

LEXIQUE

A

- AJUSTEMENT (adjustment):

Ensemble des procédures ayant pour objectif d'éliminer l'effet de certaines variables, considérées comme parasites, dans l'étude de la relation entre un facteur que l'on étudie et un critère de jugement. Le terme d'ajustement est réservé aux procédures utilisées après le recueil des données, dans l'analyse des résultats (par régression ou standardisation ou stratification).

- ALEATOIRE (random):

- Répartition aléatoire: détermination au hasard (selon un plan prédéterminé, par exemple par tirage au sort) du groupe auquel est affecté le sujet qui est inclus dans l'étude (voir randomisation).
- Variable aléatoire: variable dont les fluctuations obéissent à une loi donnée.

- ANALYSE COÛT-BENEFICE (cost-benefit analysis):

Type d'analyse visant à comparer le coût d'une stratégie et son bénéfice. Dans une telle étude, les coûts et les conséquences sont exprimés en unités monétaires.

- ANALYSE COÛT-EFFICACITE (cost-effectiveness analysis):

Type d'analyse permettant de comparer des stratégies qui diffèrent par leurs coûts et leurs effets. Elle s'exprime en unités monétaires par indicateur d'efficacité médicale (nombre d'années de vie sauvées, nombre de cas de maladie évités).

- ANALYSE COÛT-UTILITE (cost-utility analysis):

Type d'analyse comparable à l'analyse coût-efficacité, dont l'indicateur de résultat médical intègre plusieurs dimensions. L'indicateur généralement utilisé est le QALY (Quality Adjusted Life Years) dont le critère d'efficacité est le nombre d'années de vie sauvées, ou l'espérance de vie, ajusté sur un critère de qualité de vie.

- ANALYSE DE SURVIE (survival analysis)

Techniques statistiques d'étude de la survie et de recherche des effets de divers facteurs (traitement, expositions, facteurs pronostiques).

- ANALYSE EN INTENTION DE TRAITER (intention to treat analysis)

Méthode utilisée dans les essais comparatifs randomisés. Les patients de chaque groupe thérapeutique restent analysés dans ce groupe qu'ils aient ou non reçu / poursuivi le traitement.

- APPARIER (to match):

Consiste à rendre comparables deux ou plusieurs groupes de sujets en vue d'étudier les effets d'un facteur donné, et ces effets seuls, sur un critère de jugement. Cette technique vise à

neutraliser les facteurs de confusion éventuels, que sont en général l'âge, le sexe, la catégorie socio-professionnelle ou le groupe ethnique, en choisissant pour chaque individu inclus dans l'étude un ou plusieurs témoins semblables à lui pour ces caractéristiques. Dans l'appariement par paires (matching) les cas et les témoins forment des paires.

- ARBRE DE DECISION (decision tree):

Outil conçu pour déterminer la décision à rendement maximal quand différentes options sont possibles au cours d'un processus clinique à étapes (diagnostique ou thérapeutique). Il s'agit d'un graphique où à chaque étape les options possibles sont représentées par des branches, tout comme les conséquences de ces options.

Représentation graphique d'une analyse décisionnelle où chaque embranchement (ou nœud de décision) correspond à 2 (ou plusieurs) décisions possibles, chacune aboutissant à 2 (ou plusieurs) éventualités.

- ASSOCIATION (relation ou corrélation statistique) (association)

C'est une dépendance statistique entre deux ou plusieurs événements ou variables. L'association peut être positive ou négative.

- ASSORTIMENT (matching):

Voir appariement.

- AVEUGLE (blind):

Procédure destinée à éviter un biais de mesure ou un biais d'information. Synonyme: insu.

On distingue:

- les essais en simple aveugle, dans lesquels le patient ignore à quel groupe il appartient (groupe expérimental ou témoin), et par conséquent à quel type de traitement il a été soumis (par exemple, médicament ou placebo);
- les essais en double aveugle, dans lesquels le patient et l'expérimentateur ignorent tous deux si le patient appartient au groupe expérimental ou témoin. Cette procédure permet de réduire non seulement l'effet placebo dû à la subjectivité du patient, mais aussi les biais dus à la subjectivité de l'expérimentateur.
- Dans un même essai ouvert le patient et l'expérimentateur sont informés du groupe auquel appartient le patient. Dans les essais en triple aveugle, ni le patient, ni l'expérimentateur, ni le médecin chargé de surveiller le patient ne savent à quel groupe ce dernier appartient.

B

- BIAIS (bias):

Il s'agit de toute erreur systématique qui s'introduit dans une étude, au stade de la collecte ou de l'analyse des données, et de l'interprétation des résultats, et qui contribue à produire des estimations systématiquement plus élevées ou plus basses que la valeur réelle des paramètres que l'on étudie.

- BIAIS DE CONFUSION (confounding bias):

On regroupe sous ce terme toutes les erreurs qui peuvent survenir lorsque la mesure de la relation entre le facteur étudié et le critère de jugement est modifiée par un facteur, appelé facteur de confusion, lié à la fois au facteur étudié et au critère de jugement.

- BIAIS DE MESURE (measurement bias):

On regroupe sous ce terme toutes les erreurs systématiques qui peuvent s'introduire dans la mesure des phénomènes pris en compte chez les sujets qui entrent dans l'étude.

- BIAIS DE NON-REPONSE (response bias)

Erreur due aux différences de caractéristiques des personnes qui acceptent ou non d'entrer dans une étude. Les biais de non-réponse (patients inclus dans une étude ne répondant pas) peuvent représenter un biais de sélection important.

