

Méthodologie en recherche biomédicale : les études expérimentales

Yohann.Foucher@univ-nantes.fr

Equipe d'Accueil 4275 "Biostatistique, recherche clinique et mesures subjectives en
santé", Université de Nantes

Odontologie - Cours #4

Les études
expérimentales

Les tests
d'hypothèse

1. Les études expérimentales

2. Les tests d'hypothèse

Les études
expérimentales

Les tests
d'hypothèse

1. Les études expérimentales

2. Les tests d'hypothèse

Les études
expérimentales

Les tests
d'hypothèse

- L'objectif est de contrôler l'essai pour limiter les biais.
- On parle aussi d'étude contrôlée (problème sémantique avec groupe contrôle).
- Les traitements peuvent être :
 - répartis aléatoirement par **tirage au sort en groupes parallèles**
 - administrés successivement aux patients (**essais intra-individuels**)

- **Indépendance** entre l'attribution du traitement et les caractéristiques du patient.
- **Imprévisibilité** de la répartition *.
- Le seul moyen de garantir la comparabilité des groupes.
- **Il s'agit du protocole de référence pour démontrer une causalité.**
- Ne dispense pas de vérifier la comparabilité des groupes ensuite (sans faire de tests d'inférence, voir la section "tests d'hypothèse").
- En anglais : *randomized trial*.
- Inconvénients :
 - Grande variabilité des critères de jugement entre les individus.
 - Risque d'obtenir des groupes non-comparables (surtout quand les effectifs inclus sont faibles).

*. On ne peut pas deviner le traitement qui sera attribué

- Traitements administrés successivement et aléatoirement.
- Situation particulière : traitements administrés simultanément.
- Avantages :
 - Moins de sujets nécessaires que pour un essai en parallèle.
 - Comparabilité assurée des groupes de traitements.
- Inconvénients :
 - Maladie évolutive (l'état initial sera différent entre les traitements d'un même patient).
 - Traitement long (état initial différent et longueur l'étude).
 - Peu adapté lorsqu'il y a un phénomène d'apprentissage.
 - L'effet *carry-over* : la première phase de traitement peut influencer la seconde.

→ Aménagement d'une phase de sevrage intermédiaire (*wash-out*).
- 2 types d'essais intra-individuels :
 - Les **essais croisés** quand 2 traitements (*cross-over*).
 - Les **carrés** latins quand plus de 3 traitements (*latin square*).

Les études
expérimentales

Les tests
d'hypothèse

- 2 traitements A et B sont comparés.
- L'ordre, AB ou BA, est choisi aléatoirement.
- Nécessité de prendre en compte l'**effet de l'ordre**
- Possibilité de tester un effet différent du traitement selon l'ordre (**interaction**).
- Evaluer si le niveau initial du critère de jugement varie en fonction de la période (efficacité du wash-out).
- Equivalent à un carré latin 2×2 .

Les carrés latins

- Plus de deux traitements comparés.
 → Ex : Quatre traitements A, B, C et D. 4×4 *latin square*.
- Pour définir un carré latin, il est nécessaire de respecter les conditions suivantes :
 - Le nombre de sujets est égal au nombre de traitements.
 - Chaque sujet reçoit tous les traitements.
 - Chaque période inclue tous les traitements.

Patients	Période 1	Période 2	Période 3	Période 4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

- Ce schéma permet d'étudier les facteurs "sujet", "ordre" et "traitement".
- On utilise plusieurs carrés latins, la taille de l'échantillon étant un multiple du nombre de traitements.

Problème

- L'ordre des traitements peut avoir une influence
- Les séquences AB, BC, CD, DA... sont surreprésentées

Patients	Période 1	Période 2	Période 3	Période 4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

- Utilisation de **carrés latins équilibrés** (*balanced latin square*).
- Par exemple :

Patients	Période 1	Période 2	Période 3	Période 4
1	A	B	C	D
2	B	D	A	C
3	C	A	D	B
4	D	C	B	A

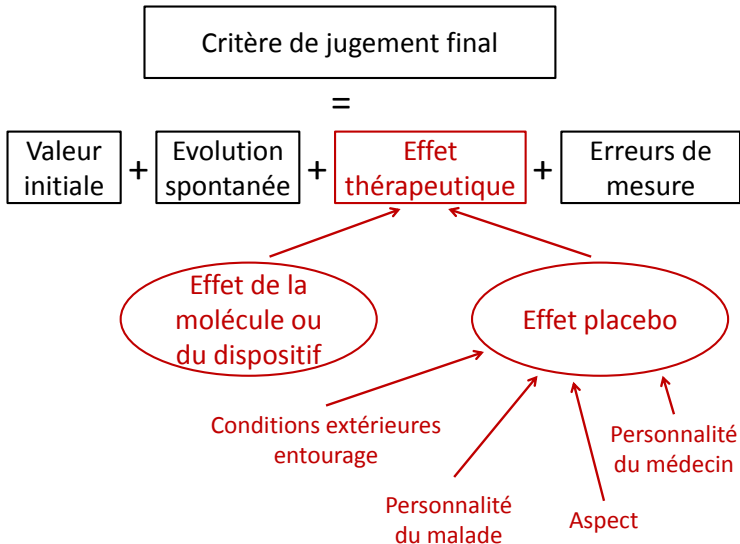
- ① L'investigateur n'est pas totalement objectif :
 - A priori favorable au nouveau traitement testé.
 - Influencé par les résultats au fil de l'essai : la préférence peut donc varier.
 - La préférence influence l'appréciation des critères de jugement :
 - vers le succès thérapeutique (efficacité surestimée).
 - vers l'absence d'effets secondaires (tolérance surestimée).

