

Glossaire des termes utiles à la lecture critique d'un article médical

Ajustement	4
Analyse de sous-groupe.....	5
Analyse de survie	5
Analyse en intention de traiter	6
Analyse en per-protocole	7
Analyse multivariée.....	8
Appariement.....	8
Aveugle.....	8
Biais	9
Biais d'attrition.....	10
Biais d'avance au diagnostic.....	11
Biais d'évolutivité	11
Biais d'incorporation	11
Biais d'indication	12
Biais de classement.....	12
Biais de confusion.....	13
Biais de sélection.....	13
Biais de surdiagnostic.....	14
Biais lié « aux travailleurs sains » (healthy worker effect).....	14
Biais protopathique.....	15
Clause d'ambivalence	15
Cohorte.....	16
Critère de jugement	16
Critère de substitution.....	17
Echantillon	17
Enquête cas-témoins.....	18

Enquête de cohorte	18
Enquête épidémiologique.....	19
Enquête exposés / non exposés	19
Enquête transversale.....	20
Equivalence.....	20
Essai clinique sur groupes parallèles	21
Essai clinique en cross-over	22
Etude contrôlée	22
Etude pronostique.....	22
Facteur de confusion	23
Facteur de risque	24
Facteur protecteur	24
Fiabilité d'un test.....	24
Incidence	24
Intervalle de confiance	25
Jugement de causalité	26
Jugement de signification.....	27
Moment du tirage au sort	28
Multicentrique.....	28
Niveau de preuve.....	28
Nombre de sujets nécessaire	29
Odds ratio (rapport de côtes).....	30
Ouvert (essai ouvert).....	31
Placebo.....	32
Population cible.....	32
Population source	32
Prévalence	33
Principe d'incertitude	33
Propriétés extrinsèques d'un test	34
Propriétés intrinsèques d'un test	34

Puissance statistique.....	34
Rapports de vraisemblance	34
Reproductibilité	35
Randomisation	35
Risque alpha, risque beta	35
Risque relatif	36
Run-in	39
Sensibilité	39
Spécificité	40
Stratification.....	40
Tirage au sort (randomisation).....	40
Valeur prédictive	41
Références	42

Ajustement

Technique d'analyse statistique visant à contrôler a posteriori l'effet d'un facteur de confusion dans la relation entre un facteur de risque et une maladie, dans les études épidémiologiques. Elle permet d'estimer un odds ratio ajusté ou un risque relatif ajusté sur le facteur de confusion (au lieu de l'odds ratio brut ou du risque relatif brut ne prenant pas en compte l'effet du facteur de confusion).

Lorsque l'ajustement porte sur un seul facteur de confusion, on parle d'analyse bivariée (exemple ci-dessous). Lorsque l'ajustement porte sur plusieurs facteurs de confusion simultanément, on parle d'analyse multivariée.

Exemple : Estimation du risque relatif de cancer du poumon chez les hommes ajusté

	K=oui	K=non	
Homme	80	120	$R_{K/H} = 80 / (80+120) =$ 0,40
Femme	50	150	

$$R_{K/F} = 50 / (50+150) =$$

sur le tabagisme

Fumeurs :

	K=oui	K=non	
Homme	70	30	$R_{K/H} = 70 / (70+30) = 0,70$ $R_{K/F} = 35 / (35+15) = 0,70$ RR brut = 0,70 / 0,70 = 1,00
Femme	35	15	

Non-fumeurs :

	K=oui	K=non	
Homme	10	90	$R_{K/H} = 10 / (10+90) = 0,10$ $R_{K/F} = 15 / (15+135) = 0,10$ RR brut = 0,10 / 0,10 = 1,00
Femme	15	135	

Au total, RR brut = 1,60 et RR ajusté sur tabagisme = 1,0

(voir biais, facteur de confusion, analyse multivariée, odds ratio, risque relatif)

Analyse de sous-groupe

L'analyse de sous-groupe est une analyse secondaire a posteriori d'un essai clinique qui consiste à diviser la population de l'essai en 2 ou plusieurs sous-groupes définis par des caractéristiques (facteurs de risque) des patients pour étudier l'efficacité du traitement dans chacun de ces sous-groupes.

Sauf lorsqu'elle est planifiée à l'avance dans le protocole (donc qu'elle résulte d'une hypothèse formulée a priori), l'analyse de sous-groupes doit être évitée. Ses principaux inconvénients sont :

- un nombre important de tests statistiques. La multiplication des tests statistiques entraîne une inflation du risque alpha (risque de conclure à tort à une différence alors qu'elle n'existe pas) ;
- des effectifs de patients limités dans certains sous-groupes, à l'origine d'un défaut de puissance des tests statistiques.

L'analyse de sous-groupe nécessite l'utilisation de tests statistiques particuliers (test d'interaction).

(voir jugement de signification, risque alpha et risque beta)

Analyse de survie

Méthodes statistiques qui permettent d'analyser le délai de survenue d'un événement (le décès le plus souvent) au cours du temps chez les patients. L'analyse de survie prend en compte la durée de suivi propre à chaque patient. La spécificité des données de survie est l'existence d'observations « censurées » caractérisées par la durée de suivi des sujets et leur statut vital (vivant) à la date des dernières

nouvelles ou à la date de point de l'étude. Grâce aux méthodes d'analyse de survie, il n'est pas nécessaire d'attendre que tous les sujets soient décédés pour estimer la survie d'une pathologie.

La fonction de survie d'un groupe de patients est estimée par la méthode de Kaplan Meier qui permet d'établir une courbe de survie (fonction du temps commençant à 100% à t0 et représentant le pourcentage de sujets en vie à différents temps de suivi). Il existe des tests de comparaison de courbes de survie univariés (test du log-rank) et multivariés (modèle de Cox).

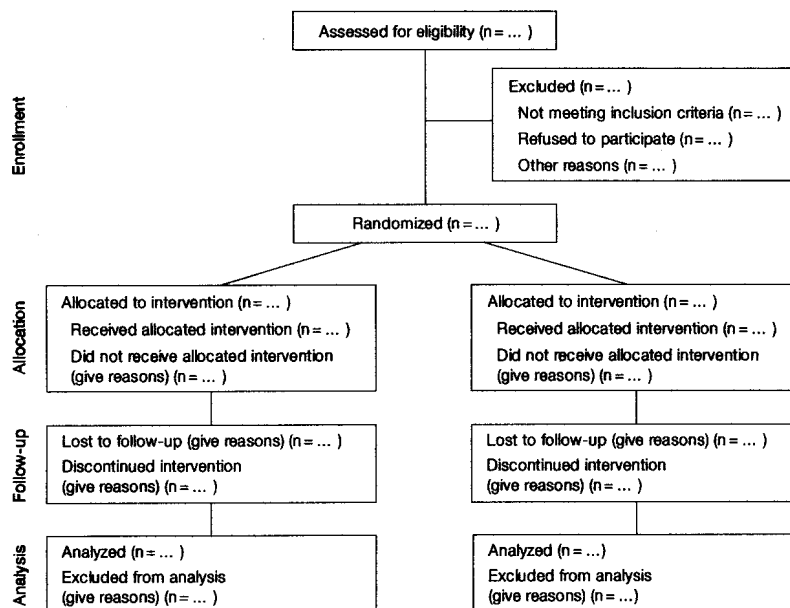
(voir analyse multivariée)

Analyse en intention de traiter

L'analyse en intention de traiter consiste à analyser tous les patients inclus dans l'essai en fonction du groupe dans lequel ils ont été randomisés, quel que soit le traitement qu'ils aient effectivement reçu. Cela signifie que les patients ayant arrêté le traitement en cours d'essai ou ayant reçu un autre traitement sont néanmoins analysés dans leur groupe d'allocation initial. L'analyse en intention de traiter est la seule méthode d'analyse susceptible de répondre à la question de l'essai sans biais. Elle doit donc être vérifiée en répondant à la question suivante :

