- ▶ Authorship

Interprétation: Fin du règne des valeurs-p ou la crise de la reproductibilité

- L'utilisation des valeurs-p pour déterminer la "significativité statistique" est très controversée depuis plusieurs années
- La reproductibilité est un débat majeur. Est-ce qu'on peut considérer les résultats robustes?



Qu'est-ce qu'une valeur-p?

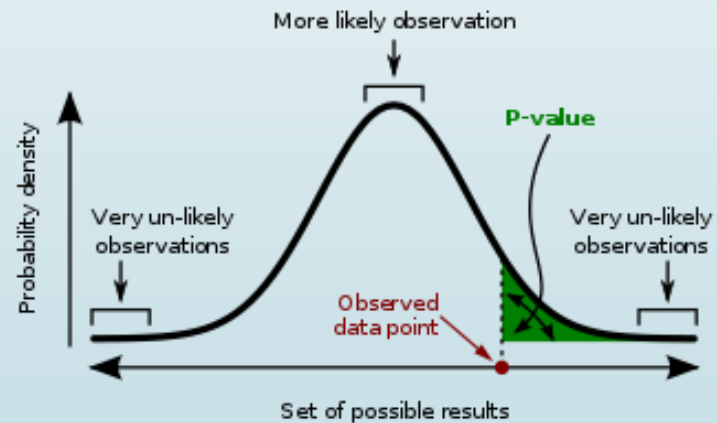
- P (observer une valeur aussi extrême si l'hypothèse nulle est vraie).

Important:

$\Pr(\text{observation} \mid \text{hypothesis}) \neq \Pr(\text{hypothesis} \mid \text{observation})$

The probability of observing a result given that some hypothesis is true is *not equivalent* to the probability that a hypothesis is true given that some result has been observed.

Using the p-value as a "score" is committing an egregious logical error:
the transposed conditional fallacy.



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Historique de la valeur-p

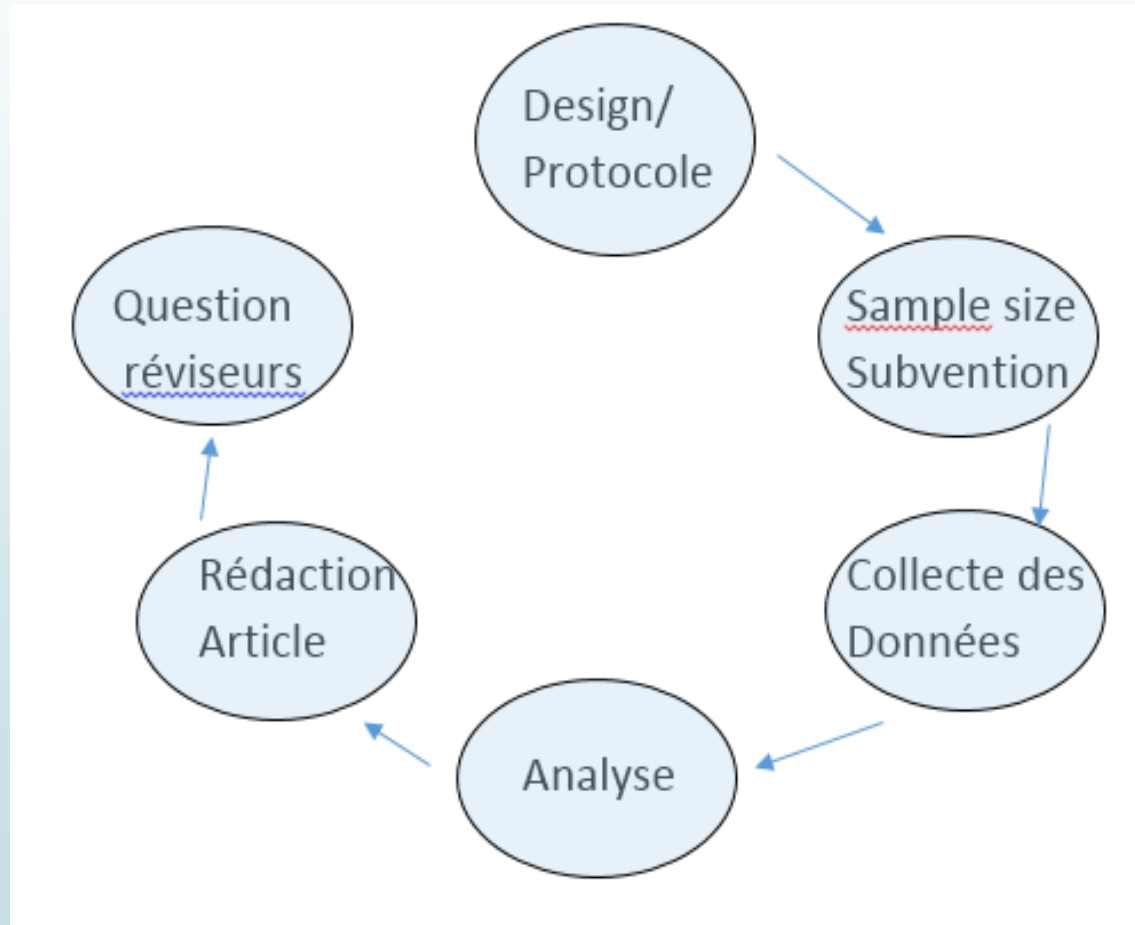
- ▶ Lorsque Fisher a introduit le concept de valeur-p, il n'a jamais souhaité que ce soit utilisé comme un test définitif.
- ▶ Il considérait celle-ci comme un moyen informatif pour évaluer si l'évidence était « significative », dans le sens plus traditionnel du terme, c'est-à-dire qu'elle méritait qu'on s'y penche une deuxième fois.
- ▶ Neyman et Pearson ont de leur côté introduit un concept alternatif incluant la puissance statistique, les faux positifs et faux négatifs, etc. (sans parler de valeur-p)
- ▶ La plupart des chercheurs qui ont écrit par la suite les manuels statistiques ont appliqué un système hybride, utilisant la valeur-p de Fisher et le système de Neyman Pearson.

