

Essais cliniques

Delphine Maucort-Boulch, MD PhD
Service de Biostatistique, HCL
UMR CNRS 5558



Université Claude Bernard



Plan

1. Introduction
2. Protocole
 - ❑ Question posée
 - ❑ Echantillon
 - ❑ Taille de l'étude
3. Conduite et analyse
 - ❑ Conduite
 - ❑ Analyse statistique
4. Réglementation
5. L'essentiel

Plan

1. **Introduction**
2. **Protocole**
 - Question posée
 - Echantillon
 - Taille de l'étude
3. **Conduite et analyse**
 - Conduite
 - Analyse statistique
4. **Réglementation**
5. **L'essentiel**

Une démarche ancienne récemment structurée

- Ancien testament (Daniel I: 12-15)
- Avicenne (980-1087), *Canon de la Médecine*: 7 règles d'évaluation de l'effet d'une drogue sur une maladie
- 1747, James Lind soumet 6 groupes de marins atteints de scorbut à différents traitements
- 1948, Sir Bradford Hill: Streptomycine et tuberculose pulmonaire

Qu'est-ce qu'un essai?

- Expérimentation planifiée dessinée
 - pour identifier l'intervention appropriée à
 - de futurs patients présentant
 - une condition médicale donnée, dans laquelle
 - les résultats d'un groupe de patients traités par l'intervention évaluée sont comparés à
 - ceux d'un groupe de patients traités par une intervention contrôle,
 - les groupes étant inclus, traités et suivis durant la même période de temps
- Essais pour expliquer l'effet d'un traitement ('explicatifs'), ou pour prendre une décision clinique ('gestion')

Un essai clinique

■ Clinique

- Clinique < grec: kline=lit
- Essai clinique: expérience qui évalue un ‘traitement’ sur l’homme
- Chez l’homme, grande diversité et fluctuations inter-individus
- Comparaison de groupes (appelés bras) d’intervention

■ Vise à estimer l’effet

- D’un traitement médicamenteux ou chirurgical
- D’un traitement non médicamenteux, une prise en charge
- De stratégies thérapeutiques
- D’actions en santé
- D’examens diagnostiques

Une démarche standardisée

- Objectif: comparer un effet dans différents groupes d'intervention
- Contexte scientifique: fondamental pour justifier l'étude, poser la pertinence
- Question posée → étude avec un schéma adapté
- Tests d'hypothèses visant à comparer les grandeurs évaluées dans les 2 bras d'étude
- Comparaison de 2 interventions → Comparaison à un groupe référent, groupe « contrôle »
- Résultat imputable à la seule intervention (\neq fluctuations d'échantillonnage)
- Déroulement de l'essai, logistique
- Analyse statistique, interprétation, conclusion
- Communication

Un travail pluridisciplinaire

- **Investigateur**

Professionnel surveillant l'a réalisation de l'étude. Professionnel responsable de l'étude est l'investigateur principal

Investigateur coordinateur coordonne la réalisation de l'essai dans les essais multicentriques

- **Promoteur**

Personne physique ou morale prenant l'initiative de la réalisation d'une étude

- **Méthodologue**

- **Gestionnaire de données**

- **Attaché de recherche clinique**

Biostatisticien

- Participe à l'écriture du protocole, réalise les analyses et aide à l'interprétation des résultats, en interaction avec les intervenants
- Traduit une question biomédicale en question statistique, proposer des réponses statistiques
- Contribue à la rédaction du protocole, plan d'analyse statistique
- Programme et réalisation les analyses statistiques des données : descriptive, inférentielle, modélisation,...
- Rédaction des rapports d'analyse
- Valorisation des résultats (articles, communication dans des congrès)

Plan

1. Introduction
2. Protocole
 - ❑ Question posée
 - ❑ Echantillon
 - ❑ Taille de l'étude
3. Conduite et analyse
 - ❑ Conduite
 - ❑ Analyse statistique
4. Réglementation
5. L'essentiel