 - ② Le sujet croit aux avancées thérapeutiques :
 - Amélioration des critères subjectifs.
 - Minimisation des effets secondaires.
- **Simple aveugle (single-blind)** :
- Seul le patient ne connaît pas le traitement.
- **Double aveugle (double-blind)** :
- Le patient et tous les autres membres de l'équipe soignante ne connaissent pas le traitement.

L'effet placebo

Les études
expérimentales

Les tests
d'hypothèse



Les études
expérimentales

Les tests
d'hypothèse

- Uniquement si la maladie est bénigne ou si aucun traitement n'est connu.
- Difficulté de recrutement.
- **Placebo impur** : traitement prescrit dans une situation où l'efficacité n'est pas prouvée.
- Même apparence que le traitement pour garantir l'aveugle.
- **Double placebo** si deux traitements A et B sont comparés mais ne peuvent pas avoir la même apparence
 - Dans le bras A, on administre le placebo de B.
 - Dans le bras B, on administre le placebo de A.

Les études
expérimentales

Les tests
d'hypothèse

- G. Bouvenot, Muriel Vray. Essais Cliniques - Théorie, pratique et critique. 2nde édition. Médecine-Science Flammarion. Collection Statistique en Biologie et en Médecine. 1996.
- A. Laplanche, C. Com-Nougué, R. Flamant. Méthodes statistiques appliquées à la recherche clinique. Médecine-Science Flammarion. Collection Statistique en Biologie et en Médecine. 1996.
- J. Bouyer et al. Epidémiologie - Principes et méthodes quantitatives. Les éditions INSERM. 1995.

Les études
expérimentales

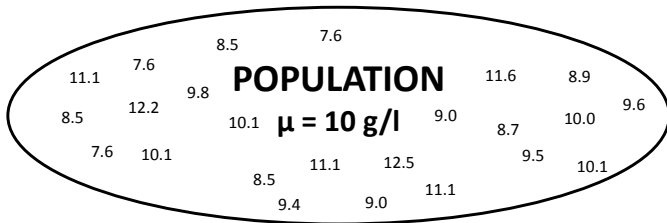
Les tests
d'hypothèse

1. Les études expérimentales

2. Les tests d'hypothèse

Les études
expérimentales

Les tests
d'hypothèse

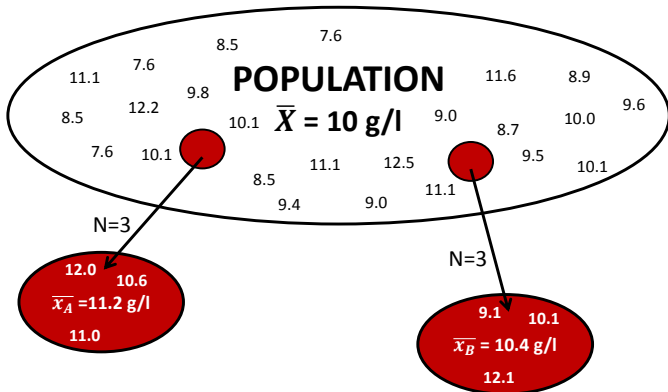


- Soit \mathcal{P} la population cible.
Ex : $\mathcal{P} =$ L'ensemble des patients anémiques.
- Soit μ le taux moyen d'hémoglobine
Ex : $\bar{X} = 10$ g/l.
- Il n'est pas possible de mesurer \bar{X} à partir de tous les patients de la population.
- On réalise un échantillon de N patients à partir desquels on observe une moyenne \bar{x} .
Ex : $N = 3$.
- Problème : si plusieurs échantillons sont réalisés, on observera autant de moyennes.