« Tous les patients inclus dans l'essai ont-ils été analysés dans le groupe dans lequel ils ont été initialement randomisés, quel que soit le traitement effectivement reçu ? »

et en s'aidant du diagramme proposé par le Consort Group :



(voir essai clinique, analyse per-protocole, biais d'attrition)

Analyse en per-protocole

L'analyse en per-protocole d'un essai consiste à analyser seulement les patients ayant effectivement reçu le traitement correspondant au groupe dans lequel ils ont été randomisés. Cela signifie que les patients ayant arrêté le traitement en cours d'essai ou ayant reçu un autre traitement sont exclus de l'analyse. L'analyse per-protocole devrait être évitée car elle tend à biaiser les résultats de l'essai en détruisant la comparabilité initiale des groupes et en occultant le fait que les arrêts de traitement en cours d'essais peuvent être liés au traitement testé (effets indésirables).

(voir essai clinique, analyse en intention de traiter, biais d'attrition)

Analyse multivariée

L'analyse multivariée est une méthode d'ajustement a posteriori sur plusieurs facteurs de confusion simultanément, dans les études épidémiologiques. Elle est réalisée par régression logistique pour estimer un odds ratio ajusté ou par modèle de Cox pour estimer un risque relatif ajusté dans une analyse de survie.

(voir ajustement, analyse de survie, biais, facteur de confusion, odds ratio, risque relatif).

Appariement

Technique qui consiste à contrôler a priori un (ou plusieurs) facteur de confusion en formant des paires homogènes quant à ce facteur de confusion, dans une étude épidémiologique. Si on considère le tabac comme facteur de confusion dans une étude cas-témoins, on appariera à chaque cas fumeur un témoin fumeur.

Inversement, on appariera à chaque cas non-fumeur un témoin non-fumeur. Pour accroître la puissance de l'étude on peut appairer chaque cas à plusieurs témoins (appariement individuel 1 cas pour 2 témoins, par exemple). L'appariement nécessite l'utilisation de tests statistiques particuliers pour séries appariées.

(voir biais, facteur de confusion, stratification)

Aveugle

Pour pouvoir porter un jugement de causalité dans un essai clinique, il est nécessaire d'assurer la comparabilité des deux groupes tout au long de l'essai clinique jusqu'au moment de l'analyse des résultats. Seul le double aveugle (ou double insu) peut maintenir la comparabilité des groupes : il est essentiel que l'administration des traitements, le suivi des malades, et l'évaluation du critère de jugement soient faits de façon identique dans les deux groupes.

L'essai est réalisé

- en double aveugle si le malade et le médecin ignorent tous deux la nature du traitement administré ;
- en simple aveugle, si seul le malade ignore la nature de son traitement.

La réalisation d'un essai en aveugle nécessite que les traitements soient indiscernables.

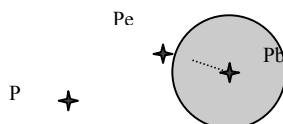
S'il s'avère impossible de réaliser une administration des traitements en aveugle, il est primordial que l'évaluation du critère de jugement soit effectuée en aveugle : l'évaluation du critère de jugement est réalisée par une tierce personne qui ignore le traitement administré à chaque patient.

(voir essai clinique, placebo, ouvert)

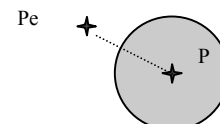
Biais

Un biais désigne une erreur systématique entre l'estimation d'un paramètre (prévalence, incidence, risque relatif, odds ratio...) sur un échantillon et la valeur de ce paramètre dans la population. Le biais est à distinguer des erreurs aléatoires (fluctuations d'échantillonnage) qui correspondent à un défaut de précision des estimations.

Biais
La prévalence observée sur l'échantillon fluctue autour d'une valeur biaisée P_b , différente de la prévalence P dans la population (erreur de visée).



Fluctuations d'échantillonnage
l'estimation P_e observée sur un échantillon s'écarte de la prévalence P dans la population (la zone grisée est celle où P_e a le plus de chance de se trouver = imprécision autour du point visé)



Dans les études épidémiologiques, on distingue trois types de biais :

- Biais de classement (voir _)
- Biais de sélection (voir _)
- Biais de confusion (voir _).

La plupart des biais, peuvent être classés dans une de ces catégories. Néanmoins, d'autres types de biais sont spécifiques des études évaluant les performances d'un test diagnostique (biais d'incorporation), des évaluations des campagnes de dépistage (biais d'avance au diagnostic, biais de surdiagnostic, biais d'évolutivité).

Biais d'attrition

Dans les essais clinique, le biais d'attrition désigne le fait que des patients randomisés aient été écartés de l'analyse. Les patients inclus mais non analysés correspondent soit à des perdus de vue, soit à des données manquantes, ce qui rend dans les deux cas le critère de jugement manquant. Pour repérer l'existence éventuelle d'un biais d'attrition, il faut se poser les questions suivantes :

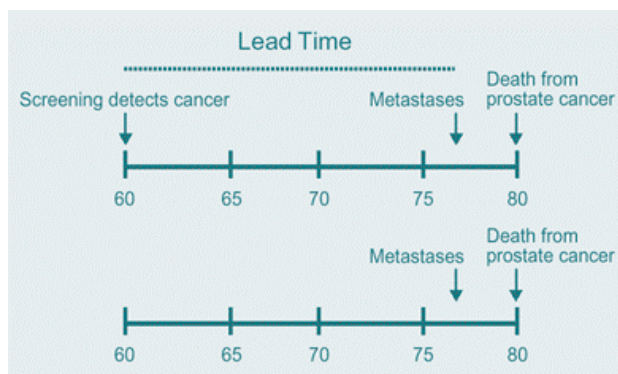
- le nombre de patients analysés est-il égal au nombre de patients randomisés ?
- quelle est la robustesse du résultat vis-à-vis de l'hypothèse du biais maximum (si dans un essai le critère de jugement principal est le décès, l'hypothèse du biais maximum sera : décès chez tous les sujets non analysés du groupe traitement et survie de tous les sujets non analysés du groupe contrôle) ?
- les auteurs ont-ils utilisé une méthode de remplacement des données manquantes ? Dans ce cas, le nombre de patients analysés est égal au nombre de patients randomisés, mais ces méthodes font des hypothèses fortes sur la nature des données manquantes.

(voir biais, intention de traiter, essai clinique)

Biais d'avance au diagnostic

Dans l'évaluation d'une campagne de dépistage, la réalisation d'un dépistage peut être associée à une augmentation apparente, mais artificielle de la survie à l'affection dépistée. Le délai entre le dépistage et le début des symptômes (intervalle latent) majore artificiellement la durée de survie.

Exemple : Schématisation du biais d'avance au diagnostic dans une campagne de dépistage du cancer de la prostate



Biais d'évolutivité

Les campagnes de dépistage identifient les formes d'évolution plus lentes de la maladie, dont la prévalence est plus élevée dans la population. La sélection de ces formes d'évolution lente peut donner l'impression d'une efficacité de la campagne de dépistage.

Biais d'incorporation

Dans l'évaluation des propriétés d'un test diagnostique, il existe un biais d'incorporation lorsque le résultat du test diagnostique évalué est inclus dans le résultat donné par la méthode de référence. Cela survient notamment lorsque la méthode de référence repose sur plusieurs examens du fait de l'inexistence d'un examen de référence unique.

