Dictionary of Pharmacoepidemiology

Bernard Bégaud

WILEY

Dictionary *of* Pharmacoepidemiology

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Dictionary *of* Pharmacoepidemiology

Bernard Bégaud
ARME Pharmacovigilance, Bordeaux, France

JOHN WILEY & SONS, LTD
Chichester · New York · Weinheim · Brisbane · Singapore · Toronto

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Baffins Lane, Chichester,
West Sussex PO19 1UD, UK

National 01243 779777
International (+ 44) 1243 779777
e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on: http://www.wiley.co.uk or
http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158–0012, USA

WILEY-VCH Verlag GmbH, Pappelallee 3,
D-69469 Weinheim, Germany

Jacaranda Wiley, Ltd., 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Asia) Pte, Ltd., 2 Clementi Loop #02–01,
Jin Xing Distripark, Singapore 129809

John Wiley & Sons (Canada), Ltd., 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

Library of Congress Cataloging-in-Publication Data

Bégaud, Bernard,
Dictionary of pharmacoepidemiology / Bernard Bégaud.
p. cm.
ISBN 0-471-80361-8 (cased)
1. Pharmacoepidemiology – Dictionaries. I. Title.

RM302.5.B443 2000
615′.1′03 – dc21

00–042263

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0–471–80361–8

Typeset in 10.5/13pt Times from the author's disks by Vision Typesetting, Manchester.
Printed and bound in Great Britain by Bookcraft (Butt) Ltd, Midsomer Norton, Somerset.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

References, Addresses and Websites

CIOMS Working Group IV (1998). Council for the International Organization of Medical Sciences (CIOMS), c/o World Health Organization, Avenue Appia 1211, Geneva 27, Switzerland (www.who-umc.org/umc.html)

Declaration of Helsinki (1964). (www.faseb.org/arvo/helsinki.htm)

EMEA (European Medicines Evaluation Agency: http://www.eudra.org/emea.html)

EU (European Union) Directive 65/65/EEC, Article 1 (for rules governing medicinal products in the EU: http://dg3.eudra.org/eudralex/index.htm)

FDA (Food and Drug Administration, USA: www.fda.org)

ICD-10. (www.who.whois/icd10/index.html)

ICH E2A. International Conference on Harmonization (ICH), *Guideline for Good Clinical Practice*, Topic E2A. ICH documents can be accessed from IFPMA, EMEA and FDA

ICH E2B. International Conference on Harmonization (ICH), *Guideline for Good Clinical Practice*, Topic E2B. ICH documents can be accessed from IFPMA, EMEA and FDA

ICH E3. International Conference on Harmonization (ICH), *Guideline for Good Clinical Practice*, Topic E3. ICH documents can be accessed from IFPMA, EMEA and FDA

ICH E6. International Conference on Harmonization (ICH), *Guideline for Good Clinical Practice*, Topic E6,

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

CPMP/ICH/135/95. ICH documents can be accessed from IFPMA, EMEA and FDA

IFPMA (International Federation of Pharmaceutical Manufacturers' Associations: www.ifpma.org/ich1.htm)

International Covenant on Civil and Political Rights, United Nations High Commissioner for Human Rights, General Assembly Resolution 2200A (XXI) of 16 December 1966, Article 7 (http://www.unhcr.ch/html/menu3/b/a__ccpr.htm)

WHO (1972). Technical Report No. 498, World Health Organization, Avenue Appia 1211, Geneva 27, Switzerland (www.who-umc.org/umc.html)

WHO Collaborating Centre for Drug Statistics Methodology, World Health Organization, Postboks 183, Kalbakken 0903 Oslo, Norway (www.whocc.nmd.no)

A

Absolute risk Risk measured in a population exposed to the factor of interest.

The absolute risk adds the risk for this population had it not been exposed (*background risk*, *reference risk*) and the risk induced by exposure (*excess risk*, *attributable risk*).

See also: **exposed, exposure, risk difference.**

Abuse Any voluntarily excessive consumption or practice (abuse of alcohol, coffee, sport, etc.).

Referring to medications, this term indicates voluntary and excessive use—chronic or episodic—which does not conform to the recommendations in the Summary of Product Characteristics or to customary medical usage. According to most commonly used definitions, abuse is thus committed by the patient (*abuser*) and not the prescriber. Abuse may also involve non-prescribed medications, if the subject disregards the recommendations of the pharmacist or those specified on the label.

See also: **proper use, Summary of Product Characteristics.**

Active Adjective applied to a drug if its administration to a living organism induces or could induce one or more pharmacodynamic effect(s).

Activity does not necessarily imply a therapeutic effect. There are substances which induce one or more pharmacodynamic effect(s), without any therapeutic action.

See also: **active ingredient, efficacy, pharmacodynamic effect, therapeutic effect.**

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Active ingredient, active principle Substance responsible for some or all of the real or expected effects of a drug.

See also: **drug, pharmacodynamic effect, therapeutic effect.**

Acute Intense and of short duration. Can be applied to an exposure (e.g. acute poisoning) or to an event (e.g. acute liver injury, when biological anomalies persist for less than 3 months).

Acute is the antonym of *chronic*.

See also: **chronic.**

Additive Adjective used when the association of two or more factors produces an effect equal to the sum of the effects of each of the factors considered alone.

This term is used in pharmacology for interactions between two or more drugs, referring to the sum of their therapeutic or undesirable effects, and in epidemiology, referring to the sum of the effects or risks induced by each of the factors.

See also: **interaction, multiplicative, synergy.**

Ad hoc Latin expression meaning 'for this purpose' and indicating a tool or approach specifically conceived for or adapted to a precise objective.

In pharmacoepidemiology an *ad hoc* study, *ad hoc* analysis, etc., refers to a strategy specifically conceived or implemented to answer a particular question.

See **ad hoc study.**

***Ad hoc* study** Study conceived specifically for a precise objective, generally to answer a single question.

Ad hoc studies are contrasted in this respect with multipurpose studies, which attempt to answer several, sometimes very different, questions and with re-analyses of data previously collected for other purposes.

See also: **exploratory study, fishing expedition, *post hoc* analysis.**

Adjustment Procedure designed to minimise or eliminate the effect of differences in the distribution of one or more

secondary variable(s) in the compared populations.

Adjustment is particularly important when those variables may act as confounders when studying the causal association between exposure to a given factor and occurrence of an event.

Unlike matching, adjustment always takes place during analysis, that is, subsequent to data collection.

In this way, an incidence rate, relative risk or odds ratio may is said to be *adjusted* for a given variable, e.g. age, sex, body mass index, if its calculation took into account possible differences in the distribution of this variable in the compared populations. Adjustment may use several statistical methods, such as the logistic regression model, or covariance analysis.

If the interfering variable has a large influence, the value of the adjusted parameter may differ markedly from the crude, unadjusted value.

See also: **confounding, confounding ratio, crude, logistic model, matching, standardization.**

Adjustment for multiple comparisons Choice of a lower statistical significance level in order to maintain the level of Type I error originally chosen (e.g. 5%), when carrying out several identical independent tests within the same study.

The construction of a statistical test relies on controlling the Type I error, that is, the risk of wrongly rejecting the tested null hypothesis. When the test is repeated, the risk of Type I error is increased. It can be shown that if a is the original Type I error, the probability of wrongly rejecting the null hypothesis in j independent repeated tests is:

$$p = 1 - (1 - \alpha)^j.$$

For example, for an a of 5%, the risk of wrongly rejecting the null hypothesis is 18.5% over four independent tests on the same sample.

In order to maintain the global risk (α) at the level fixed a

priori, each individual test is carried out with a lower significance threshold. Several methods have been proposed to adjust this significance threshold in proportion to the number of tests carried out, the best known being that of Bonferroni.

For example, if we fix the risk of wrongly rejecting the null hypothesis at 5% and carry out three tests, the null hypothesis will be rejected when the statistic corresponds to a lower probability (2.2% instead of 5%), for each of the three tests.

However, this type of correction can certainly be abused, notably in pharmacoepidemiology when it is applied to non-independent tests (for example, within an intermediate analysis).

For example, a trial in which one drug is declared more efficacious than another, because significantly superior in one of 12 independently chosen evaluation criteria, enters into the framework of multiple comparisons.

On the other hand, in a 24–month survival analysis with intermediate analyses at 6, 12 and 18 months, to apply a correction under the pretext that four tests were carried out would be an aberrant conservative attitude, and would unjustifiably increase the risk of coming to no conclusion.

See also: **Type I error.**

Admission bias, referral bias Selection bias in which a subject has a different probability of being hospitalised or admitted to a care structure participating in recruiting study subjects, according to whether the person does or does not present a characteristic linked to the parameter to be measured.

The parameter may thus be over- or under-represented in the study sample compared to the source population, which will bias the measurement. This type of bias can alter the estimates made from cohort or case-control studies; in the latter case, the term 'Berkson's bias' is sometimes used.

See also: **Berkson's bias, selection bias.**

Adverse drug reaction, adverse reaction, adverse effect, untoward effect A response to a drug which is noxious and unintended and which occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for the modification of physiological function (WHO, 1972; *see* References).

The definition excludes events related to accidental or deliberate overdose. Unlike '*adverse event*', the term '*adverse drug reaction*' implies that a causal relationship with the treatment was ascertained or is strongly suspected.

Adverse drug reactions can be categorised as expected or unexpected.

See also: **adverse event, expected adverse drug reaction, side effect, unexpected adverse drug reaction.**

Adverse event, adverse experience Any untoward medical occurrence that may present during treatment with a pharmaceutical product but which does not necessarily have a causal relationship with this treatment (WHO, 1972; *see* References).

This also applies to other types of exposure.

See also: **adverse drug reaction.**

Aetiologic fraction of the risk in the exposed *See* **attributable fraction of the risk in the exposed.**

Aetiologic fraction of the risk in the population *See* **attributable fraction of the risk in the population.**

Aetiologic study, analytic study Epidemiological study whose goal is to determine the possibly causal role of one or more factor(s) in the aetiology of a disease or the occurrence of an event (*see* Appendix 1).

Aetiologic analysis is generally based on a comparative (often cohort or case-control study). This should ideally be a prospective experimental study with randomly

allocated exposure.

See also: **case-control study, cohort study.**

Aggregated data Data resulting from the compilation or analysis of individual datasets.

If individual data are not accessible, the use of aggregated data can lead to erroneous conclusions (aggregation bias).

For example, such uses could include the comparison of the results of different clinical trials or epidemiological studies or the comparison of annual drug sales to characteristics of the presumed exposed population, such as event rate.

See also: **aggregation bias.**

Aggregation bias, ecological bias, ecological fallacy Bias resulting from the analysis of aggregated data, when an association observed between variables at the aggregate level is different from that found at the individual level.

For example, in a given country the number of deaths from severe asthma increased proportionally to the sales of a new anti-asthmatic drug. Before concluding that this drug is associated with an increased risk of death, it would be best to rule out, at the individual level, possibility of a prescription bias, i.e. that the drug was mainly prescribed to very severe asthmatics.

See also: **aggregated data.**

Alert, warning Signal indicating a possible danger and implying need for appropriate action.

In pharmacovigilance, this term is stronger than 'signal' in that it indicates the existence of a risk whose quantitative and qualitative significance is yet to be determined. An alert therefore justifies the implementation of a study or suitable decision-making.

See also: **signal.**

Algorithm Method of analysis based on a chain of interdependent operational rules.

For example, in the assessment of causality, it refers to a method based on the successive evaluation of criteria, the evaluation of each criterion being dependent on that of previous ones.

Example: Has the suspected treatment been discontinued?—yes or no; if yes, has the adverse effect regressed?—yes or no; if yes, has the treatment been resumed?—yes or no; etc.

At the end of the chain, this method produces an estimate in terms of a score, a qualifier, etc.

Algorithms are also generally used in decision-making strategies (*decision trees*).

See also: **causality assessment method, imputability.**

Alpha error *See* **Type I error.**

Alternative hypothesis *A priori* specified hypothesis which is accepted when the null hypothesis is rejected by statistical testing.

See also: **null hypothesis.**

Ambispective Adjective denoting a follow-up based on data recorded before the beginning of a study and intended to be continued into the future.

In an ambispective cohort, the population of interest is identified in the past (e.g. by means of a database), but will continue to be followed after the beginning of the study. The term '*historico-prospective*' is sometimes used to designate this type of study.

See also: **historical, prospective.**

Analytic study See aetiologic study.

Anatomical Therapeutical Chemical classification (ATC) International drug classification proposed by the World Health Organization (WHO Collaborating Centre for Drug Statistics Methodology *see* References), originally elaborated by the Nordic Council on Medicine, Oslo, Norway.

Its initial purpose was to compare drug utilization between

countries in conjunction with defined daily doses. It has since became the reference drug classification, even for other purposes.

See also: **defined daily dose.**

Antagonism *See* **interaction.**

Assessable In pharmacovigilance jargon, 'assessable' indicates a case of an adverse event for which there is sufficient qualitative and quantitative information to undertake a causality assessment.

For example, we would say that among the 120 cases of hepatic injury gathered, only 85 were assessable and considered for analysis.

Assessment, evaluation Estimating or appraising the value of a parameter, a drug or a strategy as a function of pertinent and generally pre-defined criteria, considering all of the available data.

In this way, the new drug approval (NDA) commission of a medicines agency evaluates the risk–benefit ratio of new medications on the basis of all of the pre-clinical and clinical trials and according to standardized criteria. *Assessment* also signifies the result of such an evaluation.

See also: **European Product Assessment Report.**

Association, correlation, statistical dependence Statistical relationship observed between two or more variables.

For example, there is an association if the probability of occurrence of an event (or the number of cases) varies as a function of the presence or absence of a variable or the value of that variable.

An association is said to be *positive* if the probability of an event's occurrence is greater in the presence of a factor, or if the values of two variables tend in the same direction. An association is said to be *negative* if the probability of an event's occurrence is smaller in the presence of a factor, or if

the values of two variables tend in opposite directions.

Although one may speak of *fortuitous association* (i.e. explained only by chance), the term 'association' is generally reserved for statistically significant relationships; this significance, however, does not necessarily imply the existence of a causal relationship between the variable in question and the measured parameter.

See also: **causality, coincidental.**

ATC classification *See* **Anatomical Therapeutical Chemical classification.**

Attack rate Cumulative incidence rate, usually measured over a short period of time and often used to describe epidemics.

Attributability *See* **causality assessment method, imputability.**

Attributable Adjective meaning that which may be accounted for by exposure to a risk factor, or to a medicinal treatment.

Example: *A majority of cases of agranulocytosis occurring in these patients were attributable to their treatment.*

See also: **attributable fraction of the risk in the exposed, attributable fraction of the risk in the population, attributable number of cases, causality.**

Attributable cases *See* **attributable number of cases.**

Attributable fraction of the risk in the exposed, aetiologic fraction of the risk in the exposed Proportion of exposed cases for whom the occurrence of the event of interest is attributable to exposure.

The fraction thus indicates the proportion by which the risk of the event measured in an exposed population would be reduced were the exposure eliminated. The attributable fraction of the risk (AFR) is calculated by dividing the risk difference between the exposed (R_{exp}) and the non-exposed ($R_{non/exp}$) by the risk measured in the exposed:

$$AFR = \frac{R_{exp} - R_{non/exp}}{R_{exp}}.$$

The AFR can also be calculated from the value of the relative risk (RR) or the odds ratio (OR) associated with the exposure since, by definition, $RR = R_{exp}/R_{non/exp}$:

$$AFR = \frac{RR - 1}{RR} \text{ or } \frac{OR - 1}{OR}.$$

Example: If the relative risk of agranulocytosis associated with the use of a drug is 5, the attributable fraction of the risk of agranulocytosis in the exposed subjects is $(5 - 1)/5 = 0.8$. That is, 80% of the cases of agranulocytosis observed among the patients exposed to this drug are a priori attributable to the drug.

This estimate is clearly valid only under conditions of exposure similar to those which were present for the calculation of the relative risk or the odds ratio.

The attributable fraction of the risk in the exposed is also known as the aetiologic fraction of the risk in the exposed.

See also: **attributable, odds ratio, relative risk.**

Attributable fraction of the risk in the population, aetiologic fraction of the risk in the population Proportion of cases of an event observed in a defined population (region, country, etc.) attributable to exposure of a part of this population to a given factor.

The fraction thus indicates the proportion by which the risk of the event would be reduced in this population were the exposure eliminated. The attributable fraction of the risk (AFR) is calculated by dividing the difference between the risk of the event in the entire population (R_{pop}) and the risk measured in the non-exposed subjects of the population ($R_{non/exp}$) by the risk in the entire population:

$$AFR_{pop} = \frac{R_{pop} - R_{non/exp}}{R_{pop}}.$$

In the frequent case where it is impossible to know the

values of R_{pop} and $R_{non/exp}$ precisely, this calculation requires knowledge of the value of the relative risk (RR) or the odds ratio (OR) associated with the exposure and the proportion of exposed subjects in the population of interest (E_{pop}):

$$AFR_{pop} = \frac{E_{pop}\ (RR - 1)}{1 + E_{pop}\ (RR - 1)},$$

or:

$$AFR_{pop} = \frac{E_{pop}\ (OR - 1)}{1 + E_{pop}\ (\mathrm{OR} - 1)}.$$

In a public health context, this parameter allows estimation of the proportion of cases of a given event or illness which would be avoided by the elimination of a risk factor (e.g. the withdrawal of a drug from the market).

Example: If the relative risk of agranulocytosis associated with the use of a drug is 5 and if 1% of the population is exposed to this drug, the attributable fraction of the risk in this population is:

$$\frac{0.01\ (5 - 1)}{1 + 0.01\ (5 - 1)} = 0.038.$$

This means that 3.8% of the cases of agranulocytosis occurring in this population are attributable to this drug, and would therefore be avoided if exposure to this drug was eliminated.

This type of estimate is only valid if:

- The characteristics of the exposed and non-exposed subjects in the population are identical to those of the subjects who were compared for the relative risk or odds ratio calculation.
- The definition and the conditions of exposure (dose, duration, etc.) are identical to those which were present for the relative risk calculation.

The attributable fraction of the risk in the population is also known as the aetiologic fraction of the risk in the population.

See also: **attributable, attributable number of cases, odds ratio, relative risk.**

Attributable number of cases Number of cases of an event whose occurrence in a population is related to exposure to a given risk factor.

It is calculated by multiplying the risk difference between the exposed and non-exposed by the number of subjects exposed in the population:

$$n_A = (R_{exp} - R_{non/exp}) \times n_{exp}.$$

The attributable number of cases can also be calculated from the reference risk $R_{non/exp}$ (i.e. the risk measured in the non-exposed, or in the general population if the studied factor has a weak impact on this population) and from the relative risk (or the odds ratio) estimated in a cohort or case-control study:

$$n_A = R_{non/exp} \times (RR - 1) \times n_{exp},$$

or:

$$n_A = R_{non/exp} \times (OR - 1) \times n_{exp}.$$

This allows calculation of the number of cases which would be avoided in a population if the given risk factor was eliminated (e.g. through withdrawal of a drug from the market).

Example: If the risk of agranulocytosis during treatment with a drug is 60 per million for a period of one year, while the reference risk is 5 per million per year, the number of cases attributable to this drug, for 300,000 subjects treated for one year, will be:

$$(60/1{,}000{,}000 - 5/1{,}000{,}000) \times 300{,}000 = 16.5.$$

Alternatively, if the relative risk is 12 and the reference risk is

5/100,000, the number of attributable cases to this drug for 300,000 subjects treated for one year will be:

$$5 \times (12 - 1) \times 60/100{,}000 = 16.5.$$

See also: **attributable, attributable fraction of the risk in the population, risk difference.**

Attributable risk *See* **risk difference.**

Attrition Term describing the fact that in a prospective study a certain number of subjects leave the study before the planned end of the follow-up, for various reasons such as death, lost to follow-up, intercurrent diseases, etc.

Attrition is most often expressed as a percentage.

For example, if 129 subjects out of the 4217 included in a cohort study were not followed until the end of the term anticipated in the protocol, the attrition percentage is (129 / 4217) × 100% = 3.06%.

In epidemiology, it is important to control the phenomenon of attrition which can not only lead to a loss of statistical power, in the quantitative sense, but can also introduce significant bias if the attrition is linked to the phenomenon being studied.

See also: **depletion of susceptibles, lost to follow-up.**

B

Background noise Using the broadcasting analogy of radio reception, parasitic information which hinders the recognition of a signal.

This term may apply, for example, to the incidence of an event in the general population or in a reference population. When this noise is high, the substantial number of coincidental associations expected during a follow-up may make it difficult to detect a moderately increased risk associated with an exposure. In the case of surveillance by spontaneous reporting, background noise is formed by all cases (varying greatly in type, seriousness and validity) reported for a given drug.

See also: **coincidental, signal, spontaneous reporting.**

Baseline risk, **background risk** Risk measured in a population not exposed to the factor under study.

This term is often used to designate a risk estimated in the general population when, in fact, it, more accurately represents a reference risk (measured in a population identical to the exposed population in all aspects except exposure status).

See also: **absolute risk, general population, reference risk, risk difference.**

Bayes's theorem First mathematical formulation, by the Reverend Thomas Bayes (1702–1761), of the principle of conditional probabilities allowing calculation of the probability that an event will occur or that an affirmation will be correct under certain conditions (e.g. that a test is positive or

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

negative, or that a symptom is present or absent).

For diagnostic tests, Bayes' theorem can be expressed as follows:

$$p\,(D \mid T^+) = \frac{p\;(T^+ \mid D) \times p\;(D)}{p\;(T^+ \mid D) \times p\;(D) + p\;(T^+ \mid \bar{D}) \times p\;(D)}.$$

In this notation:

- $p\,(D \mid T^+)$ is the probability that the subject is diseased if the test is positive (otherwise known as the positive predictive value of the test).
- $p\,(T^+ \mid D)$ is the probability that the test will be positive if the disease is present (otherwise known as the sensitivity of the test).
- $p\,(D)$ is the *a priori* probability that the subject is diseased before knowing the result of the test (otherwise known as the risk in the population from which the subject is drawn).
- $p\,(T^+ \mid \bar{D})$ is the probability that the test will be positive if the disease is absent, that is, the complement of the test's specificity (1 − specificity).
- $p\;(\bar{D})$ is the *a priori* probability that the subject is not diseased; $p\;(\bar{D})$ thus equals $1 - p\;(D)$.

Bayes' theorem may more easily be expressed as follows:

$$\frac{p\;(D \mid T^+)}{p\;(\bar{D} \mid T^+)} = \frac{p\;(D)}{p\;(\bar{D})} \times \frac{p\;(T^+ \mid D)}{p\;(T^+ \mid \bar{D})}.$$

In this formulation, the odds of presenting the disease given a positive test (*posterior odds*) is equal to the odds of presenting the disease estimated before analysis (*prior odds*) multiplied by the odds that the test will be positive if the disease is present. The multiplier of the prior odds is called the *likelihood ratio*. For diagnostic tests, we have seen above that this can be calculated by:

$$\frac{\text{sensitivity}}{1 - \text{specificity}}.$$

Bayes's theorem has numerous applications in clinical and epidemiological research. Among other uses, it has been applied to causality assessment in individual cases. The principle consists of fixing, before analysis, the prior odds that the drug of interest is responsible for the adverse event observed in a given patient. These odds are 1 if there is no reason to favour the hypothesis of the responsibility or non-responsibility of the drug for the adverse event. In the most favourable cases, prior odds can be estimated from the relative risk (RR) or odds ratio quantified by a previous epidemiological study (cf. aetiologic fraction of the risk in the exposed, EFR_E):

$$\text{Prior odds} = \frac{EFR_E}{1 - EFR_E} = \frac{\dfrac{RR - 1}{RR}}{1 - \dfrac{RR - 1}{RR}} = RR - 1.$$

The prior odds are then multiplied by one or more likelihood ratio(s) corresponding to available information, signs or criteria relevant to the causality analysis. Each likelihood ratio is calculated by dividing the probability of the criterion being present, if the drug is responsible for the event, by the probability of the criterion being present if the drug is not responsible.