Question clinique

- Sa pertinence est déterminée par une revue exhaustive de la littérature
 - Rapports de cas, séries de cas, études observationnelles, essais contrôlés randomisés validés ou non, méta-analyse
- Contexte scientifique
- Confirmer un résultat d'essai (biais, validation), ou réaliser un essai innovant
 - Un essai ne peut répondre qu'à une unique question
 - Cette question va conditionner le schéma d'étude = plan expérimental
 - Plan expérimental consigné dans un protocole

Exemples

1. Est-ce que la fibrinolyse pré-hospitalière est meilleure que l'angioplastie primaire dans la prise en charge de l'infarctus du myocarde en phase aigue (Etude CAPTIM)?
2. La caféine améliore t'elle la prise en charge des apnées du prématûré en terme de survie, retard mental ou déficit sensoriel?
3. La radiothérapie étendue améliore t'elle la survie des patients atteint de tumeur cérébrale en post opératoire par rapport à la radiothérapie limitée?

Critère de jugement

- Un unique critère de jugement principal
- Pertinence clinique
- Défini précisément, sans ambiguïté
 - Critère(s) qui le compose(nt)
 - Le terme au bout duquel il est recueilli
 - Les modalités d'évaluation, standardisation
- Evaluation doit être facile, fiable (préalablement évaluée)
- Formation des cliniciens
- Porte in fine le résultat du test et sera à l'origine de la conclusion sur le traitement testé

Exemples

1. Décès, réinfarctus non fatal, accident vasculaire cérébral non fatal à 30 jours
2. Décès, paralysie cérébrale, retard mental (score <85 sur l'échelle de Bayley), surdité ou cécité à un âge corrigé de 18 à 21 mois
3. Décès

Les phases de développement d'un nouveau médicament

- Pré-clinique
 - Etude de la molécule sur l'animal
 - Détermination de la DL50
- Phase I
 - Etude de la tolérance (escalade de doses), et des effets secondaires chez volontaires sains (rémunération)
 - Pharmacocinétique et pharmacodynamie
 - Études de petite taille (20 à 80 participants)
- Phase II
 - Détermination de la dose optimale et étude des effets secondaires
 - Identification de la population de patients pouvant bénéficier du traitement
 - Études de taille modérée (100 à 200 participants)

Les phases de développement d'un nouveau médicament

- Phase III
 - Étude d'efficacité
 - Comparaison entre traitement testé et soit un placebo, soit un traitement de référence
 - Étude de grande taille
 - AMM possible à l'issue
- Phase IV
 - Post AMM
 - Pharmacovigilance (effets secondaires rares, complications tardives)

Randomisation

■ Pourquoi

- On souhaite que le résultat soit imputable à la seule intervention. Seul le traitement doit différencier les deux groupes
- On souhaite que le résultat ne soit pas dû aux fluctuations d'échantillonnage
- Distribution au hasard des caractéristiques (observables ou non) dans les groupes comparés
- L'allocation du bras de traitement par une personne ou une règle établie connue pourrait introduire des biais
- Suppose l'ambivalence des investigateurs et des patients

■ Allocation aléatoire

- Seul moyen de rendre les groupes comparables en tout point
- Utilisation des tests d'hypothèse (H_0 d'absence de différence entre les groupes)
- Chaque patient a la même chance de bénéficier du meilleur traitement

Randomisation

- Correspond à une/des table(s) de nombres au hasard
- Liste établie avant le début de l'essai
- Equilibre des effectifs entre les groupes
- Allocation du bras après vérification de la possible inclusion du patient
- Procédure centralisée préférable. Les investigateurs contactent le centre en charge de la randomisation pour chaque nouveau patient à inclure (interface web ++)

Tables de nombres aléatoires

Illustration

- Crédation d'un table de nombres aléatoires (0 et 99 ici). Chaque nombre a la même probabilité d'être tiré au sort
- Choix d'un sens de lecture (e.g. de gauche à droite, de haut en bas)
- Définition d'une règle a priori à la table: les nombres entre 0 et 49 alloueront le patient au bras A et les nombres entre 50 et 99 alloueront le patient au bras B

51	17	97	69	76
70	23	93	27	92
11	79	82	12	21
21	76	40	35	38

La séquence est donc ici
BABBBBABAABABAAABAAA.
L'équilibre est, en moyenne respecté,
entre les 2 bras.