- BIAIS DE SELECTION (selection bias):

On regroupe sous ce terme toutes les erreurs qui peuvent faire que les sujets effectivement observés dans l'étude ne constituent pas un groupe représentatif des populations que l'on voulait étudier et ainsi ne permettent pas la généralisation des résultats.

C

- CAS RAPPORTES ET SERIES DE CAS (case report et case series):

Les cas rapportés, en décrivant une observation inhabituelle, constituent souvent la première étape de la reconnaissance d'une nouvelle maladie ou d'un nouveau facteur de risque. Les séries de cas représentent l'étape suivante en regroupant différentes observations similaires et en établissant ainsi l'existence probable d'une entité pathologique.

Les séries de cas et les cas rapportés traduisent avant tout l'expérience et l'observation d'un auteur et ne permettent pas de tirer de conclusions qu'on puisse généraliser à d'autres cas. Ils ne permettent pas d'établir la fréquence d'une maladie -ce qui nécessiterait une étude d'incidence ou de prévalence. Ils ne permettent pas non plus d'apprécier de manière statistique l'importance d'un facteur de risque qu'ils peuvent éventuellement suggérer.

- COHORTE (cohort):

Ensemble de sujets ayant vécu un événement semblable (par exemple, facteur étudié), et suivis dans le temps depuis la date de cette expérience. Le suivi de la cohorte est conçu pour

enregistrer un ou plusieurs événements donnés (ou critères de jugement), par exemple le décès avec sa cause.

- CO-INTERVENTION (cointervention):

Attitude particulière de l'investigateur vis-à-vis des patients inclus dans une étude, ou prise d'un traitement concomitant ayant des effets identiques, et pouvant résulter en une véritable intervention non planifiée qui peut s'avérer être la véritable raison d'une différence observée entre les groupes étudiés.

- COMPLIANCE (compliance):

Un patient est compliant lorsqu'il suit les consignes données par le corps médical sur le traitement instauré ou l'intervention proposée. On parle aussi d'adhérence au traitement. Voir aussi observance.

- CONFUSION (confounding):

Situation dans laquelle la mesure de la relation entre le facteur étudié et le critère de jugement est modifiée par un facteur, appelé facteur de confusion, lié à la fois au facteur étudié et au critère de jugement. Cette situation produit un biais, le biais de confusion.

- CORRELATION (correlation)

Il y a corrélation entre 2 variables quantitatives a et b lorsque les points représentant ces valeurs de l'une par rapport à l'autre se situent au voisinage d'une ligne droite. Ces points ne sont pas alignés, mais situés dans un nuage. La droite est construite selon la méthode des moindres carrés.

- COTE (odd):

Voir Rapport de cotes.

- COURBE ROC (Receiver Operating Characteristic curve):

Moyen graphique d'évaluation de la capacité d'un test diagnostique à différencier les sujets sains des sujets malades.

- CRITERE DE JUGEMENT (outcome factor):

Dans une étude analytique, le critère de jugement, ou facteur résultant, se définit comme la situation ou l'événement supposé être le résultat de l'influence du facteur étudié.

Les événements dignes d'intérêt, pour le patient comme pour le clinicien ou l'épidémiologiste, se répartissent en cinq catégories: la mort, la maladie, le handicap, l'inconfort, l'insatisfaction, auxquels on peut ajouter un sixième élément, la destitution, qui a une dimension plus sociale. Il peut aussi s'agir de leur inverse: survie, guérison, ...

La définition de ces critères de jugement doit être aussi précise que possible. Un critère de jugement est une variable, dite dépendante, par rapport au facteur étudié qui est la variable indépendante.

- CROSS-OVER

Essai thérapeutique constitué de deux groupes où le sujet est son propre témoin. Le premier groupe reçoit dans un premier temps le traitement A puis le traitement B alors que les patients du second groupe reçoivent le traitement B puis le traitement A. Avec pour les deux groupes avec une période de « wash-out » (sans traitement) entre les deux afin de ne pas superposer leurs effets.

D

- DEGRE DE SIGNIFICATION (level of significance):

Voir valeur de p.

- DISTRIBUTION DE GAUSS (normal distribution):

La distribution est le résumé le plus complet d'une mesure quantitative faite sur un groupe d'individus. Elle révèle combien d'individus, ou quelle proportion du groupe, présentent chaque valeur parmi toutes les valeurs possibles que la mesure quantitative peut donner.

La distribution normale, ou gaussienne, est un cas particulier. La courbe de Gauss est symétrique, en forme de cloche. Elle possède la propriété mathématique d'avoir les deux tiers des mesures observées dans un écart-type, et environ 95 % des observations dans deux écarts-types. Il est souvent admis, pour des raisons pratiques, que les mesures en clinique ont une distribution gaussienne.

NB. Le terme "normal" n'est pas à prendre dans son sens clinique de "non malade" ou son sens de "commun" ou "ordinaire", mais au sens de "conforme à une règle ou à un modèle".

E

- ECART REDUIT

C'est le rapport de l'écart observé sur l'écart-type. Voir test statistique et écart-type.

- ECART-TYPE ou DEVIATION STANDARD (standard deviation):

C'est la racine carrée positive de la variance. Il s'agit de l'unité d'écart permettant d'exprimer un écart observé dans une unité non liée au contexte. C'est une mesure de la dispersion ou de la variation exprimant la distance moyenne des valeurs d'une variable par rapport à la moyenne de ces valeurs.