MESURES POSSIBLES

- ▶ Toujours rapporter les tailles d'effet et les intervalles de confiance
- ▶ Clairement rapporter si l'étude est exploratoire ou confirmatoire
 - ▶ le, rapporter le nombre de tests, ne pas mettre en évidence un outcome non préspecifié
 - ▶ Une fois les études exploratoires réalisées, planifier l'étude confirmatoire et enregistrer le plan de l'étude dans un registre.
 - ▶ Conduire des études de reproductibilité
- ▶ Réaliser les limites de l'interprétation des statistiques conventionnelles et bien mettre en perspective le jugement scientifique: résultats d'autres études, connaissances cliniques, mécanismes biologiques, etc.
- ▶ Essayer de bannir l'utilisation du mot « signficance » et rapporter la valeur-p exacte ($p=0,024$) plutôt qu'une inéquivalence ($p < 0,05$)
- ▶ Utiliser des méthodes bayésiennes qui prennent avantage d'un angle différent (plausabilité d'un outcome, plutôt que fréquence de celui-ci)

Cycle d'une étude:

6) Question des réviseurs



6. Questions des réviseurs

- Difficulté de reprendre un projet à la toute fin
- Exemple cardio
 - Événement: temps avant le décès
 - Modèle de Cox
 - Deux variables dichotomiques + création d'une « troisième variable »
 - Interaction