Une question d'équilibre

- La randomisation assure répartition équilibrée des caractéristiques en moyenne quand $n \rightarrow \infty$
- Le hasard peut introduire des déséquilibres? en particulier lorsque n n'est pas très grand/
 - d'effectifs entre les bras
 - de caractéristiques entre les bras, problème si le déséquilibre touche un facteur de confusion
- Equilibre entre les bras au sein de chaque catégorie d'un facteur de confusion

Stratification

- Elaboration d'une liste de randomisation au sein de chaque catégorie/strate de la variable pour laquelle on souhaite un équilibre absolu
- Il y en résulte un équilibre entre les bras au sein de chaque strate
- Dans les essais multicentriques, il est recommandé de stratifié sur le centre pour tenir compte du recrutement différent possible entre les centres (pas plus de 2 ou 3 variables de stratification)
- Remarque: randomisation adaptative (à nouvelle inclusion, probabilité d'allocation dans un groupe en fonction de la distribution du facteur de confusion dans les bras)

Randomisation par blocs

- Maintenir la répartition choisie des effectifs (calcul du nombre de sujets nécessaires, condition du test,...), assurer un écart minimal d'effectif entre les bras
- Peut être combinée à une stratification
 - Si stratification: les blocs s'appliquent au sein de chaque strate
- Méthode: tirage de nombres aléatoires par un bloc puis ordonnancement des blocs
- Garantir un équilibre des effectifs entre les bras tous les m sujets inclus
 - Des blocs de 4 seront donc de la forme AABB, BBAA, ABAB, BABA, ABBA, BAAB

Insu

Insu = aveugle

- Evaluation du critère de jugement principal en aveugle pour éviter les biais d'information
- Eviter les comportements différents des patients et des médecins
- Double aveugle: patients et médecins ignorent le traitement reçu
- Simple aveugle: seuls les patients ignorent le traitement reçu
- Conserver la comparabilité des groupes jusqu'au bout
- Attention aux traitements: forme galénique, goût...
- Pas toujours possible et justifié (Traitements chirurgicaux, Traitements avec des effets secondaires caractéristiques, Essais en cancérologie)
- Si l'insu est impossible, maintien de la comparabilité des groupes par réajustement des traitements selon procédures standardisées

Groupes comparés

Schéma de comparaison

Schéma en groupes parallèles:

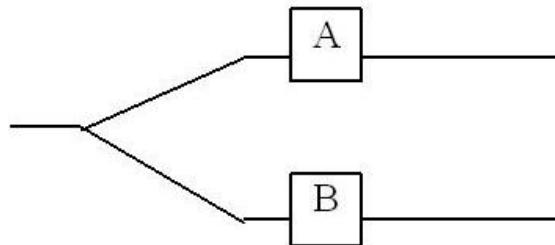
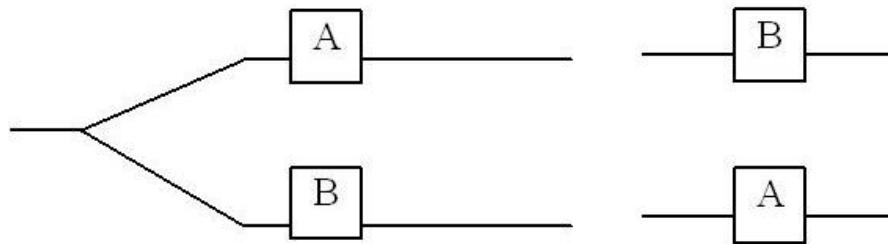


