

## LCA: Essai thérapeutique

Plan		
<ul style="list-style-type: none"> <li>- Objectif</li> <li>- Protocole</li> <li>- Recrutement</li> <li>- Randomisation</li> <li>- Evaluation</li> <li>- Analyse</li> <li>- Résultats</li> <li>- Biais</li> <li>- Validité</li> <li>- Niveau de preuve</li> </ul>		
Objectifs CNCI		
- cf liste des <a href="#">objectifs en LCA</a> du CNCI		
Recommandations	Mots-clés / Tiroirs	NPO / PMZ
<ul style="list-style-type: none"> <li>- Grille <a href="#">CONSORT</a></li> <li>- Grades <a href="#">ANAES 00</a></li> </ul>	<ul style="list-style-type: none"> <li>- Essai de supériorité / non-infériorité</li> <li>- Essai en parallèle / en cross-over</li> <li>- Nombre de sujets nécessaires (6)</li> <li>- Randomisation par bloc / stratifiée</li> <li>- Maintien du double aveugle (5)</li> <li>- Critère de jugement principal (6)</li> <li>- Intention de traiter / per-protocole</li> <li>- Analyse multi-variée = ajustement</li> <li>- NST(NNT) = 1 / RAR = 1 / (A-B)</li> <li>- Risques <math>\alpha</math> / risque <math>\beta</math> / puissance / « p »</li> <li>- Sélection / classement / confusion</li> <li>- Validité interne / validité externe (2)</li> </ul>	<ul style="list-style-type: none"> <li>- NPO d'analyser la tolérance</li> <li>- « p » sans intérêt dans la table 1</li> <li>- NPO paramètres de dispersion</li> <li>- Analyse en sous-groupe: limites</li> </ul>

Avertissement : L'usage de cette fiche est strictement personnel. Toute diffusion est donc interdite, sur support papier ou électronique. En cas de manquement, prepECN se réserve le droit d'annuler votre abonnement sans contrepartie et d'engager des poursuites conformément aux articles L-429-4 et 315-5 du Code de la Propriété Intellectuelle.



### - Objectif

#### Objectif principal

- Justification de l'essai
  - contexte / ampleur du problème / hypothèse / bénéfice attendu
- Objectif principal: précis / unique / cliniquement pertinent
  - !! Le critère de jugement principal doit reprendre l'objectif principal (vérifier !)

#### Essai de supériorité

- Objectif principal = « Tt A meilleur que Tt B ? »
- Type d'analyse: faire une analyse en intention de traiter
  - diminue la probabilité de mettre en évidence une différence
  - donc renforce la supériorité du Tt si le résultat est positif

#### Essai de non-infériorité

- Objectif principal = « Tt A est aussi bon que Tt B ? »

- Evite un essai de supériorité lorsque la différence attendue est faible (le NSN serait ↑)
- N'a d'intérêt que si le Tt expérimental est avantageux (moins cher, moins d'ES, etc)
- **Exemples:** AMM pour même classe thérapeutique ou médicament générique ++
- **Type d'analyse: faire une analyse en per-protocole eten intention de traiter**
  - augmente la probabilité de mettre en évidence la différence
  - et donc renforce la non-infériorité si le résultat est négatif
- **Marge conservatrice de non-infériorité = borne d'équivalence**
  - = valeur qu'il ne faut pas dépasser pour considérer le Tt comme non-inférieur
  - En pratique, il faut que l'ensemble de l'IC95 soit dans la borne d'équivalence +++

### !! Remarques

- **Savoir distinguer « essai de supériorité » et « essai de non-infériorité »**
  - → **regarder l'hypothèse utilisée pour le calcul du NSN**
    - si c'est une « différence attendue » c'est de la supériorité
    - si c'est une « borne d'équivalence », c'est de la non-infériorité
  - **!! Piège classique**
    - essai de supériorité non significatif interprété comme de l'équivalence...
    - non valable car il faut un NSN beaucoup plus important en non-infériorité
- **Ne pas confondre: « essai de non-infériorité » et « essai d'équivalence »**
  - Dans l'essai de non-infériorité le médicament peut être plus efficace
  - En pratique, NSN ↑ si équivalence (car zone plus étroite) donc aucun intérêt