- ECHANTILLON

- Echantillon représentatif (representative sample):

Un échantillon est représentatif d'une population s'il a les mêmes caractéristiques que cette population. Les conclusions d'une étude réalisée sur l'échantillon pourront alors être généralisées à l'ensemble de la population d'où est tiré l'échantillon.

- Echantillon aléatoire simple (random sample)

L'échantillon est dit aléatoire simple si tous les individus ont la même probabilité de figurer dans l'échantillon.

- EFFET CARRY OVER (carry-over effect)

Il s'agit de la poursuite de l'effet d'un médicament après son arrêt.

- EFFET PLACEBO (placebo effect)

Effet généralement bénéfique, attribué par les patients à un médicament ou à une intervention.

- EFFICACITE:

- Efficacité potentielle ou théorique (efficacy): avantage théorique que présente l'action de santé considérée. Il s'agit d'une approche théorique, évaluée dans des conditions idéales, expérimentales. Efficacité théorique ne signifie pas nécessairement efficacité réelle.
- Efficacité réelle (effectiveness): efficacité d'une action de santé mesurée par l'adéquation entre les résultats obtenus et les objectifs fixés lors de la mise en route de cette action de santé. Il s'agit d'une approche pragmatique, dans les conditions réelles.
- Efficience (efficiency): efficacité d'une action de santé mesurée par la comparaison des résultats obtenus aux moyens mis en œuvre pour arriver à ces résultats optimaux. C'est la notion de capacité de rendement.

- EPIDEMIOLOGIE (epidemiology)

« Etude de la distribution et des déterminants d'états de santé dans les populations données, et application de cette étude à la lutte contre les problèmes de santé. »¹

- ERREUR ALEATOIRE (random error):

Déviation d'une valeur estimée d'un paramètre à partir de l'observation d'un échantillon, par rapport à sa valeur réelle dans la population d'où provient l'échantillon. Il s'agit des fluctuations d'échantillonnage dues au hasard (par opposition aux biais).

- ERREUR DE TYPE I ou DE PREMIERE ESPECE (type I error):

C'est conclure, lorsque qu'on fait un test statistique, qu'il y a une différence entre les groupes étudiés alors qu'en réalité il n'y en a pas. La probabilité de cette erreur est le risque α .

- ERREUR DE TYPE II ou DE DEUXIEME ESPECE (type II error):

C'est conclure, lorsque l'on fait un test statistique, qu'il n'existe pas de différence entre les groupes étudiés alors qu'il en existe une en réalité. La probabilité de cette erreur est le risque β (voir puissance).

- ESSAI (trial):

¹ John M. Last. *Dictionnaire d'Epidémiologie*, Edisem inc. 2004

Expérimentation à visée interventionniste - intervention thérapeutique, de dépistage, de prévention ou d'éducation - dans laquelle l'évaluation des effets sur le groupe de sujets étudiés se fait par rapport à un groupe témoin de référence: on peut comparer une intervention à l'absence d'intervention, ou à une ou plusieurs autres interventions du même type, l'objectif étant de déterminer s'il existe une différence entre elles.

- ESTIMATION (estimate):

Estimer un paramètre (inconnu) caractéristique d'une population consiste à en proposer une valeur, appelée estimation, calculée à partir d'un échantillon. Une estimation est donnée avec son intervalle de confiance.

Exemple: la fréquence d'une maladie, observée sur un échantillon obtenu par tirage aléatoire simple, est une estimation de la fréquence de la maladie dans la population.

- ETUDE ANALYTIQUE (analytic study):

Type d'étude spécialement conçue pour vérifier une hypothèse qui a été soulevée par les résultats d'une étude descriptive, en cherchant à établir l'association possible entre un ou plusieurs facteurs étudiés et un ou plusieurs critères de jugement.

- ETUDE CAS-TEMOINS ou CAS-CONTROLES (case-control study):

Une étude cas-témoins est une étude d'observation, rétrospective, de type analytique, dans laquelle les caractéristiques de patients atteints d'une maladie (les cas), sont comparées avec celles de patients indemnes de la maladie (les témoins), ces caractéristiques étant en général l'exposition à des facteurs de risque que l'on recherche dans le passé des cas comme dans celui des témoins.

- ETUDE DE COHORTE (cohort study):

Une étude de cohorte est une étude d'observation, prospective, de type analytique, dans laquelle un groupe de sujets exposés à des facteurs de risque d'une maladie est suivi pendant une période de temps donnée. Le taux d'incidence de la maladie dans ce groupe exposé est comparé à celui d'un groupe témoin, suivi pendant le même temps, mais non exposé aux facteurs de risque.

- ETUDE DESCRIPTIVE (descriptive study):

Toute étude qui a pour but de rendre compte d'un phénomène de santé, de sa fréquence, de sa distribution et de son évolution, ainsi que des variables apparentées, au sein d'une population donnée. Ce type d'étude permet, entre autres, d'identifier certains groupes chez qui la fréquence de ce phénomène est anormalement élevée - ou anormalement faible - afin de déterminer leurs besoins en matière de santé.

- ETUDE D'OBSERVATION (observational study):

Etude caractérisée par l'absence d'intervention de la part du chercheur. Ce type d'étude vise, soit à décrire l'état de santé de la population ou ses déterminants, soit à étudier les relations entre l'état de santé d'une population et l'exposition à un ou plusieurs facteurs, celle-ci n'étant pas contrôlée. Elle ne permet pas d'établir de lien de causalité.