Schéma en cross over :



Plan factoriel :

	A	non A
B	Groupe 1	Groupe 2
non B	Groupe 3	Groupe 4

Essai contrôlé

■ Choix du groupe contrôle

- Placebo (substance inerte sur le plan pharmacologique) ou traitement de référence
- Placebo en l'absence de traitement de référence recommandé
- Définition précise des interventions pour une standardisation, en particulier dans les essais multicentriques

Pour les traitements: molécule, forme galénique, durée du traitement, posologie, horaires de prise, mode de conservation

Essai contrôlé

Choix du groupe contrôle

- Groupe historique ou groupe contemporain
 - les patients d'une époque ne sont pas comparables à ceux d'une autre époque (évolution dans le temps des modes de prise en charge d'une maladie, de la maladie elle-même). De nombreux facteurs autres que le traitement lui-même peuvent venir fausser son évaluation
 - Groupe contemporain évite les biais liés aux différences de recrutement dans le temps
- Exemple de groupe historique
 - Polyomyélite: Entre 1952 et 1953: ↓ incidence (÷2), 1954: programme de vaccination à grande échelle
 - Service de chirurgie: ↑ mortalité des patients cancéreux concomitante à arrivée d'un chirurgien de renom attirant les patients les plus graves

Inférence

- Populations et échantillon
 - On cherche à améliorer la prise en charge thérapeutique d'une population = population cible
 - Etude exhaustive impossible → échantillon
 - L'échantillon est tiré d'une population source
- Population cible ↔ Population source ↔ Echantillon
- Inférence
 - L'objectif d'un essai est à l'échelle d'une population
 - Inférer que les résultats observés sont vrais pour la population dont est issu l'échantillon (population source)
 - Extrapoler les résultats à la population pour laquelle on souhaite au final appliquer le traitement (population cible)
 - Echantillon représentatif de la population source

Critères d'éligibilité

Inclusion-Exclusion

- Critères d'inclusion

Caractéristiques de la population cible

→ Généralisation des résultats

- Critères d'exclusion (non inclusion)

Limitent l'inclusion en définissant des caractéristiques qui restreignent l'intérêt du traitement, son efficacité, une balance bénéfice-risque défavorable

→ Sécurité

Question statistique

Traduction

- Traduction de la question clinique en une question statistique
- Comparaison de grandeur estimée → réalisation d'un test statistique
 - Comparaison de fréquences
 - Comparaison de moyennes
 - Comparaison de survies
- Définition des hypothèses H_0 et H_1
- Risque α

Rappels

■ Hypothèse alternative H1, Différence minimale critique δ

- Nécessité de définir une hypothèse H1 précise correspondant à la différence minimale δ que l'on s'attend à observer
- Informations dans la littérature
- Définition d'un risque β ou d'une puissance ($1 - \beta$)

Rappels

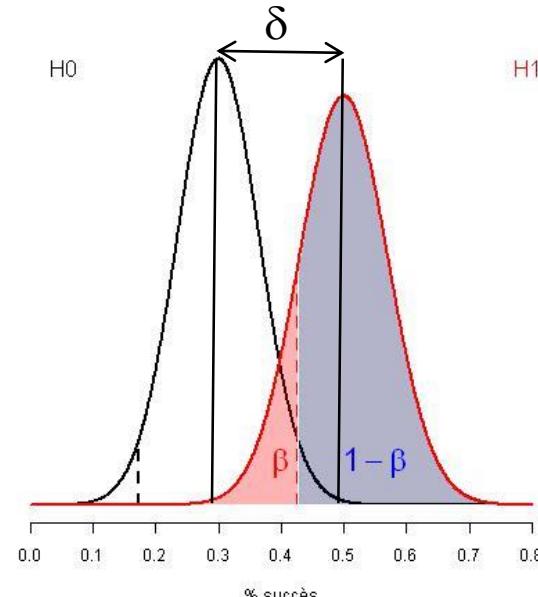
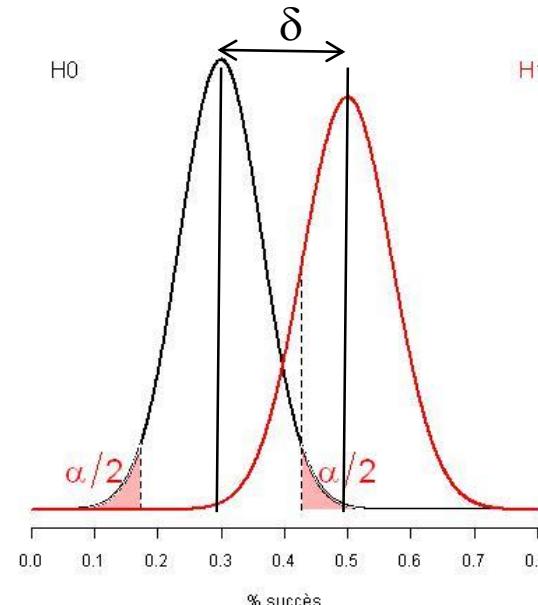
Risques consentis