Biais d'indication

Dans les enquêtes épidémiologiques d'observation, le biais d'indication désigne le fait qu'une intervention ait été prescrite plus souvent chez les sujets à risque élevé de maladie que chez les sujets à faible risque. Il s'agit d'un cas particulier de biais de confusion. Dans une enquête épidémiologique transversale conduite en milieu hospitalier, la prévalence des thromboses veineuses profondes sera plus élevée chez les sujets bénéficiant d'une prévention par héparine de bas poids moléculaire que chez les sujets sans prophylaxie : cela résulte du fait que les praticiens prescrivent préférentiellement une prophylaxie chez les sujets à risque élevé de thromboses veineuses profondes et non pas de façon aléatoire.

Biais de classement

Dans les études épidémiologiques, le biais de classement désigne une erreur systématique de mesure de l'exposition ou de détermination de la maladie. Elles conduisent à mal classer les sujets en « malades / non malades » ou en « exposés / non exposés ». On distingue :

- les erreurs de classement différentielles : on parle d'erreurs différentielles lorsque les erreurs sur l'information recueillie sur l'exposition antérieure diffèrent entre les cas et les témoins (biais d'enquêteur ou de subjectivité, biais de mémorisation, biais de prévarication). Elles conduisent généralement à une **surestimation de la mesure d'association entre l'exposition et la maladie**. Le biais de classement différentiel peut être réduit par l'interrogatoire à l'aveugle des sujets et par l'utilisation de questionnaires standardisés. De même, dans les enquêtes

exposés non-exposés, l'interprétation des examens de suivi (imagerie, anatomo-pathologie) devrait être réalisée en aveugle du statut du sujet.

- les erreurs de classement non-différentielles : la probabilité d'erreur ne diffère pas entre les deux groupes (erreur systématique de mesure due à un appareil mal réglé). Elles tendent à **sous-estimer la force d'association entre l'exposition et la maladie.**

Biais de confusion

Dans les enquêtes épidémiologiques, un biais de confusion désigne une erreur systématique de l'estimation d'une mesure d'association (odds ratio ou risque relatif) entre le facteur étudié et la maladie, du fait de l'existence d'un facteur de confusion.

Par opposition au biais de sélection et au biais de classement, il est possible de prendre en compte a posteriori un biais de confusion au moment de l'analyse des données.

(voir biais, facteur de confusion, ajustement, appariement, stratification, analyse multivariée)

Biais de sélection

Dans les enquêtes épidémiologiques, les biais de sélection affectent la constitution de l'échantillon d'enquête, c'est à dire le processus par lequel les sujets sont choisis au sein de la population. Dans la plupart des cas, il n'existe pas de liste de l'ensemble des sujets de la population au sein de laquelle l'échantillon serait constitué par tirage au sort. On doit donc constituer cet échantillon par d'autres moyens qui impliquent qu'il est issu d'une partie (une sélection) de la population qui fait l'objet de l'étude. Par exemple, l'identification des facteurs de risque de maladies

est souvent réalisée chez des patients hospitalisés. Les patients non hospitalisés font bien partie de la population cible de l'étude mais ne feront pas partie de l'échantillon. Ce phénomène se répétera à l'identique pour tous les échantillons : il s'agit donc bien d'un biais et non pas de fluctuations d'échantillonnage. Le biais d'auto-sélection désigne le biais de sélection introduit par le recrutement de sujets sur la base du volontariat. La motivation de ces sujets à participer à l'étude peut être liée à l'existence de facteurs de risque. De même, la participation à une campagne de dépistage résulte en partie de motivations liées à la présence de facteurs de risque (biais de volontarisme).

(voir biais, biais lié aux travailleurs sains)

Biais de surdiagnostic

Dans l'évaluation des campagnes de dépistage d'un cancer, le biais de surdiagnostic désigne l'accroissement du nombre de patients chez qui un cancer a été identifié sans pour autant que cette connaissance de la maladie ne modifie la survie des patients, soumis à d'autres risques. Le dépistage du cancer de la prostate conduit à une forte augmentation de son incidence apparente, mais ce surdiagnostic porte en partie sur des cas qui seraient restés latents, sans affecter la vie des personnes.

Biais lié « aux travailleurs sains » (healthy worker effect)

Il s'agit d'un cas particulier de biais de sélection. Le biais lié « aux travailleurs sains » (healthy worker effect) désigne la réduction de risque de décès chez les travailleurs comparativement à celui estimé en population générale. En effet, une population de travailleurs a en moyenne un meilleur état de santé que la population générale. Cela résulte de mécanismes de sélection (inaptitude au poste à l'embauche ou

changement de poste, voire départ de l'entreprise après l'embauche) ou d'auto-sélection (décision des sujets eux-mêmes).

(voir biais, biais de sélection)

Biais protopathique

Dans les enquêtes épidémiologiques de type cas-témoin ou de type transversal, le biais protopathique est à craindre lorsqu'il est difficile d'établir que l'exposition au facteur étudié a bien précédé la survenue de la pathologie. Dans les études cas-témoins en pharmacovigilance, le facteur suspecté (médicament) peut avoir été prescrit du fait des symptômes correspondant au début de l'affection étudiée :

« l'exposition » n'est en fait qu'un marqueur précoce (ou la conséquence) du début de la maladie et non pas la cause.

(voir biais, enquête cas-témoin)

Clause d'ambivalence

Les malades inclus dans un essai clinique doivent pouvoir recevoir indifféremment l'un ou l'autre des traitements. Dans un essai comparant un traitement chirurgical à un traitement médical, tous les malades inclus dans l'essai doivent être susceptibles de recevoir l'intervention chirurgicale. Il ne faut donc pas inclure dans l'essai les malades qui ont une contre-indication chirurgicale ou qui refuseraient l'intervention.

De même, les malades requérant une intervention chirurgicale ou qui présentent une contre-indication au traitement médical ne doivent pas être inclus dans l'essai. Le respect de la clause d'ambivalence évite d'exclure secondairement après randomisation des sujets dont on sait qu'ils ne pourront pas recevoir l'un des deux

traitements. Exclure des sujets après tirage au sort peut nuire à la comparabilité des groupes et introduire des biais dans la comparaison.

Cohorte

Une cohorte désigne un groupe de sujets ayant une caractéristique commune suivi au cours du temps. Il peut s'agir d'une génération, définie par une même année (ou période) de naissance, d'un corps de métier (les salariés d'une usine, cohorte des médecins britanniques), de personnes ayant subi une exposition particulière (sujets greffés rénaux, personnes irradiées à Nagasaki), de sujets porteurs d'une caractéristique génétique (trisomie 21), ou de sujets réunis sur une base géographique (cohorte des habitants de Framingham suivis depuis 1948).

Critère de jugement

La comparaison de deux groupes dans un essai clinique passe par le choix de critères, toujours discutables, mesurant l'efficacité et la tolérance. Un critère de jugement doit :

- être pertinent cliniquement ;
- faire l'objet d'un consensus de la communauté médicale ;
- être disponible chez tous les sujets ;
- être acceptable pour les patients (pénibilité) ;
- être de mesure simple ;
- être reproductible (fiable) ;
- être sensible (apte à détecter de petites différences) ;

- être objectif (dosage biologique fourni par un automate par exemple) mais beaucoup de critères sont subjectifs (qualité de vie, douleur) ou impliquent une interprétation (échographie)... ;
- être direct (traduisant directement le phénomène biologique ou médical étudié, c'est à dire traduisant l'activité thérapeutique) ;
- être unique. Cependant, rares sont les pathologies se résumant à un seul critère. Il est alors nécessaire de recourir à plusieurs critères de jugement qu'on hiérarchise (principal et secondaires).