### - Protocole

#### Remarque:

- Un essai contrôlé randomisé (ECR) peut être à visée thérapeutique ou diagnostique
- Le principe est le même (comparaison de 2 groupes) mais pas le protocole (cf [item LCA: Essai diagnostique](#))

#### Plan de l'essai

- **En parallèle**
  - Groupe d'intérêt et groupe contrôle distincts: comparaison en fin d'essai
- **En cross-over**
  - « chaque sujet est son propre témoin »: alternativement placebo et Tt à l'étude
  - **Conditions**
    - Maladie chronique et stable (sinon patient non comparable à lui-même)
    - Tt sans effet rémanent (après wash-out) / sans effet rebond
    - Absence « d'effet-période » : effet modifié si Tt reçu en 1er ou 2ème (ex: BZD)
  - **Avantages**
    - Elimine la variabilité intra-individuelle (patient est « son propre témoin »)
    - Donc diminue le nombre de sujets (NSN) pour une même puissance
  - **Période de lavement (« wash-out »)**
    - Période sans traitement au cross-over (en pratique: durée  $\geq 5 \times 1/2$  vie du Mdt)
    - Objectif: éliminer l'effet rémanent du Tt (= « carry-over effect »)

### - Recrutement

#### Critères de sélection

- **Critères d'inclusion = critères positifs**

- → caractéristiques que doivent présenter les patients pour être inclus
- Critères d'exclusion (ou de non-inclusion) = critères négatifs
  - → le patient ne peut être inclus s'il présente un de ces critères
  - Risques en cas de critères d'exclusion trop nombreux / restrictifs
    - Population non représentative = validité externe (applicabilité) mauvaise +++
    - Difficulté d'inclusion des sujets (donc nombre de sujets insuffisant)
    - !! ce n'est pas un biais de sélection tant que les groupes sont comparables

### Nombre de sujets nécessaires (NSN)

- Doit toujours être justifié à **priori** dans l'article +++
- Données nécessaires au calcul du nombre de sujets nécessaire (5)
  - $\alpha$  : risque de 1ère espèce (0.05 en général / !! ajustement nécessaire si plusieurs CJP)
  - $\beta$  et donc la puissance souhaitée ( $1 - \beta = 0.80$  en général)
  - $\Delta$  : différence escomptée entre les deux groupes (cliniquement significative)
  - $\sigma$  : écart-type ou variabilité inter-individuelle du critère de jugement principal
  - **P**: probabilité de survenue du critère principal dans le groupe contrôle
- !! Remarque: le nombre de sujet final tient aussi compte de (5)
  - Type d'étude: essai de supériorité ou essai de non-infériorité
  - Type de test statistique: caractère unilatéral ou bilatéral du test
  - Nombre de perdus de vue attendu (donc nombre inclus > NSN calculé)
  - Durée de l'étude: elle est fonction de P (car 1% à 1an mais 100% à 10ans..)
  - Limite de non-infériorité (à la place de  $\Delta$ ) si c'est un essai de non-infériorité

### Analyse du diagramme de flux (Flow Chart: Fig. 1) = 5 CASES, de haut en bas:

- 1 = population screenée: pas forcément éligible mais considérée pour l'étude
- 2 = population éligible: rechercher des critères d'exclusion / consentement
- 3 = population incluse (randomisée): population en intention de traiter
- 4 = population par bras après la randomisation / ceux recevant le traitement
- 5 = population analysée: doit être la même que la population incluse (ITT) +++
- !! Remarque: toujours vérifier
  - que la population analysée est égale à celle randomisée (analyse en ITT)
  - que la population incluse n'est pas trop différente de celle éligible
  - → sinon, signifie que critères trop restrictifs ou peu de consentements..

