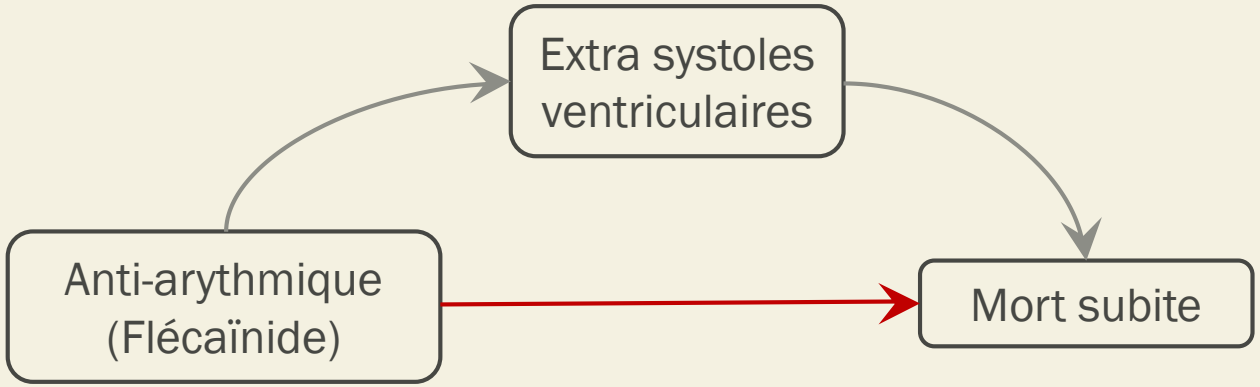
ESSAIS RANDOMISÉS

UE MRCE
27/02/2025

Léa Duchesne (lea.duchesne@univ-nantes.fr)

Médecine fondée sur les preuves

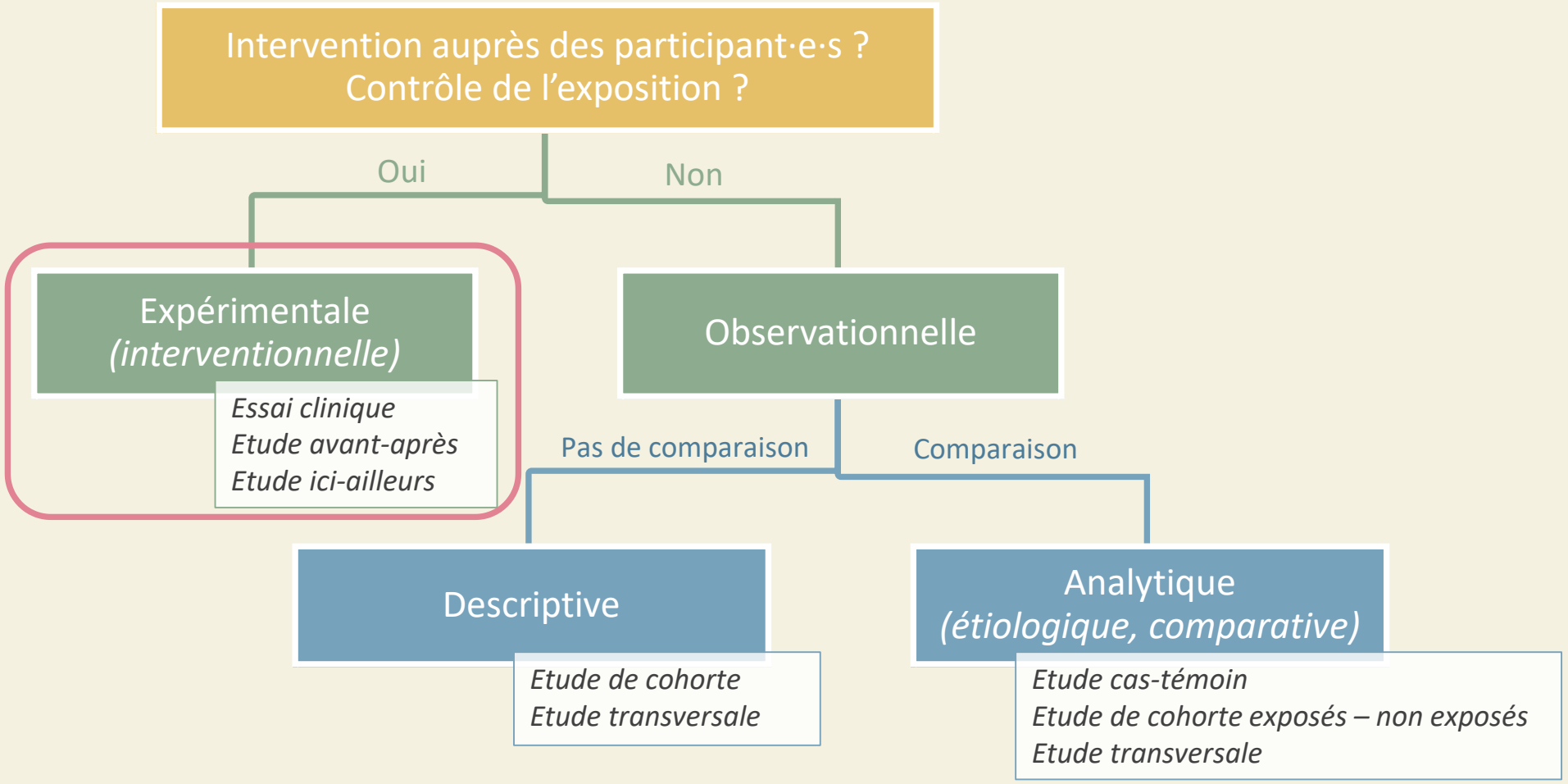
Exemple : risque de mort subite après un infarctus



- 1991 : essai randomisé CAST
- RR = 2,13 (groupe anti-arythmique versus placebo)

  Nombre de décès estimé aux Etats-Unis en raison de cela : 80 000

Essai randomisé contrôlé (ERC)



PRINCIPE D’UN ESSAI

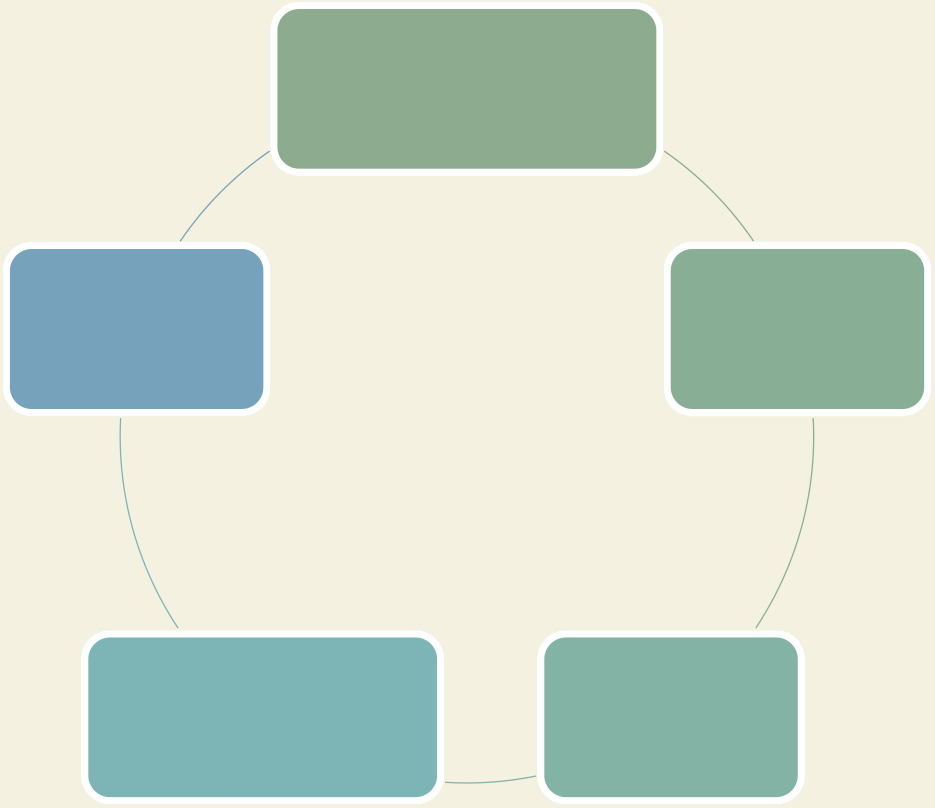
Recherche de causalité

Ampleur de l'effet (significatif ?)

Relation causale ?