Décision	Réalité H0 vraie	H0 fausse
H0 non rejetée	correct	β
H0 rejetée	α	$1 - \beta$

$\alpha = \Pr(|Z| \geq \text{seuil} | H_0 \text{ vraie})$ choisi a priori

\neq degré de signification $p = \Pr(|Z| \geq |z_0|)$

$\beta = \Pr(|Z| < \text{seuil} | H_0 \text{ fausse})$



Puissance

Paramètres qui interviennent

- Calcul du nombre de sujets dépend
 - De la nature du critère de jugement principal
 - Du test réalisé pour conclure
 - De la différence minimale cliniquement pertinente δ
 - Dispersion du critère de jugement principal
 - α, β , test uni ou bilatéral
- Il faut d'autant plus de sujets que
 - α est faible
 - β est faible
 - δ est petite
 - σ est grande

L'hypothèse H1 du calcul du nombre de sujets \neq l'hypothèse H1 du test d'analyse final

Exemples

CAPTIM, Lancet 2002

- Est-ce que la fibrinolyse pré-hospitalière est meilleure que l'angioplastie primaire dans la prise en charge de l'infarctus du myocarde en phase aiguë (Etude CAPTIM)?
- $P_{\text{fibrinolyse}} = 12.2\%$, $P_{\text{angioplastie}} = 7.2\%$
→ Différence d'efficacité minimale cliniquement pertinente $\delta = 5\%$ (absolue, 40% relative)
- Hypothèse H1 bilatérale, $\alpha = 0.05$, $\beta = 0.15$
→ 1200 sujets au total

Schmidt et al, NEJM, 2007

- La caféine améliore-t-elle la prise en charge des apnées du prématuré en terme de survie, retard mental ou déficit sensoriel?
- $P_{\text{caféine}} = 20\%$, Différence d'efficacité minimale relative = 25%
- Hypothèse H1 bilatérale, $\alpha = 0.05$, $1 - \beta = 80\%$
→ 1000 sujets dans chaque bras de traitement

Plan

1. Introduction
2. Protocole
 - ❑ Question posée
 - ❑ Echantillon
 - ❑ Taille de l'étude
3. Conduite et analyse
 - ❑ Conduite
 - ❑ Analyse statistique
4. Réglementation
5. L'essentiel

Surveillance

- Qualité des données
- Compliance des patients = cruciale, à estimer et rapporter
- Qualité du monitorage de l'essai
- Comité indépendant de surveillance, comité de validation des événements
 - Groupe d'experts évaluant la sécurité des patients et de l'efficacité du traitement lors d'analyses intermédiaires
 - Peut être amené à recommander l'arrêt de l'essai
 - Etudes internationales multicentriques pour lesquelles la surveillance peut s'avérer complexe