### Analyse des caractéristiques initiales (Table 1)

- Permet de vérifier la **comparabilité initiale** des groupes après randomisation
- Toujours chercher le test sur la comparabilité dans le texte: partie résultat ++
- En pratique: vérifier la comparabilité « à vue d'oeil » (!! ne pas sur-interpréter un p: **PMZ**)
- Vérifier surtout que les groupes sont comparables sur les facteurs de confusion probables
- !! Remarque sur les « petits p »
  - Peuvent figurer ou non dans la table 1, en pratique c'est débile car (2):
  - $\alpha = 0.05$ : par définition, 5 caractéristiques sur 100 significativement différentes
  - Echantillon important: différences statistiquement significatives mais pas cliniquement

### Période de « run-in »

- Phase pré-inclusion pendant laquelle les patients prennent un placebo (ex: 1S)

- Objectifs

- Repérer les patients « placebo-répondeurs » (pour les exclure)
- Sélectionner les observants (pour diminuer les perdus de vue)

- Randomisation

Définition

- = attribution **aléatoire** du patient à un groupe (« tirage au sort »)
- les modalités de la randomisation doivent être décrites dans l'article +++

- On distingue

- **Randomisation simple** = séquence aléatoire non contrôlée
- **Randomisation restreinte** = contrôle de certains facteurs

Objectifs

- Assurer la comparabilité a priori des deux groupes (identiques sauf Tt)
- Minimiser les biais de sélection (groupes homogènes)
- Assure une répartition équilibrée des facteurs de confusion (connus ou non)
- Permettre une analyse en intention de traiter / en sous-groupe (si stratifiée)

Méthodes de génération de séquences

- Randomisation par bloc

- Principe
  - Assurer une distribution équivalente des patients dans chaque groupe
  - Ex: « en bloc de 10'' = sur 10 patients, 5 dans contrôle et 5 dans Tt
- Objectifs
  - Permet de diminuer les fluctuations d'échantillonnage +++
  - Evite que les médecins ne devinent le groupe (**levée de l'aveugle**)
  - Ex: permet de s'affranchir d'un « **effet-temps** » : lié au moment d'inclusion
  - → BPCO avec les patients placebo inclus en été et tous les Tt en hiver..

- Randomisation stratifiée

- Principe
  - Répartition de l'échantillon en strates selon un critère (de confusion ++)
  - En pratique: autant de listes de randomisation que de strates
  - **Ne pas** multiplier les strates: sinon pas assez de patients = non significatif
- Objectifs
  - **Répartition équilibrée** des traitements dans chacune des strates
  - Rendre les groupes **comparables** vis à vis du facteur de stratification, avec
  - Facteur de stratification = facteur de confusion majeur (**effet-centre**, gravité..)
- **!! Remarque**
  - La stratification permet de renforcer la validité d'une analyse en sous-groupe
  - Mais en aucun cas ne lui donne la même validité que l'analyse principale +++
- **!! NPC stratification avec:**
  - **ajustement:** analyse multivariée ajustée sur les facteurs de confusion a posteriori
  - **appariement:** que pour étude cas-témoins / appariement au moment de l'inclusion

- Randomisation équilibrée

- Pour 1 patient inclus dans le bras placebo, 1 patient inclus dans le bras Tt
- → même nombre de patients dans chacun des groupes (≠ déséquilibre)

- **Randomisation déséquilibrée**
  - **Principe**
    - = nombre de patients dans chaque groupe différent
    - Ex: randomisation en 2/1 = 100 pour Tt et 50 pour placebo
  - **Objectifs**
    - Augmenter les données pour étude de la tolérance
    - Augmenter le recrutement si nombre important de données manquantes prévu
    - !! Ne permet pas d'augmenter la puissance (≠ étude épidémiologiques)
- **Randomisation par minimisation**
  - = à chaque nouveau patient, choix du groupe pour minimisation des biais
  - En pratique, surtout en cancérologie: pour avoir des groupe très similaires
- **Randomisation en cluster**
  - On randomise des groupes de patients (ex: centre) et non les patients eux-mêmes
  - Objectif = éviter la contamination entre les bras dans un même cluster (centre)
  - Exemple: tous les patients de l'hôpital A reçoivent le placebo et tous ceux de B l'ABT

### Méthodes d'assignation

- Doit respecter la clause d'ignorance (ou « assignation **secrète** »)
- = impossibilité pour l'évaluateur de déterminer la séquence de randomisation
- Exemple de biais d'assignation: enveloppe transparente, etc.
- **Assignation centralisée et secrète**
  - Génération de la séquence et assignation non contrôlée par évaluateur
  - Ex: serveur vocal: communique juste le groupe du patient (A ou B)