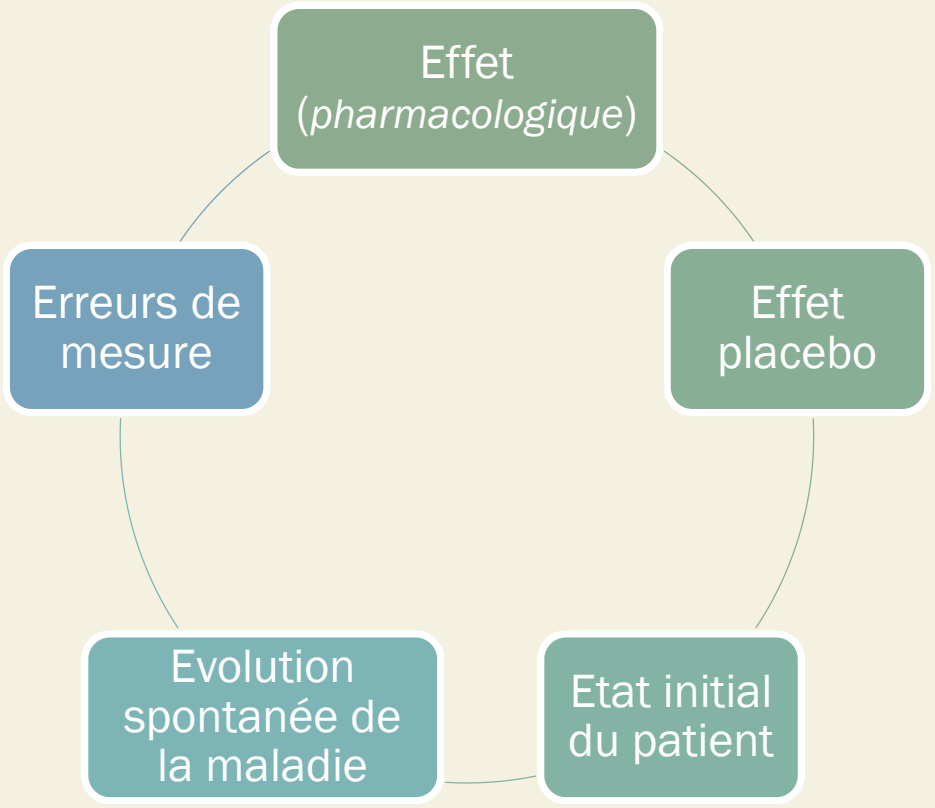
Enjeu

- *N'évaluer **que** l'action de l'intervention évaluée* ?



Enjeu

- *N'évaluer **que** l'action de l'intervention évaluée*



Comment le garantir ?

→ Comparaison

→ Design de l'étude

→ Analyse statistique

Lire les deux résumés d'article et répondre aux questions

Pour chaque article :

- Quels étaient les objectifs de l'étude ?
- Décrire la population inclue dans l'étude
- Certains groupes de personnes ont-ils été comparés dans le cadre de l'étude, et si oui lesquels ?
- Quels indicateurs ont été estimés afin de répondre aux objectifs de l'étude ?

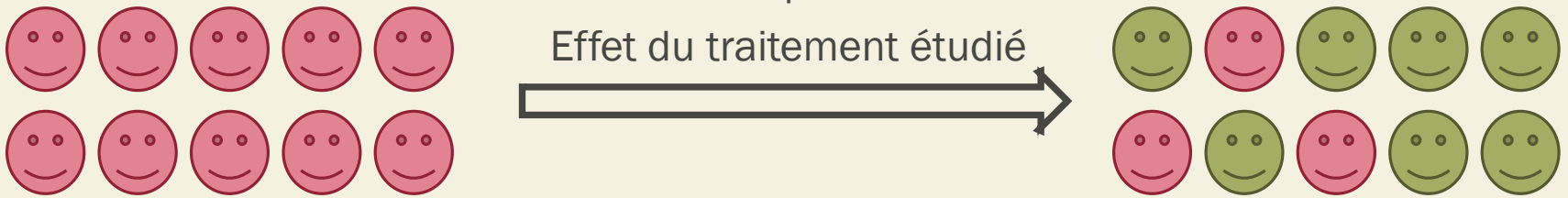
Voyez-vous des différences et/ou points communs en termes de méthodologie entre ces deux essais ?

COMPARAISON

Comparaison à un groupe contrôle

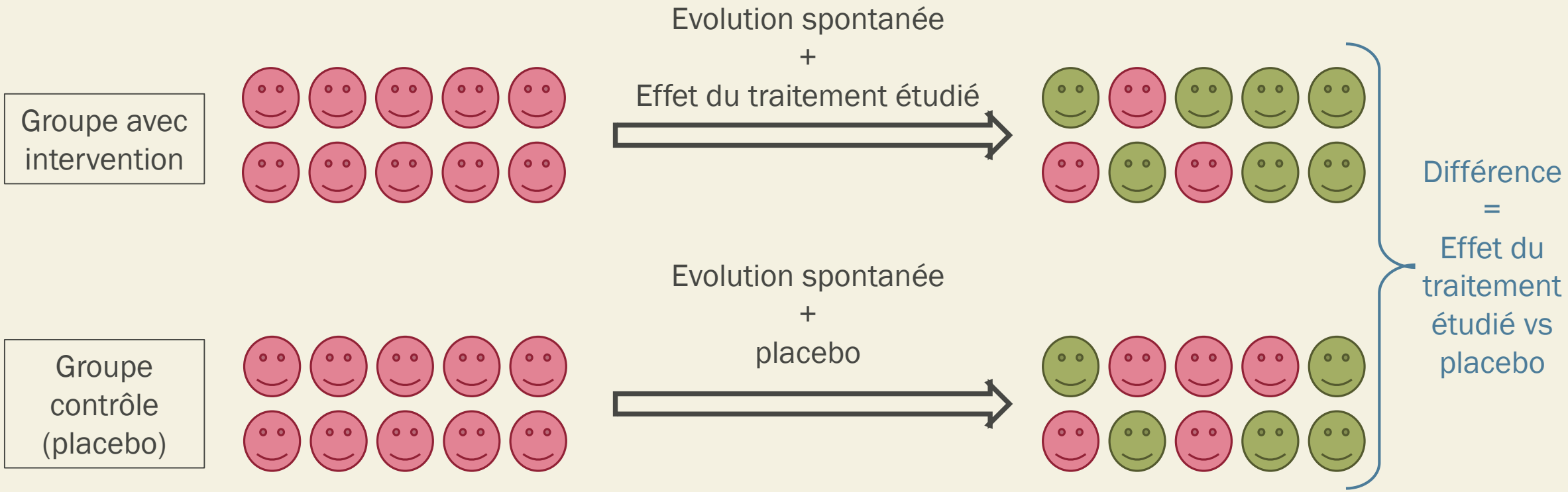
Etude avant-après

Groupe avec intervention



Comparaison à un groupe contrôle

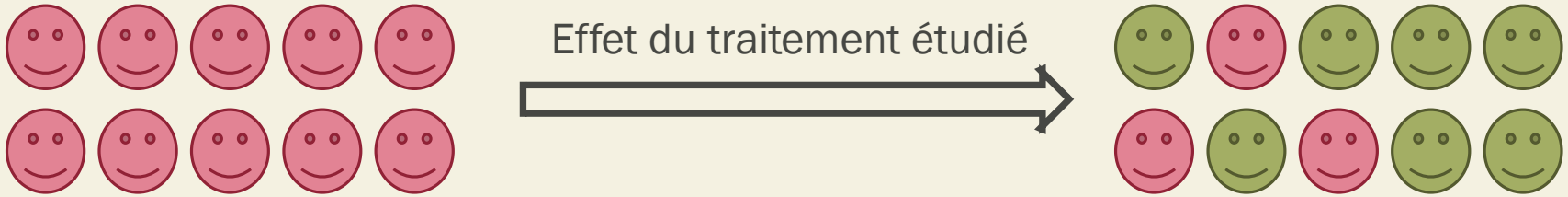
Etude avant-après



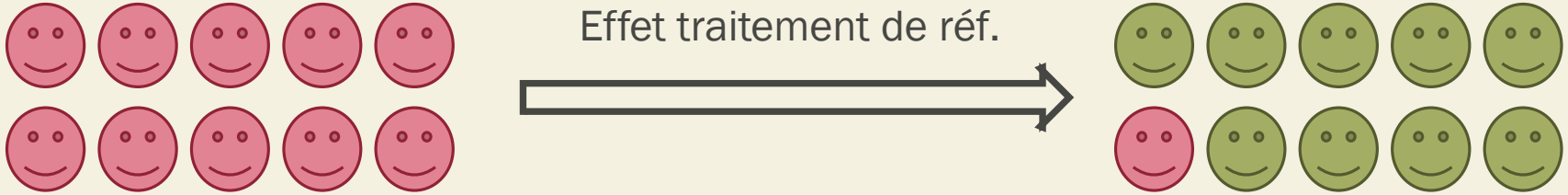
Comparaison à un groupe contrôle

Etude avec un groupe contrôle

Groupe avec intervention



Groupe contrôle (traitement de référence)