Example: If 67% of patients presenting a given adverse effect are female when the proportion of females among all users of a drug of interest is 42%, the likelihood ratio for gender will be 0.67/0.42 = 1.6 if the subject is female and 0.33/0.58 = 0.57 if the subject is male.

The multiplication of prior odds by all the likelihood ratios

relevant for the causality analysis gives the final or posterior odds that the drug is responsible for the adverse event of interest. These odds can easily be transformed into a *posterior probability*:

$$\text{Probability} = \frac{\text{odds}}{\text{odds} + 1}.$$

For example, an odds of 5.6 corresponds to a probability of $5.6/(5.6 + 1) = 0.85$.

See also: **conditional probability, imputability, odds, positive predictive value, probability, sensitivity, specificity.**

Bayesian Referring to an approach based on conditional probabilities formalized by the Reverend Thomas Bayes. *See* **Bayes's Theorem.**

Benefit Benefit usually refers to a gain (positive result) for an individual or a population resulting from an intervention, i.e. a drug treatment.

Expected benefit can be expressed quantitatively, and this would ordinarily incorporate an estimate of the probability of achieving the gain (CIOMS Working Group IV, 1998; *see* References).

Berkson's bias Particular form of admission bias that can falsify an estimate made in a case-control study.

This bias can occur if the cases are identified in a care structure and there is a higher (or a lower) probability that a case will be referred to this structure if exposed to the factor of interest. The strength of the association between exposure and an event can thus be artificially increased (or decreased).

Example: *If, in a case-control study evaluating the association between use of a drug and liver injury, the cases were selected in a hepatology department specializing in liver injury caused by medications, the probability of exposure of the cases to a particular drug may be abnormally high, which may lead to the*

conclusion of a stronger association than that which exists in reality.

The effect would be the opposite if cases of liver injury of unknown cause were preferentially hospitalised. Drugs known to be hepatotoxic could be under-represented in this sample of hospital cases.

See also: **admission bias, case-control study, selection bias.**

Beta error *See* **Type II error.**

Bias Error of reasoning or procedure leading to a false representation of reality.

In the strictest sense, bias is systematic in nature and alters an estimation in a given direction (a synonym is *systematic error*). Bias is thus different from *random error*, which results in a loss of precision but not in systematic deviation in a given direction. A particular kind of bias, called *confounding*, can be taken into account or corrected for during analysis. Other biases, known as selection and information biases, alter the relevance or validity of the collected data and cannot be subsequently controlled for in the analysis.

The principal biases mentioned in this dictionary can be grouped into three categories:

- *Selection biases*: admission bias, Berkson's bias, diagnostic bias, notoriety bias and survival bias.
- *Information biases*: interviewer bias and recall bias.
- *Interpretation biases*: aggregation bias, confounding, prescription bias (confounding by indication), protopathic bias and reverse causality bias.

Binary variable Variable which may only take two values (e.g. 0 or 1, diseased or non-diseased, alive or dead, etc.).

In order to simplify analysis, certain continuous quantitative variables (e.g. weight, height, blood glucose, serum creatinine, etc.) may be transformed into binary variables by fixing a

cut-off value. For example, the subjects of a study can be divided into two groups: those with a body mass index of less than 30 and those with a value of 30 or higher.

See also: **continuous variable, discrete variable.**

Binomial distribution, binomial law Law describing the distribution of the probability of observing x successes over n independent tests, assuming that for each test there is an identical probability p of success:

$$p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n+x}.$$

The probability of observing at least k successes being:

$$p(x \geq k) = 1 - \sum_{x=0}^{x=k-1} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

The outcome of each test must be binary (success/failure, diseased/non-diseased, etc.). The binomial law thus describes a probability distribution that is discrete or discontinuous.

In pharmacoepidemiology, the binomial distribution is used to calculate the probability of observing a given number of events x in a sample of size n; the exposure of each of the n subjects to the drug is considered as an independent test. The binomial distribution assumes that the probability p of the occurrence of the event is identical for each of the n subjects. This hypothesis can be disproved *a posteriori*.

Example: An incidence rate of 5/10,000 measured in a sample of n subjects followed over a given period essentially represents an average, and does not necessarily mean that each subject has five chances in 10,000 of presenting the event during this period. Some of the sample subjects may have a probability of presenting the event close to 1, while for others it could approach 0.

Adjustment for a different probability for each test (if the

risk varies for the different subjects) requires the use of more complex statistical models (e.g. negative binomial distribution, Bayesian approaches).

When the sample size is large ($n > 100$), calculations based on the binomial distribution become difficult because the formula includes the nth factorial. In this case, an approximation may be used:

- If the expected number of events np is also large (in practice, equal to or greater than 15), the distribution obtained with the binomial law hardly differs from a normal distribution of mean np and standard deviation $\sqrt{np(1-p)}$, making it possible to use the usual calculations based on the latter distribution.
- If n is large (> 100) and p is small (in practice, less than 0.1), the Poisson distribution provides a very good approximation.

See also: **binary variable, expected number, normal distribution, Poisson distribution.**

Bioequivalence *See* **equivalence.**

Blinding Ignorance of the fact or nature of exposure of a subject or group, which is planned and organized in the protocol of a study, in order to avoid the direct or indirect influence of such knowledge on the planned evaluation.

A standard use of blinding is in controlled clinical trials. In a single-blind clinical trial, either the patient or the researcher does not know the type of treatment used (i.e. the drug of interest or an active or inactive referent); in a double-blind clinical trial, neither the researcher nor the patient know the nature of the treatment received by the latter; in a triple-blind clinical trial, in addition, the statistical analysis is conducted before the blinding codes are opened, thus without knowledge of the nature of the compared treatments.

Implementation of blinding generally requires the preparation of pharmaceutical forms specific to the trial, perfectly identical in their presentation, whose only difference (which should be undetectable during the study) is the nature of the ingredients (e.g. active or inactive). The treatments are identified by a code known only to the organizer of the trial, which is opened at the end of the study period, or before if required by safety concerns.

Blinding is also used in epidemiology, for example during the validation leading to the inclusion of a case in a case-control study, to avoid the situation in which knowledge of the subject's exposure status influences the judgement of the experts.

See also: **active ingredient, clinical trial, controlled trial, placebo, placebo effect.**

Bradford-Hill's criteria *See* **causality.**

C

Capture-recapture Method whereby the size of a population can be estimated from two or more random and independent samplings from this population.

	Subjects captured by the second sampling	Subjects not captured by the second sampling	Total
Subjects captured by the first sampling	a	b	R_1
Subjects not captured by the first sampling	c	d	
Total	R_2		N

In this case, the total population size, N, is estimated by the product of the sample sizes divided by the number of subjects found in both samples:

$$\frac{R_1 \times R_2}{a}.$$

The confidence interval (CI) for estimating N is (using the normal approximation):

$$I_C = N \pm Z_{e-\alpha} \sqrt{\frac{R_1 \times R_2 \times (R_1 - a) \times (R_2 - a)}{a^3}}$$

The validity of the estimate requires that the samples be truly random and independent and that their sizes be not too small with respect to the total population size (otherwise, the expected

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

number of shared subjects would be near zero, making estimation impossible or very unstable).

This method was first applied (Laplace) to counting animal populations: capture–mark–release–recapture. It was used secondarily in epidemiology to estimate the total number of cases of a disease in a population.

Example: To estimate the number of cases of an infectious disease, reports to two independent surveillance systems operating in the same region were used; 127 cases were identified by the first system and 42 by the second; 12 cases (duplicates) were found by both systems. Thus the total number of cases in the population is:

$$\frac{127 \times 42}{12} = 444.$$

The two-sided 95% confidence interval for the number of cases is:

$$444 \pm 1.96\sqrt{\frac{127 \times 42 \times (127 - 12) \times (42 - 12)}{12^3}}$$
$$= [242; 646].$$

The capture–recapture approach has also been applied to spontaneous reporting, to attempt to estimate the total number of cases of an adverse event and thereby quantify the extent of under-reporting. In this case the numbers of cases reported to collection systems are used (e.g. medicines agency and manufacturer). However, the validity of this approach requires that reporting to one system or the other be a truly random and independent phenomenon, which is rarely the case in practice.

See also: **random, spontaneous reporting, under-reporting.**

Carry-over effect Error consisting of not attributing an event or effect to an earlier exposure whose influence on the organism persists despite cessation of the exposure.

The carry-over may be due to the persistence of significant concentrations of the drug in the organism or to lasting

biological modifications induced by the treatment. Carry-over effect is of particular concern in studies which compare the effect of several successive treatments, as well as in studies or trials in which the subject is his or her own control (crossover studies or trials).

This phenomenon can be avoided by separating treatment periods being compared by sufficiently long intervals without treatment (*washout periods*).

See also: **crossover trial, washout.**

Cart versus horse bias *See* **reverse causality bias.**

Case Any person or event presenting the characteristic(s) defined as the object(s) of interest of a study.

See also: **case definition.**

Case-by-case causality assessment *See* **imputability.**

Case-cohort study Epidemiological design in which the population of subjects having presented an event during the follow-up of a cohort is compared, with respect to an exposure at baseline, with a control population chosen at random from the whole population of subjects present at the beginning of this period (*see* Appendix 1).

Because of this mode of selection, the control population may include one or more subject(s) who will eventually become cases.

Case-cohort studies differ from nested case-control studies in that controls, in the former, are not matched to the cases within the cohort but selected randomly at the beginning of the follow-up. This makes statistical analysis more complex but has the advantages of facilitating control selection (since there is no prior matching), and allowing use of the same control population for other comparisons within the cohort.

Under certain conditions, the case-cohort approach can be applied to a very large population (e.g. the inhabitants of a region). This generalization is possible when we can: (1) identify all the incident cases of a disease in the region between dates t_0 and t_j, (2) randomly select a control population for the subjects

residing in the region at date t_0, and (3) ascertain that the cases identified during follow-up correspond to subjects who were present in the population at t_0.

See also: **case-control study, case-population study, cohort study, nested case-control study.**

Case-control study Epidemiological design comparing previous exposure to a risk factor of interest (e.g. use of a drug) or the presence of a characteristic in a group of subjects presenting a given event (the *cases*), to that in a group not presenting this event (the *controls*) (*see* Appendix 1).

The two groups may differ in size but, except for the presence of the event of interest, must be as similar as possible with respect to the main factors that could influence the probability of exposure to the risk factor(s). Comparability could theoretically be guaranteed by random selection of cases and controls from the same source population, but is generally attempted either through *a priori* matching, and/or *a posteriori* adjustment or stratification, considering potential confounding variables.

In order to obtain sufficient statistical power, it may be necessary to match each case to several controls, if the prevalence of exposure to the studied factor is low in the source population.

The tested null hypothesis is that there is no association between exposure to the factor of interest and the occurrence of the event defining the cases. Under this hypothesis, we expect that the proportion exposed or the odds of exposure to this factor will be the same in the cases and controls. If this is not the case (i.e. the odds ratio differs significantly from 1), we conclude that there is an association, although not necessarily a causal one, between exposure and occurrence of the event.

Thus, although a case-control study does not permit direct estimation of risk, the value of the odds ratio quantifies the strength of the association between the exposure and the occurrence of the event. Furthermore, in the absence of bias, and

if the event of interest has a low probability of occurrence, the odds ratio is a good approximation of the relative risk linked to exposure (which could otherwise only be calculated with a cohort study).

In pharmacoepidemiology, case-control studies are particularly useful in measuring the association between use of a drug and an adverse event of low probability and/or delayed occurrence. In such cases, a cohort study would require prolonged follow-up of a very large population.

See also: **adjustment, bias, community, confounding, hospital control, matching, odds ratio, population control.**

Case-crossover study Epidemiological design to evaluate a possible association between an exposure and the occurrence of an event by comparing the number of cases arising within and outside a previously defined window of exposure, in a population whose exposure status changes over time (*see* Appendix 1).

This method is applicable only under the following conditions:

- The follow-up of the population ensures the inclusion of all the cases arising during the study period.
- The exposure status is precisely known over the entire follow-up period.
- The exposure status of subjects changes during the follow-up.
- Under the hypothesis of the existence of a causal relationship between exposure and event and with reference to an occurrence mechanism, it is possible to define an exposure window during which an event induced by this exposure ought to occur.
- This time-window is short with respect to the duration of the follow-up of the cases.
- The *a priori* risk of the occurrence of the event is constant during each exposure window.

If these conditions are fulfilled, we compare the number of events occurring within the exposure window (a_E in the table below) with

the number of cases expected under the hypothesis of independence (that is, $a \times t_E/t$):

	Cases	Follow-up time
Within the exposure window	a_E	t_E
Outside the exposure window	a_{NE}	t_{NE}
Total	a	t

The relative risk can be estimated by the ratio of the observed number to the expected number, that is:

$$\frac{a_E \times t}{a \times t_E}.$$

Example: The surveillance of a population having received two injections of a new vaccine in 1 year has resulted in the identification of 22 cases of multiple sclerosis (MS). It is postulated that a vaccination can only cause the appearance of MS in the 4 weeks following the injection of a vaccine dose (first or second injection). Among the 22 identified cases, six arose during this time-window.

The total of the risk periods for the 22 subjects (for the two injections) is $22 \times 2 \times 4 = 176$ *weeks; the duration of follow-up outside the risk period is therefore* $(22 \times 52) - 176 = 968$ *weeks.*

On this basis, the relative risk is estimated by:

$$\frac{6 \times (176 + 968)}{22 \times 176} = 1.77.$$

The advantage of the case-crossover approach lies in the limiting of information gathering to only those cases appearing in the population, whatever its size. However, its practical application comes up against the above-mentioned conditions, which limit its field of application to acute exposures and events, and especially against the difficulty of defining an adequate time-window, whose value has considerable influence on the estimate of the relative risk.

See also: **expected number, time-window.**

Case definition Set of criteria which permit operational identification of the events investigated in a study.

They should be precise, unambiguous, repeatable, reproducible, relevant and pertinent to ensure the validity of measurements (e.g. an incidence rate). Case definition can be used to confirm case status in case control studies, to ascertain outcomes in cohorts or clinical trials or in the framework of spontaneous reporting. Typically, potential cases are assessed by an expert committee blind to exposure status using these criteria.

For example, thrombocytopenia can be defined as a platelet count below 100 g/l, in two measurements at least 1 week apart, with haemoglobin above 6.2 mmol/l and polynuclear neutrophils above 1.5 g/l in patients not treated with cytotoxic anticancer agents.

This definition should be established prior to the beginning of the study.

See also: **case.**

Case non-case study Method of analysing a spontaneous reporting pharmacovigilance database using internal comparisons to investigate a possible association between the exposure to a drug and the occurrence of an event (quantified by the calculation of a *relative reporting ratio*).

This method, which cannot be a substitute for classical epidemiological studies (i.e. cohort, case-control), gives relatively reliable results if the database contains an almost infinite number of different drug–event combinations.

For a given event A and drug 1, the records in the database are set out as follows:

	Event A	Other events
Drug 1	*a*	*b*
Other drugs	*c*	*d*

Under the null hypothesis of an absence of a specific association

between reporting of event A and drug 1, the odds of exposure to drug 1 are expected to be identical among the cases (presenting event A) and the non-cases (presenting another event); if this hypothesis is correct, we expect that a/c and b/d do not differ significantly at the fixed error level.

The relative reporting ratio $\frac{a/c}{b/d} = \frac{ad}{bc}$ should thus not deviate significantly from 1.

The addition to the traditional biases of the various reporting biases makes the case non-case method mostly useful as a generator of hypotheses and signals.

See also: **database, odds ratio.**

Case-population study Case-cohort approach conducted within a cohort made up of the total population of a geographical area (*see* Appendix 1).

The characteristics of cases occurring in the population (notably, their possible previous exposure to a given risk factor) are compared to those of the entire set of the subjects in this population (supposing that demographic, health or drug-use statistics for the whole population are available).

The validity of this approach assumes the existence of a surveillance system capable of identifying all the cases of an event within the population.

The null hypothesis is that, in the absence of an association between exposure and event, the odds of exposure are identical among cases and the rest of the population, barring sampling variations.

Example: To discover whether the use of a drug increases the risk of hip fracture in the elderly, all the subjects over 65 years old residing in a given geographical area are studied. If, all the cases of hip fracture occurring in this region over a given period can be identified, and drug utilization in this population is known, it would be possible to compare the odds of exposure to the drug in the cases

with the odds in the entire population over 65 for the same period.

If the odds is significantly greater than 1, it will be concluded that there is an association (causal or not causal) between the use of the drug and the risk of hip fracture.

When the comparison makes use of global population statistics, the absence of individual data hinders the control of bias, notably confounding.

See also: **aggregated data, case-cohort study, case-control study, nested case-control, odds.**

Case report The description of the clinical and laboratory, etc. data concerning the occurrence of an event in a single patient. Data elements for such regulatory transmission have been defined in International Conference for Harmonization (ICH E2B; *See* References).

See **reporting**, International Conference for Harmonisation.

Causality Aetiologic link between exposure (e.g. use of a drug) and the occurrence of an event or disease:

$$F \rightarrow E.$$

Analysis of causality consists of answering the question: *Is the factor F the cause of event E, or would it be*?

In all cases this assumes that the exposure preceded the occurrence of the event (*see* Appendix 2).

Causality may be analysed at the individual level (i.e. regarding imputability: *Did the drug of study cause the event observed in this patient*?) or at the level of the population (*Can or will the use of this drug increase the risk of occurrence of a given event*?).

A factor is called a '*necessary cause*' if the occurrence of the event requires its presence (i.e. the factor will be found in all the subjects having presented the event, but its presence does not inevitably result in the occurrence of the event). Conversely, a factor is called a '*sufficient cause*' if its presence inevitably causes the event to happen (in which case all those exposed to this factor

will present the event, although this does not necessarily mean that the factor will be found in all the subjects having presented the event). A factor is a '*necessary and sufficient cause*' if the occurrence of the event requires its presence and it is found in all cases of the event.

Much more common (especially in pharmacoepidemiology) is multifactorial causality: several factors are identified as, independently or not, increasing the risk of occurrence of the event, without any of them being necessary or sufficient causes on their own.

In the strictest sense, only a controlled trial with random allocation of treatment or exposure, and where a statistically significant association between exposure and the occurrence of an event is observed, can allow us to conclude to a causal relationship. In observational epidemiology or pharmacoepidemiology, the absence of an experimental design limits the validity of causal inference, notably because of the possible existence of confounding. According to the criteria proposed for infectious diseases by Sir Austin Bradford-Hill, a causal association is more likely if:

- The main biases (confounding, selection and information biases) which may have distorted the results have been ruled out or taken into account.
- The association between the factor of interest and the occurrence of the event, quantified by the relative risk or the odds ratio, is strong (by convention, greater than 3).
- This association is statistically significant. This requires that the probability of falsely concluding an association between exposure and event (the Type I error) is as low as possible (e.g. less than 5%); this is the case when the 95% confidence interval around the estimated value of the relative risk or odds ratio does not include 1.
- There is a relationship between intensity of exposure (with

respect to dose or duration) and the severity or the frequency of the event of interest.
- Other studies, preferably based on different designs, have also found this association.
- The relationship between exposure and event is plausible and coherent (biological, pharmacological or clinical plausibility).
- The factor of interest has one or more point(s) in common with a known and well-established aetiology of the event or disease.

See also: **bias, confidence interval, imputability, observational study, odds ratio, relative risk, Type I error.**

Causality assessment method Formalized procedure whose aim is to estimate the degree of plausibility of the conclusion that the use of a drug is the cause of an adverse event in a given patient.

See also: **algorithm, Bayes's theorem, causality, imputability.**

Censoring Absence of information, in a prospective study, about the fate of a certain number of subjects in the studied population.

Censoring concerns subjects who had not presented the studied event by the time they were no longer followed, since it is impossible to know whether or not they did later. For these subjects, the information is considered to be censored by the absence of a complete or longer follow-up.

Censoring may be the result of various phenomena: loss to follow-up, accidental or natural death unrelated to the studied event, late inclusion of the subject, early termination of the study, etc.

For example, we are studying the 5–year recurrence rate among 500 subjects receiving a new anti-cancer treatment; 220 subjects were still alive and without recurrence at the end of 5 years, 148 had a recurrence during the follow-up and 132 were not followed for the planned 5 years, but had not presented a recurrence during

the course of their incomplete follow-up. In this case it is impossible to know whether or not subjects in the last group would have presented a recurrence during the remaining part of the 5–year interval.

By extension, censoring may also apply to lack of information about a parameter of interest before a given date.

For example, in a database study of the effects of an anti-inflammatory treatment among patients suffering from rheumatoid arthritis, it is not known whether the patients treated with drug A were, or were not, treated with a drug B before the beginning of information gathering. This information is said to be censored for these subjects.

The terms 'left-censoring for information unavailable' before the beginning of a follow-up and 'right-censoring for information unavailable' after its termination are sometimes used. It may or may not be opportune to consider the available part of censored information in the final analysis, depending on the situation.

See also: **follow-up, intent-to-treat analysis, lost to follow-up.**

Chance Possibility of random occurrence.

Chance can be quantified with a probability, whether arbitrarily fixed or estimated by a calculation.

Example: A probability of 0.8 means that an event has 80 chances in 100 of occurring during a given period of time.

Until the middle of the nineteenth century, the old French word *chéance* (derived from the Latin verb *cadere*, to fall) meant the manner in which dice fall and, by extension, games of dice and chance.

See also: **coincidental, probability, random.**

Chronic Adjective designating an exposure or an illness which can persist for a certain period of time.

Chronic, in this sense, is the antonym of *acute*.

A chronic treatment is the regular and prolonged use of a drug by a subject. In the same way, a liver injury is said to be chronic if

biological anomalies persist for more than 3 months.

See also: **acute.**

Chronopharmacology Study of the variations of the effects or the pharmacokinetics of drugs according to the time of administration.

The efficacy of a drug may be increased if it is administered or taken at a particular time. The risk–benefit ratio of certain drugs can be optimized by considering chronopharmacological data, notably in oncology.

See also: **effect, pharmacokinetics, risk–benefit ratio.**

Clinical Having human beings as subjects (e.g. clinical pharmacology, clinical studies, clinical trials, clinical epidemiology, etc.).

Clinical trial Any investigation in human subjects intended to discover or verify the clinical, pharmacological and/or other pharmacodynamic effects of an investigational product(s), and/or to identify any adverse reactions to an investigational product(s), and/or to study absorption, distribution, metabolism and excretion of an investigational product(s) with the object of ascertaining its safety and/or efficacy. The terms 'clinical trial' and 'clinical study' are synonymous (ICH E6; *see* Reference).

Closed population Population (e.g. a cohort) whose members are not renewed during the study period.

Its size thus tends to decrease over time (due to deaths, loss to follow-up, etc.) and the average age of its participants tends to increase. The definition does not necessarily imply that all subjects are included at the same time, but that there is no recruitment once the pre-ordained sample size has been reached. In this sense, 'closed' is often synonymous with '*fixed*', although the latter may be used in other contexts.

See also: **cohort, dynamic population, fixed.**

Cluster, clustering Higher than expected incidence of an event in a given region (spatial cluster or place cluster), during a given period (temporal cluster or time cluster) or both (space–time or time–place cluster).

Cohort Group of subjects selected according to one or more common characteristic(s) and followed over time in order to identify, describe or quantify an event (*see* Appendix 1).

In pharmacoepidemiology, the subjects are usually identified according to their exposure to a drug (or class of drugs), and the event or outcome exposure of interest can be a therapeutic effect, an adverse effect, a behaviour or any other criterion relevant to the evaluation of the effects induced by the exposure. The term 'cohort' does not imply a given design or sample size: the study can include a small or large sample, with or without a control group (the control group being formed randomly or non-randomly, and with or without matching). In pharmacovigilance, the term 'cohort study' most often refers to a post-marketing observational study (PMOS), an observational study—without a control group—which is carried out following the market launch of a drug.

Most PMOSs are multipurpose studies; they should comply with regulatory guidelines.