- ETUDE ECOLOGIQUE ou ETUDE DE CORRELATION (ecologic study ou correlational study):

Elles établissent la comparaison entre l'importance (ou la fréquence) d'un facteur de risque supposé au sein d'une population et la prévalence ou l'incidence de la maladie supposée secondaire, à partir de données déjà disponibles au niveau de cette population. Elles sont transversales dans la mesure où elles superposent deux types de données (fréquence du facteur de risque et fréquence de la maladie) recueillies dans une même période de temps.

- ETUDE EXPERIMENTALE (experimental study):

Etude dans laquelle l'expérimentateur choisit, dans un groupe donné, les sujets qui seront soumis à l'action du facteur étudié (intervention médicale, chirurgicale...), maîtrise dans la mesure du possible l'effet d'autres facteurs concomitants et mesure les variations de l'état des sujets soumis à l'expérience afin d'établir un lien de causalité entre le facteur étudié et l'effet mesuré. Quand les exigences de l'éthique médicale la permettent, l'étude expérimentale est préférable à l'étude d'observation, car elle permet d'établir la preuve directe d'une relation causale, l'exposition au facteur étudié étant contrôlée.

- ETUDE LONGITUDINALE (longitudinal study):

Etude qui consiste à suivre une ou plusieurs cohortes à l'aide d'exams périodiques et répétés pendant une assez longue période, contrairement à ce qui se passe dans l'étude transversale où la durée est suffisamment courte pour que l'élément temps devienne une donnée négligeable.

- ETUDE PROSPECTIVE (prospective study):

Etude qui comporte la récolte de données sur des événements à venir. Elle consiste généralement à suivre un groupe de sujets exposés à un facteur de risque particulier, afin d'étudier les phénomènes de santé qui affectent ce groupe au cours du temps. Le facteur étudié est enregistré avant que se produise le critère de jugement.

- ETUDE RETROSPECTIVE (retrospective study):

Etude dans laquelle on recherche un lien possible entre un phénomène de santé (maladie ou autre) présent au moment de l'étude et des événements (facteurs de risque) survenus dans le passé. Par exemple, on examine les antécédents d'un groupe de patients afin d'identifier les sujets qui, dans le passé, ont subi une exposition à un facteur de risque particulier. L'information concernant les événements passés est obtenue au moyen de documents d'archives ou par l'interrogatoire. Le facteur étudié est enregistré rétrospectivement, après que le critère de jugement s'est produit.

- ETUDE TRANSVERSALE ou ETUDE DE PREVALENCE (cross-sectional study):

Cette étude analyse la présence d'un facteur donné ou d'une maladie particulière dans une population P à un moment précis t, sans référence au passé ou sans suivi dans le futur. Elle représente l'équivalent d'un sondage rigoureusement et scientifiquement construit, ou de l'instantané photographique d'une situation précise dans la population étudiée.

Ce type d'études est surtout effectué dans un but descriptif. Il pose des problèmes d'interprétation liés au manque d'information sur la chronologie des événements et à des difficultés de sélection.

- EXPOSITION (exposure):

Proximité ou contact avec un facteur qui peut avoir un effet nuisible ou protecteur sur le sujet exposé.

F

- FACTEUR DE CONFUSION ou CONFONDANT (confounding factor):

Toute variable liée à la fois au facteur de risque étudié et à la maladie, et susceptible de modifier la liaison statistique entre ces derniers. Il est essentiel d'identifier les facteurs de confusion potentiels et d'en tenir compte au moment de la planification de l'étude ou de son analyse.

- FACTEUR DE RISQUE (risk factor):

Caractéristique individuelle ou collective, endogène (propre à l'individu) ou exogène (liée à l'environnement), qui augmente la probabilité de survenue d'une maladie ou de tout autre phénomène de santé. Le lien est purement statistique et ne préjuge en rien de la causalité.

- FACTEUR ETUDIE (study factor):

Dans une étude analytique, le facteur étudié est l'exposition ou l'intervention supposée avoir des conséquences sur un problème de santé, une maladie ou un état clinique.

- FACTEUR PRONOSTIQUE (prognostic factor):

Un facteur pronostique est l'état, la situation ou l'événement qui, quand il est observé chez un sujet qui présente déjà un état pathologique, est associé avec une conséquence de cet état pathologique.

- FAUX NEGATIF (false negative):

Test diagnostique négatif (normal) chez un sujet qui a la maladie.

- FAUX POSITIF (false positive):

Test diagnostique positif (anormal) chez un sujet qui n'a pas la maladie.

G

- GOLD STANDARD (gold standard)

Test considéré comme indiscutable pour déterminer l'état de maladie d'un patient (indépendamment des autres tests diagnostiques) ou mesurer l'efficacité supérieure d'un médicament ou d'une thérapie.

- GROUPE CONTROLE ou GROUPE TEMOIN (control group)

Dans le cas d'un essai contrôlé, il s'agit du groupe qui reçoit le médicament de référence ou le placebo.

H

- HYPOTHESE (hypothesis):

Peut être définie comme une supposition, née d'une observation ou d'une réflexion, qui conduit à des prédictions que l'on peut réfuter. En d'autres termes, l'hypothèse est avant tout l'expression d'une opinion. Une telle opinion peut être plus ou moins forte, plus ou moins hardie. Mais quoi qu'il en soit, l'objectif d'une étude analytique est de démontrer, à partir des données, et avec des méthodes statistiques appropriées, si cette supposition est une description exacte de la relation qui existe entre les facteurs que l'on étudie, ou si elle ne l'est pas.

- HYPOTHESE ALTERNATIVE (alternative hypothesis) :

C'est l'hypothèse qui sera acceptée si l'hypothèse nulle est rejetée (voir hypothèse nulle).