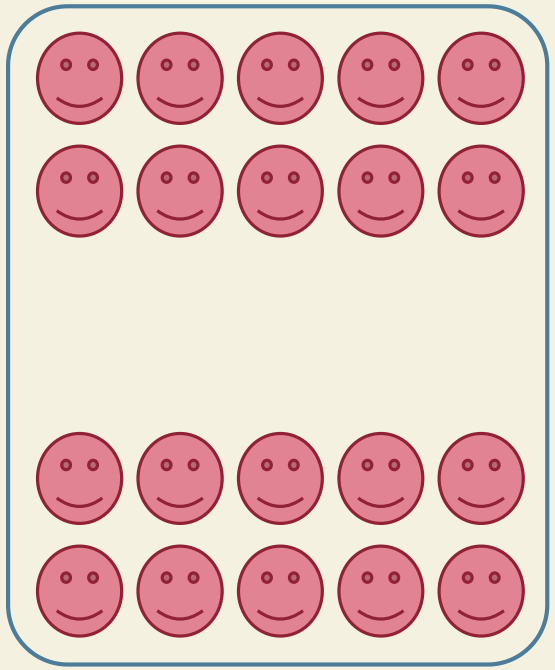


Comparaison à un groupe contrôle

Etude avec un groupe contrôle

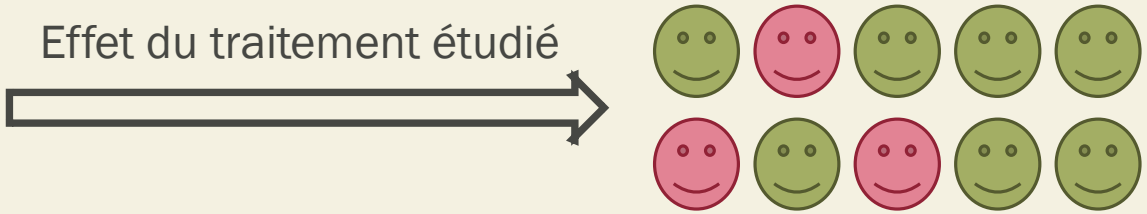
Groupe avec intervention

Groupe contrôle (traitement de référence)

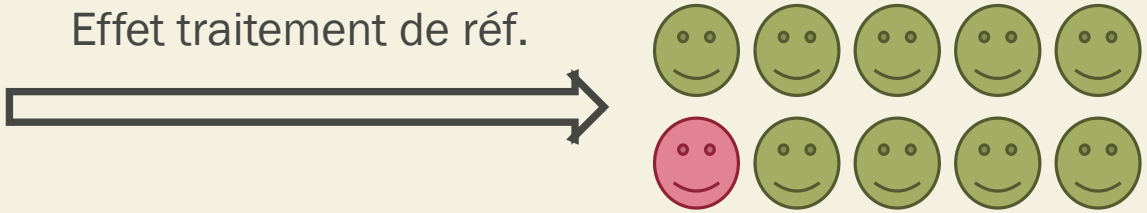


Doivent être **comparables** et **le rester tout au long du suivi**

Evolution spontanée
+
Effet du traitement étudié



Evolution spontanée
+
Effet traitement de réf.

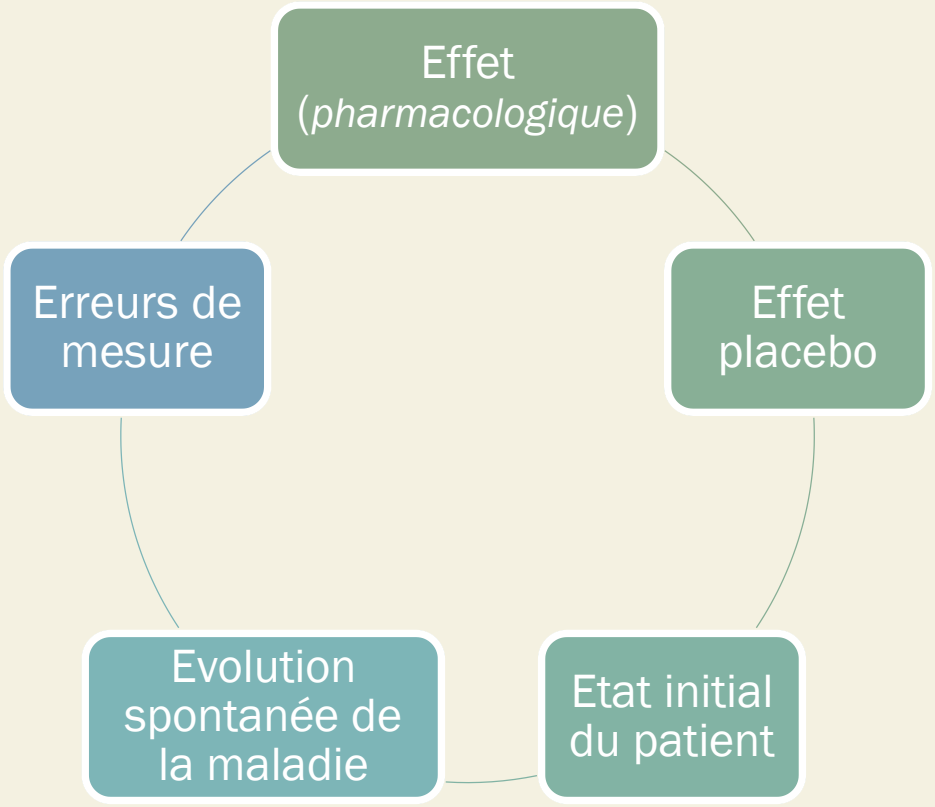


Enjeu

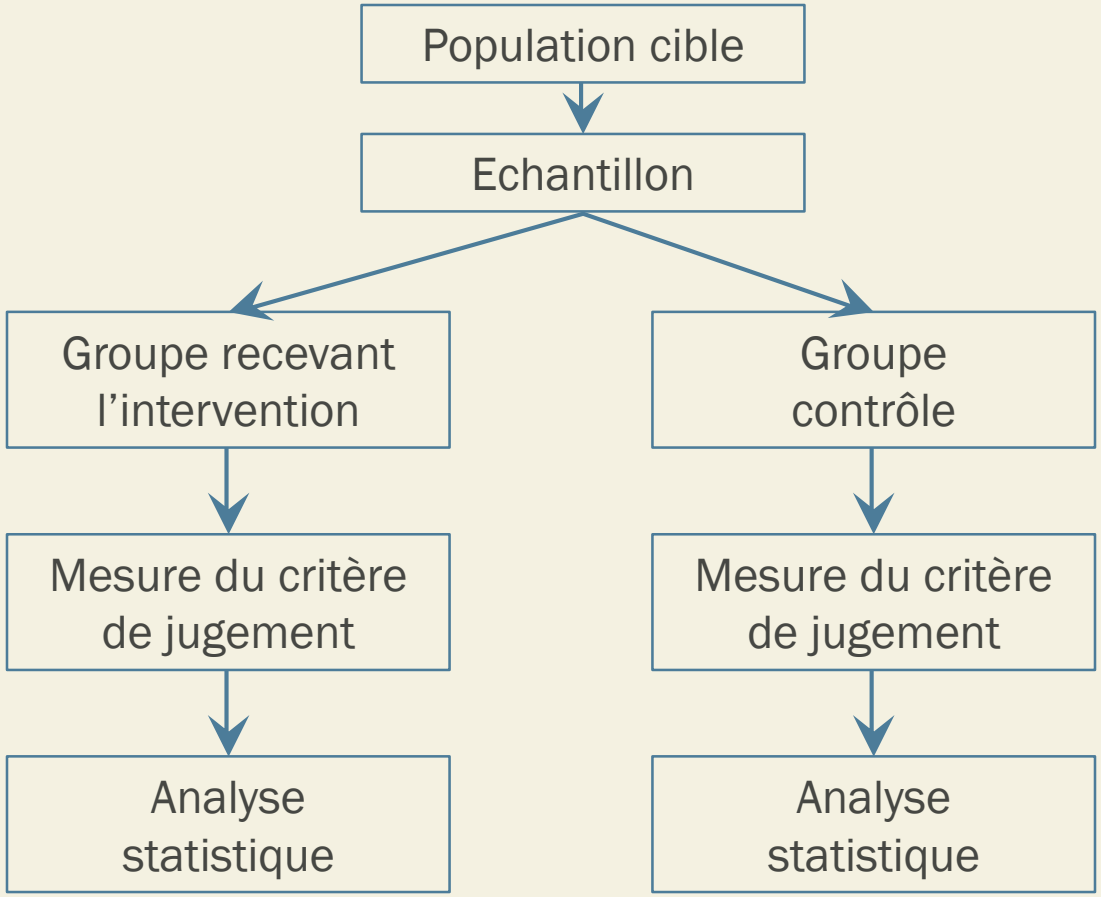
<u>Enjeux</u> : éviter les biais

- *Maintenir la comparabilité des groupes*
- *…*

→ *Design de l'étude (protocole)*



Déroulement d'un essai



Quel comparateur choisiriez-vous ici ?