The population of a cohort is said to be 'fixed' if new subjects are not included during the study, and 'dynamic' if they are. The length of follow-up can be identical for all the members of the cohort (cohort with fixed follow-up), or variable. In the latter case, a calculation using incidence density can be useful to take into account the varying contribution of individual subjects to the total follow-up time.

See also: **closed population, cohort study, epidemiological study, incidence density, person-time.**

Cohort effect, generation effect Difference(s) observed between two or more groups of individuals with respect to state of health (e.g. the incidence rate of a disease) or exposure status, linked to the fact that these individuals were not born during the same period (i.e. do not belong to the same generation).

A cohort effect thus reflects the influence of changes over time on the population containing the groups (e.g. in environment,

quality of life, state of hygiene, the health care system, etc.).

For example, the rate of cardiac valvulopathy is higher in the cohort of individuals born between 1940 and 1950 than in that of individuals born between 1960 and 1970. We know that the primary cause of valvulopathy is rheumatic heart disease during a childhood streptococcal infection, and that the treatment of childhood tonsillitis with penicillin was much more systematic after 1960.

Similarly, the term 'cohort effect' is used to denote the change in the risk of occurrence of an event during a prolonged follow-up of a non-renewed population.

See also: **cohort, depletion of susceptibles, stratification.**

Cohort study, follow-up study, prospective study Epidemiological study based on the follow-up of one or more cohort(s).

The objective of following a single cohort is generally to describe or quantify a phenomenon (i.e. a descriptive study). Comparative analysis of two or more cohorts seeks to explain the mechanism of or attribute a cause to the occurrence of an event (i.e. an analytic or aetiologic study).

See also: **clinical trial, cohort, epidemiology.**

Coincidental, fortuitous Happening by chance, in an unforeseen manner.

This term usually designates an event or an association whose occurrence at a given moment in the period of interest is due only to chance, having no association or causal link with the factor under study.

See also: **association, causality.**

Community, community control All subjects in a region living in their usual settings (i.e. not hospitalised).

Thus, in a case-control study in which the cases are recruited and interviewed at a hospital, the *community controls* will be the non-diseased subjects, recruited and interviewed outside the hospital setting. The choice between *hospital controls* and

community controls is difficult and often a source of debate. Ideally, the control for a hospitalised case should be a person from the same source population, who has the same probability of being exposed to the study factor and of being hospitalised in the same care structure (hospital or other service) if he or she had presented the event (disease or symptom) identifying the case. If the study concerns an acute disease arising outside the hospital setting, the source population for cases is the extra-hospital population (hospitalisation being justified only by the presence of the disease); in this case, the choice of community controls seems preferable. It would be different if the study concerned a disease arising during hospitalisation (the source population then being the set of patients hospitalised in the system).

The choice of type of controls is very important, as it can greatly influence the level of exposure to the factor under study and thus the odds ratio to be calculated.

For example, if we are interested in a drug used by 3% of the population outside the hospital and by 22% of the hospitalised population, and our investigation shows that 15% of the cases are exposed, the crude odds ratio will be 5.7 when comparing cases with community controls and 0.6 if we choose hospital controls.

The difficulty of settling this question leads some investigators to match several controls to each case (e.g. two hospitalised controls and two community controls to one case).

The term 'population control' is also used. These may thereby be selected from among the patients of the doctor(s) who followed the cases or from the same population sub-group (e.g. neighbours, close relatives, etc.).

Some researchers consider hospital outpatients (who are not hospitalised) as population controls.

See also: **case-control study, control, odds ratio, population control, source population.**

Comparative trial *See* **controlled trial.**

Completeness of data, consistency of data In pharmacovigilance, the informative value of collected data.

This depends on the quantity and quality (in terms of relevance and validity) of the data available.

For example, incompleteness of data precludes imputability analyses, as it is impossible or even hazardous to investigate a causal association between the use of a drug and the occurrence of an event in poorly-documented cases.

See also: **imputability.**

Compliance Respect, by a patient, of the directions for use of a therapy (medicinal or otherwise) he or she is meant to obey.

These are generally conveyed by the doctor giving the prescription, although some authors extend the definition to include recommendations made by the pharmacist dispensing the drug or those provided in the package insert. A patient is said to be *compliant* if his or her behaviour does not vary significantly from what is expected, and *non-compliant* if this is not the case. Although at times difficult to assess, awareness of a subject's compliance is fundamental to characterizing the exposure of a population and the interpretation of some therapeutic failures or adverse events.

Concurrent Term used to indicate a prospective, retrospective or cross-sectional approach to investigate events occurring after the beginning of the study. The term 'historical', on the other hand, indicates an approach using previously recorded data.

The term 'prospective' should not be used as a synonym for concurrent, being reserved to indicate a study plan based on the follow-up of a population.

See also: **directionality, historical, prospective.**

Conditional probability Probability whose value is modified according to whether a given fact or result is taken into account.

Example: What is the probability that the subject uses benzodiazepines, knowing that he or she is (or is not) being treated

with antidepressants?

The theory of conditional probabilities was formalised by Bayes's theorem, in which the *prior probability* (estimated before analysis of a file, case or situation) is multiplied by one or more *likelihood ratio*(*s*) calculated from available data.

See also: **Bayes's theorem, probability.**

Confidence Probability that an estimate, statement or conclusion reflects reality.

In the strictest sense, certainty corresponds to a confidence level of 100%. A confidence level of 95% means that a conclusion has 95 chances in 100 of being correct, or that we accept, at the most, a 5% risk of error.

See also: **confidence interval, Type I error.**

Confidence interval When measuring the value of a parameter in a sample, the set of values having a given probability (called confidence) of including the real but unknown value (which could be measured in the whole population from which the sample was drawn).

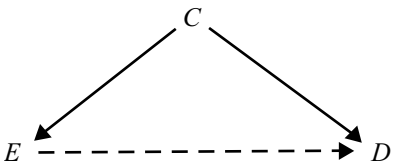
This calculation is inferential in that one concludes, at a fixed confidence level, that the real value of the parameter lies between the two limits of the interval. A confidence interval is called two-sided or two-tailed if both limits have been calculated [e.g. (0.025, 0.17)] and one-sided or one-tailed if only one limit has been calculated, the other being fixed *a priori* [e.g. (0, 0.07) or (12.9, $_{\infty}$)].

Example: If four cases of an event have been observed during the follow-up of 1800 persons treated with a drug, the two-sided 95% confidence interval calculated with the Poisson distribution around the observed proportion, 4/1800 or 0.002, lies between 0.0006 and 0.006. This means that the real risk of presenting this event during treatment has 95 chances in 100 of being between 6 and 60 per 10,000.

See also: **confidence, inference.**

Confounding, confounding variable, confounder Systematic error

resulting from the fact that a secondary variable (called the *confounder* or *confounding variable*) is linked both to the exposure and to the event of interest, which can wholly or partially explain their association found in an epidemiological study:



Thus, referring to the notation above, we falsely conclude to a direct and causal association between *E* (exposure) and *D* (disease) when they are in fact linked by the intermediate third variable acting as a confounder, *C* (*see* Appendix 2). Failure to take possible confounding variables into account can strongly bias the estimate of the association between exposure and event.

Example: In the 1980s, a study found a strong association between use of oral contraceptives and risk of malignant melanoma. It subsequently became evident that the women who used oral contraceptives, being younger, exposed themselves to the sun more often than non-users. Adjustment for duration of sun exposure substantially decreased the value of the odds ratio quantifying the strength of the association between oral contraceptives and malignant melanoma. Sun exposure, linked both to the probability of oral contraceptive use and the occurrence of malignant melanoma, acted as a confounder in this study.

Only the use of a controlled trial with randomisation allows complete control of confounding, by distributing the known and unknown potential confounders in a balanced manner between the groups. In observational epidemiology we can, at best, aspire to control the principal factors suspected to be confounders with matching or adjustment.

See also: **adjustment, matching, Simpson's paradox.**

Confounding by indication, prescription bias, prescription channelling Particular instance of confounding, leading to a distorted estimate of the association between the use of a drug and the occurrence of an event, because a given drug (or class of drugs) is preferentially prescribed to subjects who have, *a priori*, a higher or lower risk of presenting the event in question.

For example, a new non-steroidal anti-inflammatory presented as being better tolerated by the digestive system could be the cause of a particularly high number of gastrointestinal haemorrhages if it is preferentially prescribed to subjects at risk for this outcome.

Conversely, a drug correctly or incorrectly reputed to be poorly tolerated could be preferentially prescribed to subjects with a lower risk of presenting this effect.

See also: **confounding.**

Confounding ratio Ratio of the crude (unadjusted) value of a relative risk or an odds ratio to the value of this parameter after adjustment by a factor considered to be a confounder.

The value of this ratio quantifies the extent of the confounding by the factor (a ratio of 1 indicating the absence of confounding).

Example: A cohort study comparing the risk of venous thrombosis associated with two types of oral contraceptives has resulted in a relative risk estimate of 3.2. After adjustment for the duration of contraception, the value of this relative risk is 1.7. The confounding ratio, equal to 3.2/1.7 = 1.9, indicates that the duration of the contraception acts as a confounder in this study.

See also: **adjustment, confounding, crude.**

Confounding variable *See* **confounding.**

Consistency of data *See* **completeness of data.**

Constraint Any element imposed by a third party (usually the sponsor of a study) and planned beforehand (i.e. formalized by a protocol), capable of modifying usual care (e.g. regarding treatment allocation, type of surveillance or follow-up, benefits gained or risks incurred).

The existence of a constraint should entail seeking a subject's informed consent.

See also: **protocol, sponsor.**

Contingency table Method of distributing the subjects at the conclusion of a study investigating a possible association between two categorical variables, or two variables made categorical (e.g. the exposure to a risk factor and the occurrence of a disease).

The most classical distribution method (for a study of the association of two dichotomous variables) is the four cell table, or 2 × 2 (two by two) table, of the type:

	Diseased	Non-diseased
Exposed	*a*	*b*
Non-exposed	*c*	*d*

In a cohort study, the exposed and non-exposed subjects are distributed according to whether or not they presented the studied event during the follow-up.

In a case-control study, the diseased and non-diseased subjects are distributed according to whether or not they were previously exposed to the studied factor.

We can thus investigate, by means of a test statistic, whether the distribution of the number of subjects in the four cells is compatible with the null hypothesis that exposure and disease are independent.

See also: **case-control study, cohort study, odds ratio, relative risk.**

Continuity correction Correction applied to a test statistic when a continuous probability distribution is used to approximate a discrete distribution (e.g. when a test based on the normal or Chi-square distributions is used to compare the distribution of a characteristic or number of events in two small samples).

A continuity correction tends to make the test more conservative by decreasing the probability of falsely rejecting the tested null hypothesis. Continuity corrections are always makeshift, and it

is preferable to use a test based on the appropriate probability distribution (e.g. Fisher's exact test, to compare a distribution when one or more number(s) are small).

See also: **null hypothesis.**

Continuous variable Variable which can take an infinite number of values, within certain limits.

Weight, height, blood pressure, and blood glucose level are examples of continuous variables.

See also: **discrete variable.**

Contributing factor Factor whose presence is likely to increase the probability of the occurrence of an adverse effect.

This term is most often used in pharmacovigilance. In epidemiology, the terms '*risk factor*' or '*effect modifier*' are generally preferred.

See also: **interaction, risk factor.**

Control, control group Subject, or group of subjects, used as a reference in a comparison. Ideally, the controls are drawn from the same source population as the cases under study and differ from the latter only with respect to the absence of the factor whose influence is being measured.

See also: **case-control study, community, hospital control, population control, reference population.**

Controlled trial, comparative trial Prospective clinical trial in which a medication or therapeutic strategy is compared to a reference (control) (*see* Appendix 1).

The object of the comparison can be therapeutic efficacy, tolerance or any other criterion (pharmacokinetic parameters, cost, quality of life, etc.). The reference used can be an inert substance (placebo) or another strategy—medicinal or non-medicinal—recognized as being efficacious.

The method may employ a parallel group or a crossover strategy; in either case, the allocation of treatments may be randomised. Similarly, the nature of this allocation may be

unknown to the patient, the investigator, or both (single-blind, double-blind).

See also: **blinding, control, parallel groups, placebo, trial.**

Correlation *See* **association.**

Cost-benefit study, cost–benefit analysis, cost–benefit Study aimed at comparing the costs and the consequences of a therapeutic strategy, expressed in monetary terms.

Example: The cost of a vaccination campaign can be compared with the cost of treating the cases of the disease which would be or were avoided by vaccination.

See also: **cost–effectiveness study, pharmacoeconomics.**

Cost–effectiveness study Study aimed at linking the cost of a strategy to one or more indicator(s) of effectiveness expressed in physical and not monetary terms. The cost is expressed in units of result (e.g. cost per year of life gained, cost per complication avoided, etc.). The term 'cost–effectiveness' can be considered a synonym of efficiency.

See also: **cost–benefit study, effective, pharmacoeconomics.**

Cost–utility study, cost–utility Study aimed at comparing the cost of a strategy with one or more indicator(s) of value, taking into account both quantitative effects (e.g. improvement in life expectancy) and qualitative ones (e.g. improvement in quality of life).

Cox model, proportional hazards model Mathematical model used in survival analyses, allowing prediction of the probability that an event will happen to a subject at an instant t, knowing that he or she was previously unaffected, as a function of the value of a given number of variables $x_1, x_2, ..., x_j$; these variables act on the risk in a multiplicative fashion, and are independent with respect to time.

For example, if $\lambda_0(t)$ *is the incidence rate to be measured at the instant t for the null values of the variables* x_1 *and* x_2, *the incidence rate at t for the given values of* x_1 *and* x_2 *is predicted by:*

$\lambda(t) = \lambda_0(t)\ e^{\beta 1\ xi + \beta 2\ x2}$.

It is evident that when comparing two groups (assuming that they differ only by the value of a single explanatory variable: $x_1 = 1$ or 0), the instantaneous risk ratio at each interval remains constant, at $e^{\beta 1}$.

See also: **logistic model.**

Crossover trial, crossover study Controlled study where the exposure status of the subjects changes over time; each subject being successively exposed and then not exposed (or the reverse) becomes his or her own control (*see* Appendix 1).

The change of status can be imposed by the protocol (i.e. a crossover trial) or observed without intervention (i.e. an observational crossover study).

In a clinical crossover trial, the state of the subjects is compared under two treatments, A and B, possibly repeated and sometimes separated by a period without treatment (*washout period*). It is preferable to randomly allocate the treatment order to minimise the effect to the natural evolution of the studied phenomena (e.g. spontaneous improvement or worsening of the illness, influence of climate or season, etc.), or a carry-over effect from one treatment to the other.

Crossover studies are also widely used in pharmacoepidemiology. They are divided into two types of approaches:

- Those which consider the entirety of an exposed population by comparing the number of cases occurring during the exposure period to the number of cases occurring during an identical follow-up period, before or after the exposure. This, for example, is the principle of on–off studies or Prescription Event Monitoring in the UK.
- Those which only consider the subjects in the population that have presented the studied event (*case-crossover study*).

The advantages of the crossover approach include the savings

incurred by not having to collect a true comparison group, and a theoretically perfect matching, since the subjects being compared are identical by definition. This latter advantage is often diminished by the influence of temporal tendencies and co-factors (often difficult to take into account), which may—sometimes strongly—modify the probability of presenting the studied event over time.

See also: **case-crossover study.**

Cross-product ratio Synonymous with odds ratio.

See **odds ratio.**

Cross-sectional study, prevalence study Study in which the prevalence of a variable (e.g. exposure, an event, a disease) is measured in a population at a given moment; this can also be termed a prevalence study (*see* Appendix 1).

In pharmacoepidemiology, cross-sectional studies can be used to measure, for example:

- The prevalence of a disease or an event in a population.
- The prevalence of exposure to a risk factor such as the use of a drug.

Example: A study consists of interviewing 2400 people chosen at random from electoral lists on a given day; 32 of them used a given drug on this day. The prevalence of use of the drug in this region is 1.3%.

In a cross-sectional study there is, by definition, neither follow-up nor analysis of past events. It is nevertheless possible to study the potential association between an event and a characteristic (e.g. an exposure); however, the absence of temporal analysis of the exposure/event relationship can lead to errors in interpretation, such as *reverse causality bias* (also known as '*cart versus horse bias*').

For example, in a cross-sectional study of a population, we find greater use of aspirin in hypertensive subjects than in those with normal blood pressure. This difference being statistically significant

after adjustment for age and gender, we conclude that aspirin use is a risk factor for arterial hypertension. In fact, the use of aspirin by hypertensives can be partly explained by preventive use, since stroke can be induced by systemic hypertension. An approach based on the analysis of the chronological exposure/event sequence would have prevented this interpretative error.

See also: **epidemiological study, prevalence, prevalence rate, reverse causality bias.**

Crude Adjective denoting a value resulting from a calculation in its simplest form, which does not take into account the possible influence of certain factors that may be present in the studied population, and that may have biased the measurement or rendered it inapplicable to a differently composed population.

In the calculation of an incidence rate, the standardization method makes it possible to eliminate the effects of certain factors (e.g. age) which are distributed differently in the study and comparison populations. Similarly, in a cohort study or a case-control study, we speak of a crude relative risk or odds ratio when the result does not take into account certain potentially confounding factors.

See also: **adjustment, confounding, confounding ratio, standardization.**

D

Database Set of logically-connected data or files accessible with specialized software.

The data can be classified according to geographical origin (*population database*), by exposure to a risk factor (e.g. the subjects treated with a given drug), by the presence of a disease, etc.

Large population databases (such as the General Practice Research Database (GPRD) in the UK, which gathers information on 4 million people collected from 550 medical practices) are very useful sources of observational data for pharmacoepidemiology.

Days of treatment *See* **incidence density, person-time.**

Decision tree *See* **algorithm.**

Defined Daily Dose (DDD) Dose of a drug for 1 day of treatment under standardized conditions.

The DDD is determined by the WHO Collaborating Centre for Drug Statistics Methodology (*see* References), from the approved dosages in different countries and validated by an international expert panel, with possible revisions over time. The DDD is part of the ATC/DDD scheme for international comparison of drug utilization.

DDD can thus differ from RDD (recommended daily dose), as found in the product authorisation and the Summary of Products Characteristics (SPC), and from PDD (prescribed daily dose) obtained from the observation of real prescriptions.

Example: If the defined daily dose for an anti-hypertensive

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

drug is 1.5 g and if 12,500 packages of 30 tablets of 1 g of this drug have been sold in a given region in one year, this is equivalent to a total of (12,500 × 30) / 1.5 = 250,000 DDD or 250,000 treatment-days. This figure gives no information about the number of patients treated during the same period: e.g. it could be 250,000/30.4 = 8224 patients treated for 1 month or 250,000/365 = 685 patients treated for 1 year.

A calculation in terms of the DDD can be used in this way to estimate person-time (population time).

For comparison purposes, DDD estimates can be standardized for the size of the population.

Example: In the above example, if the study took place in a region of 625,000 inhabitants, the number of DDDs per day for 1000 inhabitants is:

$$(250{,}000 \times 1000) / (625{,}000 \times 365) = 1.1.$$

This means that if sales are constant, an average of 1.1 inhabitants per 1,000 (0.11%) are treated with this drug each day in this region.

The DDD is useful mainly as a basis for comparison between drugs or countries, since the number of units sold is expressed in the form of a common reference. This is, however, the principal limitation of the DDD, which allows study and comparison of consumption but does not always reflect a population's real level of exposure (especially if the daily dose used, or the duration of treatment differ between countries being compared).

See also: **ATC classification, person-time, prescribed daily dose.**

Dependence, pharmacodependence Totality of behavioural, cognitive and physiological phenomena, of varying intensity, in which the use of one or more psychoactive substance(s) becomes a high priority.

The essential characteristics are the obsessive desire to procure and take the causal substance, and a constant pursuit of it.

The determining factors of dependence and the ensuing problems may be biological, psychological or social. The state of dependence is not necessarily harmful in itself, but may result in self-administration of the causal substance in doses producing physical or behavioural effects which constitute social and public health problems.

The term 'pharmacodependence' is more and more often being replaced by 'dependence'.

Dependent variable Variable whose value is modified by the presence or the value of another variable.

See also: **independent variable, variable.**

Depletion of susceptibles More or less rapid disappearance of a sub-group of subjects at increased risk during the follow-up of a population, if they are no longer followed or are no longer at risk after having presented the adverse event.

For example, if an event has a tendency to occur more often at the beginning of an exposure, the hazard rate measured in a fixed (non-renewed) exposed cohort will tend to decrease over time. This phenomenon can introduce a significant bias if the risk is expressed as an incidence density (diluting this rate if the follow-up is long), or if the follow-up began after the period of depletion.

See also: **cohort, hazard rate, incidence density, survival bias, time-window.**

Descriptive study Epidemiological study whose goal is to identify a phenomenon, describe it, measure its frequency, and/or study its development and distribution in different sub-groups of a population, but not to determine its cause or determinants (*see* Appendix 1).

Even though the distinction between a descriptive study

and an aetiologic study may be somewhat artificial, descriptive studies generally use a prospective approach (cohort study), without a comparison group.

See also: **aetiologic study, prospective.**

Design, research design Plan, protocol, strategy or scheme of a study.

Determinant Factor which influences the probability of the occurrence of an event or disease, or the state of health of a population.

This factor may act alone or in association with other factors; however, the term 'determinant' (as opposed to '*risk factor*') implies a causal association with the disease.

See also: **causality, risk factor.**

Diagnostic bias Selection bias in which a case has a different probability of being diagnosed according to whether he or she is exposed to a risk factor of interest or not.

Example: In a cohort study designed to assess the risk of peripheral neuropathy associated with a drug, a physician might be inclined to seek this symptom more actively if he or she knows that the patient is treated by this drug.

See also: **selection bias.**

Directionality Direction of reference to time in an epidemiological study (i.e. *prospective*, *retrospective* or *cross-sectional*).

See also: **epidemiological study.**

Discrete variable Variable which may only assume a certain number of integer values within its limits of variation (e.g. number of children per couple, number of episodes of an illness, etc.), unlike a continuous variable.

See also: **binomial distribution, continuous variable.**

Double-blind *See* **blinding.**

Drug, medication, medicine, medicinal product Any substance or combination of substances which may be administered to human beings or animals with a view to making a medical

diagnosis or to restoring, correcting or modifying physiological functions (EU Directive 65/65/EEC; *see* References).

This term thus applies not only to substances intended to treat a disease or symptom, but also to vaccines, contrast media and other agents administered for diagnostic purposes, intravenous solutions, blood products, oral contraceptives, etc.

See also: **active ingredient.**

Drug prescription study Study seeking to describe—qualitatively and quantitatively—the population of health professional prescribing a given drug (or class of drugs) and/or the frequency of, reasons for and conditions (e.g. dosage, duration, recommendations, surveillance, etc.) of this prescription.

As in *drug utilization studies*, such a study is only relevant if it is truly observational and does not risk modifying the prescription behaviour.

See also: **drug utilization study, observational study.**

Drug safety monitoring *See* **pharmacovigilance.**

Drug utilization study Study designed to describe—qualitatively and quantitatively—the population of users of a given drug (or class of drugs) and/or the conditions of use (e.g. indications, duration of treatment, dosage, previous or associated treatments, compliance, etc.).

The study of the drug utilization pattern is only relevant within the framework of an observational study that does not modify the conditions of use.

See also: **drug prescription study, observational study.**

Duplicate Case erroneously used twice during a count. The risk of this happening should, for example, be taken into account in surveillance by spontaneous reporting, when a physician observing an adverse event has several different means of reporting. Similarly, the same case may be reported by several health professionals. Only adequate completeness of data can ensure detection of such duplications.

See also: **completeness of data, spontaneous reporting.**

Dynamic Adjective designating a population whose composition changes during the follow-up period because some subjects have left it while others have joined the population.

Thus, on any given date, the effective duration of follow-up or the time elapsed since the beginning of the exposure can differ to some extent for each member of the population.