### Vérifier qu'une randomisation est bien faite

- = regarder le tableau 1 (caractéristiques initiales des patients) +++
- **Bonne randomisation si**
  - Aucune différence significative pour chacun des critères
  - Nombre de patient équivalent dans les deux groupes
- **!! Remarque:**
  - différence ok si elle porte sur une caractéristique rare (peu de patients)

### - Evaluation

#### Groupe contrôle

- Systématique pour essais de phase III (≠ phase II)
- **Deux types de contrôles possibles**
  - Tt de référence (ou Gold Standard si essai diagnostique)
  - Placebo (si aucun Tt connu: sinon ce n'est pas éthique !)
- **Objectif**
  - Minimiser les biais de confusion +++
  - Les deux groupes doivent être exactement identiques sauf pour le Tt
- **Conditions éthiques pour utiliser un placebo +++**
  - Il n'y a pas de traitement de référence
  - Il n'y a pas de perte de chance: les symptômes sont réversibles

#### Aveugle (insu)

- **Essai ouvert**

- médecin et sujet connaissent la nature du traitement
- Simple aveugle (ou simple insu)
  - Le sujet ne connaît pas la nature du traitement
  - Mais le médecin la connaît
- Double aveugle (ou double insu)
  - Ni la patient ni le médecin ne connaissent la nature du Tt
  - → permet d'éviter les biais de classement (= biais de mesure ou d'évaluation)
- Critères assurant le maintien de l'aveugle tout au long de l'essai(5)
  - Enveloppes d'allocation de traitement scellées
  - Séquences de traitement identiques (ex: bolus puis perfusion)
  - Présentation identique (même aspect des flacons ou des comprimés)
  - Validation des critères de jugement en aveugle par un **comité d'adjudication**
  - Analyse statistique par statisticiens en aveugle des groupes
  - Technique du double placebo
    - indispensable si impossible de faire des médicaments identiques
    - chaque patient reçoit deux « seringues/pilule » d'aspects identiques 2 à 2

### Critères de jugement

- Critère de jugement principal: doit toujours être: (6)
  - **unique** / simple / consensuel / standardisé / reproductible / cliniquement pertinent
  - « consensuel » = critère **validé** par des publications (définitions internationales)
  - **!! Remarque**
    - Un critère de jugement principal doit être évalué à **un moment** précis (ex: 12S)
    - S'il est évalué à 3 temps différents il y a 3 critères et donc biais (!!  $\alpha$  à ajuster)
    - Si on veut se ramener à un seul critère il faut faire une analyse de survie +++
- Critères de jugement secondaires
  - Comprend toujours la **tolérance**: effets indésirables (même si implicite !) (PMZ)
  - Ne pourront en aucun cas être mis au 1er plan de l'analyse +++
  - → une autre étude en mettant le critère II en critère principal sera nécessaire
  - **!! Ne doivent pas être trop nombreux sinon « biais des comparaisons multiples »**
    - «  $\alpha = 0.05$  » signifie qu'il y a 5% de résultats significatifs du seul fait du hasard
    - Plus on multiplie les tests statistiques, plus on en trouvera un positif par hasard
    - Correction possible par pondération du risque (correction de type Bonferroni)
- Critère « décès toute cause » toujours préférable à décès « spécifique » car
  - évite le biais de classement car difficulté à diagnostiquer une cause précise de décès
  - prend en compte les décès non spécifiques mais liés au Tt (ex: AVP sur somnolence)
  - ne diminue pas la probabilité de montrer une différence car randomisation +++
- Critère composite souvent préférable à un critère unique
  - **Avantages**
    - Augmente le nombre d'évènements et donc la puissance statistique
    - Correspond à la situation clinique réelle (ex: IDM, revascularisation..)
    - Permet de regrouper des facteurs d'efficacité et de tolérance
  - **Inconvénients**
    - Regroupe des évènements de gravité différente (ex: TVP et EP)

- Sous-estime la gravité d'une suite d'évènement (ex: TVP puis EP)
- Pour savoir si un critère composite est cliniquement pertinent
  - Regarder si la gravité des sous-critères est globalement comparable
  - Si « décès » inclus, il faut des complications graves (mais ça reste valable)