- Essai sur l'efficacité d'un nouveau traitement contre le diabète de type 1
  - Insuline
  - Placebo

Comparateur

- **Que doit recevoir le groupe contrôle ?**
  - **Placebo** si aucune possibilité thérapeutique préexistante

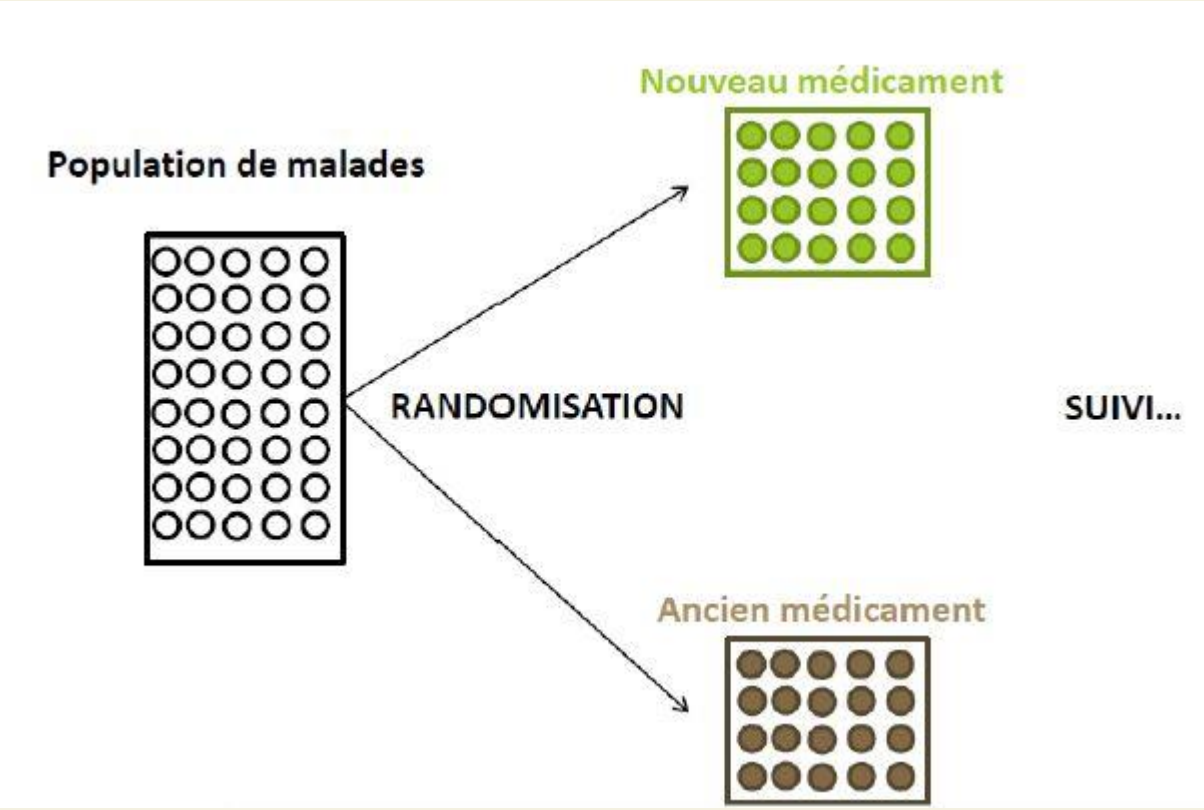
    → Bénéfice thérapeutique au delà de l'effet placebo
  - **Stratégie de référence** sinon