This term is also used to designate any event or variable whose nature or value changes over time.

See also: **cohort, fixed.**

Dynamic cohort *See* **cohort.**

Dynamic population Population whose composition changes over time because of a partial or complete renewal of its subjects.

This term is to be preferred to 'open population' although the latter term has the same meaning.

See also: **dynamic.**

E

Ecological Any study or analysis carried out not on one or more population(s) with individuals selected according to a protocol, but on global population data (*aggregated data*), available at the regional or national level, and generally not collected for this purpose (*see* Appendix 1).

This approach is extremely useful because it allows rapid exploration of a large number of hypotheses, without requiring the implementation of a specific data collection. However, the absence of information about the characteristics of individuals and the conditions of their possible exposure to the risk factor of interest makes this approach particularly subject to numerous biases and interpretative errors.

Example: In New Zealand, a research team observed an increase in mortality from asthma proportional to the number of prescriptions of a new anti-asthmatic drug, using national statistics. This could have corresponded to a particular risk associated with the drug but also to an unrelated increase in the incidence of severe asthma in the country, leading to increased prescription of the drug.

See also: **aggregated data, aggregation bias.**

Ecological bias *See* **aggregation bias.**

Ecological correlation *See* **ecological.**

Ecological fallacy *See* **aggregation bias.**

Ecological study *See* **ecological.**

Effect, reaction Modification of a previous state which can reasonably be attributed to an exposure, particularly to a

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

drug (e.g. *therapeutic effect*, *adverse reaction*).

If the causal relationship with the exposure (the drug) is not established or at least strongly suspected, the term '*event*' should be used instead (since '*effect*' implies that the event was induced by the exposure).

See also: **event.**

Effect modification *See* **interaction.**

Effect modifier *See* **interaction.**

Effectiveness, effective Demonstration of the therapeutic effect of a drug under real conditions of prescription and use.

See also: **cost-effectiveness study, efficacious.**

Efficacy, efficacious Ability to produce the expected effect.

A drug is deemed '*efficacious*' if it can induce a therapeutic effect. A distinction should be made between the terms '*active*' and '*efficacious*'. The first most often applies to a substance or drug which causes one or more pharmacodynamic effect(s) that may induce a therapeutic effect (e.g. an active principle, pharmacological activity). The term '*efficacious*' is applied to a drug whose therapeutic effect has been demonstrated under standardized or experimental conditions such as clinical trials.

'*Efficacy*' differs from '*effectiveness*', the latter referring to the demonstration of a drug's therapeutic effect under real conditions of prescription and use.

See also: **active, clinical trial, effective.**

Efficiency, efficient Comparison of the effects achieved by a therapeutic strategy of known efficacy and effectiveness with the cost and time required for the implementation of the strategy.

In this way, one treatment can be more efficient than another if the same therapeutic benefit can be obtained with less time and/or expenditure.

EPAR *See* **European Product Assessment Report.**

Epidemiological study Study whose objective is to describe

the characteristics, behaviour or state of health of a population (i.e. a *descriptive study*), or to investigate the associations (causal or otherwise) which may exist within this population between a characteristic (e.g. the exposure to a risk factor) and the occurrence of an event (i.e. an *aetiologic study*).

Three major types of epidemiological study can be described, according to the *directionality* of the analysis (*see* Appendix 1):

- *Prospective or cohort studies*, in which a population is followed after the beginning of an exposure, in order to investigate how much this modifies the probability of the occurrence of an event (the causal inference being from exposure to event).
- *Retrospective studies* (e.g. case-control studies), in which an exposure is sought for the previous history of subjects having presented an event (the causal inference being from event to exposure).
- *Cross-sectional studies*, which analyse the state of health or the characteristics of a population at a given moment, with neither a follow-up nor an analysis of the past.

See also: **case-control study, cohort study, cross-sectional study.**

Epidemiology Study of the relationships between diseases or any other biological phenomenon and various factors (e.g. lifestyle, environment or social setting, individual traits, etc.) which can influence their frequency, distribution and evolution.

Several somewhat arbitrary distinctions can be made:

- *Descriptive epidemiology*, in which the objective is to describe a population (e.g. drug utilization studies),
- *Analytic epidemiology*, in which the objective is to study the associations (causal or otherwise) that can exist, within a population, between the occurrence of an event and a given characteristic (e.g. exposure to a risk factor).

See also: **epidemiological study, pharmacoepidemiology.**

Equivalence State which can be considered identical, or differing at most by a negligible amount.

Two therapeutic strategies can thus be considered equivalent if they induce—both qualitatively and quantitatively—the same effects. Equivalence may apply to both therapeutic and adverse effects. The term '*bioequivalence*' is used to indicate the fact that two drugs induce the same active plasma concentrations in an organism during a period that covers the absorption and elimination phases, by which the identical (qualitative and quantitative) biological, pharmacological and therapeutic effects may be expected. In this instance equivalence is defined by regulatory criteria.

See also: **equivalence study.**

Equivalence study Study seeking to demonstrate the absence of a difference between two magnitudes or two strategies based on measurements taken from one or two samples.

Testing the equivalence of two parameters X_1 and X_2 consists firstly of defining a difference. If $|X_1 - X_2|$ is smaller than Δ, X_1 and X_2 will be considered equivalent. Given measurements x_1 and x_2 taken from two samples, we will test the null hypothesis, H_0, that X_1 and X_2 differ by at least the amount Δ. Rejection of H_0 allows us to conclude (at the chosen risk of Type I error) that X_1 and X_2 are equivalent, that is, equal or differing by at most an amount Δ. The same reasoning can apply to the ratio X_1/X_2.

The procedure for an equivalence test is thus the reverse of that in tests seeking to find a difference. In the latter, we test a null hypothesis of equality, the rejection of which leads us to conclude the existence of a difference (at the chosen risk of error). In both cases, the test statistic used may be one- or two-sided.

See also: **alternative hypothesis, equivalence, null hypothesis,**

Type I error.

Estimation Procedure consisting of determining, as precisely as possible, the unknown value of a parameter in a population, based on a measurement made on a sample taken from this population.

Estimation is, by definition, an inferential procedure. The term '*estimate*' indicates the result of this procedure.

See also: **confidence interval, inference, sample.**

European Product Assessment Report (EPAR) Expert report resulting from the marketing authorisation of a centrally approved drug. It is produced by the European Medicines Evaluation Agency (EMEA: *see* References) and available on its website when the drug is approved.

Evaluation *See* **assessment.**

Event Any phenomenon which can be observed or studied in a subject or population.

In pharmacovigilance, 'event' designates a desirable or undesirable manifestation, without presuming that it is or not related to the use of a drug.

This term is also used each time the numbers of cases occurring in two populations, of which one is not exposed, are compared; it would be incorrect to speak of an '*effect*' in a population which has not been exposed to a risk factor (such as the use of a drug).

See also: **adverse event, effect.**

Excess risk *See* **risk difference.**

Expected See expected adverse drug reaction, expected number, Poisson distribution.

Expected adverse drug reaction, expected adverse effect Harmful and undesirable manifestation attributed to a drug, whose occurrence is apparently related to a known pharmacological property of the drug.

The term '*expected*' implies that the knowledge of the

pharmacological properties of a drug makes it possible to foresee such adverse events in a certain proportion of treated patients.

Examples: *Kaliuresis with loop diuretics*, *prolonged bleeding time with aspirin*.

To accept that this effect is the expression of a pharmacological property of the drug implies that it will be found in an appreciable proportion of treated patients, that it is generally dose-dependent and will be reproducible, in whole or part, in animals. In this sense, an expected adverse effect is synonymous with '*side effect*', and is sometimes also called a 'Type A reaction'. By extension and somewhat abusively, an expected adverse drug reaction is considered the same as a *labelled effect*. Indeed, current international definitions designate an expected adverse effect as a manifestation clearly mentioned in the Summary of Product Characteristics (SPC) or in the investigator's brochure for a clinical trial.

See also: **side effect, Summary of Product Characteristics.**

Expected number Number of cases of an event expected to be observed in a given population during a given period.

This is calculated by multiplying the number of surveyed units n in this population (e.g. number of subjects, number of risk periods, etc.) by the probability p of occurrence of the event for each unit considered.

Example: If the annual incidence of Guillain–Barré syndrome is 30 per million inhabitants, the expected number per month among 500,000 subjects (assuming that the incidence rate is constant) is $(30/1{,}000{,}000) \times 500{,}000 \times 1/12 = 1.25$.

The observed number of cases may differ more or less significantly from the expected number. This variability can be predicted by an adequate probability model. For example, the *prediction interval* constructed around the expected number of events defines the set of values for the number of cases with

a given probability of being observed. The expected number of cases (designated as m or λ) is the basis of probability calculations using the Poisson formula.

See also: **expected-observed ratio, Poisson distribution, prediction interval.**

Expected–observed ratio Approach used to investigate a possible association between an exposure and an event, by comparing the number of cases observed in an exposed population with the number to be expected in this population, under the hypothesis that the factor does not modify the probability of the event occurring (*see* Appendix 1).

The null hypothesis of independence between exposure and event is rejected if the ratio of the number of observed cases to the number of expected cases differs significantly from 1.

The expected–observed comparison differs from classic epidemiological approaches (cohort studies, case-crossover studies) in that the reference value (the expected number) is not taken from an *ad hoc* comparison group, but estimated from aggregated data, not specifically gathered for the purposes of the study. As in all ecological approaches, this may be the cause of interpretative errors.

For example: The 1–year follow-up of a population of 5700 subjects treated with a new anti-hypertensive has identified 12 cases of suicide. On the basis of an annual incidence rate of 2.5 per 10,000 in the general adult population, the expected number of suicides in the exposed population is $5700 \times 2.5/10{,}000 = 1.42$. *The observed number of cases is therefore* $12/1.42 = 8.45$ *times greater than the expected number. A supplementary calculation (using the Poisson approximation) shows that this difference is significant at the 5% error level. This conclusion must nevertheless be tempered because of the impossibility of controlling for a certain number of biases (e.g. the baseline suicide risk in this population of patients may not be the same as that in the general*

population).

The expected–observed ratio is a valuable to test a hypothesis quickly or to devise an alert strategy, by virtue of its simplicity; it cannot, however, replace a comparison based on a true reference group.

See also: **aggregated data, ecological, expected number, reference population.**

Experimental Adjective designating a strategy in which the researcher intentionally alters the natural course of events for the purposes of research. The classical experimental design consists of artificially creating two groups (e.g. by randomly allocating exposure) that are identical with respect to the variables that can influence the planned measurement, and differ only in their exposure status. Under these strict conditions, any statistically significant difference observed between the two groups (not explained by sampling variation) can be attributed to the studied factor.

In clinical research, these modifications are formalised in advance by a protocol. A typical experimental design is a Phase III clinical trial in which groups are made comparable by randomisation.

See also: **interventional, naturalistic, observational study, randomisation, research.**

Experimentation, experiment Research of an experimental nature.

The term 'experimentation' can also more restrictively mean direct intervention on the human body (e.g. a trial of a new surgical technique).

Explanatory trial Experimental controlled trial.

An explanatory trial compares (under standardised conditions) two or more randomly constituted groups, which are therefore deemed to differ only with respect to the status of the factor under study (*see* Appendix 1).

The explanatory trial is an undisputable reference for any aetiologic analysis. It is used as much in clinical pharmacology (e.g. controlled trials with randomisation) as in analytic epidemiology.

See also: **controlled trial, epidemiology, intent-to-treat analysis, per protocol analysis, Phase III clinical trial, pragmatic trial.**

Explanatory variable *See* **independent variable.**

Exploratory study Study in which the objective of the research is not precisely defined beforehand.

The term '*exploratory*' may apply to studies which are perfectly scientifically justified when, for example, nothing is known about the effects of a new drug, or if the objective of the study is specifically to learn whether it would be appropriate to undertake a particular study or to refine its protocol (i.e. a *feasibility study*).

See also: ***ad hoc*** **study, fishing expedition.**

Exposed Subject or population presenting a precisely defined characteristic that may increase the probability of occurrence of an event or disease.

This characteristic could be, for example, past or present contact with a pathogen (e.g. virus, radiation, etc.).

In pharmacoepidemiology, the term '*exposed*' designates a patient treated by one or more drug(s) under conditions (such as dose or duration) which place him or her at risk for a given adverse effect. This term also applies to indirect exposure (e.g. exposure *in utero* or an infant's exposure to a drug via breast-feeding). The risk period does not necessarily correspond to the treatment period: for an effect with delayed onset (e.g. cancer), subjects are considered at risk (i.e. exposed) even though their treatment dates back several years. In other respects, not every treated subject is necessarily at risk of presenting the effect: if it cannot occur until after 3 months of

treatment, subjects treated for 1 week are not at risk and should be considered non-exposed.

See also: **exposure, incidence rate, treated.**

Exposed/not-exposed study Prospective aetiologic study design based on the follow-up and comparison of two populations, of which only one is exposed to the factor whose influence on the occurrence of an event under study.

If exposure is determined by chance, an exposed/not-exposed study is the same as an explanatory or experimental controlled clinical trial. With respect to observational epidemiology, an exposed/not-exposed study is equivalent to a cohort study with a reference group.

See also: **cohort, controlled trial, experimental, observational study.**

Exposure Fact of a subject's or a population's contact with a factor likely to increase the probability of occurrence of an event or disease.

This factor can be infectious (e.g. a virus, bacteria, etc.), physical (e.g. ionizing radiation, electrical fields, etc.) or chemical (e.g. pesticides, vitamins, drugs, etc.). For some authors, the definition also includes individual characteristics such as age, race, a given phenotype, etc.

In pharmacoepidemiology, exposure is generally represented by the use of one or more drug(s) under conditions (e.g. dose, duration, etc.) which allow us to expect a therapeutic or adverse effect.

See also: **exposed.**

F

False negative Negative diagnostic test result in a subject who has the disease under study.

By analogy, this term is also used when a surveillance method does not detect a case corresponding to the definition of the event of interest.

See also: **sensitivity, spontaneous reporting, true negative.**

False positive Positive diagnostic test result in a subject who does not have the disease under study.

By analogy, this term is also used when a surveillance method detects a case which is not related to the event of interest, or wrongly attributes the occurrence of the event to the risk factor of interest (e.g. consumption of a particular drug).

See also: **specificity, true positive.**

Field study Study which is based on information gathered in the field, specifically for its purposes, and not on information already recorded (e.g. in a database).

Thus, we can use the terms '*field cohort study*' or '*field case-control study*'.

See also: **database.**

Fishing expedition, data dredging, data trawling Study in which neither the objective nor the type of information to be investigated are precisely defined beforehand.

We thus speak of a '*fishing expedition*' for a cohort study in which an arbitrarily fixed number of subjects is followed for an arbitrarily determined period, to investigate possible adverse effects whose type, frequency, severity and conditions

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

of occurrence are, *a priori* unknown.

Another example is the systematic exploration of databases for possible correlation of any event and any drug exposure.

The term '*fishing expedition*' has a slightly negative connotation which 'exploratory study' does not.

See also: **exploratory study, false positive.**

Fixed Regulated in a precise and definitive manner.

In pharmacoepidemiology, this term can describe:

- A population whose members are not renewed by the inclusion of new subjects (e.g. a *fixed cohort*).
- A follow-up period that is identical for all of the members of a population (i.e. a cohort with a fixed follow-up).

See also: **cohort study.**

Fixed cohort *See* **cohort.**

Follow-up Prospective surveillance set up to identify the occurrence of an event in an individual or population.

The term applies only when the surveillance method is effective and relevant to the event of interest.

See also: **cohort study, dynamic, fixed, lost to follow-up, prospective.**

Force of morbidity *See* **instantaneous incidence rate.**

Force of mortality *See* **instantaneous incidence rate.**

Fortuitous *See* **coincidental.**

Frequency Quantification of the total number of events in a population (i.e. *prevalence*) or only those cases occurring during the study period (i.e. *incidence*).

The term is often used to designate relative frequency, that is, the number of cases of an event compared to the size of the study population and the duration of the follow-up. In this case, frequency is synonymous with prevalence or incidence rate. These terms, being more precise, are preferable.

See also: **incidence, prevalence, rate.**

G

Gamma error *See* **Type III error.**

Gaussian distribution *See* **normal distribution.**

Gender ratio *See* **sex ratio.**

General population Geographically defined population (e.g. the inhabitants of a city, region or country).

This population is usually heterogeneous and can include subjects having very different characteristics. A risk, incidence rate or prevalence rate estimated in the general population can thus be very different from that which would be measured in an *ad hoc* reference group.

See also: **reference population, reference risk.**

Generation effect *See* **cohort effect.**

Generic drug Drug which has the same qualitative and quantitative composition in active ingredient(s) and the same pharmaceutical form, as the reference drug whose bioequivalence has been demonstrated by appropriate studies.

The marketing of a generic drug is only possible after the end of the period of patent protection of the copied innovation (e.g. active principle, manufacturing process, form, new association, etc.).

See also: **equivalence, equivalence study.**

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

H

Hazard *See* **risk.**

Hazard function, risk function Representation of the evolution of the probability of occurrence of an event as a function of time.

Ideally, this evolution in risk is estimated from repeated measurements of the instantaneous incidence rate, carried out at regular and sufficiently close intervals in a population adequately followed over a period of time.

For some authors, a hazard or risk function also denotes a mathematical model allowing prediction of the instantaneous risk at a given time during the follow-up.

See also: **cohort, instantaneous incidence rate, instantaneous incidence, risk.**

Hazard rate, attack rate Probability that a subject, as yet unaffected, will present the event of interest during a very short time interval, approaching zero. It is estimated from an instantaneous incidence rate.

See also: **hazard function, instantaneous incidence, instantaneous incidence rate.**

Healthy worker effect Bias that has been mainly described for occupational studies.

As employed persons are generally healthier than unemployed (severely ill and disabled being excluded from employment), using the last as a reference group can lead to an underestimation of excess risk associated with a particular occupation. This refers to the lack of comparability between groups and to a different probability of being diseased

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

independently of exposure.

Haemovigilance Application of surveillance methods, e.g. spontaneous reporting, to the specific problem of labile blood products, instituted after the HIV-related blood scandals with a special emphasis on *traceability* of products.

See also: **spontaneous reporting, traceability.**

Historical Adjective denoting a study which uses pre-recorded data.

The events of interest that have occurred prior to the implementation of the study are investigated, for example, using a database. In a *historical cohort study*, the subjects are followed beginning at a date in the past, towards the present, with the aim of finding events which may have occurred. If the follow-up continues into the future (beyond the date of implementation of the study), the term '*ambispective cohort*' is sometimes used.

See also: **ambispective, database, historical cohort.**

Historical cohort Cohort whose population is identified and followed-up based on pre-recorded data, usually collected for other purposes (e.g. in a population database).

A study based on a cohort of this type remains a prospective study insofar as the possible occurrence of an event after the beginning of the subjects' exposure to the risk factor of interest (even if the events occurred and were recorded prior to the implementation of the study) is investigated.

See also: **cohort study, historical, prospective.**

Historico-prospective *See* **ambispective.**

Hospital control In a case-control study, a control selected from among the population of patients present in the care structure or hospital from which the case was identified.

This term is thus the opposite of *community control*. The choice between community and hospital controls is not easy. The ideal control is a person who would have been hospitalised

in the same system as the case, and who would have had the same probability of having been exposed to the studied factor, if he or she had presented the disease identifying the case.

See also: **case-control, community control, control, source population.**

Hypothesis Proposition accepted provisionally before being tested.

I

Iatrogenic Induced by a doctor or medical treatment.

'*Iatrogenesis*' covers all of the consequences of medical intervention which are proved to be, or may potentially be, harmful to health (e.g. diagnostic errors, maladapted prevention or prescription, complications of a therapeutic act).

'Medication iatrogenesis' refers to all the consequences of medication (whether prescribed or used in self-medication) which are proved to be, or may potentially be, harmful to health. Its scope is therefore wider than that of expected or unexpected adverse effects alone, since it also includes the confirmed ineffectiveness of a poorly adapted treatment, the effects induced by the context of the prescription (*nocebo effect*) and the creation of a state of dependence.

See also: **dependence, self-medication, unexpected adverse drug reaction.**

Imputability, attributability, case-by-case causality assessment Estimate of the possible causal link between a drug treatment and the occurrence of an adverse event in a given patient.

The aim of imputability analysis is to determine the degree of plausibility of the assumption that a given drug caused the adverse effect presented by a patient. This assessment is thus strictly individual and does not refer to the potential danger of the drug in the absolute, or the significance of the risk for a population. Various methods have been published or proposed to standardize the process of imputation and thereby reduce

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

inter- and intra-observer variability. These are based either on algorithms (classic imputability methods) or on theories of conditional probabilities (e.g. Bayesian methods). The former are used more often in routine surveillance, and make use of the sequential evaluation of a given number of criteria, allowing calculation of an imputability score.

See also: **algorithm, Bayes's theorem, causality, intrinsic imputability.**

Imputability criteria *See* **algorithm, imputability.**

Imputable Adjective synonymous with attributable.

See **attributable.**

Imputation *See* **imputability.**

Incidence Number of new cases of an event identified in a population over a given period (e.g. incident cases, annual incidence, incidence rate).

In everyday language, this term is frequently used to designate an incidence rate (e.g. an annual incidence of 3 per 1000).

See also: **case, incidence density, incidence rate, instantaneous incidence, prevalence.**

Incidence density, incidence density rate Incidence rate expressed as number of cases of an event divided by person-time, where the denominator represents the sum of the follow-up or exposure times of the subjects in the population.

Example: Six cases of liver injury are identified in a cohort study of 278 persons during which 120 subjects have been followed for 1 month, 80 subjects for 2 months, 65 subjects for 3 months and 13 subjects for 6 months. The sum of follow-up times is $(120 \times 1) + (80 \times 2) + (65 \times 3) + (13 \times 6) = 553$ *months. The risk of liver injury expressed as an incidence density is thus 6 for 553 person-months, that is, 10.8 per 1000 person-months or 0.13 per person-year.*

Incidence density is particularly useful in the expression of

the relationship between a risk and the exposure time.

In the preceding example, a risk of 10.8 per 1000 person-months allows us to expect 11 cases to occur in a population of 1000 subjects during a 1–month follow-up, and 32 cases during a 3–month follow-up.

Incidence density is also a convenient method of calculation when the length of follow-up or exposure is not the same for the different subjects in a study.

On the other hand, calculation using incidence density relies on comparing the number of events observed during a study period with a sum of individual follow-up durations that can vary greatly from subject to subject. This can lead to errors of interpretation if the risk of presenting the event is not constant over time.

For example, the rate of 10.8 per 1000 person-months calculated above does not necessarily imply that the risk is 10.8 per 1000 for each month of exposure.

See also: **depletion of susceptibles, hazard function, incidence rate, person-time, risk.**

Incidence rate Number of new cases of an event occurring in a population during a given period, divided by the number of subjects at risk in the population.

The incidence rate represents the speed of production of new cases in the population, per unit of time. It is the fundamental parameter for estimating the risk of an event or disease in a population, or for a subject or sub-group of this population.

The number of cases appearing in the population can be compared to the number of subjects followed during the period of interest or to the sum of follow-up or exposure times of the subjects making up the population (*incidence density*).

In pharmacoepidemiology the study population is usually an exposed population, due to treatment with a drug, but can

also be an untreated population (e.g. a reference or general population). For an incidence rate, subjects having presented the event (constituting the numerator) are a sub-group of the total studied population (the denominator); each subject in the population theoretically has the same baseline probability of presenting the event.

Subjects with abnormally low or zero probability of presenting the event (e.g. those treated for too short a period to be at risk, or subjects physiologically unable to present the event of interest) should not be included in the calculation of the denominator, neither should those who have not been surveyed in a manner which will guarantee the identification of the event of interest.

See also: **incidence, rate, risk, standardization, stratification.**

Incident case *See* **incidence.**

Independent variable Variable whose status or value affects that of another variable.

For example, in a regression model, the value of a dependent variable is predicted by the value(s) of one or more independent (or explanatory) variable(s).