Projet à long terme: Exemple

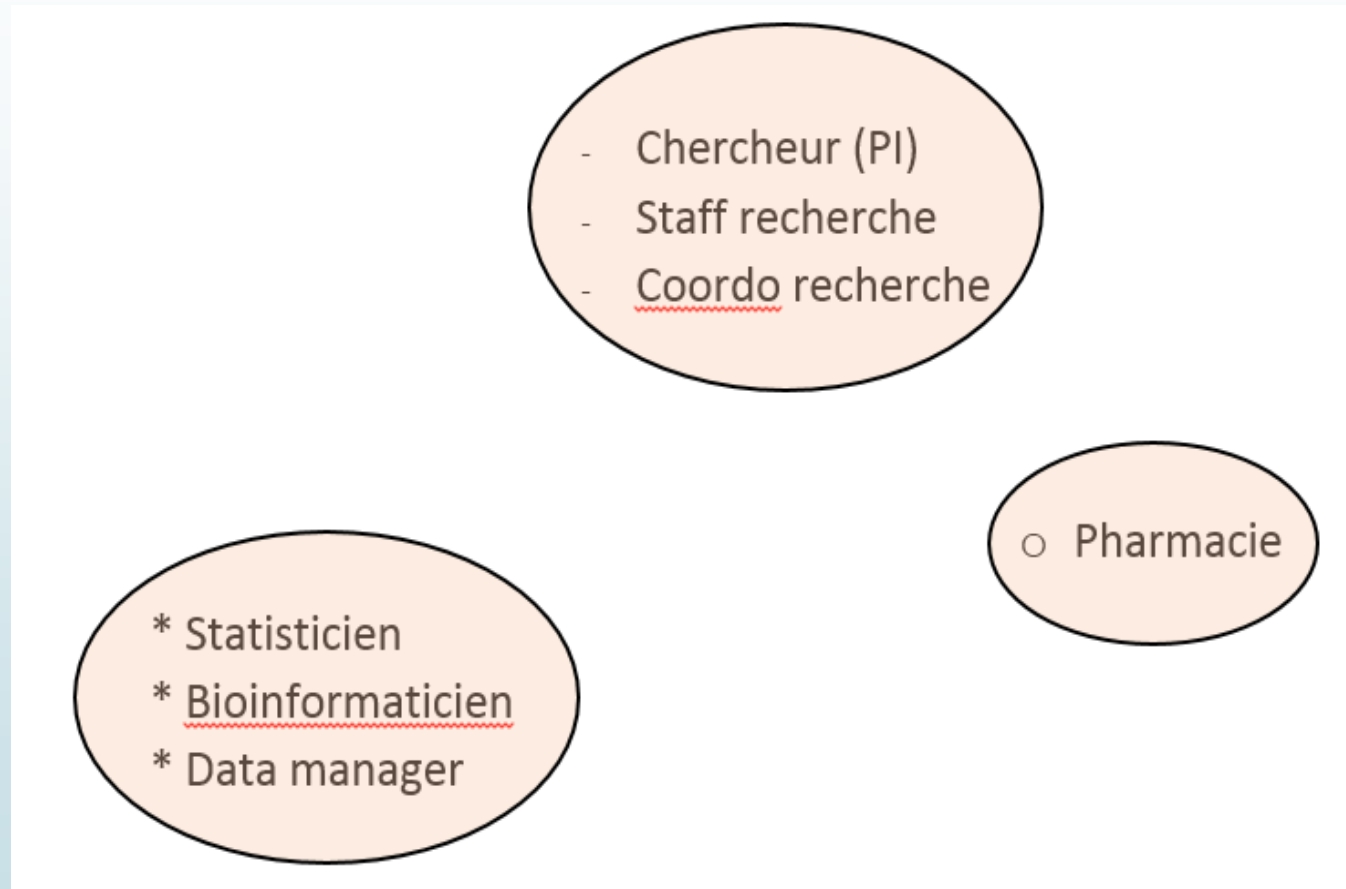
- Population: poqués
- Intervention: Médicament A
- Comparateur: Placebo
- Outcome: Rechute du problème

Groupe 1 ----- o

Groupe 2 -----x

- Time to event
- Étude clinique
- Assez longue durée (patients suivi 1 an, durée de l'étude 3 ans)

Projet à long terme: qui fait quoi



Projet à long terme:

- ▶ Double aveugle
 - Mise en place de l'insu
 - Levée d'insu d'urgence

- ▶ Randomisation
 - Par strate

- ▶ AE

- ▶ DMC: Data monitoring committee

- ▶ Ethique

- ▶ Santé Canada

Projet à long terme: à quoi sert le statisticien?