    → Bénéfice au-delà de ce qu'apporte l'intervention médicale de référence

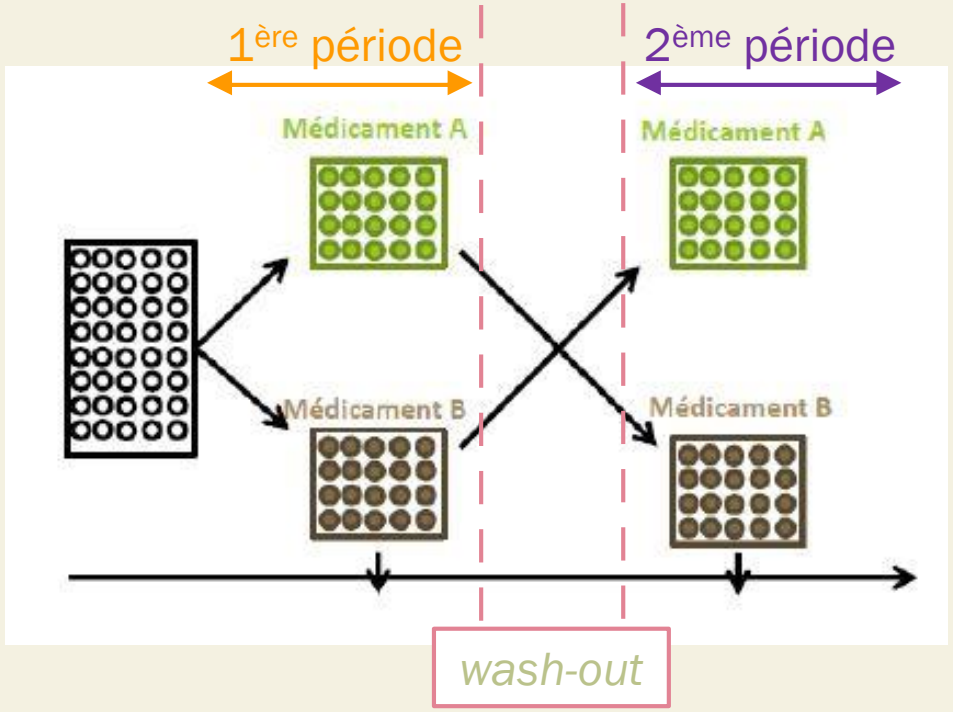
    → Potentiellement plus éthique

Plans d'expérience

Bras parallèles



Plans d'expérience
Cross-over

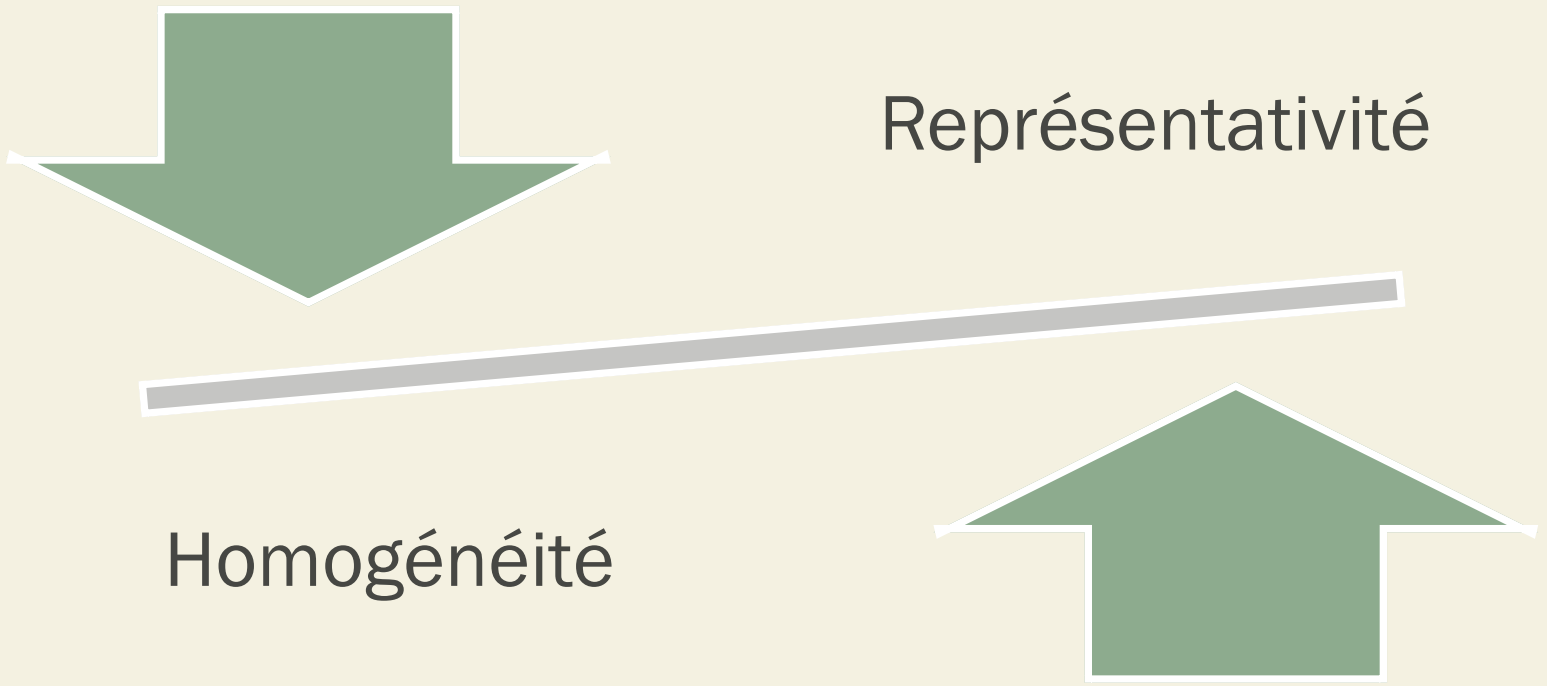


Hypothèses :

- Effet du 1er traitement administré ne se prolonge pas en 2ème période
- Retrouver en 2ème période un état de base de la maladie identique à celui observé en 1ère période
- Absence d'interaction entre les traitements et les périodes

POPULATION D'ÉTUDE

Critères d’éligibilité



Critères d’éligibilité
Définitions

Critères d’inclusion

(patients souhaités dans l’essai)

Critères de non-inclusion

(patients répondants aux critères d’inclusion mais non souhaités dans l’essai)

D’après vous, pourquoi les critères suivants ont été choisis ?

- Critères d’inclusion :
  - Proven chronic HCV infection (third-generation ELISA IgG test) with genotype 1, 2 or 4 and HCV-RNA >12 IU/ml
  - Agreement on using a contraceptive method 1-month prior treatment initiation and 7 months after treatment completion
- Critères de non-inclusion :
  - Daily alcohol consumption >20 g

Critères d'éligibilité
Définitions

Critères d'inclusion (patients souhaités dans l'essai)	○ Critères démographiques (âge, sexe, poids, etc.) ○ Critères définissant le diagnostic de la maladie
Critères de non-inclusion (patients répondants aux critères d'inclusion mais non souhaités dans l'essai)	○ Patients protégés par la loi (mineurs, femmes enceintes, etc.) ○ Contre-indications à l'un des traitements (**clause d'ambivalence**) ○ Pathologies associées (patients avec comorbidités) → risques particuliers

Critères d'éligibilité

Caractéristiques

Spécifique

Sensible

Reproductible

Critères d’exclusion

- Critères entrainant une sortie de l’étude (rares)
  - Consentement puis refus du patient
  - Critère évalué avant l’inclusion mais résultat connu qu’après l’inclusion
  - …

RANDOMISATION

Groupes comparables
Comment ?

Randomisation : tirage au sort de l'intervention reçue par les participants

La répartition ne doit dépendre

- Ni du patient
- Ni de la maladie
- Ni du médecin
- Ni du médicament

Principes à respecter pour **garantir une randomisation imprévisible et équilibrée**

- Clause d'ambivalence
- Clause d'ignorance
- Séquence de randomisation décidée selon une règle non systématique

Que pensez-vous de ces méthodes de randomisation ?

- Randomisation reposant sur une caractéristique du malade (ex : initiale du prénom, date de naissance …)
- Randomisation reposant sur le jour d'inclusion
- Enveloppe contenant le bras de randomisation tirée au hasard lors de la visite d'inclusion

Groupes comparables
Types de randomisation

- Randomisation simple
  - Probabilité de recevoir l'une des interventions est fixée *(ex : si 2 interventions, 50% pour chacune)*

    → Tirage à pile ou face
  - **Imprévisibilité ?**

    → Garantie
  - **Equilibre ?**

    → Non garanti, surtout en cas de petits échantillons

Groupes comparables
Types de randomisation

- Randomisation par bloc
  - Bloc de k patients

    → Chaque k participant, la moitié reçoit A, l'autre moitié reçoit B
  - Imprévisibilité ?

    → Garantie sauf si essai ouvert
  - Equilibre ?

    → Quasiment garanti (cas où un bloc n'est pas complet)

# patient	Intervention
1	A
2	A
3	B
4	B
5	A
6	B
7	A
8	B
9	B
10	B
11	A
12	A

Groupes comparables
Types de randomisation

# patient	Intervention
1	A
2	A
3	B
4	B
5	A
6	B
7	A
8	B
9	B
10	B
11	A
12	A

- Randomisation stratifiée
  - Si facteurs pronostiques importants connus (+ si multicentrique : sur le centre)

    → 1 liste de randomisation par strate de chaque facteur

Exemple :

Stade non avancé

Stade avancé

→ *66% de stade avancé dans le groupe B (4/6)*

→ *16% de stade avancé dans le groupe A (1/6)*

Groupes comparables
Types de randomisation

# patient	Intervention
1	A
2	A
3	B
4	B
5	A
6	B
7	A
8	B
9	B
10	B
11	A
12	A

- Randomisation stratifiée
  - Si facteurs pronostiques importants connus (+ si multicentrique : sur le centre)

    → 1 liste de randomisation par strate de chaque facteur

Exemple :

Stade

Stade

→ *66% de sta*

→ *16% de sta*

# patient	Intervention
1	
2	
3	
4	
5	
6	
7	
8	

B (4/6)

A (1/6)

# patient	Intervention
1	
2	
3	
4	
5	
6	
7	
8	

Groupes comparables
Types de randomisation

# patient	Intervention
1	A
2	A
3	B
4	B
5	A
6	B
7	A
8	B
9	B
10	B
11	A
12	A

- Randomisation stratifiée
  - Si facteurs pronostiques importants connus (+ si multicentrique : sur le centre)

    → 1 liste de randomisation par strate de chaque facteur

Exemple :

Stade

Stade

→ *66% de sta*

→ *16% de sta*

# patient	Intervention
1	A
2	A
3	B
4	B
5	B
6	B
7	A
8	A

B (4/6)

A (1/6)

# patient	Intervention
1	A
2	B
3	B
4	B
5	A
6	A
7	B
8	B

Groupes comparables
Délai entre la randomisation et l'intervention

- À minimiser afin d'éviter que ce produise :
  - Evénement
  - Décès
  - Retrait de consentement
  - Evolution du patient entrainant le statut de « non incluable »

Groupes comparables
Résultat de la randomisation

- Comparaison des caractéristiques des participants à l'inclusion
  - Etape primordiale
  - Pas de tests statistiques !
  - Examen de la répartition des facteurs pronostiques connus
- Si déséquilibre constaté à la fin de l'essai ?

  → Analyses statistiques appropriées (ajustement, analyses stratifiées en sous-groupes, etc.)

Table 1 – Patient characteristics at baseline.

Characteristic, measure	Complete cases sample	
	Maintenance arm (n = 54)	Spacing arm (n = 44)
Age (y), mean ± SD	56.0 ± 10.6	53.6 ± 9.9
Female, n (%)	38 (70.4)	35 (79.5)
Disease duration (y), mean ± SD	11.7 ± 11.2	7.2 ± 4.5
IgM RF positivity, n (%)	30 (66.7)	28 (70.0)
ACPA positivity, n (%)	34 (79.1)	34 (85.0)
Erosive disease on x-rays, n (%)	48 (88.9)	37 (84.1)
Swollen joint count (28 joints), mean ± SD	0.2 ± 0.6	0.3 ± 0.7
Tender joint count (28 joints), mean ± SD	0.5 ± 1.0	1.1 ± 2.2
ESR (mm/first hour), mean ± SD	11.1 ± 8.5	11.6 ± 7.7
CRP (mg/L), mean ± SD	3.0 ± 1.9	2.6 ± 1.6
DAS28, mean ± SD	1.7 ± 0.5	1.9 ± 0.6
DAS44, mean ± SD	0.9 ± 0.4	1.1 ± 0.5
EQ-5D-3L VAS, mean ± SD	81.2 ± 17.4	76.6 ± 23.0
Previous RA treatment		
Number of previous DMARDs, mean ± SD	3.0 ± 2.0	3.0 ± 2.0
Patients with previous biological DMARD, n (%)	16 (29.6)	13 (29.5)
Ongoing treatment		
Adalimumab/etanercept, n (%)	26 (48.1)/28 (51.9)	20 (45.5)/24 (54.5)
Monotherapy/combination, n (%)	13 (24.1)/41 (75.9)	8 (18.2)/36 (81.8)

Groupes comparables
Limites de la randomisation

- Assure qu'en moyenne les groupes sont comparables

  → Effet du hasard peut entrainer un déséquilibre (problème si facteur pronostic ou prédictif)

- Principe d'ignorance

  → Acceptabilité ?