Critère de substitution

Un critère de substitution est un critère de jugement d'une étude épidémiologique ou d'un essai clinique dont l'évaluation est davantage compatible avec les délais et les effectifs habituellement requis que ne le serait le critère d'intérêt principal. Un critère de substitution n'est envisageable que s'il est fortement corrélé au critère d'intérêt principal et qu'il repose sur un rationnel physiopathologique. Dans les essais cliniques de cancérologie, les promoteurs sont parfois tentés de substituer à la durée de survie un critère de régression tumorale sous l'hypothèse que les sujets répondant bien au traitement (forte régression tumorale) auront des délais de survie en moyenne plus longs.

Echantillon

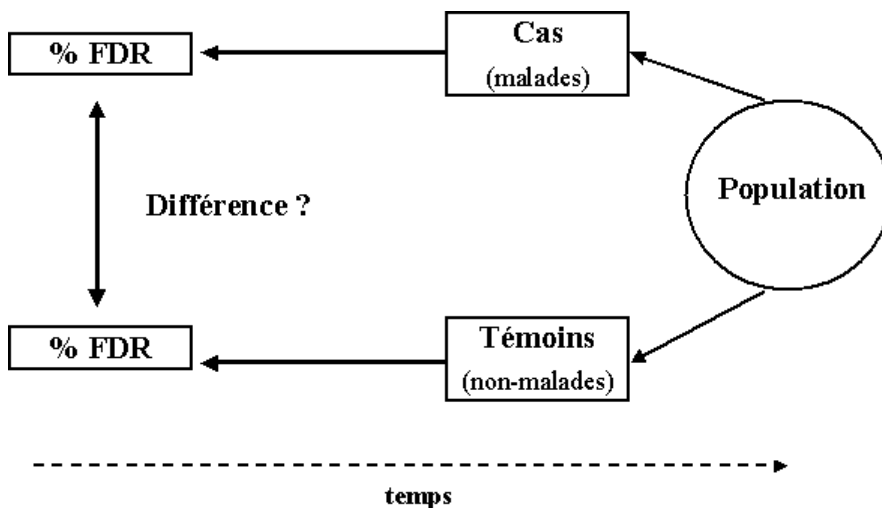
Un échantillon est un sous-ensemble de la population source. Un échantillon est dit représentatif lorsqu'il est constitué par sondage aléatoire simple (tirage au sort) où tous les individus statistiques de la population source ont la même probabilité d'être sélectionnés.

(voir population source, population cible)

Enquête cas-témoins

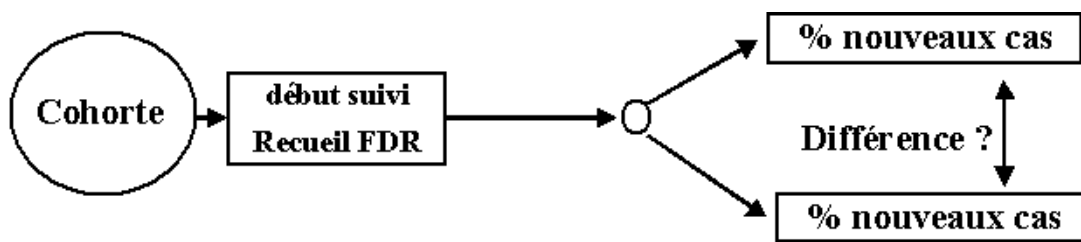
Enquête épidémiologique étiologique qui consiste à comparer la fréquence d'exposition antérieure à un (ou plusieurs) facteur de risque dans un groupe de « cas » atteints de la maladie étudiée, et dans un groupe de « témoins » indemnes de la maladie étudiée. La mesure du facteur de risque étudié est effectuée **rétrospectivement**, en général par entretien ou par auto-questionnaire.

Le recrutement des cas peut porter sur les cas présents au début de l'étude (cas prévalents) ou sur les nouveaux cas (cas incidents – étude cas-témoin qualifiée à tort de « prospective » : en fait c'est le recrutement qui est prospectif ; l'évaluation des facteurs de risque demeure rétrospective).

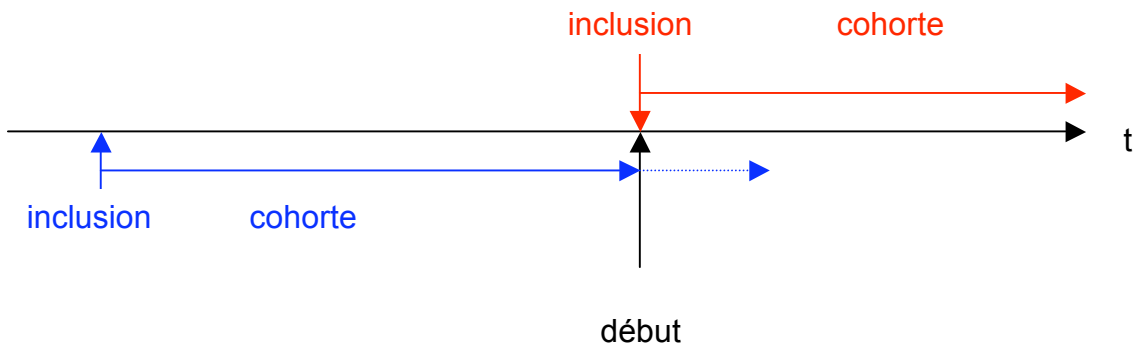


Enquête de cohorte

Enquête épidémiologique étiologique qui consiste à comparer la proportion de malades (ou de décès) observée entre deux ou plusieurs groupes de sujets initialement indemnes de la pathologie et différant par leur niveau d'exposition à un ou plusieurs facteurs de risque soupçonnés de cette maladie.



Une enquête de cohorte peut être prospective ou rétrospective :



(voir cohorte, enquête exposés / non-exposés)

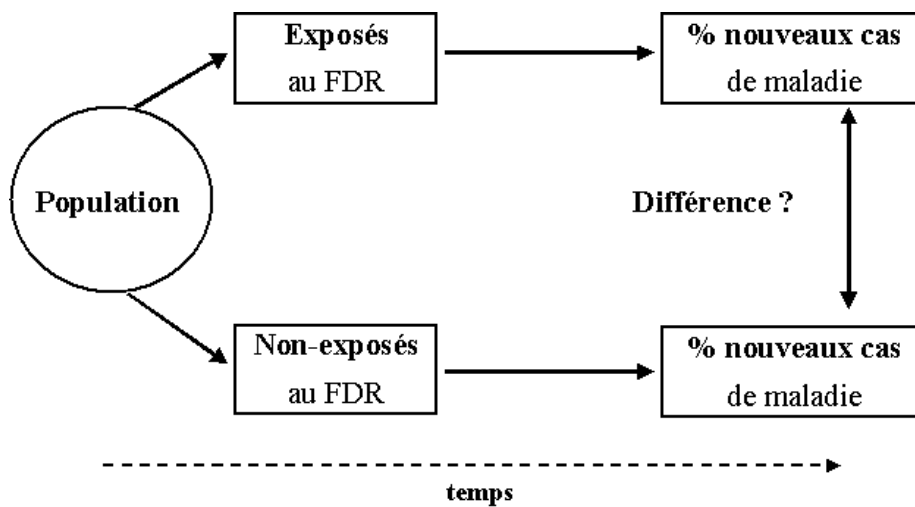
Enquête épidémiologique

Observation organisée de la population destinée à étudier la distribution d'un état de santé ou d'un facteur de risque (enquête descriptive), l'association entre un facteur de risque et un état de santé (enquête étiologique), ou l'association entre une stratégie de santé et un état de santé (enquête évaluative).