## - Analyse

Quel que soit le type d'étude, toujours 3 étapes dans l'analyse +++

- Analyse descriptive
  - Paramètres de position: moyenne / médiane (si valeurs extrêmes)
  - Paramètres de dispersion: écart-type / min-max ou Q1-Q3 (PMZ)
  - Survie: courbe de Kaplan-Meier (en escalier)
  - !! Il faut toujours donner les deux dans les résultats (et résumé)
- Analyse univariée
  - = comparaison des résultats sur les CJP/CJS entre les deux groupes
  - En général **suffisante dans ETR** car randomisation (sauf si stratification)
  - Si pourcentages = test du Chi-2 (Fischer si cas particulier)
  - Si moyennes = t-test de Student (Wilcoxon si cas particulier)
  - Si courbes de survie = test du log-rank
- Analyse multivariée
  - **Ajustée** sur les facteurs de confusion qui sortent dans l'analyse en sous-groupe
  - Si elle est faite: il faut que sa réalisation potentielle ait été annoncé avant l'étude
  - !! Remarque
    - En théorie, pas d'analyse multivariée (= ajustée) dans un essai thérapeutique
    - Car une randomisation bien faite permet la comparabilité des groupes +++

## Analyse en intention de traiter

- L'analyse statistique doit prendre en compte la **totalité** des patients randomisés au départ
- même si: interruption du Tt / Tt non pris / modification du rythme ou de la posologie
- Mode d'analyse le plus rigoureux car (3) +++
  - Evite le **biais d'attrition: conserve la comparabilité** initiale issue de la randomisation
  - Option pragmatique: se rapproche des conditions de la vie réelle (arrêt sur ES..)
  - Hypothèse du biais maximum: renforce le niveau de preuve d'une différence
- !! Seulement 2 situations où l'ITT n'est pas suffisante
  - Essais de non-infériorité (d'équivalence) et évaluation de la tolérance
  - → l'hypothèse du biais maximum est celle du per-protocole: faire les deux +++
- « ITT modifiée »
  - Ex: une ITT qui n'inclue en réalité que les patients ayant pris au moins 1 dose du Tt
  - Gestion des non inclus: utiliser la valeur initiale du CJP comme valeur finale d'analyse

## Analyse en sous groupe

- Consiste à comparer le Tt dans une partie seulement de l'effectif total (ex: patients >65ans)
- !! NPC: n'est **pas** une analyse multivariée mais univariée dans chacun des sous-groupes
- Limites de l'analyse en sous-groupe (3) +++
  - Perte de la **comparabilité** (biais de sélection) en l'absence de stratification préalable
  - Petits effectifs donc IC très larges et **puissance insuffisante**
  - **Comparaisons multiples** donc il faudrait ajuster (diminuer) le risque  $\alpha$

- Conclusion sur une analyse en sous-groupe
  - Résultats descriptifs et **exploratoires**: « il semble que... »
  - Mieux si **stratification**: répartition équilibrée donc sous-groupes comparables
  - Mais dans tous les cas ne permet **jamais** de trancher de façon définitive
  - → il faudra refaire une étude spécifique dans le sous-groupe considéré +++ (PMZ)
- Diagramme de Forrest
  - Façon la plus élégante de représenter une analyse en sous-groupe
  - ligne d'identité:  $RR = 1$  / si l'IC touche la ligne d'identité: non significatif

#### Analyse de survie +++

- = analyse **temporelle** du délai de survenue d'un critère principal binaire (décès, récurrence..)
- !! suppose que le risque de décès soit **constant** tout au long de l'étude ++
- Courbe de Kaplan-Meier: analyse descriptive / !! peut être « inversée » :  $1-S(t)$ 
  - !! à  $t_0$ , on doit retrouver la population en intention de traiter +++
  - Effectif exposé doit rester comparable et assez grand, sinon extrapolation impossible
- Test du Log-rank: analyse univariée: comparaison des surfaces sous les courbes
- Modèle de Cox: analyse multivariée: ajustement sur les facteurs de confusion