  Exemples :

  - *Patient sans optique de traitement*
  - *Patient au pronostic jugé mauvais par son médecin*
  - *Essai contre placebo, choix du patient*

AVEUGLE

Aveugle

- Désigne le fait qu'une partie ou l'ensemble des parties prenantes d'un essai (patient, médecin, statisticien, évaluateur, etc.) ne connaissent pas quels patients reçoivent l'une ou l'autre des interventions

- Plusieurs niveaux : ouvert, simple, double, triple

- Objectif : éviter les modifications de comportement des participants

  → *Maintien de la comparabilité*

  → *Egalité d'appréciation du critère de jugement*

Aveugle

- Mise en œuvre
  - Si comparaison d'une **intervention médicamenteuse**
    - *Placebo*

      →*Même forme et mode d'administration que le traitement évalué*

      →*Mêmes effets indésirables que le traitement évalué (difficulté)*
    - *Traitement de référence*

      →*Mêmes critères que précédemment*

Aveugle

- Mise en œuvre
  - Si comparaison d'une **intervention non médicamenteuse**

    →*Intervention la plus semblable possible*

    - Pas toujours faisable ou éthique
    - Alors **évaluation du critère de jugement en insu** par une tierce personne ignorant la nature du traitement

Comment mettre en place l'aveugle ici ?

On cherche à mettre en place un essai pour comparer l'efficacité du traitement A au traitement B.

- Intervention A : piqûre 1 fois/ jour

- Intervention B : 2 pilules/ 2 fois par jour (matin et soir)

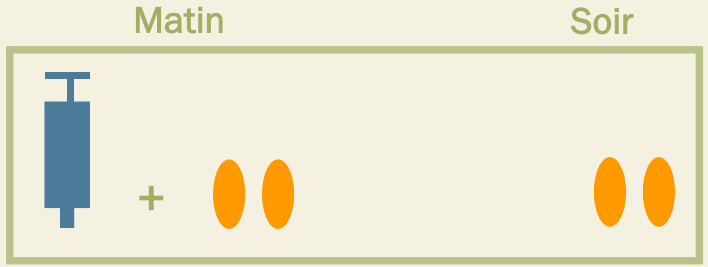
Aveugle

Exemple de **double placebo** :

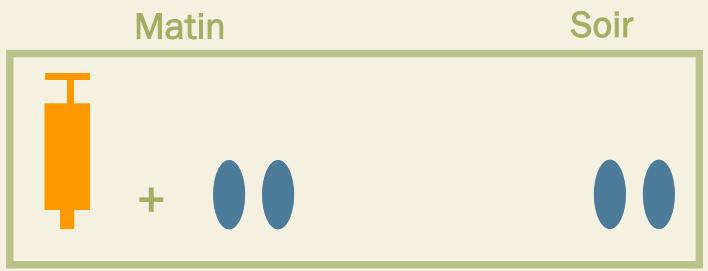
- Intervention A : piqûre 1 fois/ jour
- Intervention B : 2 pilules/ 2 fois par jour (matin et soir)

⇨ Chaque patient reçoit l'intervention correspondant à son bras de traitement + 1 placebo :

Intervention A +
placebo de l'intervention B



Intervention B +
placebo de l'intervention A



CRITÈRE DE JUGEMENT

Définition

- **Critère principal** : 1 seul (normalement, sinon les limiter)
- Plusieurs **critères secondaires/exploratoires** possibles

  → Tous définis dans le **protocole** en amont de l'essai

- Différents types de critères :
  - Qualitatifs / quantitatifs
  - Objectifs / subjectifs *(douleur, consommation, …)*
  - Combinaison de mesures *(IMC, ratio, …)*

Que pensez-vous de ce critère de jugement ?

Etude de mars 2020 de Didier Raoult sur l'évaluation de l'hydroxychloroquine comme traitement de la COVID-19

→ Critère principal : diminution de la charge virale

Qualités recherchées

- Doivent être :
  - Exacts (valide)
  - Précis (fiable)
  - Reproductible (variabilité intra-individuelle)
  - Sensible
  - Spécifique
  - Acceptable
  - Facile à réaliser
  - Peu coûteux

Qualités intrinsèques

Qualités médicales (doit être cliniquement pertinent, et consensuel si possible)

Qualités opérationnelles

Moment de mesure

- A l'issue de l'intervention
- Avant l'initiation de l'intervention et à l'issue de cette dernière

  → Si étude de la variation du critère

  - Variation absolue
  - Variation relative
- Avant l'initiation de l'intervention et de façon répétée pendant la période de suivi
  - Complexifie l'analyse
  - Plus d'information (cinétique)

ANALYSE STATISTIQUE

Calcul du nombre de sujets nécessaires

- Intérêt ?
  - **Statistique** : *risque β dépend de la taille de l'échantillon*
    - ⇨ S'assurer d'être capable de mettre en évidence une différence intéressante si elle existe
  - **Ethique**
    - Si sous-dimensionné : perte de **puissance**
      - ⇨ Inclusion d'individus de manière inutile
      - ⇨ Erreur de conclusion peut avoir des conséquences graves
    - Si surdimensionné : complexification inutile
      - ⇨ Retard des résultats
      - ⇨ Risque d'abandon

Calcul du nombre de sujets nécessaires

- Prérequis :
  - Définition du test (**unilatéral** / **bilatéral**)
  - **Bénéfice minimum cliniquement intéressant** ou **différence attendue** (Δ) sur le critère principal
  - **Variabilité du critère principal** dans la population (σ)
  - Définir les **risques d'erreur** souhaités
    - Risque de 1ère espèce (α)
    - Risque de 2ème espèce (β)

Calcul du nombre de sujets nécessaires

Influence de α, β, σ et Δ sur le NSN ?

- $\Delta \nearrow \Rightarrow$ NSN $\searrow$
- $\sigma \nearrow \Rightarrow$ NSN $\nearrow$
- $\alpha \nearrow \Rightarrow$ NSN $\searrow$
- $\beta \nearrow \Rightarrow$ NSN $\searrow$

Question de recherche

- Supériorité
  - ⇨ Souhait de démontrer la supériorité de la nouvelle intervention vis-à-vis de la référence
  - ⇨ Définition d'une différence que l'on veut mettre en évidence
- Equivalence
  - ⇨ Souhait de démontrer l'équivalence d'efficacité de la nouvelle intervention vis-à-vis de la référence
  - ⇨ Définition d'une marge d'équivalence
- Non-infériorité
  - ⇨ A minima, l'efficacité n'est pas moins bonne par rapport à celle de la référence

Evaluation aussi sur des **critères secondaires**

⇨ Bénéfice par rapport à l'intervention de référence ?