Enquête exposés / non exposés

Enquête épidémiologique étiologique qui consiste à suivre deux cohortes, l'une constituée de sujets exposés au facteur de risque étudié et l'autre constituée de

sujets non-exposés au facteur de risque, et de comparer le risque d'apparition d'une (ou plusieurs) maladie. Elle peut être prospective ou rétrospective.



Enquête transversale

Les enquêtes transversales descriptives (enquêtes de prévalence) étudient la distribution d'un facteur de risque ou d'un état de santé dans une population à un instant donné (c'est l'équivalent d'un instantané photographique). Les enquêtes transversales étiologiques consistent à mesurer simultanément l'exposition actuelle (ou antérieure) à un facteur de risque et l'état de santé.

Equivalence

On ne peut pas conclure à l'équivalence de deux traitements sur l'absence de différence statistiquement significative d'un test d'hypothèse classique. En effet, l'absence de différence statistiquement significative entre deux traitements résulte en général d'un défaut de puissance de l'essai (effectif de patients insuffisant) : la différence d'effet entre deux traitements n'est nulle que très rarement dans la réalité. La mise en évidence d'une équivalence de deux traitements nécessite donc la mise en place d'un essai spécifique qualifié « d'essai d'équivalence ». Ce type d'essai est conçu pour montrer que la différence entre deux traitements est inférieure à une

borne en deçà de laquelle elle peut être considérée comme négligeable sur le plan clinique.

(voir puissance statistique, risque alpha et risque beta).

Essai clinique sur groupes parallèles

Il consiste à comparer sur deux (ou plusieurs) groupes parallèles de patients, les effets de deux (ou plusieurs) traitements, mesurés par un critère de jugement.

L'essai repose sur trois principes généraux :

- la comparaison d'un groupe recevant le traitement étudié et d'un groupe ne le recevant pas ;
- le tirage au sort qui est le seul moyen d'obtenir des groupes de malades comparables pour toutes les caractéristiques connues ou inconnues, autres que le traitement étudié ;
- le double aveugle qui est la seule façon de garantir le maintien de la comparabilité des deux groupes tout au long de l'essai et de garantir une égalité d'appréciation du critère de jugement.

On distingue les essais explicatifs des essais pragmatiques :

	Essai explicatif	Essai pragmatique
<i>Objectif</i>	<i>Connaître l'effet du traitement dans des conditions très précises dans un but d'explication</i>	<i>Documenter le bénéfice qu'apportera le traitement aux patients qui seront traités en pratique</i>
<i>Population visée</i>	La plus homogène possible afin de contrôler au maximum l'influence d'autres facteurs	Représentative des futurs patients qui seront traités
<i>Critères</i>	Précis définissant parfaitement les patients recherchés. Faisant appel à des techniques sophistiquées	Simple correspondant au critère utilisé en pratique pour décider un traitement. Faisant appel à des examens de routine

(voir tirage au sort, aveugle, placebo, critère de jugement)

Essai clinique en cross-over

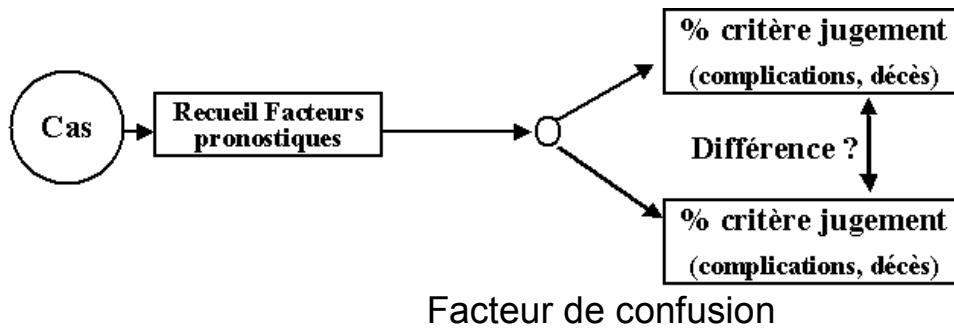
Dans l'essai clinique en cross-over, le sujet est pris comme son propre témoin. Il n'y a plus qu'un seul groupe de patients (par opposition à l'essai sur groupe parallèles). Le sujet reçoit successivement le traitement testé et le traitement de référence (placebo) dans un ordre déterminé aléatoirement.

Etude contrôlée

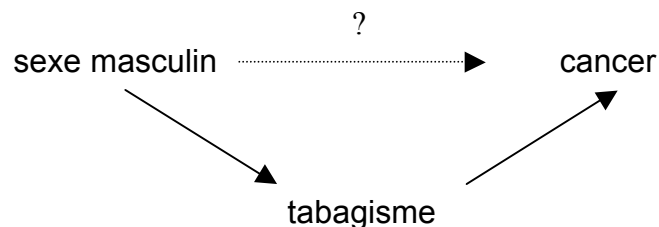
Anglicisme : étude (essai) comparative dans laquelle le groupe d'intérêt (traitement testé) est comparée à un groupe témoin (groupe contrôle) .

Etude pronostique

Etude épidémiologique (observationnelle) dont le but est d'identifier les facteurs prédictifs (pronostiques) de la survenue d'un événement (décès, complications) en suivant un groupe de cas de la maladie.



Un facteur de confusion est un facteur lié à la fois à l'exposition et à la maladie étudiées. Par exemple, le tabagisme est un facteur de confusion dans l'étude de la relation entre le sexe et le cancer bronchique. Les conclusions d'une analyse brute seraient : les sujets qui ont un cancer sont plus souvent des hommes (sexe masculin = facteur de risque). En fait, les hommes sont plus souvent fumeurs que les femmes. Seul le tabac est un facteur de risque de cancer bronchique. La relation entre sexe masculin et le cancer du poumon n'existe que par le biais du tabagisme (tabac = facteur de confusion dans la relation entre le sexe et le cancer bronchique).



Le facteur de confusion peut être contrôlé par :

- appariement individuel ;
- stratification ;
- ajustement.

(voir biais de confusion, biais, ajustement, appariement, stratification, analyse multivariée)

Facteur de risque

Un facteur de risque est une caractéristique associée à une probabilité plus élevée de maladie. C'est une définition statistique qui ne permet pas d'affirmer le rôle causal.

(voir jugement de causalité, jugement de signification, facteur protecteur)

Facteur protecteur

Un facteur protecteur est une caractéristique associée à une probabilité réduite de maladie. C'est une définition statistique qui ne permet pas d'affirmer le rôle causal.

(voir jugement de causalité, jugement de signification, facteur de risque)

Fiabilité d'un test

(voir reproductibilité)

Incidence

Le taux d'incidence est le rapport du nombre de nouveaux cas de maladie (incidents) au nombre de personnes susceptibles d'être atteintes dans la population, pendant une période donnée.

$$TI = \frac{M}{(P1 + P2) / 2}$$

M = nombre de nouveaux cas (malades) pendant la période d'étude

P1 = population en début de période

P2 = population en fin de période

On parle par exemple d'un taux d'incidence annuel de 10 cas pour 100 000 sujets par an.

$$TI = \frac{M}{PT}$$

La densité d'incidence est le nombre de nouveaux cas apparus par unité de temps d'exposition.

M = nombre de nouveaux cas (malades) pendant la période d'étude

PT : nombre de personnes temps.