- Révision du protocole/Design
- Révision des questionnaires: outcomes principaux!
- Critères d'éligibilité
- Production de la liste de randomisation
- Coordination avec la pharmacie pour la mise en place de l'insu
- DMC: Data safety Monitoring Comitee
- Analyses

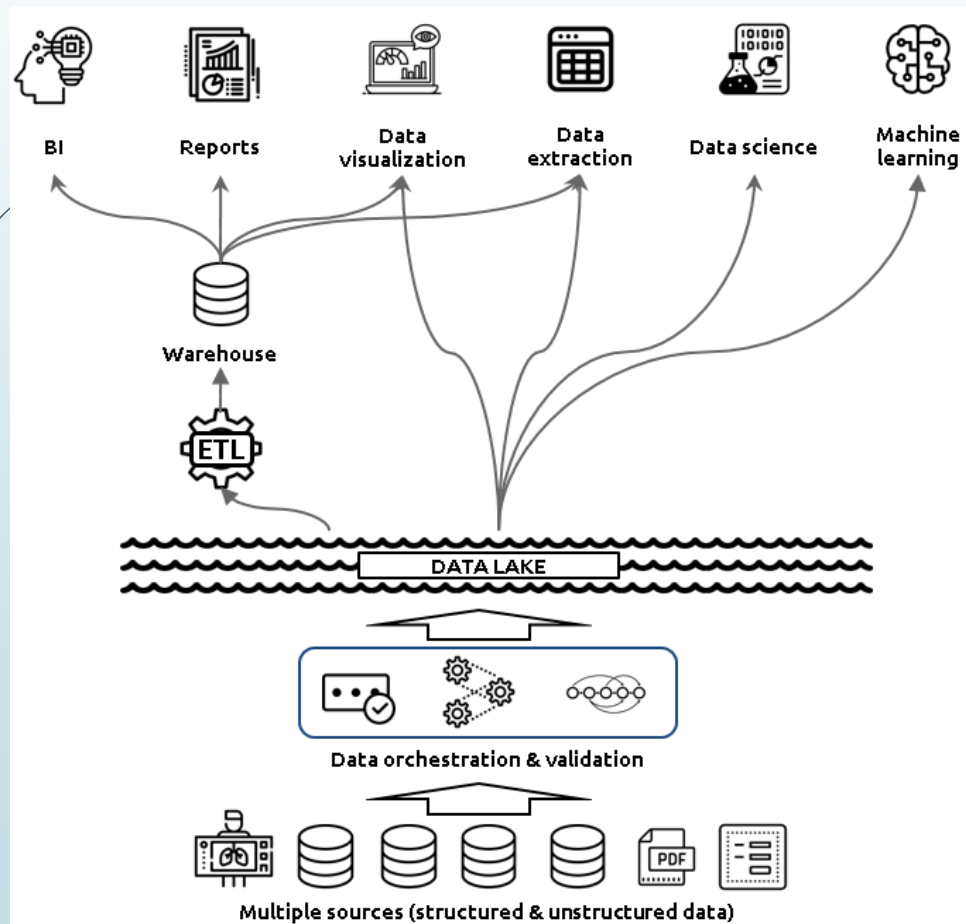
Qu'est-ce que ça implique?

- Meilleure qualité des données
- Développement d'une relation avec les équipes
- Implication d'un méthodologiste à toutes les étapes du projet

- Plus cher pour le chercheur
- Du côté du statisticien: plus de gestion

Projets hospitaliers et futurs projets

Infrastructure CITADEL



- Intégration de diverses données
- Jeux de données synthétiques
- Analyse distribuée
- Optimisation des flux de patients
- Plateforme automatisée d'essais cliniques

Comité scientifique

- Lecture de protocoles
- Mise en contact avec énormément de sujets
- Protocoles pharmaceutiques
- Point de vue des médecins
- Éthique

Pour finir.....

The best thing about being a statistician is that you get to play in everyone's backyard. John Tukey.

Merci!

Questions?

Un merci particulier à:

Kip Brown, Officier scientifique à CITADEL

Benoit Masse,URCA

Michaël Chassé, directeur scientifique à CITADEL

Références

1. Shmueli, Galit. "To explain or to predict?." (2010): 289-310.
2. Schumi, J., & Wittes, J. T. (2011). Through the looking glass: understanding non-inferiority. *Trials*, 12, 1-12.
3. Jennison, C., & Turnbull, B. W. (1999). *Group sequential methods with applications to clinical trials*. CRC Press.
4. http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf
5. Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., ... & Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC medical research methodology*, 10(1), 1-10.
6. Genders, T. S., Spronk, S., Stijnen, T., Steyerberg, E. W., Lesaffre, E., & Hunink, M. M. (2012). Methods for calculating sensitivity and specificity of clustered data: a tutorial. *Radiology*, 265(3), 910-916.