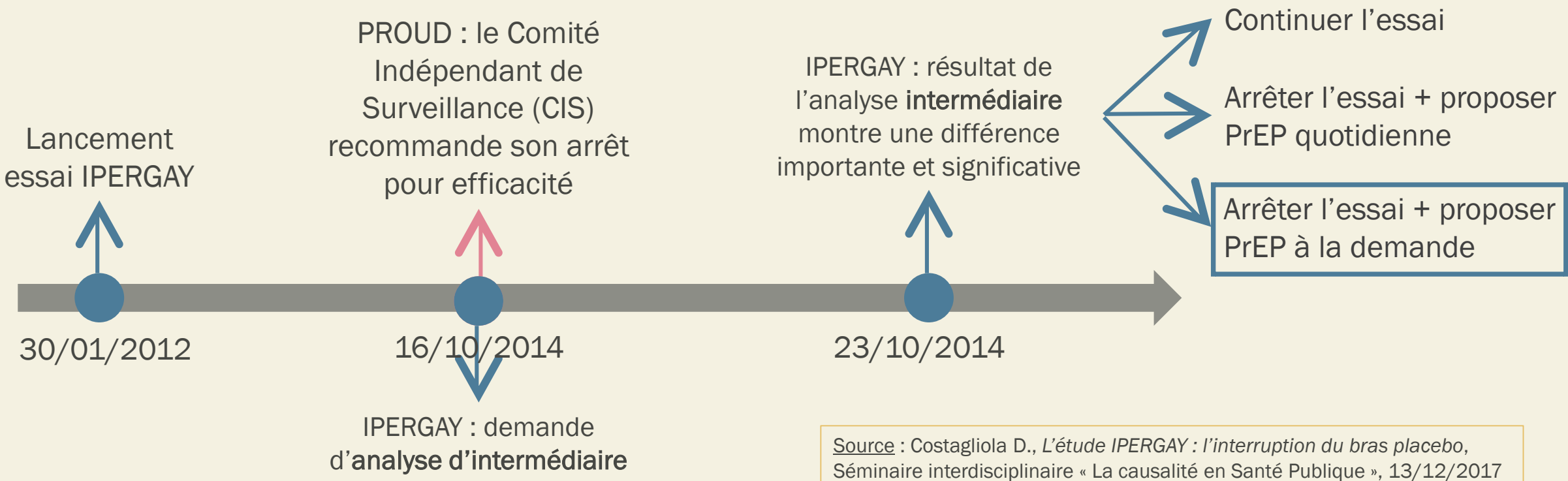
Ce que l'on planifie

- Analyses sont décrites en amont dans le **protocole**
- Spécificités des **analyses intermédiaires** :
  - Doivent aussi être définies en amont
  - Peuvent entraîner l'arrêt de l'essai
    - Arrêt pour **toxicité**
    - Arrêt pour **efficacité**
    - Arrêt pour **futilité**

D'après vous, pourquoi pourrait-on vouloir arrêter un essai en cours de route ?

Arrêt pour efficacité
Essai IPERGAY

	Population	Intervention	Design
PROUD	HSH+ transgenres ayant des rapports anaux sans préservatifs	PrEP immédiatement vs dans 12 mois	Essai ouvert
IPERGAY	Idem	PrEP à la demande vs placebo	Essai en double aveugle



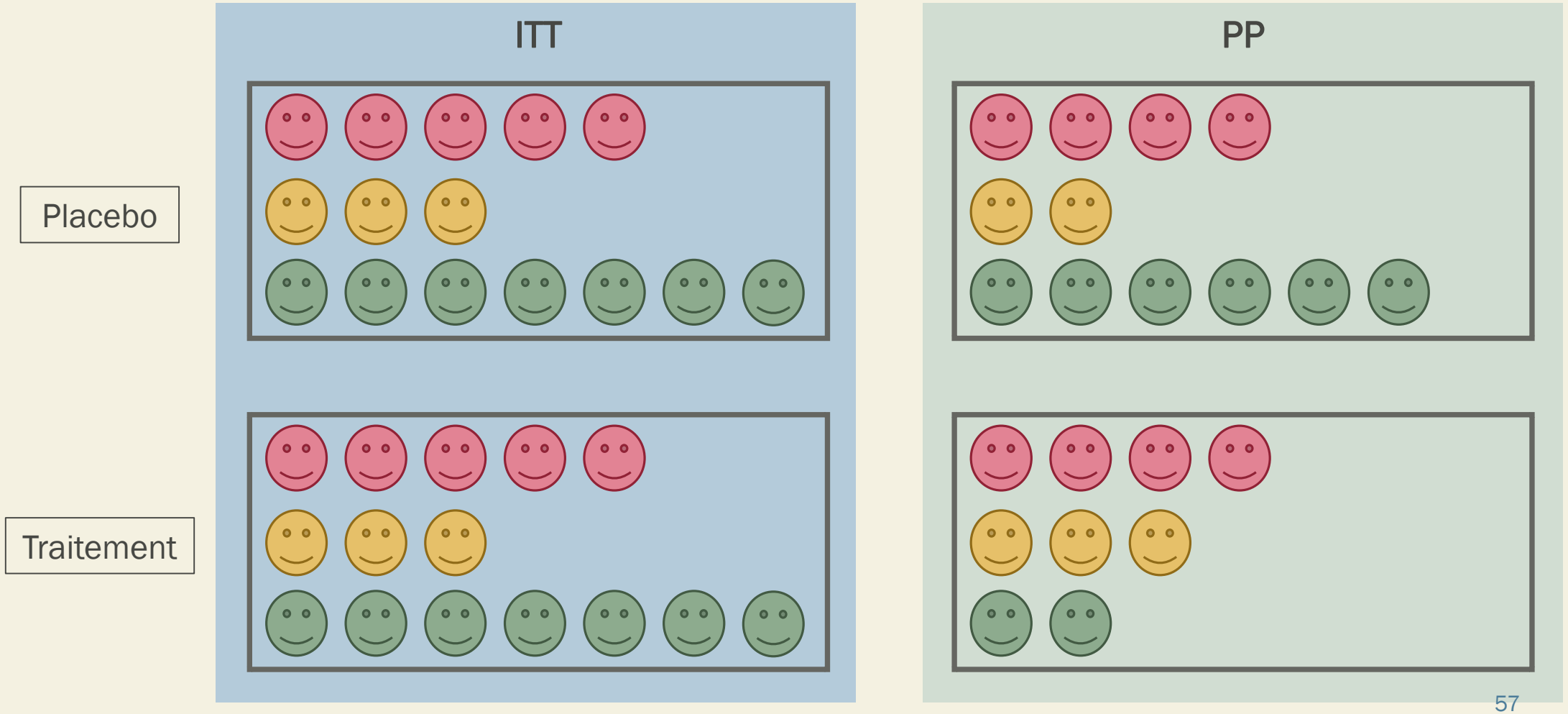
Source : Costagliola D., *L'étude IPERGAY : l'interruption du bras placebo*, Séminaire interdisciplinaire « La causalité en Santé Publique », 13/12/2017 https://causalitesantepublique.wordpress.com/references

Ce que l'on ne planifie pas

Patients inclus à tort	○ Non respect de la clause d'ambivalence ○ Raisons éthiques
Attrition	○ Perdus de vue ○ Retrait de consentement en cours d'étude
Non respect de l'intervention allouée par la randomisation	○ Patients non traités ou ayant reçus le mauvais (pas celui désigné par la randomisation) ○ Patients ayant reçus la moitié de la dose prévue ○ Patients prenant un traitement concomitant interdit par le protocole

→ Comment analyser ces patients ?