Rappels : fluctuation d'échantillonnage

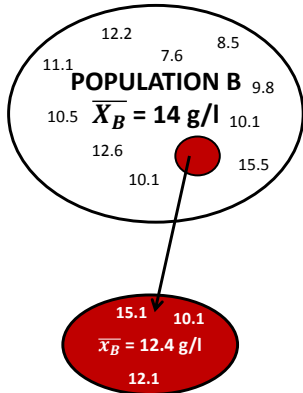
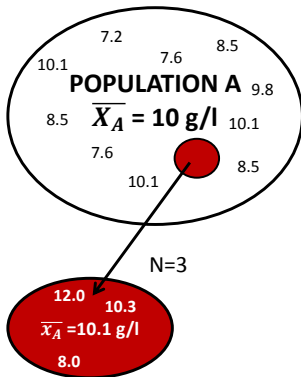
Les études
 expérimentales

Les tests
 d'hypothèse



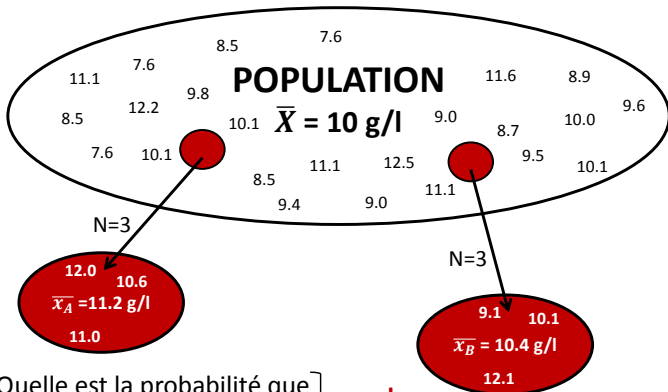
Deux hypothèses bien contrastées

Hypothèse alternative (H_1): le traitement a un effet sur le taux d'hémoglobine, c'est à dire $\bar{X}_A \neq \bar{X}_B$



Deux hypothèses bien contrastées

Hypothèse nulle (H_0): le traitement n'a pas d'effet sur le taux d'hémoglobine, c'est à dire $\bar{X}_A = \bar{X}_B = \bar{X}$



Quelle est la probabilité que cette hypothèse soit vraie ? } **p-value**

- J'observe une moyenne \bar{x}_A à partir d'un échantillon A et \bar{x}_B à partir d'un échantillon B (et leur variance).
- Est ce que la différence que j'observe est due à la fluctuation d'échantillonnage ?
- Deux hypothèses :
 - ① La différence entre les moyennes observées est minime. Elle est due au fait que trop peu de sujets de la population ont été inclus. Si tous les sujets des deux populations avaient été inclus, on aurait observé une égalité.
→ **Hypothèse nulle** : $H_0 (\bar{X}_A = \bar{X}_B)$.
 - ② La différence observée est importante. Elle ne peut pas être due au fait que tous les sujets n'ont pas été inclus. Il est évident que si on avait inclus tout le monde, les moyennes observées auraient été différentes.
→ **Hypothèse alternative** : $H_1 (\bar{X}_A \neq \bar{X}_B)$.

$$H_0 (\mu_A = \mu_B).$$

$$H_1 (\mu_A \neq \mu_B).$$

- Probabilité de se tromper si on rejette H_0 :
 - Risque de première espèce.
 - On l'estime à partir des tests statistiques : **probabilité critique**.
 - En anglais : *p-value ou type I error*
- Probabilité de se tromper si on rejette H_1 :
 - Risque de seconde espèce (*type II error*).
 - 1 moins la *puissance (power)*.

Les études
expérimentales

Les tests
d'hypothèse

- 1 Définition a priori d'un seuil de décision :
 - Risque de première espèce maximal accepté.
 - Ex : $\alpha = 0.05$
- 2 Calcul de la p-value.
 - Ex : comparaison de deux moyennes et grands échantillons : t-test.
- 3 Conclusions :
 - Si la p-value est trop importante, c.-à-d. supérieure à α : tendance au non-rejet de H_0 .
 - Si la p-value est faible, c.-à-d. inférieure à α : rejet de H_0 .

- Si la p-value est supérieure à 5%, **une conclusion du type "il semble qu'il n'ait pas d'effet du traitement" est fausse !**
 - Le risque d'erreur lié au rejet de H_1 est le risque de 2nd espèce.
 - Le test statistique ne donne pas cette probabilité.
 - Ex :
 - Quand les effectifs des deux échantillons \searrow alors la p-value \nearrow .
 - On rejette rarement H_0 pour des effectifs très faibles.
 - Cela ne veut pas dire qu'il n'y a pas de différence en réalité.
 - Manque de puissance.
 - Tests d'équivalence ou de non-infériorité.
- **Toujours inclure une notion d'incertitude car les populations ne sont jamais observées.**

Les études
expérimentales

Les tests
d'hypothèse

- Pour tous les tests d'inférence, il est nécessaire que les effectifs soient d'autant plus importants que :
 - la différence attendue et intéressante cliniquement est petite.
 - le critère de jugement est variable.
 - les risques de première et seconde espèces sont faibles.
- Les effectifs nécessaires dépendent aussi des tests statistiques utilisés.
- Les effectifs nécessaires sont calculés a priori.