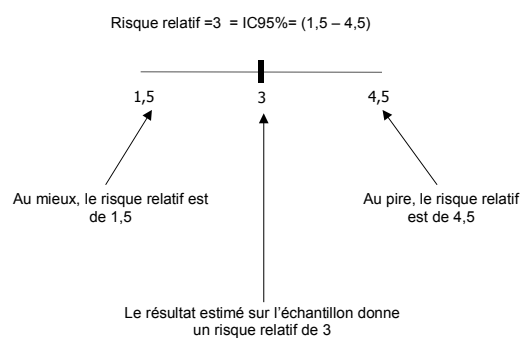
On parle par exemple d'une densité d'incidence de 10 cas pour 100 000 personnes-années.

Intervalle de confiance

L'intervalle de confiance traduit la précision de l'estimation d'un paramètre. Le but de l'estimation est de connaître à partir d'un échantillon la valeur d'un paramètre dans une population, par exemple le risque relatif d'apparition d'une pathologie.

Cependant, l'estimation ponctuelle de ce paramètre sur l'échantillon peut être assez éloignée de la valeur dans la population, du fait des fluctuations aléatoires d'échantillonnage (c'est à dire du fait du hasard). L'intervalle de confiance permet de prendre en compte cette incertitude aléatoire dans la présentation des résultats. Si on constitue 100 échantillons et qu'on calcule sur chacun d'entre eux un intervalle de confiance (IC) à 95%, alors 95 de ces échantillons contiendront la valeur du paramètre de la population. Avec un peu moins de rigueur, il est possible d'admettre

que l' IC à 95% représente la fourchette de valeurs à l'intérieur de laquelle nous sommes certains à 95% de trouver la vraie valeur recherchée. L'intervalle de confiance est donc l'ensemble des valeurs raisonnablement compatibles avec le résultat observé (estimation ponctuelle). Il donne une expression formelle de l'incertitude rattachée à une estimation ponctuelle du fait des fluctuations d'échantillonnages.



Jugement de causalité

Le jugement de causalité consiste à conclure ou non que l'association mise en évidence entre un facteur de risque et une pathologie correspond à une relation de cause à effet. La question de la causalité ne se pose que si on a objectivé une association statistiquement significative (jugement de signification).

Dans un essai clinique, le jugement de causalité ne peut être posé que si on a assuré la comparabilité initiale des deux groupes (tirage au sort) et son maintien tout au long de l'essai (double aveugle).

Dans une enquête épidémiologique d'observation, on ne peut pas d'emblée porter un jugement de causalité. Il faut donc examiner les arguments qui plaident en faveur de la causalité de la relation observée :

Critères internes à l'étude :

- séquence temporelle : l'exposition doit précéder l'apparition de la pathologie ;

- force de l'association : risque relatif ou odds ratio élevé ;
- relation de type « dose-effet » entre l'exposition et la fréquence de la pathologie ;
- spécificité de la cause et de l'effet : la relation est spécifique si le facteur de risque est présent chez presque tous les malades et seulement chez eux, et ce pour cette seule maladie ;
- cohérence interne de l'étude : prise en compte la plus complète possible des facteurs de confusion et minimisation des biais de sélection ou de classification.

Critères externes à l'étude :

- constance de l'association et reproductibilité dans diverses situations (différentes périodes, régions ou populations) ;
- cohérence des résultats avec les résultats d'études publiées ;
- plausibilité biologique : modèles expérimentaux chez l'animal ;
- cohérence avec les connaissances générales et les hypothèses qui ont conduit à la réalisation de l'enquête ;
- parallélisme de la distribution (dans l'espace et dans le temps) du facteur de risque et de la pathologie.

Jugement de signification

Le jugement de signification consiste à conclure, sur la base d'un test statistique, si l'on rejette ou non l'hypothèse nulle (H_0) d'absence d'association entre un facteur de risque et une pathologie dans une enquête épidémiologique, ou d'absence d'association entre le traitement et le critère de jugement dans un essai clinique.

Une différence observée entre deux groupes d'un essai clinique ou d'une enquête épidémiologique peut soit refléter une différence entre les deux populations dont sont issus les deux échantillons, soit être due au hasard (fluctuations d'échantillonnage).

Le test statistique permet d'évaluer la réalité statistique de cette différence observée.

Le risque de conclure à tort à un effet du traitement devant une différence due au hasard est le risque alpha (erreur de première espèce). Un résultat statistiquement significatif au seuil de 5%, limite ce risque à moins de 5%.

Pour s'assurer de la réalité statistique d'un effet du traitement (ou d'une association entre un facteur de risque et une pathologie), il faut répondre aux questions :

- le résultat est-il statistiquement significatif à un seuil inférieur ou égal à 5% ?
- le test statistique utilisé est-il adapté ?

- peut-on considérer que le risque alpha a été contrôlé (analyses intermédiaires protégées contre l'inflation du risque alpha, absence d'analyse en sous-groupes) ?

(voir jugement de causalité, risque alpha et risque beta)

Moment du tirage au sort

Le moment du tirage au sort désigne le moment dans l'histoire du malade où celui-ci est affecté à l'un des deux groupes de l'essai. Le tirage au sort est effectué après avoir vérifié les critères d'éligibilité du patient et obtenu la signature de l'accord de participation. Par ailleurs, le tirage au sort devra avoir lieu le plus tard possible dans l'histoire du sujet, c'est à dire juste avant l'administration du traitement qu'il devra recevoir. Ceci permet de réduire au maximum le délai séparant le tirage au sort du début du traitement (si ce délai était long, des événements empêchant l'administration du traitement risqueraient de survenir (décès du patient) et il serait tentant lors de l'analyse de ne pas prendre en compte ces sujets).

Multicentrique

Un essai clinique ou une étude épidémiologique multicentrique est réalisée simultanément dans plusieurs centres. Les avantages sont l'augmentation de la capacité de recrutement (lorsque le recrutement par un seul centre entraînerait une durée excessive de recrutement) et la possibilité de généralisation des résultats (variabilité inter-centre).

Niveau de preuve

Niveau de preuve scientifique apporté par une étude

(adapté de l'American College of Chest Physicians)

Niveau	méthodologie
1	Essai contrôlé randomisé avec résultats indiscutables (méthodologiquement)
2	Essai contrôlé non randomisé bien mené
3	Essai prospectif non contrôlé bien mené (suivi de cohortes par exemple)
4	Etude cas-témoin, essai contrôlé présentant des biais
5	Etude rétrospective, cas cliniques (série de malades), toute étude fortement biaisée

Nombre de sujets nécessaire

La détermination du nombre de sujets à inclure dans un essai résulte d'un calcul prenant en compte :

- l'importance de la différence minimale δ attendue du critère de jugement entre les deux groupes comparés (différence cliniquement intéressante) ;
- la variabilité du critère de jugement (qui gêne la mise en évidence de la différence δ attendue entre les deux groupes) ;
- le risque alpha consenti, qui est la probabilité de conclure à tort à une différence δ qui n'existe pas dans la réalité (différence due aux fluctuations d'échantillonnage), on prend habituellement $\alpha = 5\%$;
- la puissance souhaitée de l'essai, c'est à dire la probabilité de mettre en évidence la différence δ entre les deux groupes, sachant que cette différence existe dans la réalité. On prend habituellement $(1 - \text{risque } \beta) = 80\%$ voire plus.

Odds ratio (rapport de côtes)

L'odds ratio quantifie l'association entre l'exposition à un facteur de risque et la maladie étudiée. L'odds ratio est une approximation du risque relatif dans les enquêtes cas-témoins, sous réserve que la maladie soit rare dans la population (<5%, soit la majorité des cas en épidémiologie),

Calcul de l'odds ratio dans une enquête cas-témoins

	Exposés	Non-exposés
Cas	a	b
Témoins	c	d

$$\text{Odds ratio : OR} = (axd) / (bxc)$$

L'odds ratio s'interprète comme le risque relatif.