See also: **dependent variable, model, regression analysis.**

Index date *See* **time-window.**

Inference Extrapolation of the results drawn from a measurement or an estimate made from a study sample to a larger population (usually, from which the sample was drawn).

See also: **confidence interval, selection bias, target population.**

Information bias, observational bias Distortion of the estimate of the association between a risk factor (e.g. use of a drug) and the occurrence of an adverse event, due to a systematic difference in the way information concerning the measured parameter is collected for the groups being compared.

Recall bias and *interviewer bias* are types of information bias.

See also: **interviewer bias, recall bias.**

Informed consent Free and formally expressed agreement of a person to participate in a clinical research.

In any research on human beings, each potential subject must be adequately informed of the aims, methods, anticipated benefits and potential hazards of the study and the discomfort it may entail. He or she should be informed that he or she is at liberty to abstain from participation in the study and that he or she is free to withdraw his or her consent to participation at any time. The physician should then obtain the subject's freely-given informed consent, preferably in writing [The Declaration of Helsinki (*see* References), adopted by 18th World Medical Assembly, Helsinki, Finland, 1964, and revised by 29th World Medical Assembly, Tokyo, Japan, 1975, 35th World Medical Assembly, Venice, Italy, 1983, and 41st World Medical Assembly, Hong Kong, 1989].

Instantaneous incidence Number of cases occurring in a population at a given point in time or during a very short time period (approaching zero).

This measurement is usually divided by the number of subjects at risk during this period in order to calculate an *instantaneous incidence rate*, sometimes called *force of morbidity* or *force of mortality* (the latter if fatalities are involved). The measurement of instantaneous incidence serves to describe or predict the risk of occurrence of an event in a population at a given moment.

See also: **hazard function.**

Instantaneous incidence rate, force of morbidity, force of mortality Incidence rate measured in a population during a very short time interval, approaching zero.

This rate estimates the hazard rate of an event or a disease in the population and is sometimes called the *force of morbidity* or *force of mortality*.

See also: **hazard function, hazard rate, instantaneous incidence.**

Intense, intensity *See* **severity.**

Intent-to-treat analysis Statistical analysis taking into account all evaluable patients included in a study of efficacy or tolerance.

Contrary to analyses which retain only those subjects treated according to the protocol, an intent-to-treat analysis considers all subjects from the time of their inclusion in the study (e.g. once they have been allocated a randomisation code in a parallel group trial), even if it turns out that a subject never took the prescribed treatment or modified its timing, dosage or duration, or if he or she left the study or is deceased.

This type of approach has the great advantage of better reflecting the actual (or future) conditions of the doctor–patient and patient–medication relationships, but tends to be more conservative (in diminishing the chances of observing a significant statistical difference between the groups, by increasing the inter-subject variability).

Intent-to-treat analyses can be used in clinical trials, both controlled and non-controlled, and with or without blinding.

The intent-to-treat approach may also be used in prospective (cohort) pharmacoepidemiological studies seeking to confirm an efficacy or to describe a risk or usage.

See also: **cohort study, explanatory trial, per protocol analysis, pragmatic trial.**

Interaction In pharmacology, term designating a modification in effect (whether beneficial or adverse) obtained or expected from a drug, as the result of the concomitant use of another substance, whether medicinal or not (*see* Appendix 2) or of the presence of another factor (e.g. age).

An interaction is called '*pharmacokinetic*' if the second drug modifies—at whatever level—the pharmacokinetics, and thus the tissue concentration, of the first (e.g. by accelerating or slowing down hepatic metabolism or renal elimination). In a '*pharmacodynamic*' interaction, the second drug does not

affect the pharmacokinetics of the first, but interacts with its pharmacological action by inhibiting it ('*antagonism*') or reinforcing it ('*synergism*').

In epidemiology, the term 'interaction' (or *effect modification*) is used when the presence of a factor or the value of a variable modifies the effect produced by an exposure. This is the case when the strength of the association between an exposure and the occurrence of an event (e.g. quantified by the relative risk) is not the same in the presence or absence of another factor, called an '*effect modifier*'. Observing an interaction sometimes allows identification of groups at risk.

Example: In a case-control study assessing the association between the use of a drug and the occurrence of acute pancreatitis, the odds ratio calculated for the entire group of subjects is 3. After stratification according to age, the odds ratio is 1.8 for subjects less than 65 years old and 5.2 for subjects 65 years and over. Age is thus a modifier of the risk; subjects over 65 years are a group at risk for this adverse effect.

See also: **additive, multiplicative, pharmacodynamic, pharmacokinetics, synergism.**

Intermediate endpoint Judgement criterion which predicts *a priori* a relevant clinical criterion, and is used in place of the latter to temporarily assess the efficacy or tolerance of a therapeutic strategy.

This means that the efficacy or tolerance will only be considered to have been truly assessed on the grounds of a study, or subsequent analysis, based on the relevant clinical criterion. On the other hand, a *surrogate endpoint* can be used instead of the relevant clinical criterion, once the validity of its predictive value has been established.

See also: **efficacious, surrogate endpoint.**

Intermediate factor Factor which does not constitute a cause in itself, but which represents a crucial step in the exposure-to-

effect chain (*see* Appendix 2).

For example, certain medications can result in liver injury, not by a direct effect, but by inducing low blood pressure causing hypoperfusion and thus hepatic anoxia; in this case, low blood pressure acts as an intermediate factor.

See also: **causality.**

International Classification of Diseases (ICD) Classification of diseases proposed by the World Health Organization (WHO).

It is used in most databases to code diseases or reasons for practitioner consultation. The available classification is the 10th edition (ICD-10: *see* References). In some domains, definitions and criteria for the use of terms exist (e.g. psychiatry).

International Conference on Harmonization (ICH) Process implicating the European, US and Japanese regulatory authorities, and pharmaceutical industry associations, to promote the development of harmonious regulation for drug registration.

There are three main areas (quality, safety, efficacy) with many subtopics, many of which have been incorporated into national or community regulations. Although mainly focused on pre-marketing activities, there are some topics that concern pharmacovigilance.

The ICH topics and guidelines can be found on the websites of the European Medicines Evaluation Agency (EMEA; *see* References) and the International Federation of Pharmaceutical Manufacturers Associations (IFPMA; *see* References).

Intervention *See* **constraint.**

Interventional, non-observational Adjective characterizing a situation in which the normal course of events or the therapeutic relationship is voluntarily modified for the purposes of research, generally in order to create artificial situations or

groups for the purposes of comparison under optimal conditions.

This term is the antonym of '*observational*'. An *experimental design* is the most interventional of studies.

See also: **constraint, experimental, observational study.**

Interviewer bias Information bias in which the interviewer questions the subjects in a systematically different way according to whether they belong to one or the other of the groups being compared.

This bias is classic in case-control studies when the interviewer knows the hypothesis being tested or is interested in its validation. He or she can, intentionally or not, question cases more precisely and insistently than controls regarding their exposure to the suspected risk factor.

See also: **information bias.**

Intrinsic imputability Analysis of the degree of responsibility of a drug for the occurrence of an adverse event in a given patient, not accounting for the degree of notoriety of this putative adverse effect. The analysis thus only takes into account the information drawn from the observed case.

See also: **Bayes's theorem, causality assessment, imputability.**

Investigator Person or organisation undertaking research according to a pre-ordained protocol on behalf of, and usually at the request of, a *sponsor*.

In multicentre researches involving many investigators in different sites, one of these investigators must be identified as the *main investigator*.

See also: **protocol, sponsor.**

J

Joint population Population actually benefiting from a public health action (e.g. the population of a country in which a diagnostic test has been carried out).

In pharmacoepidemiology, this term usually designates the population actually treated by a drug; this can be different from the *target population* recognized by the officially approved indications of the drug.

See also: **target population.**

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

L

Labelled, unlabelled By international convention, an adverse effect which is clearly mentioned in the most recent version of the Summary of Product Characteristics (SPC) or, in the case of a clinical trial, in the investigator's information brochure.

Conversely, an effect is said to be *unlabelled* if it is not clearly mentioned in these documents, even if it has been the subject of one or more scientific publication(s).

See also: **expected adverse drug reaction, Summary of Product Characteristics, unexpected adverse drug reaction.**

Latency period, latent period, latency Interval of time separating the beginning of exposure to a risk factor and the occurrence of an event.

This term usually has a more restrictive meaning than '*onset delay*' in that it tends to imply a causal relationship between exposure and event. Indeed, '*latency*' designates the quiescent character of a process set off by exposure but appearing only much later, such as for cancer.

See also: **onset delay.**

Likelihood ratio *See also*: **Bayes's theorem, imputability.**

Logistic model Mathematical modelling of the relationship between a dependent variable y and one or more explanatory variable(s) x, of type:

$$y = \frac{1}{1 + e^{-(\alpha + \beta x)}}.$$

The logistic model is widely used in epidemiology, for various

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

reasons:

- The logistic function describes a sigmoidal curve, between 0 and 1, which adequately represents a great number of relationships observed in practice (such as the relationship between the exposure density and the probability of occurrence of the disease).
- The explanatory variables may be qualitative dichotomous, or polydrotomous, or quantitative.
- The coefficient β quantifies the strength of association between the dependent and the explanatory variable(s); for example, in analyses studying the probability of occurrence of the event as a function of the value of the exposure (x), it can be shown that the coefficient β is equal to the logarithm of the odds ratio, thus:

  $$\beta = \mathrm{L}n \text{ (odds ratio), and odds ratio} = e^{\beta},$$

- The logistic model lends itself easily to multivariate analyses:

  $$p(y) = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_j x_j}}.$$

See also: **continuous variable, discrete variable.**

Longitudinal study A somewhat ambiguous term used to define any study based on the follow-up of one or more group(s) of subjects.

A longitudinal study may be considered a synonym for a *prospective study*, and, according to some authors, for a cohort study. In practice, the term '*longitudinal*' is often used for descriptive studies based on the observation of a single cohort.

Lastly, according to other authors, the term '*longitudinal*' may indicate any prospective or retrospective study in which, contrary to cross-sectional studies, a phenomenon is studied or a possible association between exposure and illness is investigated over a sufficiently long time period.

See also: **cohort study, cross-sectional study, descriptive study, prospective study.**

Lost to follow-up In a prospective study (e.g. a clinical trial or cohort study), subjects who have not been followed for the full extent of the planned term because there is no knowledge of outcome beyond a certain date for unknown reasons.

In the strictest sense, it is also advisable to consider as lost to follow-up all subjects for whom information about their subsequent fate is available, but where the exhaustiveness and reliability of this information do not correspond to the quality criteria stated in the protocol.

See also: **attrition, censoring, follow-up.**

M

Mandatory reporting Legal obligation for a health professional who has observed an adverse drug reaction to notify a pharmacovigilance system.

It has not been demonstrated that making reporting mandatory changes reporting rates or patterns.

See also: **serious, spontaneous reporting, unexpected adverse drug reaction.**

Masked Synonymous with '*blinded*'.

To avoid confusion with the medical meaning of '*blind*', some authors prefer to use '*masked*' rather than '*blinded*'.

See **blind.**

Marketing authorisation, NDA, AMM Official authorisation by a competent administrative organisation for the marketing of a brand-name pharmaceutical product, under given conditions and for given indications.

See also: **Summary of Product Characteristics.**

Matching Procedure rendering two groups (e.g. an exposed group and a reference group in a cohort study, or a diseased group and a non-diseased group in a case-control study) as comparable as possible with respect to the distribution of certain variables capable of biasing the results of the proposed statistical analysis by acting as confounders.

Matching consists of selecting, for each subject included in a study, one or several controls comparable with respect to the match variable(s).

Example: If in a case-control study we think that age could

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

be a confounder (by being linked to both the disease which defines a case and the probability of exposure to the factor of interest), the two groups (cases and controls) should be rendered comparable with respect to age distribution. This may be achieved by selecting one or more control(s) for each recruited case of the same age as the case.

Matching only concerns known prognostic variables, and cannot achieve the effectiveness of randomisation. The latter is the only method capable of ensuring that known and unknown prognostic variables will be equally distributed among the groups being compared.

Matching on a variable makes it impossible to study the effect of this variable on the studied outcome.

See also: **confounding, over-matching.**

MedDRA (Medical Dictionary for Drug Regulatory Activities) Standardized international medical terminology selected and developed by the International Conference on Harmonisation (ICH).

This thesaurus has thereby become officially adopted by most regulatory authorities. It includes all the terms from the previous adverse reaction terminologies (WHOART and COSTART) and a number of terms from the International Classification of Diseases (ICD). It also includes terms for medical or surgical procedures and for life events.

See also: **International Conference on Harmonization.**

Medication *See* **drug.**

Medicinal product *See* **drug.**

Medicine *See* **drug.**

Meta-analysis Statistical method in which the results of several trials or studies devoted to the same topic are combined, with the goal of obtaining greater statistical power or more accurate estimates.

Meta-analysis has been most frequently applied to clinical

trials (most notably within the Cochrane Collaboration), but can be useful in pharmacoepidemiology (in cohort or case-control studies) to obtain, for example, a more accurate estimate of a relative risk or an odds ratio.

Meta-analysis has, however, a number of significant limits, linked to the difficulty in accessing the data sources of the studies, the validity of the data, and the fact that an unknown number of studies on the subject of interest may not have been published.

Metabolite Substance produced by the *in vivo* transformation of a drug.

A drug is generally transformed in the organism into several metabolites with different chemical structures. The metabolite can variously be active (i.e. retaining the pharmacological effect), inactive, reactive or toxic.

See also: **drug, pharmacodynamic effect.**

Method Procedure adopted by a researcher to test a hypothesis or to investigate or solve a problem.

This term also designates the formalization of this procedure to allow other researchers to arrive at the same result or apply it to a different context.

See also: **methodology.**

Methodology Study of scientific methods, in particular the set of methods available in a given field (e.g. the methodology of clinical trials).

The term '*methodology*' is often wrongly used in the place of '*method*' (e.g. we should say, 'the method adopted for this study' and not 'the methodology of this study').

See also: **method.**

Misuse Prescription or use of a drug that does not conform to the approved recommendations.

Misuse is thus the opposite of '*proper use*', and, contrary to '*abuse*', may concern the prescriber as well as the user.

See also: **abuse, proper use.**

Model, modelling Simplified representation of a generally complex process or system.

In this way, research on new drugs often leads to trials with *animal models* that are supposed to faithfully represent the human disease to some degree.

The term 'model' is also given to mathematical formulae used to predict the probability of appearance of a disease or event based on the value of different variables (e.g. logistic model, Cox model).

See also: **Cox model, logistic model, Poisson process.**

Months of treatment *See* **incidence density, person-time.**

Multicentre Designates research (e.g. a study or trial) carried out at different sites by more than one investigator, but following the same protocol.

See also: **investigator.**

Multiple comparison techniques, multiple testing *See* **adjustment for multiple comparisons.**

Multiplicative Adjective characterizing an association of two or more factors producing an effect of intensity greater than the sum of the effects of each of the factors considered in isolation.

See also: **additive, synergy.**

Multipurpose study Antonym of '*ad hoc* study'.

See also: ***ad hoc* study, cohort study.**

N

***n* of 1 trial** Specific trial method where the effect (positive or negative) of a drug is tested in a single individual by the use of multiple random order repeated crossover design.

Naturalistic Said of a study which does not interfere with the normal care.

See **normal care relationship, pragmatic.**

NDA Acronym for 'new drug approval'.

Negative predictive value Probability that a subject for whom a diagnostic test is negative is not diseased.

The negative predictive value is calculated by dividing the number of *true negatives* (*d*) by the total number of times the test results are negative ($c + d$):

	Diseased	Non-diseased	
Positive test	*a*	*b*	$a + b$
Negative test	*c*	*d*	$c + d$

See also: **false negative, positive predictive value, true negative.**

Nested case-control, nested case-control design, nested case-control study Case-control study carried out within the population of a cohort, whether constructed for this purpose or not (*see* Appendix 1).

All the subjects in the cohort having presented the event of interest during the follow-up (the *cases*) are compared, regarding exposure to the studied risk factor, with subjects in the cohort not having presented this event (the *controls*) with possible matching or adjustment.

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

As opposed to classical case control studies, absolute risk can be estimated.

See also: **case-cohort study, case-control study, cohort study, confounding.**

Network A group of geographically distinct entities, persons or structures collaborating on a common objective.

In pharmacovigilance or pharmacoepidemiology, the term can designate a group of clearly identified, voluntary observers (physicians or pharmacists) regularly participating in a project coordinated by a single organization. The aim could be to identify or quantify a risk, or to describe the prescription or use of one or more drug(s).

Nocebo effect *See* **placebo effect.**

Non-compliance *See* **compliance.**

Non-controlled study, non-controlled trial Study in which the value of a measurement is not compared to a measurement taken from another sample nor to a reference value, known or fixed in advance (*see* Appendix 1).

A non-controlled strategy is frequently used in descriptive epidemiology to measure the frequency of a phenomenon, to describe it, or study its evolution with a longitudinal study.

The terms '*open study*' and '*open trial*' have the same meaning, but can lead to confusion by implying that the sample population changes during the study period (the term '*open population*' often being used in epidemiology as a synonym for '*dynamic population*').

See also: **controlled trial, crossover trial, dynamic, longitudinal study, open study.**

Non-experimental study *See* **observational study.**

Non-participant, non-responder Person contacted for or included in a study, or a poll, etc., and who does not actively participate for whatever reasons.

Non-participation can be the cause of a *selection bias*, i.e. non-representativeness of the studied sample. It is thus crucial to

compare the characteristics of non-participants to those of participants, at least with regard to variables possibly linked to the studied parameter.

See also: **selection bias.**

Normal approximation Use of the normal, instead of the binomial, distribution as a reference for a probability calculation or a statistical test for a binary variable.

This approximation is valid when the probability of occurrence of an event is close to 0.5 or when the study sample size is large ($\geq$ 100) and the number of events occurring or expected is itself large (in practice, greater than 15).

See also: **binomial distribution, normal distribution.**

Normal care relationship Therapeutic relationship in which contact with the patient, choice of treatment, expected benefit, incurred risks and surveillance methods are neither modified nor organised by external intervention.

The physician takes only the interests of the patient into account, and not those of a third party interested in the results of a study, e.g. a sponsor; he or she is guided by the present state of medical knowledge and the recommendations of the scientific community or health authorities. The maintenance of the normal care relationship between the physician-researcher and the patient defines the conditions of an observational study.

See also: **observational study.**

Normal distribution, normal law, Gaussian distribution Extremely important law in probability and statistics, describing the distribution of a variable, the value of which results from the influence of a number (considered to be infinite) of independent factors acting in an additive manner (e.g. the distribution of the height of the inhabitants of a region or country; the level of blood glucose or creatinine in a sufficiently large population).

The normal distribution describes the frequency distribution of a variable x, from $-\infty$ to $+\infty$, according to the following

equation:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)2/2\sigma^2}.$$

The bell-shape of the resulting distribution depends solely on the value of the parameters μ (the mean) and σ (standard deviation). This distribution is continuous and symmetrical with respect to $x = \mu$.

In addition, a normal distribution with mean $\mu = np$ and standard deviation $\sigma = \sqrt{n\ p\ (1-p)}$ is a satisfactory approximation of the binomial distribution of parameters n and p as soon as np becomes large (greater than 15).

The normal distribution is the probability distribution used most often for statistical tests and calculations (e.g. calculations of confidence intervals, comparisons of means, calculations of sample sizes, etc.). The validity of these calculations and the resulting inferences requires, however, that the variable under consideration is distributed according to the normal law in the studied population. In the case of an approximation (e.g. of the binomial distribution by the normal distribution), it is advisable to verify the validity of this approximation (i.e. the sample size is sufficiently large and the expected number of events is not too small).

See also: **binomial distribution, Poisson distribution.**

Notification *See* **reporting.**

Notoriety bias Selection bias in which a case has a greater chance of being reported if the subject is exposed to the studied factor known to cause, thought to cause, or likely to cause the event of interest.

Example: This bias is classic in case-control studies of digestive bleeding. A physician could be inclined to preferentially report cases for which exposure to a non-steroidal anti-inflammatory drug has been noted.

See also: **selection bias.**

Null hypothesis, test hypothesis Tested hypothesis which may possibly be rejected at the end of a study or calculation, in favour of another, called the *alternative hypothesis*. Both hypotheses should be clearly stated *a priori*.

Example: Comparison of two proportions p_1 *and* p_2 *involves testing the null hypothesis of equality of the proportions* H_0*:* $p_1 = p_2$ *against one, and only one, of the following alternative hypotheses:* H_1*:* $p_1 \neq p_2$ *(two-sided statistical test),* H_2*:* $p_1 < p_2$ *or* H_3*:* $p_1 > p_2$ *(one-sided tests).*

Type I error or α error consists of wrongly rejecting the null hypothesis when it is true. Type II error or β error consists of not rejecting the null hypothesis when it is false.

See also: **power, Type I error, Type II error.**

O

Observational bias *See* **information bias.**

Observational study, non-experimental study Approach in which the researcher merely observes a population or phenomenon without interfering in any way with the natural course of events (*see* Appendix 1).

In this way, 'observational' is equivalent to the term '*non-interventional*' used by some researchers. In clinical pharmacology, pharmacovigilance and pharmacoepidemiology, a study can be described as '*observational*' if all aspects (e.g. allocation and length of treatment, surveillance method, risk incurred, etc.) are carried out as if no study existed.

The main value of such studies is that they are based on real-life situations, in order to describe them most faithfully. However, the absence of an experimental design increases the possibility of bias (notably confounding), and makes causal analysis difficult.

See also: **experimental, interventional, normal care relationship.**

Odds Probability that an event will occur divided by the probability that it will not occur:

$$Odds = \frac{probability}{1 - probability}.$$

Example: If the odds in favour of an event are 12, this means that the event is 12 times more likely to happen than to not happen during a given period.

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Similarly, it is possible to calculate the odds of a characteristic or of exposure to a risk factor, e.g. the probability that a subject is (or has been) exposed divided by the probability that he or she is not (or has not been) exposed, during a given period.

Example: If 31 of 120 subjects have been exposed to a drug, the exposure odds of the drug for this population is 31 / (120 − 31) = 0.35.

From the value of an odds, it is easy to calculate the corresponding probability or proportion:

$$\textit{Probability, proportion} = \frac{odds}{odds + 1}.$$

In the preceding examples, an odds of 12 means that the event has a probability of 12 / (12 + 1) = 0.92, or 92 chances in 100 of occurring during the period of interest. Similarly, an exposure odds of 0.35 corresponds to an exposure proportion of 0.35 / (0.35 + 1) = 0.26 or 26%.

See also: **odds ratio.**

Odds ratio Ratio of two odds, each being the ratio of one probability to its complementary probability.

Thus, if p_1 and p_2 are two proportions (e.g. of diseased subjects, of exposed subjects, etc.), each measured in a different group of subjects, the corresponding odds ratio (OR) is:

$$OR = (p_1 / (1 - p_1)) / (p_2 / (1 - p_2))$$

When p_1 and p_2 are small, $(1 - p_1)$ and $(1 - p_2)$ are very close to 1; in this case, the odds ratio is practically the same as the ratio of the proportions p_1/p_2.

The odds ratio is used often in epidemiology and pharmacoepidemiology, for three main reasons:

- Unlike the relative risk, it can be calculated equally well in a prospective (cohort) study and in a retrospective

(case-control) study.

- It has mathematical properties which permit analysis by common models, such as the logistic model.
- When the probability of occurrence of the event (the risk) is low, the odds ratio is practically equal to the relative risk, and thus constitutes a satisfactory or even very good approximation of the latter (*see* below).

In a case-control study, we calculate the ratio of the odds of exposure to the factor of interest in the cases and controls:

	Cases	Controls
Exposed	a	b
Non-exposed	c	d

$$OR = (a/c) \,/\, (b/d) = \frac{ad}{bc}.$$

This term quantifies the strength of the association between exposure and the presence of the characteristic or disease defining a case; moreover, if the event of interest is rare, it is a good estimate of the relative risk. We can see in the table above that in a cohort study, the relative risk (RR) would be calculated by:

$$RR = \frac{a \;/\; a \times b}{c \;/\; c + d}.$$

If both the compared risks are low, a and c are small relative to b and d, and the formula can be rewritten as:

$$RR \cong \frac{a \;/b}{c \;\; d} = \frac{ad}{bc}.$$

In this situation, we can conclude the equivalence of the relative risk and the odds ratio.