- HYPOTHESE NULLE (null hypothesis):

Correspond à l'hypothèse selon laquelle il n'y a pas de différence entre les groupes étudiés pour une variable donnée. Dans un test statistique, il s'agit de l'hypothèse qui est "soumise à épreuve". Si l'hypothèse nulle est rejetée, alors on retient l'hypothèse alternative avec un risque d'erreur α .

I

- INCIDENCE (incidence):

Nombre de nouveaux cas d'une maladie donnée ou de personnes qui sont atteintes de cette maladie, pendant une période donnée, dans une population déterminée.

- INCLUSION - NON INCLUSION (CRITERES D') (inclusion - exclusion criteria):

- Les critères d'inclusion sont les caractéristiques que les sujets doivent obligatoirement présenter pour être inclus dans une étude donnée.
- Les critères de non inclusion (parfois appelés critères d'exclusion), sont les critères qui interdisent l'inclusion d'un sujet, bien qu'il satisfasse aux conditions de la première liste.

Tout sujet entrant dans l'étude devra satisfaire à tous les critères d'inclusion, et ne présenter aucun des critères de non-inclusion. Les critères d'exclusion sont valables durant toute la durée de l'étude.

- INDICATEURS DE SANTE (health indicator):

Descripteur quantitatif d'un phénomène donné. Dans la pratique ce sont plutôt des critères de « non-santé ». Exemple : le nombre de jours d'arrêt de travail pour maladie, le taux de mortalité infantile.

- INSU (blind):

Voir aveugle.

- INTERACTION (interaction)

Action réciproque de 2 ou plusieurs causes pour produire ou prévenir un effet. L'interaction entre deux causes peut être synergique ou antagoniste.

K

- KAPPA (COEFFICIENT DE CONCORDANCE):

Un des coefficients les plus utilisés pour mesurer le degré de concordance entre des juges classant les mêmes individus (par exemple, des patients) dans une ou plusieurs catégories préalablement définies (par exemple, des diagnostics). On peut également calculer un kappa sur plusieurs mesures du même phénomène évaluées par un même juge (kappa intra-observateur).

M

- MEDIANE (median):

Valeur centrale qui sépare la population en deux parties d'effectifs égaux. Dans le cas d'effectifs impairs, la médiane est la moyenne des deux valeurs.

- MESURE (measurement):

Quantification d'un phénomène en utilisant un instrument ou en appliquant une échelle de mesure standard sur une variable ou un ensemble de valeurs.

- MODE (mode):

Valeur la plus fréquente pour une variable discrète.

Classe correspondant au maximum de l'histogramme pour une variable continue.

- MOYENNE (mean):

Caractéristique de valeur centrale, pour une distribution observée ou théorique. Si n est le nombre de valeurs observées, la moyenne arithmétique est la somme des valeurs observées divisée par le nombre de valeurs observées.

$$x_1, x_2, \dots, x_n: \frac{x_1 + x_2 + \dots + x_n}{n}$$

O

- OBSERVANCE (compliance):

Le fait de se conformer, pour le sujet inclus dans l'étude, aux prescriptions du corps médical, et pour le chercheur au protocole de recherche. Synonyme: compliance.

P

- PERCENTILE (centile):

Voir quantile.

- PERDU DE VUE (lost to follow-up):

Sujet dont on n'a plus de nouvelles et dont on a perdu la trace au cours d'une étude prospective et pour lequel les informations le concernant ne seront pas disponibles au moment de l'analyse.

- PHASE I (phase I):

Les essais de phase I visent à préciser la sûreté et la tolérance; ils sont faits chez un petit nombre de sujets.

- PHASE II (phase II):

Les essais de phase II précisent l'efficacité optimale du traitement.

- PHASE III (phase III):

Les essais de phase III établissent l'efficacité du traitement, le plus souvent grâce à des essais thérapeutiques comparatifs, idéalement randomisés.

- PHASE IV (phase IV):

Les essais de phase IV, après la commercialisation, visent à préciser les effets du traitement à long terme.

- PLACEBO (placebo):

Procédure ou substance inactive ou inerte que l'on substitue à la vraie procédure ou à la vraie substance médicamenteuse pour contrôler ou induire les effets psychologiques accompagnant la médication.

- POPULATION (population):

En statistique, une population est un ensemble d'unités susceptibles d'être observées (qui peuvent ne pas être des personnes physiques).

En épidémiologie, il convient de distinguer la population examinée, que l'on peut également qualifier d'échantillon, de la population d'où provient cet échantillon (à condition qu'on puisse la définir).

Dans l'un et l'autre cas, la population doit être, dans la mesure du possible, définie et décrite de manière précise.

- PRECISION (precision):

Propriété d'une mesure rendant compte de la dispersion des valeurs qu'elle fournit autour de sa moyenne pour une grandeur donnée. La mesure sera d'autant plus précise que la dispersion des valeurs autour de la moyenne sera faible.

- PREVALENCE (prevalence):

Nombre de cas d'une maladie donnée ou de personnes atteintes de la maladie ou de tout autre événement morbide, dans une population déterminée à un moment donné ou au cours d'une période donnée, sans distinction entre les cas nouveaux et les cas anciens.

- PREVENTION PRIMAIRE (primary prevention):

Ensemble de mesures destinées à diminuer l'incidence d'une maladie dans une population (en évitant cette maladie à des sujets qui n'en sont pas atteints). Il peut s'agir de mesures qui visent à augmenter la résistance des individus ou à éliminer les facteurs de risque de la maladie.