Données manquantes

- A envisager dès le protocole
- Perdus de vue → manque le critère de jugement principal pour certains participants
- Perdus de vue indépendamment de l'essai i.e. motifs identiques dans les 2 bras
 - biais de sélection non différentiel (perte de puissance)
- Perdus de vue pour une raison liée au traitement
 - biais de sélection différentiel (sus ou sous estimation de l'effet traitement)

Données manquantes

- Méthodes statistiques pour prendre en compte les données manquantes
- Méthode du biais maximum pour les critères de jugement dichotomiques

Remplacer les critères de jugement manquants par des succès dans un groupe, des échecs dans l'autre et analyser puis inversement. Si les conclusions sont concordantes alors on peut conclure de même

Analyse en Intention de Traiter

- Tous les participants à l'étude restent étudiés dans le groupe auquel ils ont été assignés. Aucune exclusion a posteriori
- Assure le maintien de la comparabilité des groupes
- Biais d'attrition (non comparabilité des groupes au moment du recueil du critère de jugement principal)
- Exclusion les patients qui n'ont pas commencé le traitement pour une raison indépendante du bras de randomisation = analyse en intention de traiter modifiée
- Exclusion de l'analyse des patients pour lesquels le protocole initial n'est pas totalement respecté = analyse per protocole