See also: **case-control study, odds, relative risk.**

One-sided test, one-tailed test Statistical test in which the rejection

area of the null hypothesis lies on only one side of this hypothesis.

Example: In order to compare the risks p_1 and p_2 associated with two drugs, the null hypothesis $H_0 : p_1 = p_2$ is tested against a single pre-determined alternative hypothesis, $H_1 : p_1 > p_2$ or $H_2 : p_1 < p_2$. This alternative hypothesis will be accepted at a fixed risk of Type I error if H_0 is rejected. However, non-rejection of H_0 does not allow the conclusion that the risks are equal.

The prior choice of a one-sided test, rather than a two-sided one, generally allows a reduction in the sample size required to guarantee a given probability (power) of rejecting the null hypothesis when it is false. However, this imposes the choice of only one alternative hypothesis (in the example above, $p_1 > p_2$ or $p_1 < p_2$), which is not always possible or convenient.

See also: **equivalence study, power, two-sided test, Type I error, Type II error.**

One-tailed test *See* **one-sided test.**

Onset delay, time to onset Interval of time separating the beginning of exposure to a risk factor (e.g. treatment by a drug) and the occurrence of a phenomenon (e.g. an adverse event).

Unlike the term '*latency period*', onset delay may be used even in the absence of an association between the two episodes, and is a particularly useful term in the absence of a causal association.

See also: **latency period.**

Open population *See* **dynamic population.**

Open study, open trial According to difficult authors, *open study* has different meanings:

- A *non-controlled*, non-comparative study.
- A study based on a *dynamic*, i.e. renewed, sample.
- A non-blinded, controlled or non-controlled, study.

See **blind, dynamic, non-controlled study.**

Outcome Expected or measured result of exposure to a causal factor or of a therapeutic intervention.

Over-matching In comparing two populations, over-matching occurs if the matching procedure:

- Is needlessly complex and, for example, considers variables which can in no way act as confounding factors (this can restrict the feasibility of the study, especially regarding the possibility of recruiting appropriate controls).
- Affects one or more variable(s) associated with the measured parameter, i.e. the probability of occurrence of the studied event (in a cohort study) or the probability of exposure (in a case-control study). The purpose of matching is to control the influence of variables which could act as confounders, by distributing them in a balanced manner among the groups being compared. A confounder is, by definition, linked to both the status of the subject (i.e. exposed/non-exposed in a cohort study or diseased/non-diseased in a case-control study) and to the analysed variable (i.e. the occurrence of the disease in a cohort study, the exposure in a case-control design). To match on a factor linked only to the analysed variable may result in partial or complete masking of an association between exposure and disease.

For example, in a case-control study investigating the association between the use of anorectic agents and the occurrence of pulmonary arterial hypertension, the cases and controls were matched by weight, among other factors. In a situation (unfortunately only theoretical) where the anorectic agents are systematically taken by overweight subjects, and only by them, the fact of matching by weight would be the same as matching by the studied factor (the use of these agents), which would mask their possible association with the disease.

See also: **case-control study, confounding, matching, odds ratio.**

P

Panel Sample of subjects (usually representative of a larger population) selected for the purposes of an inquiry, generally on a prolonged or permanent basis.

For example, in pharmacoepidemiology, panels of users or prescribers are used to study the utilization or prescription of drugs.

See also: **drug utilization study, selection bias, source** population, survey.

Parallel groups Strategy wherein two or more groups, which differ with respect to the presence of a studied factor, are formed and are followed in an adequate and sufficiently prolonged manner in order to compare them as to the frequency of a biological, clinical or behavioural effect, or any other parameter deemed to be pertinent (e.g. quality of life, cost of care, etc.).

Controlled clinical trials, pragmatic trials, and cohort studies with a reference group are all parallel group strategies.

See also: **clinical trial, cohort study, controlled trial, pragmatic trial.**

Per protocol analysis Statistical analysis based on patients having strictly conformed to the protocol in a study of efficacy or tolerance. This approach is preferably used in explanatory trials.

See also: **explanatory trials, intent-to-treat analysis.**

Peri-approval study, peri-approval trial Study or trial conducted immediately before or after drug approval is obtained with

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

the aim of completing Phase III clinical trials, particularly with respect to tolerance.

Such investigations almost always take the form of cohort studies.

Personal therapeutic benefit Possible benefit to the health or well-being of a patient expected due to his or her participation in biomedical research.

Biomedical research (e.g. a clinical trial) can accordingly be described as being with or without personal therapeutic benefit. Research is generally said to have personal therapeutic benefit if at least some of the participating subjects can expect a benefit to their state of health or well-being. Thus, a blinded, randomised trial that compares the efficacy of a drug with that of a placebo may be deemed to have a personal therapeutic benefit even though, by definition, some of the patients in the trial, unknown in advance, will not receive the active treatment.

Person-days *See* **person-time.**

Person-months *See* **person-time.**

Person-time Sum of the follow-up times of each of the members of a population, generally used as the denominator of an incidence rate.

This unit is particularly useful when the follow-up duration differs between subjects. It can be expressed in person-years, person-months, person-weeks or person-days. When using data from spontaneous reporting in pharmacovigilance, it is customary to consider the duration of treatment, roughly estimated from sales data, rather than that of follow-up, and speak in terms of treatment-days (days of treatment), treatment-months (months of treatment) or treatment-years (years of treatment).

For example: Twelve cases per 100,000 treatment-months means that, for the period and region of interest, 12 cases of an

event were reported for a cumulative duration of drug treatment of 100,000 months. This conveys no information regarding the duration of each individual treatment: 100,000 treatment-months could correspond to 100,000 subjects treated for 1 month, 3.04 million subjects treated for 1 day or 8333 subjects treated for 1 year; it is most likely that the treatment durations varied from subject to subject.

See also: **incidence density, spontaneous reporting.**

Person-years *See* **person-time.**

Pharmacodependence *See* **dependence.**

Pharmacodynamic Relative to the action of a drug.

The objective of the field of *pharmacodynamics* is to describe, quantify and explain the set of effects induced by a drug in a living organism.

See also: **pharmacodynamic effect, pharmacology.**

Pharmacodynamic effect Organic or biological, quantifiable, reproducible and usually dose-dependent manifestation related to the administration of a drug to a living organism.

The observation of a pharmacodynamic effect does not necessarily imply the existence of a therapeutic effect.

See also: **therapeutic effect.**

Pharmacoeconomics Study of evaluation of the medico-economic consequences attributable to the use of a drug.

This assessment can be made before placing a drug on the market (predictive studies) or afterwards, possibly including comparison with another strategy. The assessment may make use of predictive models, explanatory or pragmatic clinical trials or observational studies.

See also: **cost–benefit study, cost–effectiveness study.**

Pharmacoepidemiology Study of interactions between drugs and populations.

More restrictively, '*pharmacoepidemiology*' can be defined as the study of the therapeutic effect(s), risk and use of drugs,

usually in large populations, using epidemiological methods and/or reasoning.

Although experimental studies are sometimes preferable, the goal of the pharmacoepidemiological approach is generally to describe the situation in real life, avoiding as much as possible any modification to this situation caused by the study itself.

This applies to descriptive approaches (e.g. describing a phenomenon or estimating a risk in real circumstances) as well as to aetiologic approaches (e.g. determining the extent to which a drug is likely to increase a risk under real conditions of use).

See also: **epidemiology, experimental, observational study.**

Pharmacogenetics Study of the variation of the effects or pharmacokinetics of drugs as a result of the genetic characteristics of an living organism.

Since the majority of the sites of action and transformation of drugs are genetically determined, polymorphism may cause variations between individuals in the metabolism of, or response to, a given drug.

Example: An hydroxylase involved in the oxidative metabolism of a drug can be of high or low activity, making it possible to identify those subjects in a population who are rapid or intensive hydroxylators and others who are slow hydroxylators. The administration of the same dose will produce, a priori, *a weaker and/or shorter effect in the former subjects.*

See also: **effect, metabolite, pharmacokinetics.**

Pharmacokinetics Study of the qualitative and quantitative aspects of the fate of a drug within a living organism (human or animal).

Pharmacokinetics is concerned with the evolution of concentrations of a drug and its possible metabolites in the blood and certain biological tissues after or during

administration.

See also: **pharmacology.**

Pharmacology Study of interactions between drugs and living organisms.

The term '*interaction*' should be understood to encompass the influence of a drug on a living organism (*pharmacodynamics*, *therapeutic effect*) and that of a living organism on a drug (*pharmacokinetics*, *metabolism*).

Clinical pharmacology is concerned with the fate and effects of drugs in human beings.

See also: **drug, pharmacodynamic, pharmacokinetics.**

Pharmacosurveillance *See* **pharmacovigilance.**

Pharmacovigilance The detection, evaluation, understanding and prevention of adverse drug reactions.

This term should be used instead of 'post-marketing surveillance' and 'pharmacosurveillance'. The pharmacovigilance approach can be clinical, epidemiological, experimental (e.g. attempts to reproduce an adverse effect in animals in order to understand the mechanism involved) or diagnostic (e.g. imputability methods). The ultimate aim of pharmacovigilance is the optimization of the risk–benefit ratio of marketed drugs at the individual level (i.e. the choice of the most suitable treatment for a given patient) or at the population level (i.e. maintenance or removal of a drug from the market, informing prescribers of its potential risks, etc.).

See also: **imputability, risk-benefit ratio.**

Phase I clinical trial First trial of a new active ingredient in humans.

This type of trial is usually conducted on healthy volunteers, and its purpose is to assess tolerance as a function of dose and to carry out the first studies of clinical pharmacokinetics.

See also: **volunteer.**

Phase II clinical trial Therapeutic pilot study designed to

demonstrate activity and to assess short-term safety of an active ingredient in patients suffering from a disease or condition for which this active ingredient is intended.

This type of trial is performed on a limited number of subjects and often, at the end of this phase, with a comparative protocol (e.g. against a placebo). Phase II trials are also used to determine the appropriate doses, dosages and regimens and, if possible, to clarify the dose–response relationship, in order to provide an optimal basis for the design of therapeutic trials with a greater number of subjects.

A distinction is sometimes made between Phase II_A trials, which characterise the pharmacodynamic effects of drugs and study tolerance, and Phase II_B trials, focused on studying therapeutic activity as a function of dose.

Phase III clinical trial Trial involving larger and, if possible, more diversified groups of patients than those involved in Phases I or II in order to determine the short- and long-term risk–benefit ratio, and to assess the overall and relative therapeutic value of one or more formulation(s) of an active ingredient.

These trials permit investigation of the type and profile of the most frequent adverse effects, as well as the clinically relevant drug interactions and the influence of factors such as age on the results. The trial protocol will preferably be randomised and double-blind, but other designs may be acceptable, especially for long-term safety studies. Generally, the conditions of a Phase III trial should be as close as possible to normal conditions of use.

Phase IV Designating all studies or trials of a drug conducted after its marketing approval, and under the conditions (e.g. indication, dose, duration, etc.) specified by this approval.

The term 'Phase IV', unlike Phases I, II and III, does not imply a particular method or objective, but refers to an

administrative situation, namely the post-approval period. It is thus preferable to be more specific and speak of a 'Phase IV study' or 'Phase IV trial', instead of simply 'Phase IV'.

See also: **Phase IV trial, trial.**

Phase IV trial, Phase IV study Post-marketing trial or study of a drug conducted on the basis of the information (indications, dose, etc.) contained in the Summary of Product Characteristics.

They can be useful for pharmacovigilance purposes or to assess therapeutic utility or cost of a strategy. Although prevailing conditions should be taken into account, Phase IV trials must be conducted with the same methodological rigour as pre-marketing clinical trials and, at the very least, include a precise protocol.

From a regulatory viewpoint, all post-approval trials designed to obtain an extension of the approval (e.g. for a new indication, new dosage, etc.) are Phase II or Phase III trials. For studies conducted immediately after approval and possibly before launching, the terms '*peri-approval*' or '*post-approval studies*' are sometimes used.

See also: **observational study, post-marketing surveillance, study.**

Placebo Substance, pharmaceutical preparation or any other treatment presented as being therapeutically efficacious when it is in fact devoid of intrinsic activity, but not necessarily of therapeutic effect.

Such a substance can cause an improvement in the state of health or well-being of a patient ('*placebo effect*') or provoke an adverse event ('*nocebo effect*').

When it is possible, the use of a placebo in one of the groups in a clinical trial (called a '*placebo group*') allows estimation of the significance of the effects induced by the prescription, the prescription environment or the simple fact of participation in a clinical study; an optimal reference is

thus obtained for the quantification of the studied drug's effects.

See also: **controlled trial, placebo effect.**

Placebo effect Modification to the state of health or well-being of a subject not explained by a pharmacological property of a drug, but apparently related to the psychological context surrounding its prescription, its environment or the patient's confidence in the efficacy of the drug.

This effect can also be observed with non-medical treatments. The etymology of the word '*placebo*' (from the Latin, I will please), restricts its usage to beneficial effects expected by the patient. For adverse manifestations, the term '*nocebo* effect' is generally used.

See also: **placebo.**

Point prevalence Number of prevalent cases found in a given population during a very short time interval approaching zero.

See also: **cross-sectional study.**

Poisson approximation Use of the Poisson, instead of the binomial, distribution as a reference for a probability calculation or a statistical test for a binary variable.

This approximation is valid when the study sample size is large ($\geq$ 100) and the probability of occurrence of an event approaches 0 (in practice < 0.1).

See also: **binary variable, Poisson distribution.**

Poisson distribution, Poisson law Law describing the probability of observing x successes when m are expected over the course of a large number of trials:

$$p(x) = \frac{e^{-m}m^x}{x!}$$

where e is the base of natural logarithm, i.e. 2.7183...

The cumulative probability of observing at least k successes is:

$$p(x \geq k) = 1 - \sum_{x=0}^{x=k-1} \frac{e^{-m} m^x}{x!}.$$

The Poisson formula is derived from a simplification of the binomial formula when the number of trials, n approaches infinity (in practice, more than 100) and the probability of success, p (assumed to be constant for each trial) approaches 0. Under these conditions, the results obtained with the Poisson formula are very close to those produced by the binomial formula, the latter being difficult to use when n is large (due to the calculation of the nth factorial). The expected number of successes, m, is both the mean and the variance of the distribution.

In pharmacoepidemiology, m usually represents the expected number of events in the population during a given time period, and is the product of the population size n and the average probability of the occurrence of the event p.

Example: If the risk of an adverse event in treated patients is 0.5%, the expected number of events during the follow-up of 1200 patients will be $1200 \times 0.005 = 6$. *According to the formulae above, the probability of observing eight cases of this event is 0.103 or 10.3% and the probability of observing eight cases or more is 0.256 or 25.6%.*

When $m = np$ is large ($\geq$ 15), the distribution does not differ significantly from a normal distribution with mean np and standard deviation $\sqrt{np}$.

The Poisson distribution is widely used in pharmacoepidemiology when low risks and large populations are involved (e.g. to calculate the probability of occurrence of a given number of cases of an event, to calculate a confidence interval for a rate, to estimate required sample sizes, etc.). The Poisson equation is also the basis for several models used in epidemiology (e.g. Poisson regression).

See also: **binomial distribution, expected number, normal distribution, Poisson approximation, Poisson process, Poisson regression.**

Poisson process Situation in which the probability of occurrence of an event can be described by the Poisson distribution.

Independent of the general conditions under which the Poisson distribution applies (i.e. number of trials approaching infinity and probability of success approaching zero), any phenomenon which satisfies the following three conditions corresponds to a Poisson process:

- During each unit time interval in the observation period, the event of interest can occur only once, or not at all.
- The probability that the event will occur during a time interval is independent of whether or not it occurred during the previous intervals (memory-free process).
- The probability that the event will occur in a given time interval is independent of the chronological order of the interval.

See also: **Poisson distribution.**

Poisson regression Procedure consisting of predicting the value of a dependent variable (generally, a risk) from the value(s) of one or more explanatory variable(s), by using the mathematical model derived from the Poisson equation.

See also: **Poisson distribution.**

Poll *See* **survey.**

Population control *See* **community.**

Population-time Sum of follow-up or exposure times for a population. This term is sometimes used to designate the amount of person-time used to calculate an incidence rate.

See also: **incidence density, person-time.**

Positive predictive value Probability that a subject for whom a diagnostic test is positive is diseased.

The *positive predictive value* is calculated by dividing the number of *true positives* (*a*) by the number of times the test results are positive ($a + b$):

	Diseased	Non-diseased	
Positive test	*a*	*b*	$a + b$
Negative test	*c*	*d*	$c + d$

See also: **negative predictive value, true positive.**

Post-approval study *See* **Phase IV trial.**

Posterior odds *See* **Bayes's theorem.**

Posterior probability *See* **Bayes's theorem.**

***Post hoc* analysis** An analysis that was not anticipated or described before data collection.

See also: ***ad hoc* study.**

Post-marketing Observational Study (PMOS) Process of monitoring and evaluating the safety of marketed medicines using a variety of methods.

Their purpose is generally to study large populations of users in ordinary practice with the aim of identifying adverse drug reactions that had been missed, before marketing because clinical trials were too small, too short or did not always reflect real life. To some extent these studies have now fallen into disrepute, both on scientific grounds and because they have often been thought to be a thinly-disguised promotional exercise. They should comply with existing guidelines such as those of the MCA (Medicines Control Agency, UK) and ABPI (SAMM) (Association of the British Pharmaceutical Industry) or the CPMP (Committee for Proprietary Medicines Products) Notice to Applicants (all can be accessed via EMEA: *see* References),

See **cohort.**

Post-marketing Surveillance (PMS) *See also*: **pharmacoepidemiology, pharmacovigilance.**

Potentialization, potentializing effect *See* **synergy.**

Power, statistical power Probability that a statistical test will reject the null hypothesis when it is false in favour of a specified alternative hypothesis.

Statistical power is the complementary probability of the Type II or b error (*power* $= 1 - \beta$).

The statistical power defines, for example, the capability of a statistical test to reveal a given difference between two treatments, if this difference really exists.

Example: If the aim is to detect a difference of at least a factor of 2 between two risks, the choice of a power of 0.8 ($\beta = 0.2$) means that there is at least 80% chance of detecting this difference, if it exists.

Power depends on the size of the difference we wish to detect, the chosen risk of Type I error and the number of subjects included in the study.

See also: **required sample size, Type II error.**

Pragmatic That which is adapted to conditions in the real world, can have practical applications, concerns everyday life.

In the case of a drug assessment, a judgement criterion, trial or study is termed '*pragmatic*' if it approaches normal or usual conditions of care.

Example: In a clinical trial, inclusion or assessment criteria are called 'pragmatic' if they are based on the clinical examination or check-up usually carried out at the time of prescription of the studied drug.

This term is not synonymous with '*observational*' in the sense that a pragmatic assessment can be made within an experimental design, as is the case in a pragmatic trial, for example.

See also: **observational study, pragmatic trial.**

Pragmatic trial Trial using pragmatic criteria chosen *a priori*, in order to assess the overall therapeutic value of a treatment compared to an already existing strategy (*see* Appendix 1).

This kind of trial thus differs in its principle and objectives from an *explanatory clinical trial*, which seeks to demonstrate

the intrinsic efficacy and/or tolerance of a treatment under standardized conditions. Ideally, a pragmatic trial is conducted after Phase III trials, when the treatment and strategies to compare have already been in use for some time and are well-known to the medical community.

Treatments are allocated at random, as in an explanatory trial, but the nature of these treatments is known to both patient and physician, which allows the normal therapeutic relationship to be taken into account. In this way, the prescriber can be guided in making his or her therapeutic decisions with full knowledge of the prescription context. The absence of blinding may, however, be a limitation if the assessment of the judgement criterion is subjective.

In the strictest sense, this type of trial does not leave open the possibility of not reaching a conclusion, the treatment with the best results is adopted (balancing the advantages and disadvantages) without a statistical test or the possibility of non-decision (which corresponds, in the extreme, to accepting a Type II error of 0% and a Type I error of 100%). On the other hand, minimizing Type III error (γ error), an error which would consist of adopting the worse of the two options, requires calculation and recruitment of a sufficient number of subjects.

See also: **Phase III clinical trial, simplified clinical trial, Type II error, Type III error.**

Prediction interval Represents the set of possible values of a given parameter having a given pre-chosen probability of being observed in a sample of size n, assuming knowledge of the true value of this parameter.

Example: If the probability of occurrence of a neutropenia over the course of a treatment is 1%, 10 cases can be expected if a sample of 1000 persons is followed. A calculation using the Poisson approximation shows a 95% chance that the observed number of cases will be between 4 and 18. These values represent the

two-sided 95% prediction interval for the expected number of events in this sample.

The method of calculation is similar to that of a confidence interval except that the reasoning is reversed: a confidence interval is constructed around an observed parameter and is assumed to include (with a given probability) the true, unknown value of this parameter.

See also: **confidence interval, random error.**

Prescribed Daily Dose (PDD) Average dose of a drug prescribed per day for a given therapeutic indication (in mg, units of administration, etc.).

It can be estimated from prescriber panels, drug utilization study or databases. It can differ from the recommended daily dose (RDD).

See also: **defined daily dose, drug utilization study, recommended daily dose.**

Prescription bias *See* **confounding by indication.**

Prescription channelling *See* **confounding by indication.**

Prescription Event Monitoring (PEM) A post-marketing surveillance scheme in the UK, based on a follow-up of the first prescription of selected new drug.

Prevalence Number of cases of an event recorded in a population during a given period of time.

The prevalence of a disease includes cases occurring during the study period (*incident cases*) and the cases occurring prior to this period that have yet not recovered. Prevalence measured in a population is thus always higher than incidence except when the duration of the disease is very short; in this case, the prevalence and incidence tend to be identical. In common parlance, the term 'prevalence' is often used to designate a prevalence rate (e.g. a prevalence of 8 per 1000).

See also: **case, cross-sectional study, point prevalence, prevalence rate.**

Prevalence rate Number of cases reported in a population during a given period, compared to the number of individuals at risk in the population.

This rate includes cases appearing during the period of interest (incident cases) and the cases occurring prior to this period that have not yet recovered.

See also: **prevalence, standardization.**

Prevalence study *See* **cross-sectional study.**

Prevalent case *See* **prevalence.**

Preventability Term fixed by usage, referring to the avoidable character of particular adverse effects of a therapeutic strategy, usually a medicinal treatment.

An adverse effect that would not have occurred had the therapeutic attitude conformed to the most commonly recognized recommendations (e.g. regarding whether or not to treat, choice of treatment, dose, duration, etc.) is said to be '*avoidable*' or '*preventable*'. A study of preventability seeks to identify, enumerate and describe the adverse effects that can be prevented, and to propose actions to ensure this prevention in order to diminish the risk associated with medicinal treatments.

See also: **adverse drug reaction.**

Preventable fraction of the risk in the population Proportion by which the risk of an event is, or would be, reduced in a population due to exposure of a portion of the subjects of this population to a factor thought to prevent the occurrence of the event.

The calculation of the preventable fraction is identical to that of the attributable fraction of the risk in the population (*see* the notation in the latter definition):

$$FP_{\text{pop}} = \frac{R_{\text{non/exp}} - R_{\text{exp}}}{R_{\text{non/exp}}},$$

$$FP_{\text{pop}} = \frac{E_{\text{pop}}\ (1 - RR)}{1 + E_{\text{pop}}\ (1 - RR)},$$

or:

$$FP_{\text{pop}} = \frac{E_{\text{pop}}\ (1 - OR)}{1 + E_{\text{pop}}\ (1 - OR)}.$$

where E_{pop} represents the proportion of the population exposed to the preventive factor.