- PREVENTION SECONDAIRE (secondary prevention):

Ensemble de mesures destinées à diminuer la prévalence d'une maladie dans une population (en réduisant la durée de la maladie). Assurée par les organismes de santé publique et par le corps médical, elle se fonde sur le dépistage des maladies à un stade précoce, pendant que celles-ci peuvent être facilement curables.

- PREVENTION TERTIAIRE (tertiary prevention):

Ensemble des mesures destinées à diminuer les complications dégénératives d'une maladie ou les incapacités chroniques. Elle vise à diminuer la prévalence des incapacités chroniques.

- PROBABILITE A *POSTERIORI* OU PROBABILITE POST-TEST (posterior probability):

C'est la probabilité de la maladie étant donné le résultat du test (voir valeurs prédictives).

- PROBABILITE A *PRIORI* ou PROBABILITE PRE-TEST (prior probability):

C'est la fréquence de la maladie dans la population dont est issu le patient soumis à l'examen (c'est la prévalence lorsqu'il s'agit d'une population importante).

- PROTOCOLE (protocol):

Description de l'ensemble des étapes nécessaires pour la réalisation d'une étude. Dans une étude, la mise au point du protocole est au moins aussi importante que le recueil des données, car la validité des résultats dépend de cette mise au point.

- PUISSANCE (power):

La puissance d'un test (ou d'une étude) est la probabilité de conclure, à partir de l'échantillon, à une différence ou à un effet statistiquement significatif lorsque cette différence ou cet effet existe réellement dans la population (puissance = $1 - \beta$). Une étude manque de puissance lorsque la conception de son protocole ou les conditions de sa réalisation ne permettent pas de mettre en évidence un effet ou une différence, alors que ceux-ci existent réellement dans la population.

Q

- QUANTILE (quantile):

Les quantiles d'une variable quantitative permettent de diviser la population en groupes d'effectifs égaux. Les quartiles divisent la population en quatre groupes d'effectifs égaux, les quintiles en cinq, les déciles en dix et les percentiles en cent.

- QUESTIONNAIRE (questionnaire):

Ensemble de questions prédéterminées pour récolter des données d'ordre clinique, social, comportemental ou de santé publique.

R

- RANDOMISATION ou REPARTITION ALEATOIRE (randomization, random allocation):

Répartition d'un échantillon ou d'une population en deux ou plusieurs groupes comparables, à l'aide d'une méthode fondée sur le hasard. La répartition est effectuée par tirage au sort, notamment à l'aide des tables de nombres au hasard.

C'est la meilleure solution pour contrôler l'influence des facteurs de confusion, qu'ils soient connus ou inconnus, en répartissant également ces variables dans les différents groupes, de sorte que leur effet sur le critère de jugement s'annule et qu'ainsi, si une différence est observée entre les deux groupes, elle puisse être attribuée à l'effet du facteur étudié.

- RAPPORT DE COTES (odds ratio):

C'est une estimation indirecte du risque relatif, ce dernier n'étant pas disponible dans les études cas-témoins. Il correspond au rapport de deux cotes qui sont respectivement le rapport des exposés sur les non exposés chez les malades et le rapport des exposés sur les non exposés chez les non malades.

Cote (odd) d'exposition parmi les cas (les malades):

$$\frac{\text{Proportion de cas exposés}}{\text{Proportion de cas non exposés}} = \frac{a / (a+b)}{b / (a+b)} = \frac{a}{b}$$

Cote (odd) d'exposition parmi les témoins (les non-malades):

$$\frac{\text{Proportion de témoins exposés}}{\text{Proportion de témoins non exposés}} = \frac{c / (c+d)}{d / (c+d)} = \frac{c}{d}$$

$$\text{Rapport de cotes} = \frac{\text{Cote d'exposition parmi les cas}}{\text{Cote d'exposition parmi les témoins}} = \frac{a / b}{c / d} = \frac{ad}{bc}$$

Le rapport de cotes exprime le risque relatif lorsque la maladie est rare, et se calcule très simplement par le rapport des produits en croix de la table 2×2 .

- **RATIO (ratio):**

Rapport, parfois exprimé sous forme de pourcentage, entre deux grandeurs, deux quantités, ayant chacune une nature distincte. Exemple: sex-ratio. Le ratio est différent du taux dont le dénominateur contient le numérateur.

- **REGRESSION (regression):**

Modèle mathématique permettant de décrire une variable dépendante comme une fonction d'une ou plusieurs variables indépendantes. Le modèle le plus commun est le modèle linéaire. Lorsque la variable dépendante a deux modalités, on utilise un modèle de régression logistique.

- **REPRODUCTIBILITE (reproductibility):**

Un test est reproductible s'il donne les mêmes résultats dans des situations identiques. La quantification de la reproductibilité d'un test est fondée sur le coefficient KAPPA pour les variables quantitatives.

- **RISQUE (risk):**

Probabilité de survenue d'un événement pendant une période donnée. En épidémiologie, cette probabilité varie en fonction de certaines caractéristiques endogènes (âge, sexe, hérédité, ...) ou exogènes (milieu, profession, ...) qui peuvent constituer des facteurs de risque. La mesure du risque ne permet ni un diagnostic, ni une évaluation certaine du pronostic individuel.

- **RISQUE ABSOLU (absolute risk):**

Probabilité de survenue d'un événement exprimée sans référence à une autre probabilité. S'il s'agit de la probabilité d'être atteint d'une maladie, le risque absolu sera souvent estimé d'après le taux d'incidence de la maladie au sein de la population.

- **RISQUE ATTRIBUABLE A UN FACTEUR DONNE (attributable risk):**

Différence entre le taux d'incidence d'un phénomène de santé au sein d'une population exposée à un facteur donné, et le taux d'incidence de ce phénomène au sein de la population non exposée.