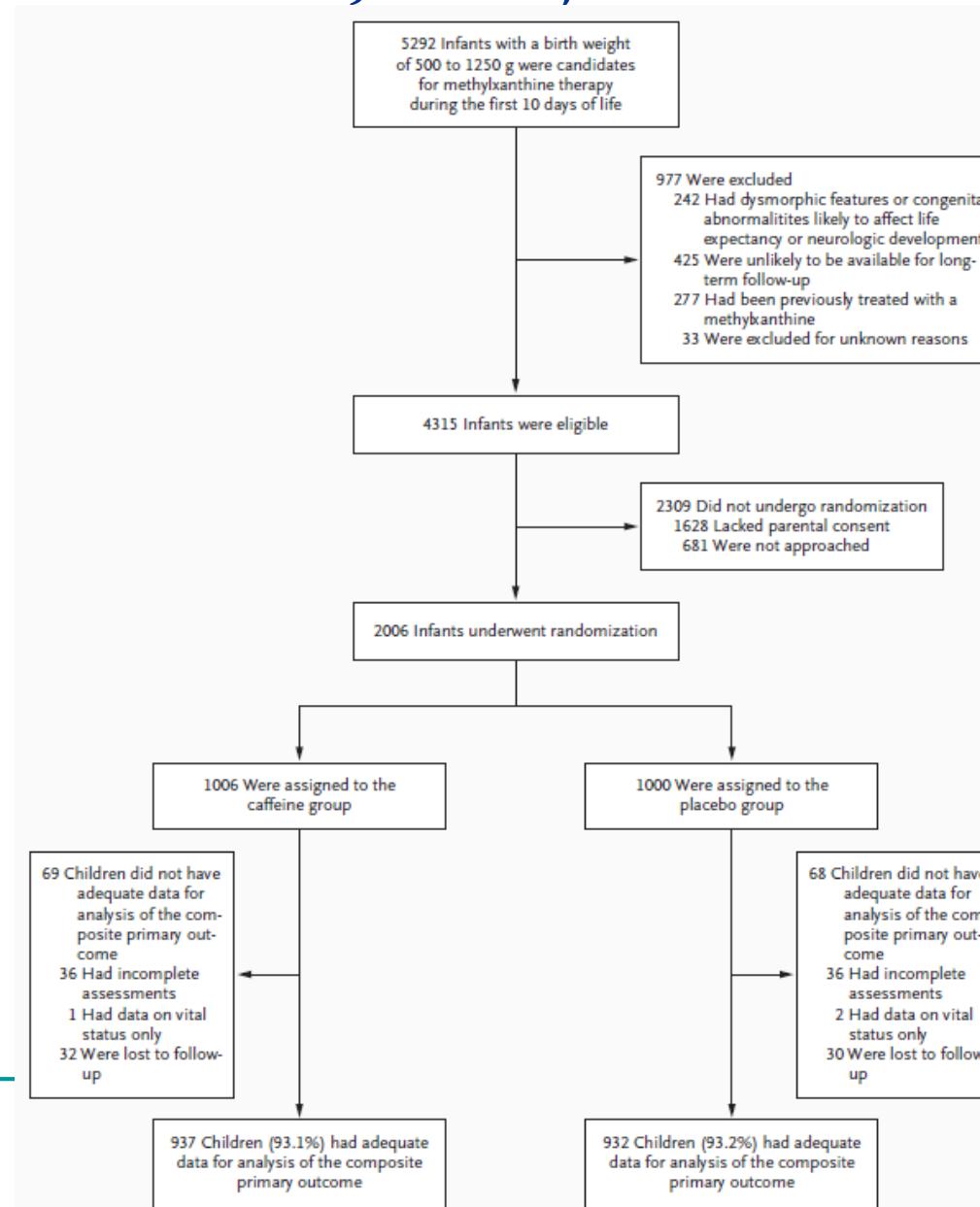
Analyses statistiques

- Analyses intermédiaires prévues dès le protocole
- Flow chart, description des caractéristiques globale et par les bras, données manquantes
- Risque relatif ($RR=R_1/R_0$), différence de risque ($DR=R_1-R_0$)
→ Nombre Nécessaire à Traiter ($NNT=1/|DR|$)
- Taux relatifs (critère de survie)
- Intervalles de confiance ++

Le rejet de l'hypothèse H_0 ne signifie pas l'acceptation de l'hypothèse H_1 utilisée pour le calcul du nombre de sujets nécessaire

Le non rejet de l'hypothèse H_0 ne signifie pas l'équivalence d'efficacité

Flow chart, Schmidt, NEJM 2007



Plan

1. Introduction
2. Protocole
 - ❑ Question posée
 - ❑ Echantillon
 - ❑ Taille de l'étude
3. Conduite et analyse
 - ❑ Conduite
 - ❑ Analyse statistique
4. Réglementation
5. L'essentiel

Lois

- Expérimentation humaine
- Déclaration Universelle des Droits de l'Homme (ONU 1978): liberté, sécurité, égalité, dignité humaine
- 4 grands principes
 1. Principe de respect de la personne humaine (consentement)
 2. Principe de bienfaisance (balance bénéfice-risques)
 3. Principe de justice
 4. Principe du respect de la dignité de la science et de ses propres exigences méthodologiques (Bonnes Pratiques, méthodologie de la recherche)

Lois

- Loi du 20 décembre 1988 ‘relative à la protection des personnes qui se prêtent à des recherches biomédicales’ (Loi Huriel-Serusclat)
- Loi Jardé, 2012 (promulguée 2016): passage en CPP pour les recherches impliquant la personne humaine de type interventionnelle
- Comités de Protection des Personnes (CPP)
- Autorisation de Agence nationale de sécurité du médicament et des produits de santé (ANSM)
- Base internationale des essais cliniques ClinicalTrials.gov

Plan

1. Introduction
2. Protocole
 - ❑ Question posée
 - ❑ Echantillon
 - ❑ Taille de l'étude
3. Conduite et analyse
 - ❑ Conduite
 - ❑ Analyse statistique
4. Réglementation
5. L'essentiel

L'essentiel

Comprendre

- Projet collectif impliquant de nombreuses interactions entre différents acteurs
- Tout est consigné dans un protocole
- Protocole et déroulement conditionnent les conclusions

Retenir

4 caractéristiques classiques

- Comparatif (oui/non)
- Contrôlé (oui/non)
- Randomisé (oui/non)
- En aveugle (oui/non)

L'hypothèse H1 du calcul du nombre de sujets n'est pas l'hypothèse alternative du test réalisé, n'est pas l'hypothèse en faveur de laquelle se fait la conclusion