For example, a preventive treatment for osteoporosis would diminish the risk of vertebral stress fracture in menopausal women by a factor of 5 (i.e. RR = 0.2). If 10% of the population at risk for this injury is treated with the drug, the preventable fraction is:

$$FP_{\text{pop}} = \frac{0.1\ (1 - 0.2)}{1 + 0.1\ (1 - 0.2)} = 0.074.$$

This means that the risk (or number of cases) of fracture in the whole population would be reduced by 7.4% by this treatment of 10% of the population.

The validity of this type of estimate depends on the same conditions expressed at the end of the definition for the attributable fraction of the risk in the population.

See also: **attributable fraction of the risk in the population.**

Prior odds *See* **Bayes's theorem.**

Prior probability *See* **Bayes' theorem.**

Probability Relative frequency defining the number of times an event is expected to occur or a hypothesis is expected to be verified over a large number of attempts, approaching infinity.

A probability is generally expressed as a proportion between 0 and 1, sometimes converted to a percentage from 0 to 100%.

See also: **frequency, risk.**

Pro-drug Drug of weak or non-existent activity but which gives rise *in vivo* to an active metabolite responsible for all or most of the therapeutic effect.

See also: **metabolite.**

Prolective Seldom used adjective designating a study of any type

(e.g. cohort or case-control) concerned with events likely to occur after the beginning of the study.

See also: **historical, prospective, retrolective, retrospective.**

Prompted reporting *See* **spontaneous reporting.**

Proper use Prescription or use of a drug conforming to recommendations.

Proper use concerns the prescriber (respecting the Summary of Product Characteristics and medical references as to indications, contraindications, dosage, length of treatment, etc.), as well as the patient (respecting oral or written recommendations regarding both prescribed and non-prescribed medications).

See also: **abuse, misuse, Summary of Product Characteristics.**

Proportion Number of subjects presenting a given characteristic divided by the total number of subjects in the study population.

A proportion can vary from 0 to 1, and can be converted into a percentage by multiplying it by 100% or into a rate by multiplying by a power of 10.

For example, 16 cases observed during one year in a sample of 827 subjects is equivalent to a proportion of 16/827 = 0.0193, 1.93% or an annual incidence rate of 19.3 per 1,000.

See also: **frequency, rate.**

Proportional hazards model *See* **Cox model.**

Prospective That which concerns the future.

This term is applied to studies in which a population is followed from a given date in order to detect events likely to occur after this date. A cohort study is thus prospective. The seemingly redundant term '*prospective cohort*' is sometimes used to designate a cohort whose subjects are identified and included after the beginning of the study and followed into the future. Such a cohort is thus the opposite of an '*historical*' or '*ambispective cohort*'. For many authors, the term '*concurrent cohort*' is preferred.

See also: **ambispective, cohort, concurrent, historical, prolective.**

Prospective study Any epidemiological study in which the subjects

are followed after their inclusion in order to study the possible appearance of a phenomenon (*see* Appendix 1).

This term is thus synonymous with 'cohort study' and can be based on the descriptive analysis of a single cohort or an aetiologic analysis comparing two or more cohorts.

See also: **cohort study, prospective.**

Protective effect Diminution of the risk of occurrence of an adverse event related to exposure to a factor or the presence of a characteristic.

In controlled studies with a reference group, a protective effect can be shown when a relative risk or odds ratio is significantly lower than 1.

See also: **odds ratio, relative risk.**

Protocol Document describing the rationale, objectives, approach and statistical methods of a trial and defining the conditions under which the trial will be carried out and analysed.

This definition also applies to epidemiological studies, including observational approaches. The existence of a protocol (which formalizes the procedure to make it reproducible by other observers) does not necessarily imply the introduction of constraints. This would be the case if the protocol imposed an additional element as a requirement for the study that would affect the normal conditions of care or the natural course of events.

See also: **constraint.**

Protopathic bias Distortion of the estimate of the association between the exposure to a risk factor (e.g. the use of a drug) and the occurrence of an adverse event linked to the fact that the exposure actually started after the occurrence of the event, or at least after the first appearance of its manifestations.

In the most pronounced and restrictive form of this bias, the use or prescription of the drug may even be motivated by the appearance of the first symptoms of the disease of interest.

Example: A cardiovascular drug could be wrongly accused of

causing pulmonary fibrosis if it is frequently prescribed for dyspnoea which is, in fact, the first sign of incipient fibrosis.

See also: **confounding by indication.**

Q

QALYs Acronym for 'Quality Adjusted Life Years'.

Expression in QALYs balances the number of years of life gained due to the implementation of a therapeutic strategy against the quality of the subject's life during this period.

In this way, a strategy ensuring a survival duration which is shorter but of better quality than another could be considered superior in a comparison expressed in QALYs.

For many types of ailments, notably in cardiology and oncology, scores allow quantification of the discomfort felt by the patient, and thus a conversion of survival duration into number of QALYs.

Expression in QALYs is particularly used in cost–utility studies.

See also: **cost–utility study.**

Qualitative variable Variable which can take only discrete values that have no quantitative relationship to one another (e.g. sex, blood type, race, etc.). These are often permanent and characterize an individual.

See also: **continuous variable, discrete variable.**

Quantitative variable *See* **continuous variable, discrete variable.**

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

R

Random Situation in which an outcome is not systematically the result of an intervention (e.g. the prescription of a given drug) but is subject to chance and can be described by an appropriate probability model.

Examples: random selection of the controls for a study, random occurrence of an adverse event, etc.

See also: **binomial distribution, randomisation.**

Random error Deviation of the value of a parameter estimated in a sample from its true value.

As opposed to bias, this deviation is not systematically in a given direction, but assumes a random character and may be described by an appropriate probability model.

Thus, the measurement of a parameter in a randomly chosen sample may assume a certain number of values compared to the true value that would be measured in the entire source population. This set of possible values, reflecting the sampling variation, can be predicted by an appropriate probability distribution (depending on the case, the binomial, Poisson or normal distribution), and constitutes the *prediction interval* for the measurement of this parameter.

Example: If the risk of agranulocytosis associated with a medical treatment is 2 per 1000, four cases can be expected to occur in a follow-up of 2000 treated patients. A calculation using the Poisson distribution shows that the number of cases that will be observed in this sample of 2,000 patients will be between 1 and 10, 95 times out of 100 (two-sided 95%

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

prediction interval).

In the same way, the construction of a confidence interval around the value measured in a sample minimises the risk of erroneous conclusion with respect to the true value of the parameter.

See also: **bias, binomial distribution, confidence interval, normal approximation, Poisson approximation, prediction interval, random, source population.**

Randomisation Decision or designation depending wholly upon chance.

In a clinical trial, for example, randomisation is used to decide with a table of random numbers or a computerised procedure which of the compared treatments each subject will receive, without subjective intervention.

Randomisation is the only method which guarantees a balanced distribution among groups of known or unknown variables likely to bias the planned analysis. It is only within this framework that statistical tests are strictly reliable.

See also: **random.**

Rate Quantification of the frequency of an event in a population per unit of time.

The definition of a rate always includes a reference to time (e.g. an incidence rate of six cases per 10,000 subjects per year). When follow-up time is not the same for all subjects in the studied population, the number of cases of the event is often compared to a sum of person-times sometimes called *population time* (e.g., five cases per 1000 person-years). This denominator is obtained by summing the follow-up or exposure times of all of the subjects in the population.

An important characteristic of a rate is the fact that the cases (e.g. the subjects having presented the event) forming the numerator come from the population represented by the denominator. In other words, only those subjects considered

a priori susceptible to presenting the event and observed in a manner that guarantees identification of the event should be included in the denominator.

See also: **incidence density, incidence rate, person-time, prevalence rate.**

Ratio Division of one quantity by another.

This very general term can, for example, express the relative size of two populations (e.g. ratio of elderly subjects/young subjects, sex ratio of males/females). Unlike a rate or proportion, a ratio can refer to two measurements carried out on different populations: incidence rate ratio (comparing two incidence rates), risk ratio (comparing two risks), odds ratio (comparing two odds), etc.

See also: **odds ratio, sex ratio.**

Rationale Reasoning justifying the implementation of a study or the performance of an analysis.

Reaction *See* **effect.**

Recall bias Information bias in which subjects are likely to recall the event or exposure of interest with a different degree of accuracy according to the comparison group to which they belong.

Example: In a case-control study investigating the association between a given drug and a congenital malformation, mothers of malformed children (the cases) will tend, during an interview, to have a better recollection of drugs taken at the beginning of their pregnancy than mothers of normal children (the controls). This can falsely increase the value of the odds ratio estimated in this study.

Similarly, in a cohort study, subjects treated with a given drug could have better recollection of an adverse event than non-exposed subjects.

Selection bias only occurs when the recollection deviates systematically in a given direction. If the interviewed subject's

memory is imprecise or poorly guided by the interviewer, but not in a systematic manner, the result is random error which leads to a loss of precision but not a recall bias.

See also: **information bias, random error.**

Re-challenge Term used in imputability to designate the voluntary or involuntary resuming of a subject's treatment with a drug (under similar conditions of dose, duration, etc.) suspected of having previously caused an adverse event in this subject.

Re-challenge is termed *positive* if the adverse event recurs, *negative* if it does not and *inconclusive* if the conditions (e.g., dosage, associated treatments, disease evolution, etc.) are not the same for the repeat of the treatment.

See also: **intrinsic imputability.**

Recommended daily dose (RDD) *See* **defined daily dose.**

Record linkage Process of assembling the information contained in two or more complementary files which are usually computerised, in such a way as to allow epidemiological analyses.

For example, the data from a disease registry can be used to select subjects having presented a given illness, and their possible prior exposure to a drug can be ascertained from prescription records. Conversely, prescription records can be used to identify persons being or having been treated with a drug, and hospital records can serve to identify persons hospitalised for a given event.

The technique of record linkage attempts to complete the data available on each of the study subjects. It is thus important that the records refer to the same source population and that there is a way to clearly identify an individual in the various sets of records.

See also: **database, registry, source population.**

Reference *See* **control, reference population, reference risk.**

Reference date *See* **time-window.**

Reference population, reference group Set of subjects not exposed to the risk factor of interest and, ideally, similar to the exposed group in all other respects.

The reference group is used to estimate a *reference risk* which will then be compared to the *absolute risk* measured in the exposed subjects. The ideal reference group is formed by random selection in the same source population from which the studied group arose. If this is not possible, which is generally the case in pharmacoepidemiology, the comparability of the two groups must be verified with respect to the main variables capable of biasing the measurement and proposed comparison.

See also: **relative risk, risk difference, selection bias.**

Reference risk Risk measured in a population, called the *reference population*, which resembles the exposed population in all respects except that its members have not been exposed to the factor under study.

This is, for example, the risk measured in the control (comparison) group in a clinical trial or cohort study. The reference risk is a fundamental parameter in pharmacoepidemiology, since it is the only one that quantifies the risk for a subject who has not been exposed to the drug of interest. The reference risk can be very different from the risk measured in the general population.

See also: **absolute risk, baseline risk, reference population, relative risk, risk difference.**

Referral bias *See* **admission bias.**

Registry, register System of regularly collecting and recording, in an organized manner, all cases of a given event (usually a disease) in a given geographical area (e.g. town, region, country, etc.). The creation of a registry presupposes the implementation of means to guarantee completeness of data

collection.

In pharmacoepidemiology, registries (e.g. of congenital malformations, cancer, etc.) are valuable sources of data to measure incidence rates or conduct a case-control study, for example.

Regression analysis Statistical technique consisting of finding the mathematical model that most adequately predicts the value of a dependent variable y given the value(s) of one or more independent variable(s) $x1$, $x2$, . . . , xj.

The simplest model is the linear type, $y = bx$ or $y = a + bx$. Data analysis in epidemiology and pharmacoepidemiology often uses other models, such as the logistic or Cox models.

See also: **Cox model, logistic model, Poisson regression.**

Regression coefficient Coefficient affecting the value of an independent variable in a regression model in order to best predict the value of another, dependent, variable.

For example, if the value of the dependent variable y is predicted by that of the independent variable x according to the equation:

$y = a + bx$,

b is the regression coefficient of the variable x.

See also: **logistic model, regression analysis.**

Relative reporting ratio *See* **case/non case study.**

Relative risk Multiplicative factor, applied to the reference risk, associated with an exposure.

This is calculated by dividing the risk measured in an exposed population (the *absolute risk*) by that measured in a reference population similar in all respects except that its members were not exposed (the *reference risk*).

Example: If the risk of aplastic anaemia measured in a population treated by a drug is 12.6/100,000 and the reference risk in a non-exposed population for the same period is 7/1,000,000, the relative risk is (12.6/100,000) / (7/

1,000,000) = 18. We deduce that the use of this drug during the period under consideration multiplies the risk of aplastic anaemia by 18.

The relative risk measures the strength of the association between the exposure to a risk factor (e.g. a drug treatment) and the occurrence of an event. A value not significantly different from 1 indicates the absence of association, a value significantly greater than 1 indicates an increased risk linked to exposure, and a value significantly less than 1 indicates a protective effect due to exposure. The more the value of the relative risk differs from 1, the stronger the association.

It is always preferable to give the corresponding one- or two-sided confidence interval with the value calculated for the relative risk.

Example: In the above example, the two-sided 95% confidence interval for the calculated value of 18 is (0.4; 31.2); this means that the true value of the relative risk associated with the exposure has 95 chances in 100 of being included within these two limits.

See also: **absolute risk, causality, confidence interval, reference population, reference risk, risk ratio.**

Reporting Signalling of an adverse event by an observer (usually a health professional) to a surveillance system (e.g. with a report or case report).

See also: **case report, spontaneous reporting.**

Representativeness *See* **sample, sampling, selection bias, survey.**

Required sample size Number of subjects which should be included in an aetiologic or controlled study (e.g. clinical trial, cohort or case-control study), to guarantee at least a given probability (called '*statistical power*') of rejecting the tested null hypothesis with a test statistic, if this hypothesis is proved to be false.

The calculation of required sample size presupposes precise

specification of the tested null hypothesis and alternative hypothesis, as well as the type of test statistic envisaged and the accepted level of Type I error.

The choice of a one-sided comparison test generally results in a smaller sample size than a two-sided test.

The calculation of required sample size is most often derived from statistical tests based on the normal distribution. However, if the expected number of events in one of the groups is small (less than 15), a calculation based on the binomial distribution or the Poisson approximation is preferable.

See also: **binomial distribution, expected number, normal distribution, Poisson distribution, power, Type I error.**

Research General term applied to any activity with the goal of increasing knowledge.

Studies, trials and experiments are researches.

See also: **experimentation, study, trial.**

Research design *See* **design.**

Retrolective Seldom used adjective designating a study of any type (e.g. cohort or case-control) concerned with events which have occurred before its implementation.

The study thus uses data already recorded, and the value of the research is dependent on the quantity and quality of these data. The term '*retrolective*' is the antonym of '*prolective*' and can be considered synonymous with '*historical*'.

See also: **historical, prolective.**

Retrospective Concerning the past, or looking to the past.

This term is the antonym of '*prospective*'. The distinction between the two terms is based on the temporal directionality of the analysis and not on the period concerned (i.e. a prospective study can be based on pre-recorded data).

A prospective analysis begins with exposure to a risk factor, and investigates the possible later occurrence of a

given event. A retrospective analysis begins with the occurrence of an event and investigates the subject's possible prior exposure to a given risk factor.

See also: **case-control study, cohort study, prospective, retrospective study.**

Retrospective cohort study, trohoc study Incorrect and ambiguous term (since a cohort is by definition prospective) sometimes used to designate a historical cohort, that is, a study concerned with exposure and events that have occurred in the past. The term '*trohoc*' denotes the retrospective nature of the study by spelling 'cohort' backwards. It was first used for case-control studies.

See also: **historical, historical cohort, retrospective.**

Retrospective study Epidemiological study in which the possible cause or determinants of an event (a disease, symptom, etc.) are investigated in the past history of subjects having presented this event (*see* Appendix 1).

This type of study may be controlled (a case-control study) or non-controlled. It is therefore incorrect to consider 'retrospective study' and 'case-control study' as being synonymous terms, as some authors suggest. Similarly, it is incorrect to use the term '*retrospective study*' to designate any study based on pre-recorded data. Retrospective denotes the temporal directionality of the analysis (beginning with the observed event and investigating possible previous exposure to a given risk factor) and not the fact that the event of interest has already occurred.

See also: **case-control study, retrospective.**

Reverse causality bias, cart versus horse bias Bias consisting of falsely concluding a causal relationship of the type $A \rightarrow B$ when the relationship, if it exists, is of type $B \rightarrow A$.

Reverse causality bias is to be feared when the data do not permit precise analysis of the chronology of events, such as in

a cross-sectional study.

See also: **cross-sectional study, protopathic bias.**

Risk, hazard Probability that an event will occur at a time t when it has not occurred at time $t - 1$.

In pharmacoepidemiology, this term designates the probability that a subject (whether exposed to a drug or not) will present an event at any given time, knowing that the subject had not presented it in the preceding time interval. The risk is often estimated from the incidence rate for the period of interest.

Example: If the annual incidence rate of an event measured in a population is 2.5 per 1000, the risk for a subject during a 2–month period will be $(2.5/1000) \times (2/12) = 0.00042$ *or 4.2/10,000.*

Inferring an individual estimate (the risk) from a population estimate (the incidence rate) can lead to errors of interpretation if possible variations in the risk over time or among subjects are not taken into account.

See also: **depletion of susceptibles, hazard rate, incidence rate, inference, probability, specific rate.**

Risk–benefit ratio Consideration of therapeutic benefit when judging the acceptability of the risk associated with a medicinal or non-medicinal therapeutic strategy.

The concept can be applied either at the level of the individual (i.e. the acceptability of the strategy for a given patient) or the population, by comparing, in a group of subjects, the expected or observed results in terms of therapeutic benefit and the risk of adverse effects. In practice, the term '*ratio*' is slightly incorrect, because of the difficulty of expressing benefit and risk in coherent and comparable units.

See also: **ratio, therapeutic benefit.**

Risk difference, excess risk, attributable risk Difference between the risk measured in an exposed population (the *absolute risk*) and that measured in a non-exposed population (the *reference*

risk). This difference quantifies the increase in risk associated with exposure.

Example: If the risk of thrombocytopenia measured in a cohort exposed to a given drug is 18/10,000, while the risk measured in a non-exposed cohort is 3/10,000, the risk difference for thrombocytopenia is 15/10,000.

For many authors, the terms '*excess risk*' and '*attributable risk*' should not be used synonymously with '*risk difference*', since the first two imply that the increase in risk observed in the exposed subjects can only have been induced by the exposure. Such a conclusion would be permissible within the context of an experimental study comparing two populations which are identical with respect to the factors that could influence the measurement; this is not necessarily the case in an observational study.

See also: **attributable number of cases.**

Risk factor Characteristic associated with an increased probability of occurrence of an event or disease.

This characteristic can be inherent to the individual (e.g. age, sex, genetic trait, etc.), or linked to a disease, environmental factor, diet, drug use, etc. The term '*risk factor*' does not imply the existence of a causal relationship between the presence of the factor and the occurrence of the event or disease. A risk factor can only be validly identified with a controlled study (i.e. if this factor is over-represented among the subjects presenting the event compared with those not presenting it).

See also: **association, causality, interaction.**

Risk function *See* **hazard function.**

Risk ratio Relationship between two risks, generally estimated in different populations.

The ratio of the risk estimated in the exposed (*absolute risk*) to the risk measured in the non-exposed (*reference risk*) represents the particular case of the *relative risk*.

See also: **absolute risk, ratio, reference risk, relative risk.**

Rule of three Rule which states that in order to guarantee, with 95% confidence, the occurrence of at least one event whose probability of occurrence is p the sample size n must be equal to at least $\frac{3}{p}$ (i.e. $np \geq 3$).

This rule is derived from the Poisson law, which shows that the probability of observing at least one case of an event is 0.95 when the expected number, np, is 2.996, that is, approximately 3; the same calculation shows that this probability is 0.90 for $np = 2.30$ and 0.80 for $np = 1.61$.

Example: If we think that the probability of a neutropenia occurring over the course of a treatment is about 1/1000, the observation of a sample of 3000 patients will guarantee a 95% probability of observing at least one case.

Similarly, if no cases of an event have been observed in a sample of size n, the probability of occurrence of this event lies between 0 and $3/n$, 95 times out of 100.

Example: During the observation of 5200 subjects, no cases of liver injury were observed. It can be concluded, if surveillance has been correctly carried out, that the risk of liver injury associated with this drug has 95 chances in 100 of being between 0 and 3/5200 or 0 and 5.8/10,000.

See also: **expected number, Poisson distribution.**

S

Sample Sub-group of a population selected for the purposes of a study.

If the results of a measurement are to be extrapolated to a larger population, this sub-group should be as representative as possible of this population, at least with respect to the principal characteristics that can influence the parameter to be measured. Ideally, a sample should be formed by drawing subjects at random from the entire source population. Since this is rarely feasible in pharmacoepidemiology, the possibility of selection bias should be kept in mind.

See also: **random, sampling, selection bias, source population, target population.**

Sample distortion bias *See* **selection bias.**

Sampling Operations leading to the selection of a sample of subjects from a source population.

The validity of statistical inference requires that sample subjects be representative of the population to which the results of the measurement(s) are to be extrapolated. From this point of view, there are three major sampling methods:

- *Random sampling*, which consists of sampling at random from the source population. If the sample size is sufficiently large, the influence of sampling variation becomes negligible, which guarantees the representativeness of the study sample and eliminates the possibility of selection bias.
- *Systematic* (or *pseudo-random*) *sampling*, in which the

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

selection of subjects is not random, but follows a system based on a criterion deemed to be independent of the studied phenomenon.

For example, it is possible to select one subject in five during a day of consultation of all neurological services of a given country.

- *Non-random sampling*, in which the selection of subjects obeys neither a random nor a systematic process (e.g. selection of patients who agree to answer a questionnaire). While this method is simpler to put into practice, it rarely ensures the representativeness of the sample and is therefore susceptible to selection bias.

See also: **random, random error, sample, selection bias, source population.**

Sampling fraction Ratio of the size of a sample (usually selected for a survey) to that of the population from which the sample was drawn, and to which the results of the measurement will generally be extrapolated.

The sampling fraction is most often expressed as a percentage or a rate (*sampling rate*).

For example, a survey of the use of psychotropic medications consisted of interviewing 2040 adult subjects living in France. On the basis of a total population of 40 million adults, the sampling fraction is: (2040 / 40,000,000) × 100% = 0.005%.

The sampling fraction is a fundamental parameter for judging the statistical stability of inferences made from surveys; notably, it is used in the calculation of a confidence interval around an estimate.

See also: **sampling, source population, target population.**

SCT See simplified clinical trial.

Seeding study Pharmacoepidemiological study presented as

being conducted for scientific purposes, when the principal (or even the only) objective for its sponsor is an increase in the number of prescriptions or the market share of the drug concerned.

The term *seeding* applies principally to studies of drugs already on (or about to enter) the market, which therefore enter into the framework of Phase IV studies (hence the pejorative character sometimes—wrongly—attributed to Phase IV studies).

In seeding or implantation studies, the recruitment, with honoraria, of a certain number of clinician-researchers is supposed to induce a somewhat lasting habit of preferential prescription.

Except in a few obvious cases, it is not always easy to recognize whether a study has a promotional objective or whether it is a scientific study that may lead to increased market share through its positive outcomes.

See also: **Phase IV.**

Selection bias, sample distortion bias Distortion of the estimate of the association between a risk factor (e.g. use of a drug) and the occurrence of an event, resulting from measurements made in a sample which is not representative of the population to which the results are to be extrapolated.

A selection bias occurs when the event of interest is over- or under-represented in the sample as the result of a systematic error in selection. This is the case, for example, in a cohort study if the protocol leads to preferential selection of subjects (from the source population) that have a greater or lesser risk of presenting the event. In a case-control study, selection bias causes the subjects (cases or controls) to be non-representative of the source population with respect to exposure to the factor of study.