L'estimation ponctuelle de l'odds ratio doit être encadrée d'un intervalle de confiance (à 95% le plus souvent) qui permet de juger de la précision de cette estimation et de porter un jugement de signification (on peut aussi se servir de la p-value du test du χ^2 effectué sur le tableau de contingence pour porter un jugement de signification) :

- OR supérieur à 1 et [IC_{95%}] excluant la valeur 1 (ou $p_{\chi^2} < 0.05$) : risque significativement accru de maladie chez les sujets exposés comparativement aux sujets non-exposés ;

- OR inférieur à 1 et [IC_{95%}] excluant la valeur 1 (ou $p\chi^2 < 0.05$) : risque significativement réduit de maladie chez les sujets exposés comparativement aux sujets non exposés (effet protecteur)
- [IC_{95%}] incluant la valeur 1 (ou $p\chi^2 > 0.05$) : on ne met pas en évidence de relation statistiquement significative entre le facteur étudié et la maladie.

(voir facteur de risque, risque relatif, jugement de signification, intervalle de confiance)

Ouvert (essai ouvert)

Dans certaines situations, la réalisation d'un essai clinique en double aveugle est impossible pour des raisons éthiques ou pratiques. L'évaluation du critère de jugement doit être réalisée en aveugle du traitement reçu, par un investigateur indépendant des médecins ayant en charge les patients. En dehors des situations suivantes, la non-réalisation d'un essai en double-aveugle est insatisfaisante :

- Un des traitements comparés est une intervention chirurgicale ou invasive (radiologie interventionnelle comme une angioplastie).
- Un des traitements comparés nécessite un appareillage lourd dont il est impossible de faire un simulacre comme la radiothérapie.
- Un des traitements comparés s'accompagne d'effet indésirable ou d'une toxicité évocatrice qui laisse deviner la nature du traitement dans presque tous les cas : chute de cheveux dans des chimiothérapies anticancéreuses.
- Les traitements comparés sont des stratégies de prise en charge : traitement à domicile versus traitement hospitalier.
- Un des traitements comparés concerne une prise en charge améliorée : stroke unit, kinésithérapie, aide à domicile, etc.
- Le traitement factice risque d'avoir un effet : faux massage, placebo de chewing-gum pour l'arrêt du tabac, etc.
- Un des traitements comparés délivre son action de façon évidente et non dissimulable. Il est donc impossible d'en faire un simulacre sans effet : (chirurgie) dans une certaine mesure kinésithérapie, cure thermale, physiothérapie (chaleur), etc.
- D'une manière générale, toutes les situations où la réalisation d'un traitement « placebo » ayant la même apparence que le traitement étudié s'avère trop compliquée à réaliser ou illusoire, par exemple, quand l'action du traitement est directement visible (comme la chirurgie, le recours à une aide humaine, etc.).

(voir aveugle)

Placebo

Un placebo est un produit dénué d'action pharmacologique, ressemblant en tous autres points au produit actif, en particulier en ce qui concerne sa présentation, son mode d'administration et sa posologie. L'effet placebo résulte de phénomènes d'auto-suggestion par le malade et des phénomènes d'hétéro-suggestion par l'entourage, notamment médical. Il est admis que des symptômes tels que la douleur, les troubles du sommeil, la fatigue peuvent être traités avec succès par un placebo (le placebo peut aussi engendrer des effets indésirables). Lors d'un essai comparant un traitement à un placebo, la différence observée entre les deux groupes rendra compte du seul effet du produit actif contenu dans le traitement, éliminant les composantes propres à l'effet placebo. L'essai contre placebo n'est éthique qu'en l'absence de traitement standard ayant fait la preuve de son efficacité.

Population cible

La population cible est la population à laquelle on souhaiterait extrapoler les résultats de l'étude

(voir population source)

Population source

La population source est la population d'où est extrait l'échantillon d'étude. Si on réalise une enquête en population générale (population cible) par interrogatoire téléphonique, les sujets non abonnés au téléphone ne font pas partie de la population source.

(voir population cible)

Prévalence

Proportion de cas recensés dans une population à un instant donné t. C'est une mesure de risque ou probabilité (donc compris entre 0 et 1).

$$P = \frac{M}{N}$$

M = nombre de cas (malades)

N = nombre total de sujets (malades + non-malades)

La prévalence intègre :

- la durée de la maladie. Plus la maladie dure longtemps, plus le nombre de malades présents est grand. La prévalence d'une pathologie chronique peut augmenter sous l'effet de progrès thérapeutiques qui allongent la durée de survie. A l'inverse une réduction de la prévalence peut résulter d'une réduction de la durée de la maladie (par guérison ou par décès)
- la vitesse d'apparition des nouveaux cas au sein de la population. Plus celle-ci est grande plus la prévalence est grande.

Principe d'incertitude

Le principe d'incertitude guide l'inclusion des sujets de plus en plus souvent dans les essais cliniques modernes. Ce principe repose sur l'idée fondamentale que toute décision thérapeutique doit respecter les règles de l'éthique. Il peut s'énoncer ainsi : « si pour un patient vous ne savez pas lequel des deux traitements comparés dans l'essai serait le meilleur, incluez le patient dans l'essai et tirez au sort son traitement. En revanche, si pour un autre patient, vous avez suffisamment d'arguments pour penser que l'un des traitements serait meilleur que l'autre, alors n'entrez pas le patient dans l'essai et donnez-lui ce traitement ». L'avantage de cette attitude est

que la population d'étude correspondra à la population de patients pour laquelle les cliniciens se posent réellement la question que pose l'essai clinique.

Propriétés extrinsèques d'un test

Les propriétés (ou performances) extrinsèques d'un test sont la valeur prédictive positive et négative.

Elles varient en fonction de la prévalence de la maladie.

(voir valeur prédictive, propriétés intrinsèques d'un test)

Propriétés intrinsèques d'un test

Les propriétés (ou performances) intrinsèques d'un test sont la sensibilité et la spécificité. Elles ne dépendent pas de la prévalence de la maladie.

(voir sensibilité, spécificité, propriétés extrinsèques d'un test)

Puissance statistique

(voir risque alpha, risque beta)

Rapports de vraisemblance

Dans l'évaluation des performances d'un test diagnostique, le rapport de

$$L = \frac{P(T+ / M+)}{P(T+ / M-)} = \frac{Se}{1 - Sp}$$

vraisemblance positif est défini par

Un test diagnostique performant a un rapport de vraisemblance positif très supérieur à un.

Le rapport de vraisemblance négatif est égal à

$$\lambda = \frac{P(T- / M+)}{P(T- / M-)} = \frac{1 - Se}{Sp}$$

Un test diagnostique performant a un rapport de vraisemblance négatif très inférieur à un (proche de 0)

(voir sensibilité, spécificité)

Reproductibilité

La reproductibilité d'un test (radiologique, anatomo-pathologique....) est sa propriété à aboutir à un même résultat lorsqu'il est appliqué chez un même sujet par deux observateurs indépendants au même moment (concordance inter-observateurs) ou à deux moments différents par un même observateur (reproductibilité intra-observateur), sans se préoccuper de la validité du résultat du test (le résultat peut être juste ou erroné).

Randomisation

(voir tirage au sort, moment du tirage au sort)

Risque alpha, risque beta

Une différence observée entre deux groupes d'un essai clinique ou d'une enquête épidémiologique peut soit refléter une différence existant entre les deux populations dont sont issus les deux échantillons, soit être due au hasard (fluctuations d'échantillonnage). Le test statistique permet d'évaluer la réalité statistique de la différence observée.