#### Analyse per protocole

- Ne prend en compte que les patients ayant effectivement suivi tous les points du protocole
- → seuls les patients ayant reçu le Tt attribué lors de la **randomisation** sont analysés +++
- Meilleure estimation de l'effet propre du Tt mais biais de sélection: perte de la comparabilité
- En pratique, per-protocole valable dans 2 situations
  - essai de **non-infériorité**: renforce la non-infériorité si absence de différence
  - étude de tolérance: permet de mieux mettre en évidence des effets secondaires

#### Analyse intermédiaire

- = analyse en cours d'étude / sur données partielles / peut interrompre l'étude si effet évident
- A pour effet d'augmenter le nombre de sujets nécessaire (puisque effet souhaité précoce)
- Permet d'interrompre le recrutement en cas d'effet ou ES ou d'absence d'effet évident
- En pratique, analyse intermédiaire valable (et même éthique !) si
  - Elle est **prévue** à l'avance dans le protocole de l'étude (et décrite)
  - Le risque  $\alpha$  (ou le p) est **ajusté** a priori: diminution de la valeur seuil +++

#### - Résultats

##### Risques statistiques

- Hypothèses
  - Soit  $H_0$  l'hypothèse nulle: il n'existe pas de différence entre les 2 groupes
  - Soit  $H_1$  l'hypothèse alternative: il existe une différence entre les 2 groupes
- Risque  $\alpha$  (risque de 1ère espèce)
  - $\alpha = p$  (rejet  $H_0 / H_0$ ) = probabilité de trouver une différence si elle n'existe pas
  - $\alpha$  augmente si le nombre de tests (= critères de jugement) augmente
  - En pratique, seuil de signification:  $\alpha < 5\%$
- Risque  $\beta$  (risque de 2nde espèce)
  - $\beta = p$  (rejet  $H_1/H_1$ ) = probabilité de ne pas trouver de différence si elle existe
  - En pratique,  $\beta = 20\%$  (selon la puissance désirée)
- Puissance =  $1-\beta$

- $1-\beta$  = probabilité de conclure à une différence entre les traitements si elle existe
- la puissance augmente lorsque le nombre de sujets augmente
- **p = degré de signification**
  - p = probabilité que la différence trouvée soit due entièrement au hasard
  - Remarque:  $p \neq \alpha$  : p est fixé de manière arbitraire à l'avance /  $\alpha$  dépend de l'étude

### Résultats

- **Réduction absolue du risque (RAR)**
  - A = survenue (%) du critère principal pour le placebo (ou Tt de référence)
  - B = survenue (%) du critère principal pour le Tt expérimental
  - **RAR = A-B** (!! aucun test de significativité associé au RAR)
- **Nombre de sujets à traiter (NNT) +++**
  - nombre de sujets à traiter pour éviter la survenue d'un événement
  - **NNT = 1 / RAR = 1 / (A-B)** (= number needed to treat: NNT)
  - ex: A = 20% et B = 10%: NNT =  $1/0.10 = 10$  patients
  - **!! Remarque**
    - Le NNT se calcule sur une durée donnée (ex: NNT à 1an sur la mortalité totale)
    - Donc il faut prendre la RAR correspondant à ce délai (décès à 1an et pas à 6M)
- **!! NPC avec réduction relative du risque (RRR)**
  - $RRR = (A-B) / A$  (!! aucun test de significativité associé au RRR non plus)

### Interprétation

- **Si pas de différence significative sur le critère principal, cela signifie:**
  - soit que le Tt est réellement moins efficace que placebo / Tt de référence
  - soit que la différence n'a pas pu être mise en évidence (= manque de puissance)
  - soit que la validité interne de l'étude est défectueuse: présence de biais
- **!! Cela ne signifie PAS que les deux traitements sont équivalents (PMZ)**
  - Pour cela il faut faire un essai de non-infériorité (ou un essai d'équivalence)

### - Biais

#### Biais de sélection

- **Définition:** concerne la comparabilité des groupes
- **Comprend:**
  - Biais de recrutement (type de maladie, type de centres)
  - Biais des perdus de vue
  - Biais d'attrition: les patients analysés ne sont pas les patients randomisés
  - Effet-centre: lié à la non représentativité du centre. Ex: trop spécialisé, etc.
- **Mesures préventives**
  - **Randomisation** correcte (stratifiée si analyse en sous-groupe / par bloc, etc)
  - Analyse en intention de traiter: ITT (biais d'attrition)
  - Etudes multicentriques (effet-centre dans toute étude monocentrique)