- **RISQUE RELATIF (relative risk):**

Taux d'incidence d'une maladie donnée au sein d'un groupe exposé à un facteur de risque donné, rapporté au taux d'incidence de cette maladie au sein d'un groupe non exposé. Ce rapport est généralement considéré comme une mesure de la force de l'association existant entre le facteur et la maladie. Il répond à la question: combien de fois les sujets exposés ont-ils plus de chance que les sujets non exposés de contracter la maladie ?

S

- **SELECTION (selection):**

Procédure qui consiste à choisir les sujets à inclure dans l'étude.

- SENSIBILITE (sensitivity):

C'est la probabilité d'avoir un test diagnostique positif (anormal) quand on a la maladie. Cela correspond au taux de vrais positifs, nombre de vrais positifs rapportés à l'ensemble des malades.

- SERIE DE CAS (case series):

Voir cas rapporté et série de cas.

- SPECIFICITE (specificity):

C'est la probabilité d'avoir un test diagnostique négatif (normal) quand on n'a pas la maladie. Cela correspond au taux de vrais négatifs, nombre de vrais négatifs rapportés à l'ensemble des non-malades.

- STANDARDISATION (standardization, adjustment):

Méthode qui rend comparables des taux bruts dans des groupes qui diffèrent par la distribution d'une autre variable (âge, sexe, niveau socio-économique, ...). On distingue:

- la standardisation directe: pour l'âge, on calcule le taux que l'on observerait dans la population étudiée si elle avait la même structure d'âge qu'une population de référence;
- la standardisation indirecte: on calcule le nombre d'événements attendus (maladies, décès, ...) dans la population étudiée, en appliquant à chaque classe d'âge les taux spécifiques d'une population de référence.

- STATISTIQUE (statistics):

C'est l'art de faire la collecte, la synthèse et l'analyse des données qui sont sujettes à des variations aléatoires.

- STRATIFICATION (stratification):

Répartition des sujets en sous-groupes ou strates, en fonction d'une ou plusieurs caractéristiques, de sorte que chaque strate soit homogène pour cette ou ces caractéristiques. On peut ainsi stratifier les individus en fonction de l'âge, du sexe, des deux, du niveau socio-économique, du niveau de risque, ...

Cette méthode permet de rendre comparables, en procédant strate par strate, deux groupes différents.

- SUIVI (follow-up):

C'est l'observation pendant une période de temps donnée d'un individu, d'un groupe ou d'une population dont les caractéristiques ont été définies, afin de pouvoir observer et enregistrer tout changement ou tout événement concernant la santé de ces individus.

T

- TABLE 2×2 (2×2 table):

Mode de représentation, sous forme de tableau, de la relation entre deux variables qualitatives: chaque ligne représente une modalité d'une variable, chaque colonne une modalité de l'autre variable. La relation entre le résultat d'un test diagnostique et la survenue de la maladie, comme la relation entre les résultats d'un test statistique et la vraie différence entre les deux groupes, peuvent être représentés sous la forme d'une table 2×2 .

- TAILLE DE L'ECHANTILLON (sample size):

C'est le nombre de sujets à inclure nécessairement dans l'étude afin que celle-ci ait une puissance suffisante (possibilité de conclure à une différence lorsque celle-ci existe réellement).

- TAUX (rate):

Dans un contexte épidémiologique et démographique, il s'agit:

- de l'effectif d'une sous population n , dénombré à une date déterminée et rapporté à l'effectif de la population N dont cette sous-population fait partie, soit n/N ;
- du nombre des événements (maladies, décès, ...) n survenus dans une population donnée durant une période déterminée, rapporté à l'effectif de cette population p pour la même période, soit n/p .

- TAUX BRUT (crude rate):

- Taux qui rend compte uniquement d'une fréquence d'événements au sein d'une population donnée, au cours d'une période déterminée. Il permet seulement d'évaluer l'intensité d'un phénomène de santé dans une population, mais ne permet pas de comparer l'impact de ce phénomène dans des populations dont les structures (d'âge, de sexe, ...) sont différentes.
- Taux calculé pour l'ensemble d'une population ou d'une fraction de cette population, exprimant la fréquence de la mortalité ou de la morbidité attribuables à une cause déterminée (par exemple, taux de mortalité par cause, taux d'incidence ou de prévalence d'une maladie donnée).

- TAUX D'ATTAQUE (attack rate):

Dans le contexte d'une épidémie, nombre de cas d'une maladie donnée survenus pendant une période d'observation déterminée, rapporté à la population exposée à cette maladie au début de la période. Celle-ci représente en général la totalité du temps pendant lequel la maladie peut se produire.

- TAUX SPECIFIQUE (specific rate):

Taux calculé pour une fraction d'une population déterminée par une ou plusieurs caractéristiques (par exemple, taux spécifique par tranche d'âge et par sexe, par catégorie socio-professionnelle, par ethnie, ...).

- TAUX STANDARDISE (standardized rate; adjusted rate):

Taux calculé par des méthodes particulières (standardisation) permettant d'éliminer l'effet de certains facteurs (par exemple, l'âge, les facteurs socio-culturels, ...) distribués différemment entre deux ou plusieurs populations. On peut ainsi comparer l'intensité d'un phénomène donné (la mortalité par exemple) dans des populations de structures différentes.

- TEMOIN (control):

Sujet ou groupe de sujets dont les caractéristiques servent de référence pour l'estimation de l'association entre un ou plusieurs facteurs étudiés et un ou plusieurs critères de jugement dans la population que l'on étudie.