Le risque de conclure à tort à un effet du traitement devant une différence en réalité due au hasard est le risque alpha (erreur de première espèce). Un résultat statistiquement significatif au seuil de 5%, limite ce risque à moins de 5%.

Le risque de ne pas conclure à un effet du traitement alors qu'il existe une différence entre les deux populations dont sont issus les deux échantillons est le risque beta (erreur de deuxième espèce).

		décision	
		rejet de H0	non-rejet de H0
Populatio n	H0 est vraie	erreur de 1 espèce α	$1-\alpha$
	H1 est vraie	$1-\beta$	erreur de 2 ^{ème} espèce β

La puissance statistique d'une étude quantifie la probabilité de mettre en évidence une différence entre les deux groupes comparés, sachant que cette différence existe entre les deux populations. La puissance correspond à (1- risque β).

(voir jugement de signification, nombre de sujets nécessaires)

Risque relatif

Le risque relatif quantifie l'association entre l'exposition à un facteur de risque et la maladie étudiée. Il désigne le risque de présenter la maladie chez les sujets exposés au facteur par rapport aux non-exposés à ce même facteur (« le risque de développer la maladie est RR fois supérieur chez les sujets exposés

comparativement aux sujets non-exposés »). Il ne peut être estimé que dans les enquêtes de cohortes (enquêtes exposés / non-exposés par exemple).

Calcul du risque relatif dans une étude de cohorte

	Malades	Non-malades
Exposés	a	b
Non-exposés	c	d

Risque de maladie chez les sujets exposés : $Re = a / (a+b)$

Risque de maladie chez les sujets non-exposés : $Rne = c / (c+d)$

Risque relatif = rapport du risque de maladie chez les exposés sur le risque de maladie chez les non-exposés : $RR = Re / Rne$

L'estimation ponctuelle du risque relatif doit être encadrée d'un intervalle de confiance (à 95% le plus souvent) qui permet de juger de la précision de cette estimation et de porter un jugement de signification (on peut aussi se servir de la p-value du test du χ^2 effectué sur le tableau de contingence pour porter un jugement de signification) :

- RR supérieur à 1 et [IC_{95%}] excluant la valeur 1 (ou $p_{\chi^2} < 0.05$) : risque significativement accru de maladie chez les sujets exposés comparativement aux sujets non-exposés ;
- RR inférieur à 1 et [IC_{95%}] excluant la valeur 1 (ou $p_{\chi^2} < 0.05$) : risque significativement réduit de maladie chez les sujets exposés comparativement aux sujets non exposés (effet protecteur)
- [IC_{95%}] incluant la valeur 1 (ou $p_{\chi^2} > 0.05$) : on ne met pas en évidence de relation statistiquement significative entre le facteur étudié et la maladie.

Dans les enquêtes cas-témoins, on utilisera l'odds ratio pour approximer le risque relatif.

(voir facteur de risque, odds ratio, jugement de signification, intervalle de confiance)

Run-in

Période de latence sous placebo avant randomisation et introduction du principe actif testé prévue dans le protocole d'un essai clinique pour s'assurer des critères d'éligibilité et éliminer les sujets répondeurs au placebo

(voir essai clinique, placebo).

Sensibilité

La sensibilité d'un test est la proportion de tests positifs chez les personnes malades (probabilité d'avoir un test positif si on est malade)

La sensibilité d'un test est donnée par la formule : $Se = VP / (VP + FN)$

Sa valeur est comprise entre 0 et 1.

Résultats du test	Malades (M+)	Non malades (M-)
Test positif (T+)	VP	FP
Test négatif (T-)	FN	VN
Total	(VP + FN)	(FP + VN)
Indices	$Se = VP / (VP + FN)$	$Sp = VN / (VN + FP)$

(voir spécificité, valeur prédictive, propriétés intrinsèques d'un test)

Spécificité

La spécificité d'un test est la proportion de tests négatifs chez les personnes indemnes de la maladie (probabilité d'avoir un test négatif si on est indemne de la maladie). La spécificité est donnée par la formule : $Sp = VN / (VN + FP)$

Sa valeur est comprise entre 0 et 1. (voir sensibilité, valeur prédictive, propriétés intrinsèques d'un test)

Stratification

Technique qui consiste à répartir un échantillon en sous-groupes de patients (les strates) homogènes pour un ou plusieurs facteurs de confusion.

(voir ajustement, appariement, facteur de confusion)

Tirage au sort (randomisation)

Dans un essai clinique, pour pouvoir porter un jugement de causalité, il est nécessaire que les deux groupes soient comparables. Le seul moyen d'assurer la comparabilité initiale des deux groupes sur tous les facteurs de confusion connus ou inconnus est d'attribuer les traitements par tirage au sort. Le tirage au sort doit être équilibré : on dit d'un tirage au sort qu'il est équilibré tous les 8 sujets (bloc de 8), si après l'inclusion de 8 patients supplémentaires, 4 sont affectés à chacun des deux groupes. L'ordre dans lequel les traitements sont attribués doit être imprévisible par le médecin qui détermine l'éligibilité des patients (il ne doit donc pas disposer de la liste qui détermine l'attribution des traitements, ni connaître la taille des blocs).

(voir moment du tirage au sort, aveugle).

Valeur prédictive

La valeur prédictive positive est la probabilité qu'a un sujet d'être effectivement malade s'il a un test positif. La valeur prédictive positive est donnée par la formule :

$$VPP = VP / (VP + FP)$$

La valeur prédictive négative est la probabilité qu'a un sujet d'être effectivement indemne de la maladie s'il a un test négatif. La valeur prédictive négative est donnée par la formule :

$$VPN = VN / (VN + FN)$$

Résultats du test	Malades (M+)	Non malades (M-)	Total	Indice
Test positif (T+)	VP	FP	(VP + FP)	$VPP = VP / (VP + FP)$
Test négatif (T-)	FN	VN	(FN + VN)	$VPN = VN / (FN + VN)$

Les valeurs prédictives dépendent à la fois des propriétés intrinsèques du test (sensibilité et spécificité) et de la prévalence de la maladie dans l'échantillon. En effet, on voit que la répartition entre malades et non-malades intervient dans le calcul de la valeur prédictive positive et négative. Par conséquent, les valeurs prédictives positive et négative n'ont de sens que si elles ont été estimées sur un échantillon représentatif de la population pour laquelle le test est destiné. Une erreur courante est d'estimer les valeurs prédictives d'un test sur des patients hospitalisés et d'en déduire la performance de ce test pour une utilisation en population générale.

Références

Bouvenot G, Vray M. Essais cliniques : théorie, pratique et critique. Paris : Flammarion, 1999.

Bouyer J, Hémon D, Cordier S & Al. Epidémiologie : principes et méthodes quantitatives. Paris : éditions INSERM, 1995

Cucherat M. La lecture critique des essais cliniques. <http://www.spc.univ-lyon1.fr/lecture-critique/>

Czernichow P, Chaperon J, Le Coutour X. Epidémiologie : connaissances et pratiques. Paris : Masson, 2001.

Haynes RB, Sackett DL, Guyatt GH, Tugwell P. Clinical epidemiology. How to do clinical practice research. 3rd Ed. New York: Lippincott Williams & Wilkins, 2006.

Jolly D, Ankri J, Chapuis F, Czernichow P, Guillemin F. Lecture critique d'articles médicaux. Paris : Masson, 2005.

Laplanche A, Com-Nougué C, Flamant R. Méthodes statistiques appliquées à la recherche clinique. Paris : Flammarion, 1996.

Rumeau-Rouquette C, Blondel B, Kaminski M. Epidémiologie : méthodes et pratique. Paris : Flammarion, 1997.