#### Biais de classement (= biais de mesure)

- **Définition:** concerne la mesure du critère de jugement principal
- **Comprend:**
  - Biais de suivi (ou biais de performance): suivi / observance  $\neq$  selon le groupe
  - Biais d'évaluation: toujours une part subjective même dans un critère « objectif »

- Biais de réalisation: soins prodigués de façon différente
- Mesures préventives
  - **Double aveugle**
  - Critère d'évaluation principal unique / standardisé / validé / reproductible
  - Comité d'adjudication / évaluation double / évaluateurs indépendants
  - Investigateurs différents pour diagnostic / randomisation / acte thérapeutique / suivi

### Biais de confusion

- **Définition:** concerne l'existence de facteurs de confusion
  - !! Ne sera un biais que si la distribution des facteurs diffère entre les 2 groupes
- **Comprend:**
  - effet placebo du médicament
  - caractéristiques intrinsèques des patients (âge, sexe, atcd, etc)
  - effet non médicamenteux de la prise en charge
  - évolution naturelle de la maladie
- Mesures préventives
  - **Groupe contrôle**
  - **A priori:** stratification / restriction / randomisation / (appariement si cas-témoin)
  - **A posteriori:** ajustement par l'analyse multivariée

### - Validité de l'étude

#### Validité interne

- **Définition:** validité méthodologique des résultats obtenus
- **Evaluation = vérifier l'absence de biais**
  - **Biais de sélection:** randomisation bien faite / perdus de vue / analyse en ITT
  - **Biais de classement:** double aveugle / critère principal unique, standard..
  - **Biais de confusion:** groupes comparables / stratification..

#### Validité externe; comprend

- **l'applicabilité (ou généralisabilité) = pertinence clinique**
  - **définition:** généralisabilité des résultats à la population « réelle »
  - **évaluation**
    - **critères d'inclusion:** patients **représentatifs** / pas trop restrictifs
    - **taux de refus:** limité / !! doit figurer dans le diagramme de flux
    - **centres:** multicentrique / mixtes / similaires à la pratique courante
    - **critère d'évaluation:** pertinence clinique (objectif thérapeutique)
    - **groupe contrôle:** existence d'autres Tt que celui choisi / placebo justifié ?
    - **résultats:** taille du bénéfice pertinente / rapport bénéfice-coût-tolérance
- **la cohérence externe**
  - **définition:** concordance des résultats avec les connaissances préalables
  - **évaluation**
    - Résultat confirmé par au moins un autre essai
    - Cohérence avec les données physiopathologiques
    - Cohérence avec une éventuelle méta-analyse

### - Niveau de preuve (ANAES oo)

Niveau de preuve	Type d'étude	Grade de recommandation
Niveau 1	<ul style="list-style-type: none"> <li>▪ Essais comparatifs randomisés de forte puissance</li> <li>▪ Méta-analyse d'essais comparatifs randomisés</li> <li>▪ Analyse de décision basée sur les études bien menées</li> </ul>	<p style="text-align: center;">A Preuve scientifique établie</p>
Niveau 2	<ul style="list-style-type: none"> <li>▪ Essais comparatifs randomisés de faible puissance</li> <li>▪ Études comparatives non randomisées bien menées</li> <li>▪ Études de cohorte</li> </ul>	<p style="text-align: center;">B Présomption scientifique</p>
Niveau 3	<ul style="list-style-type: none"> <li>▪ Etudes cas-témoins</li> </ul>	<p style="text-align: center;">C Faible niveau de preuve scientifique</p>
Niveau 4	<ul style="list-style-type: none"> <li>▪ Études comparatives comportant des biais importants</li> <li>▪ Études rétrospectives</li> <li>▪ Études épidémiologiques descriptives</li> <li>▪ Série de cas</li> </ul>	
non défini	<ul style="list-style-type: none"> <li>▪ Opinions d'experts</li> <li>▪ Accords professionnels</li> </ul>	non défini