- TEST STATISTIQUE (statistic test):

Lorsque l'on observe une différence, soit entre un paramètre observé sur un échantillon et une valeur théorique, soit entre les valeurs prises par ce paramètre sur plusieurs échantillons, se pose la question de savoir si l'écart observé est seulement dû au hasard. On ne peut jamais rejeter, de façon absolue, l'hypothèse nulle selon laquelle l'écart observé est dû au hasard, mais on sait estimer le risque d'observer un tel écart simplement par le fait du hasard, par le calcul de la statistique de test. Si ce risque est faible (en pratique inférieur à 5 %), on rejette l'hypothèse nulle.

La statistique de test est l'expression qui transforme un écart observé, tout à fait lié aux circonstances (en particulier à l'effectif de l'échantillon), en un écart exprimé dans une unité non liée au contexte, permettant la comparaison à des valeurs tabulées. Il s'agit de l'écart réduit.

- THEOREME DE BAYES (Bayes' theorem):

Moyen mathématique d'estimer un jugement correct (valeurs prédictives) sur une maladie donnée, en fonction de la prévalence de la maladie, de la sensibilité et de la spécificité du test.

V

- VALEUR DE p (p value):

La valeur de p est l'expression, en termes quantitatifs, de la probabilité que les différences observées dans une étude puissent être observées par la chance seule.

Une autre façon d'exprimer la valeur de p est de répondre à la question: "s'il n'y a en réalité pas de différence en termes d'effet de l'intervention étudiée et que l'étude était répétée de nombreuses fois, combien de fois conclurait-on cependant qu'un effet existe ?"

Il est habituel de donner une importance spéciale aux valeurs de p situées au-dessous de 0,05, parce qu'il est généralement admis qu'une chance sur vingt est un risque acceptable.

Les effets associés à une valeur de p inférieure à 0,05 sont dits statistiquement significatifs. Il faut cependant souligner que cette limite de 0,05 est totalement arbitraire. On pourrait exiger une valeur différente, plus élevée ou plus basse, selon les conséquences que l'on accepte de conclusions faussement positives. Pour $p = 0,05$, on considère que la probabilité de 1 sur 20

est trop faible pour que le résultat observé soit survenu par chance seule. L'intervention ou l'exposition étudiée explique le résultat observé.

Pour briser cette alternative, $< 0,05$ ou $> 0,05$, on peut se référer à la valeur exacte de p (par exemple, 0,00843 ou 0,07). L'interprétation de ce qui est statistiquement significatif est alors laissée au lecteur. Cette valeur numérique de p est appelée degré de signification.

- VALEUR PREDICTIVE NEGATIVE (negative predictive value):

Probabilité de ne pas avoir la maladie quand le test diagnostique est négatif (normal).

- VALEUR PREDICTIVE POSITIVE (positive predictive value):

Probabilité d'avoir la maladie quand le test diagnostique est positif (anormal).

- VALIDITE (validity):

Possibilité d'une méthode (enquête, examen, test diagnostique, dépistage) de fournir une valeur exacte de ce qu'elle est censée mesurer.

Dans le cas d'un test diagnostique, c'est la proportion des mesures qui représentent la valeur réelle de l'objet mesuré. Une bonne validité suppose un degré élevé de sensibilité et spécificité.

Dans le cas d'un instrument de mesure, on étudie plusieurs aspects de la validité, les deux premiers s'appliquant aux questionnaires:

- pertinence apparente du questionnaire (face validity): caractère sensé et rationnel des questions
- exhaustivité du questionnaire (content validity): capacité du questionnaire à mesurer tous les domaines qu'il est sensé mesurer et uniquement ceux-là
- comparaison d'un instrument de mesure à une "référence" ou "étalon" (criterion validity)
- lorsqu'aucun examen de référence n'existe, on peut évaluer l'exactitude d'un instrument en le comparant à un faisceau d'arguments ou à une mesure différente en se fondant sur l'hypothèse qu'ils mesurent la même chose ou au moins des caractéristiques très semblables (construct validity)

- VALIDITE EXTERNE (external validity):

Possibilité de généralisation des observations issues de l'échantillon à la population générale. La validité externe est menacée par le biais d'échantillonnage, lorsque les conclusions fondées sur un échantillon sont généralisées à d'autres groupes qui ne sont pas semblables à l'échantillon.

- VALIDITE INTERNE (internal validity):

Capacité intrinsèque de la méthode à fournir des observations exactes pour les sujets ayant participé à l'étude, dans les conditions particulières de cette étude. La validité interne est menacée par les erreurs systématiques (biais) et aléatoires (chance).

- VARIABLE (variable):

En général, toute grandeur susceptible de prendre n'importe quelle valeur dans un ensemble préalablement défini.

Dans un sens plus restreint, propriété ou caractère servant à distinguer les individus d'une population et pouvant être soit qualitatif (par exemple, le sexe) soit quantitatif (par exemple, l'âge).

Parmi les variables quantitatives, on distingue:

- les variables discrètes (ou discontinues) : le nombre d'enfants;
- les variables continues: le poids, la taille.

Parmi les variables qualitatives, on distingue:

- les variables nominales: le sexe, le groupe sanguin;
- les variables ordinales, à modalité ordonnée: classification TNM pour les cancers.

- VARIANCE (variance):

Caractéristique de dispersion, pour une distribution théorique ou observée.

Les auteurs remercient Peggy Falgon, Jérôme Mollard et Adeline Roux qui ont relu ce glossaire.