Analyse en intention de traiter versus per protocole



BIAIS ET MÉTHODOLOGIE D'UN ESSAI

En résumé

Biais

Randomisation

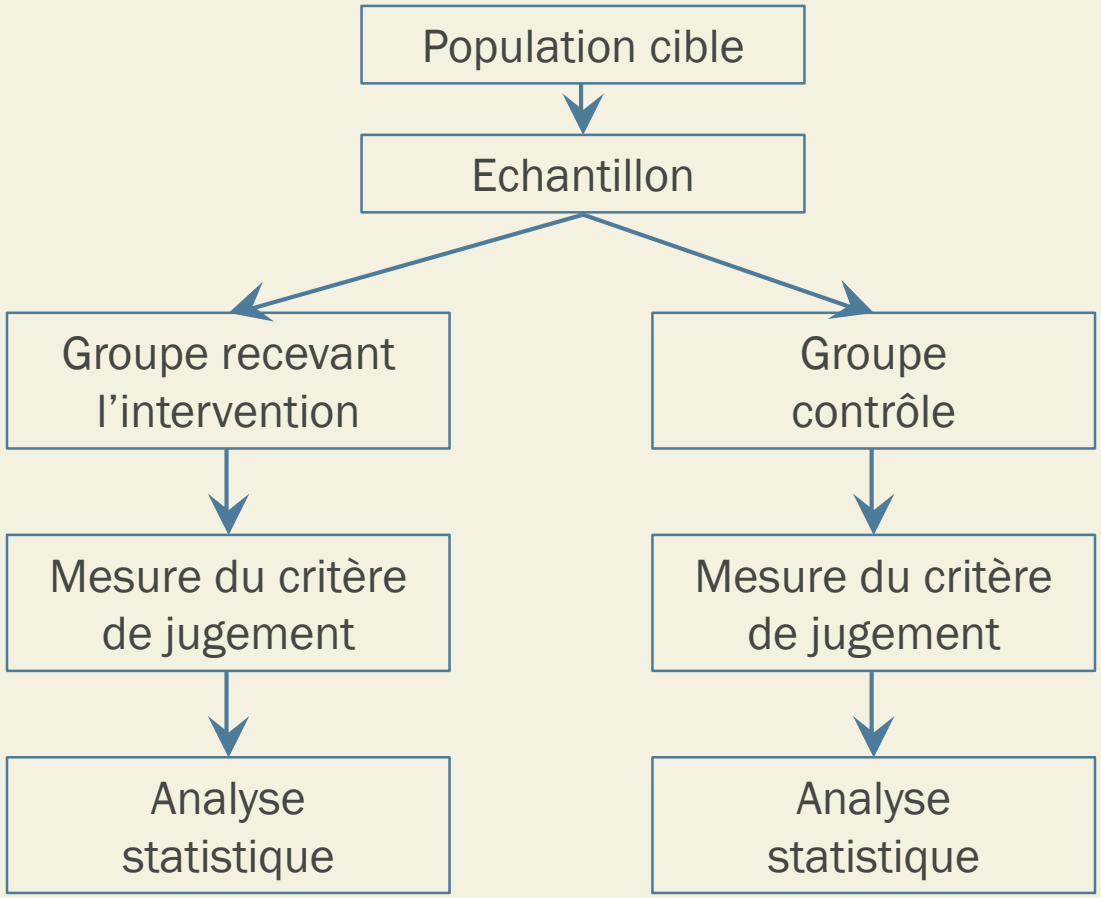
Evaluation en insu

Définition CJP

Double aveugle

Analyse en ITT

Critères d'éligibilité



Biais de suivi

Biais de sélection

Biais de confusion

Biais d'attrition

Biais de classement

Biais

Définition CJP

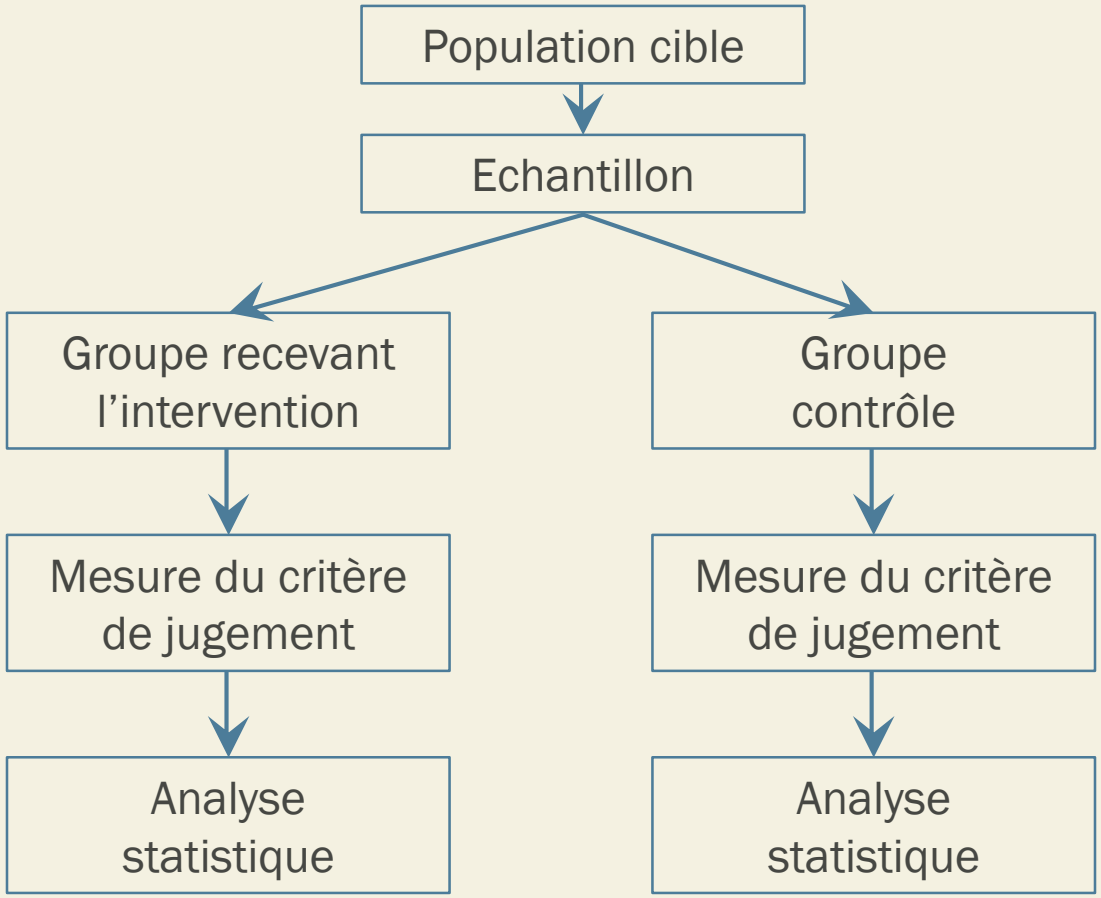
Critères d'éligibilité

Randomisation

Double aveugle

Evaluation en insu

Analyse en ITT



Biais de classement

Biais de sélection

Biais de confusion

Biais de suivi

Biais de classement

Biais d'attrition

QUELQUES REMARQUES

Monitorage

- Contrôle des données d'un essai (en temps réel)
- Respect des **Bonnes Pratiques Cliniques** (BPC)
- **Attaché de recherche clinique** (ARC)
    - Vérification :
        - Respect des critères d'éligibilité
        - Recueil des consentements éclairés
        - Cohérence des données avec le protocole
        - Validation des événements
        - …

Et après l'essai clinique ?

- Généralisation des résultats :
    - Dépend de la **population analysée** vs **population cible**
    - **Méta-analyses**
- Diffusion de l'intervention auprès des patients
- Pharmacovigilance
    - Surveillance de routine
    - Essai de phase IV

Ethique

- **ERC : contradiction entre l'intérêt du patient et l'évaluation ?**

  Exemple : étude Ipergay : question du recours à un placebo

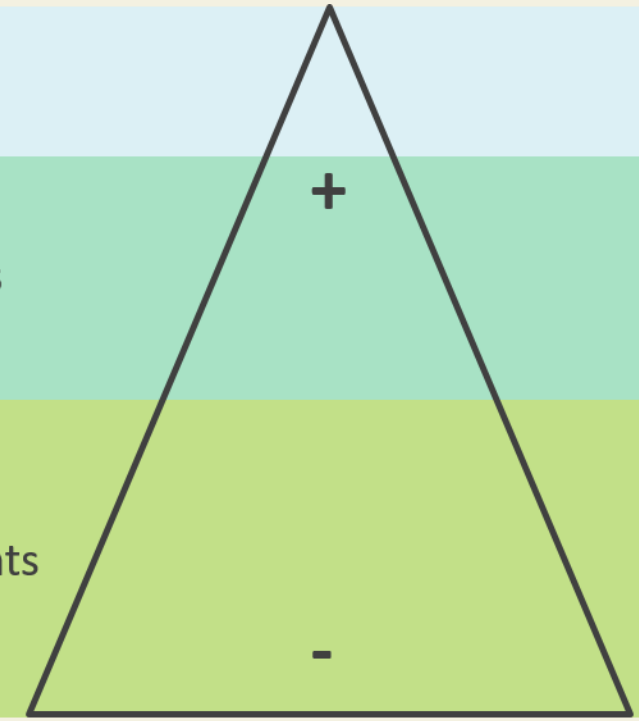
  - Au début de l'essai : plusieurs organisations font savoir qu'elles sont contre le bras placebo car l'efficacité de la PrEP serait déjà prouvée
  - Etudes contradictoires à l'époque
  - Si arrêt bras placebo :
    - Pas de système de mise à disposition de la PrEP à ce moment-là en France
    - Perte de l'accompagnement préventif

  → Décision de continuer l'essai par le CIS

« Le plus haut niveau de preuve »

- Hiérarchie niveaux de preuve apportés selon le schéma d'étude (classification HAS)

Essais comparatifs randomisés de forte puissance Méta-analyse d'essais comparatifs randomisés	**Grade A** (fort niveau de preuve)
Essais comparatifs randomisés de faible puissance Etudes comparatives non randomisées bien menées Etudes de cohortes	**Grade B** (niveau intermédiaire de preuve)
Etudes cas-témoins Etudes comparatives comportant des biais importants Etudes épidémiologiques descriptives Séries de cas	**Grade C** (moindre niveau de preuve)



! Le niveau de preuve apporté repose aussi sur la **qualité** de l'étude

Essai randomisé : pas toujours possible

- Problème éthique
  - Exemple : comparer l'efficacité des greffes rénales par rapport à la dialyse
- Critère de jugement à trop long-terme
  - Exemple : intérêt des traitements contre l'hypertension dans la mortalité cardio-vasculaire
- Phénomène de contamination suite à la randomisation
  - Exemple : évaluation de l'intérêt d'une formation auprès de médecins

→Recours aux essais non randomisés individuellement (randomisation par grappes [clusters]), aux essais non randomisés, aux études non interventionnelles… (pose la question de la comparabilité des groupes)