Admission bias, *diagnostic bias*, *notoriety bias* and *survival*

bias are some of the many forms of selection bias.

See also: **admission bias, diagnostic bias, notoriety bias, survival bias.**

Self-medication Use of one or more medication(s) without the prior advice of a doctor or pharmacist, with an intent to treat.

This definition excludes *abuse* or *misuse* with the goal of voluntary intoxication or that related to *drug addiction.*

See also: **abuse.**

Sensitivity For a diagnostic test, the proportion of *true positives* (i.e. diseased patients for whom the test is positive).

This is calculated by dividing the number of *true positives* (a) by the number of times ($a + c$) the test has been administered to patients suffering from the disease:

	Diseased	Non-diseased
Positive test	a	b
Negative test	c	d
	$a + c$	$b + d$

A test whose sensitivity is 1 (100%) is positive every time the disease is present, that is, gives no *false negatives.* By analogy, the sensitivity of a method of detection of an adverse effect is the proportion of cases of the effect identified by this method in a population.

Example: If 12 cases of liver injury associated with a drug were spontaneously reported when 87 cases actually occurred, the sensitivity of the surveillance system is 12/87 or 0.14.

See also: **false negative, true positive.**

Sensitivity testing, sensitivity analysis Statistical method that uses a model to study the effect of changing the value(s) of one or more variable(s) on the final results.

Such testing is frequently used in decision-making strategies, for example when calculating the attributable number of cases or the number of subjects to include in a study, under different

hypotheses of the value of the risk in the exposed persons.

See also: **attributable number of cases.**

Serious According to the regulatory definition, adjective usually applied to an event or adverse effect that is fatal, or life-threatening, that causes hospitalisation or prolongation of hospitalisation, permanent or temporary functional disability or incapacity, cancer or congenital anomaly.

This term thus appraises the consequences of the event or effect, whereas *severe* quantifies its intensity.

Definition of *serious* can be found in ICH E2A (*see* References) and in EU Directive 65/65/EEC (*see* References), proposing the following consequences of an event to be deemed serious:

- Death.
- Threat to the patient's life at the moment of occurrence of the event.
- Necessity for hospital admission or prolongation of hospitalisation.
- Significant and lasting disability (meaning the impossibility of carrying out normal tasks of daily living).
- Congenital abnormality or perinatal crisis.
- Obvious risk to the patient, requiring treatment or intervention in order to prevent any of the five complications noted above.

See also: **severity.**

Severity The intensity of an adverse event without, unlike the *seriousness*, pre-supposing its consequences in terms of mortality or morbidity.

An adverse event or symptom is described as:

- *Mild* if it does not affect the daily activity of the patient.
- *Moderate* if this activity is disturbed.
- *Severe* if this activity is made impossible.

See also: **serious.**

Sex ratio, gender ratio Ratio of the number of men to the number of women in a sample or population.

Example: If a group of 1250 subjects comprises of 527 men and 723 women, the sex ratio in this population is 527 / 723 = 0.73.

See also: **ratio.**

Side effect Unwanted manifestation related to a pharmacological property of a drug.

This may be a particular expression of its main pharmacological effect (used therapeutically) or a subsidiary property. A side effect may or may not have harmful consequences.

See also: **expected adverse drug reaction.**

Signal Event exceeding a pre-determined frequency or intensity level, recognized as worthy of attention during surveillance.

In practice, this term is used when the value of a parameter (e.g. number of cases of an event, incidence rate, etc.) differs from expected or normal values. After validation (e.g. *signal strengthening*), a signal is an *alert* which should lead to decision-making or the implementation of an appropriate study.

In pharmacovigilance, '*signal*' designates reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously. Usually more than a single report is required to generate a signal, depending upon the seriousness of the event and the quality of the information [WHO Collaborating Centre for Drug Statistics Methodology (*see* References), September 1991].

See also: **alert, background noise.**

Sign error *See* **Type III error.**

Simplified Clinical Trial (SCT) Experimental design conserving the random allocation of compared treatments, but having less rigorous criteria for subject selection, surveillance procedures and control of recorded data, in order to simplify the implementation of the trial, reduce its cost, eventually include a greater number of subjects and approximate the usual conditions of care.

Simplified clinical trials cannot replace the clinical trials required to obtain new drug approval, and are generally undertaken after this has been obtained.

Certain types of simplified trials are similar to a *pragmatic trial*. In all cases, they are prospective studies based on the observation of one or more cohort(s).

See also: **cohort, pragmatic trial, prospective.**

Simpson's paradox Particular and extreme form of confounding in which the direction of an association changes according to whether or not a confounder is taken into account.

Example: In a case-control study of 200 cases and 200 controls, 23 cases and 39 controls were exposed to a drug. The odds ratio or OR is: (23 × 161) / (39 × 177) = 0.54. This apparently protective effect of exposure, however, does not exist in the sub-group of women (150 cases and 50 controls): OR = (8 × 48) / (2 × 142) = 1.35, or in that of men (50 cases and 150 controls): OR = (15 × 113) / (37 × 35) = 1.31. This is explained by the fact that, in this study, gender is a confounding variable, strongly linked to both the disease (75% of cases were women) and the exposure (exposure rate was 30% for men and 5.3% for women).

Simpson's paradox (which is not really a paradox at all but only the exaggerated and expected expression of major uncontrolled confounding) is also to be suspected in a meta-analysis of epidemiological studies which have not accounted for the same confounding variables.

See also: **confounding.**

Single-blind *See* **blinding.**

Source population Population from which the study subjects are selected.

The source population can be defined by geographical area (e.g. city, region, country, etc.), in which case it corresponds to the general population, or by a sociodemographic characteristic (e.g. the presence of a disease, the fact of treatment with a given

drug, etc.).

When the results of an estimate made in a group of study subjects are extrapolated to the entire population, the source population corresponds to the target population.

See also: **general population, selection bias, survey, target population.**

SPC *See* **Summary of Product Characteristics.**

Specific rate Incidence or prevalence rate calculated in a given sub-group of a population.

In *stratification* and *standardization* methods, for example, the specific rate is calculated for each of the sub-groups defined according to the value of a given variable, or the presence or absence of a given criterion.

See also: **rate, standardization, stratification.**

Specificity For a diagnostic test, the proportion of *true negatives* (i.e. non-diseased patients for whom the test is negative).

This is calculated by dividing the number of *true negatives* (d) by the number of times ($b + d$) that the test has been administered to subjects not suffering from the disease.

	Diseased	Non-diseased
Positive test	a	b
Negative test	c	d
	$a + c$	$b + d$

A test whose specificity is 1 (100%) is always negative when the disease is absent, that is, gives no *false positives*. By analogy, the specificity of a surveillance method is its capacity to detect only the events corresponding to the case definition, without false positives.

See also: **false positive, true negative.**

Sponsor Individual or organization supporting a research project (study or trial) and interested in its results.

See also: **investigator.**

Spontaneous reporting Method of passive surveillance based on the collection, usually on a national level, of cases of an adverse effect associated with marketed drugs.

The surveyed population theoretically includes all of the patients treated by the drug in the geographic region considered.

In spontaneous reporting, the observer of an adverse effect is expected to report it to a clearly identified and, if possible, unique surveillance system. This assumes that the observer identifies the adverse effect, is aware of the existence of the surveillance system and is convinced of the need to report adverse effects that he or she observes.

The method thus rests on the motivation of the observer (hence the term '*spontaneous*') and is therefore subject to numerous biases, notably selection bias.

Although the reporting of adverse effects is compulsory in many countries (e.g. France), this method rarely ensures exhaustive collection of all the adverse events that have occurred. Moreover, it gives no direct information on the size or characteristics of the population treated with the drug, or the treatment conditions (dose, duration, associated drug use, etc.), which makes it difficult to calculate an incidence rate or to detect risk factors. Spontaneous reporting is, however, quite irreplaceable as a means to detect previously unknown adverse effects (i.e. as an alert system).

In certain situations the surveillance method can be based on less random information gathering, for example, by sending reporting forms to prescribers, making regular calls or visits to these prescribers or setting up a computerised network. Such a system is called *facilitated* or *prompted reporting*.

The term *spontaneous reporting* should, however, be retained to denote any prospective population surveillance method which does not ensure exhaustive collection of cases and which gives no information on the size or characteristics of the exposed

population.

See also: **under-reporting.**

Standard incidence ratio *See* **standardisation.**

Standardisation Procedure by which two or more rates estimated in different populations can be made comparable with respect to a variable distributed differently in these populations.

For example, standardisation by age allows minimisation of the influence of a different distribution of ages when comparing mortality rates measured in two populations.

Direct standardisation, or the *standard population method*, consists of calculating the rate that would be observed in the studied population if this population had the same distribution of a variable as in a reference population. This is done by dividing the studied population and the reference population into sub-groups according to the variable of interest (e.g. age-groups). The expected number of events is then calculated for each sub-group of the reference population. This expected number is the product of the size of each sub-group and the corresponding specific rate in the studied population. The *standardised rate* for the variable of interest is obtained by dividing the sum of the expected number of events by the total size of the reference population.

Indirect standardisation, or the *standard rate method*, consists of calculating an expected number of events for each of the sub-groups of the studied population. This expected number is the product of the size of the sub-group and the corresponding rate in the reference population. The ratio between observed and expected number of events is called the *standardised incidence ratio* (SIR) or *standardised mortality ratio* (SMR).

See also: **adjustment, rate, specific rate.**

Standard mortality ratio *See* **standardisation.**

Standard population method *See* **standardisation.**

Standard rate method *See* **standardisation.**

Statistical power *See* **power.**

Statistical relationship *See* **association.**

Stratification Approach consisting of dividing a sample or population into several sub-groups according to the distribution of a given criterion (e.g. age, socio-economic status, number of associated drugs, etc.).

The calculation of *specific rates* for each of these sub-groups, defined according to the presence of a factor or the value of a variable, can be used to assess the influence of this factor or the value of this variable on the measurement.

Example: Stratification according to age can reveal the fact that an excess mortality measured in a treated population is actually due to an over-representation of older subjects in this population (if a comparison with the equivalent age-groups in a reference population does not show a significant difference).

Stratification also makes it possible to detect the possible modifying effect of a given variable.

Example: Stratification according to degree of renal insufficiency shows that the risk of ototoxicity associated with an aminoglycoside is higher in patients with severe renal insufficiency than in those with moderate renal insufficiency or normal renal function. The degree of renal impairment acts as a modifier of the ototoxic effect of this drug.

See also: **interaction, specific rate.**

Study General term designating different types of evaluation (e.g. animal study, clinical study, epidemiological study).

In pharmacoepidemiology, this term is more and more often used to denote observational evaluations rather than experimental designs of the clinical trial type.

Summary of Product Characteristics (SPC) The Summary of Product Characteristics forms an intrinsic and integral part of the marketing authorisation. It contains information on the following items: the name of the product, its qualitative and

quantitative composition, the pharmacological properties of the active ingredient(s), therapeutic indications, dosage and method of administration, known adverse effects, possible drug interactions, possible contraindications, and specific precautions for use and storage.

Surrogate endpoint Judgement criterion which predicts the evolution of another more robust criterion, and used as a substitute when the latter proves difficult or inconvenient, notably for reasons of sample size, duration of study or ethics.

This may be a biological criterion (e.g. blood glucose level, CD_4 count), a test (e.g. blood arterial pressure, electromyography), an imaging method (e.g. X-ray, scan or MRI results, visualization of a tumour), etc.

For example, hormone replacement therapy in menopausal women seeks to reduce the risk of fracture in this population, notably due to osteoporosis. To judge the efficacy of a new treatment using the robust criterion of a lessening of the risk of compression fractures would require a follow-up of more than 10 years. It is permissible, at least at the beginning, to judge this efficacy by the comparative evolution of the bone mineral density, which allows for a considerably shorter study.

A surrogate endpoint is only acceptable if:

- Its evaluation is reliable, unambiguous and reproducible.
- There is a clear and well-established relationship between the surrogate endpoint and the relevant clinical criterion.
- It has been demonstrated that its evolution is correlated, in a reciprocal manner, with that of the relevant clinical criterion (strong predictive value); this implies that a favourable or unfavourable evolution of the surrogate endpoint predicts an evolution in the same direction (and if possible of the same amplitude) of the relevant clinical criterion, and this correlation has been demonstrated in the population under study (ICH E3; *see* References).

Surrogate endpoints are also used in pharmacovigilance studies (e.g. electromyogram for peripheral neuropathy, specific activity of transaminases for hepatitis, etc.).

See also: **efficacious, intermediate endpoint, positive predictive value.**

Survey Selection of a given number of individuals from a population (the source population) in order to study their characteristics, state of health or well-being, behaviour or opinion about a given problem (outside the field of epidemiology, such a strategy is known as a *poll*).

Surveying is an inferential procedure in that measurements taken from the sample are usually extrapolated to the entire source population from which the subjects were gathered. This requires that the study sample is representative of the source population regarding the main variables likely to bias the measurement or analysis.

There are numerous methods for achieving representativeness: random selection of subjects from the entire source population (simple *random sampling*), selection of a certain number of group of individuals (families or collective groups) from the population (*cluster sampling*) or surveying in sub-groups chosen from the population according to a given parameter such as age (*stratified sampling*).

In all cases, the relation of the sample size to the source population (*survey fraction* or *rate*) determines the precision of the estimate and the validity of the inference.

See also: **inference, randomisation, sampling, sampling fraction, source population, target population.**

Survey fraction *See* **sampling fraction, survey.**

Survival bias Selection bias in which the fact of being exposed to the factor of interest modifies the duration of survival, and thus a subject's probability of being included in a study. As a result, exposure can be over- or under-represented among survivors,

which risks biasing estimation.

Example: If a drug causes cardiac arrhythmia with an immediately fatal outcome, the exposed cases (treated with this drug) will have a lower probability of living long enough to be interviewed than those who were not exposed. This will likely result in the under-estimation of exposure among cases and thus the risk associated with the drug. The consequences would be the reverse if the cardiac arrhythmia induced by this drug had a lower early mortality than that induced by other causes.

See also: **depletion of susceptibles, selection bias.**

Survival function, survival distribution Mathematical function used to predict the proportion of the initially unaffected subjects of a population who, at a given moment t, will still be alive or will not have presented the event in question.

The survival function is defined by:

$$S_{(t)} = 1 - F_{(t)}.$$

Switch Term implying change and denoting the fact that a subject (called a *switcher*) is successively exposed to two or more risk factors during a follow-up, without this change being imposed by the study protocol.

A switch describes the fact that a patient changes, for example, from one therapeutic strategy to another with or without medical advice.

Example: Treatment substitutions are frequently observed in studies of non-steroidal anti-inflammatory drugs; these switches must be taken into account when evaluating the risk associated with a given drug, because of the reasons for the switch and because of a possible carry-over effect.

See also: **depletion of susceptibles, exposed.**

Synergism, synergy Type of interaction in which the effect of one factor on the studied variable is increased by the presence of another factor.

In pharmacology, we generally make the following distinctions:

- *Additive effect*, in which the resulting effect is equal to the sum of the effects of each of the drugs considered separately (also called additive effect).
- *Partial additive effect*, in which the resulting effect is greater than the effect of the drug considered in isolation, but less than the sum of the effects of each of the drugs administered.
- *Synergism*, in which the resulting effect is greater than the algebraic sum of the effects of each the drugs considered in isolation.

In the strictest sense, '*additive effect*' refers to two drugs acting by the same mechanism, while for *synergism* they act at different sites.

In epidemiology, '*synergism*' or '*synergy*' usually indicates a multiplicative model in which the resulting effect is greater than the sum of the effects of each of the factors considered separately.

See also: **additive, interaction, multiplicative.**

Systematic error *See* **bias.**

T

Target population Term with two different meanings, used to designate:

- A population addressed by a public health intervention, particularly one which is likely to receive a given treatment. This target population does not necessarily correspond to the population actually treated, called the *joint population*.
- The population to which the results of a study or estimates made in a sample can be legitimately extrapolated.

See also: **inference, joint population.**

Test hypothesis *See* **null hypothesis.**

Test of equivalence *See* **equivalence study.**

Therapeutic benefit Expected or measured improvement in the state of health or well-being of a subject or population, related to the implementation of a therapeutic strategy, whether or not this involves a medical treatment.

This improvement can be immediate or delayed, transient or permanent, but must be relevant in terms of individual benefit or public health, and assessed according to sufficiently robust criteria (e.g. decreased mortality or morbidity, improved quality of life, etc.).

See also: **therapeutic effect.**

Therapeutic drug monitoring The use of plasma drug concentrations to adjust the dosage of drug in individual patients, based on a range of target values, or on mathematical models derived from individual patients' pharmacokinetic parameters, or from population pharmacokinetic studies.

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Therapeutic effect Measurable improvement in the state of health or well-being of a subject associated with the use of a drug and apparently explainable by one or more of its pharmacological properties.

This improvement can be immediate or delayed, transient or permanent.

See also: **pharmacodynamic effect, therapeutic benefit.**

Therapeutic index Ratio of the dose of a drug considered toxic to the dose considered therapeutically efficacious for a given indication.

Theoretically, the greater the ratio, the easier the drug is to use.

For drugs with a small therapeutic index, the toxic dose is very close to the dose shown to be efficacious in a majority of subjects; dosage should therefore be adjusted with great precision.

See also: **toxic effect.**

Therapeutic trial Controlled trial whose goal is to assess the value (e.g. in terms of efficacy, tolerance) of a therapeutic strategy, particularly one involving a medication. In the latter case, this term is essentially synonymous with a Phase III clinical trial.

See also: **Phase III clinical trial.**

Time to onset *See* **onset delay.**

Time-window Interval of time considered relevant in the design or analysis of an epidemiological study.

In case-control studies, it is essential to define *a priori* the period during which possible exposure to the risk factor of interest will be investigated in the previous history of both cases and controls.

Example: If it is thought that liver injury cannot occur before the end of the first week nor after 2 months of continuous drug treatment, the only subjects considered to be exposed to the

drug will be those whose treatment began more than 7 days but less than 61 days before the occurrence of the first symptoms of liver injury.

The time-window used to characterize an exposure is always defined with respect to an *index* or *reference date*, which can be, for example, the date of occurrence of the event or the first symptoms for the cases. An equivalent date must be defined for the controls.

Similarly, in a cohort study, the time-window defines the period after the beginning of the exposure during which the occurrence of the event of interest will be investigated.

Example: A x-ray contrast media product is suspected of inducing hepatocarcinoma. Based on data available in the literature, this event may occur between 5 and 15 years after the injection of the drug. This interval forms an adequate time-window for comparing the incidence of this cancer in exposed and non-exposed populations.

The *a priori* choice of a time-window is tricky if the pathogenetic mechanism or the hazard function of the event of interest is poorly documented. An inappropriate time-window can strongly bias the estimate of risk.

See also: **depletion of susceptibles, exposure, hazard function, incidence rate, index date.**

Toxic effect, toxic reaction Harmful manifestation resulting from the administration of a toxic, supra-therapeutic dose of a drug. This effect can be the result of a large dose over a short or extremely short period (*intoxication* or *acute toxicity*) or repeated doses resulting in an excessive cumulative dose (*cumulative* or *chronic toxicity*).

Examples: nephrotoxicity of analgesics, toxicity of aminoglycosides for the inner ear or the kidney, neurotoxicity of vinca alkaloids.

The term more generally applies to manifestations resulting from too great a concentration of an active ingredient or one of its metabolites in a target tissue or organ. The therapeutic safety window of a given drug depends on the relationship between toxic and efficacious doses; this ratio is sometimes termed a '*therapeutic index*'.

See also: **therapeutic index.**

Traceability Capability of following the history of a package or batch of medication from its manufacture until its storage, dispensation or administration to a subject.

Ensuring the traceability of a product requires implementation of means by which all packages and their present or past users can be found at any time. Traceability has become a legal requirement for blood products and drugs derived from human blood.

See also: **hemovigilance.**

Treated Term applied to a subject using a drug under conditions (e.g. of indication, dosage, duration of treatment, etc.) which allow the expectation of a therapeutic benefit.

These conditions do not necessarily correspond to those characterizing the exposure, which place the subject at risk of presenting a given adverse effect. A patient can thus either be treated and exposed, treated and not exposed, or exposed and not treated (if the dose and/or duration of the treatment are not sufficient to expect a therapeutic benefit).

See also: **exposed, therapeutic benefit.**

Treatment-cohort Cohort of subjects identified according to a common treatment which they are receiving or have received.

See also: **cohort.**

Treatment-days, days of treatment *See* **incidence density, person-time.**

Treatment-months, months of treatment *See* **incidence density, person-time.**

Trial Experimental research project having a pre-defined strategy which is imposed by a protocol.

See also: **research.**

Triple-blind *See* **blinding.**

Trohoc study *See* **retrospective cohort study.**

True negative Subject free from a disease that a test is supposed to detect, and for whom this test is negative.

See also: **negative predictive value, specificity.**

True positive Subject with a disease that a test is supposed to detect, and for whom this test is positive.

See also: **positive predictive value, sensitivity.**

Two-sided test, two-tailed test Statistical test in which the rejection area of the null hypothesis lies on either side of this hypothesis.

Example: In order to compare the risks p_1 and p_2 associated with two drugs, the null hypothesis H_0: $p_1 = p_2$ is tested against the alternative hypothesis H_1: $p_1 \neq p_2$. If H_0 is rejected it can be concluded, at a fixed risk of Type I error, that the two risks are different. However, the non-rejection of H_0 does not permit the conclusion that the risks are equal.

The choice of a two-sided test is conservative in the sense that it requires inclusion of more subjects than a one-sided test, in order to have a given probability (*statistical power*) of detecting a given difference between the variables being compared. This test is nevertheless preferable whenever we are interested in exploring both sides of the rejection zone, that is, either $p_1 < p_2$ or $p_1 > p_2$, in the above example.

See also: **equivalence study, one-sided test, power, Type I error, Type II error.**

Two-tailed test *See* **two-sided test.**

Type A reaction *See* **expected adverse drug reaction.**

Type B reaction *See* **unexpected adverse drug reaction.**

Type I error, alpha error Probability of wrongly rejecting the null hypothesis in favour of a pre-stated alternative hypothesis.

When comparing two therapeutic strategies, for example, Type I error is the maximum risk we will accept of falsely finding a difference which does not in fact exist. The complement of the Type I error $(1 - \alpha)$ represents the *confidence level* of the chosen conclusion.

See also: **alternative hypothesis, confidence interval, null hypothesis.**

Type II error, beta error Probability of not rejecting the null hypothesis, although it is false, in favour of a pre-stated alternative hypothesis.

This is, for example, the risk of not detecting a given difference which in fact exists. The complement of this probability $(1 - \beta)$ represents the power of the statistical test or study.

See also: **alternative hypothesis, null hypothesis, power, required sample size.**

Type III error, gamma risk In a comparison of two values or strategies using a two-sided statistical test, the risk of concluding that one of them is superior to the other, when it is in fact inferior (*sign error*).

Example: To compare the efficacy of two treatments A and B, we test the null hypothesis that these two treatments are equivalent ($H_0 : A = B$). Rejection of this hypothesis leads to the conclusion that a difference exists between these two treatments and to the adoption of the alternative hypothesis (H_1: $A \neq B$). This difference could be $A < B$ or $A > B$. A sign error would consist of wrongly concluding the correctness of one of these possibilities. Type III error represents the maximum accepted probability of making such an error.

Type III error is often associated with *pragmatic trials*,

for which the possibility of non-conclusion is not considered by adopting the strategy which is shown to be superior on the basis of certain criteria.

See also: **pragmatic trial, Type I error, Type II error.**

U

Under-reporting, under-reporting coefficient Failure to report to a surveillance system a proportion of the cases of an adverse effect that have occurred during a given period in a given area.

This may be caused by numerous factors: absence of motivation or time on the part of the observer, absence of diagnosis of the event or non-attribution of the event to a drug.

The effectiveness of spontaneous reporting can be quantified by the proportion of reported cases (the number of cases of adverse effects reported divided by the total number of cases occurring in the population during the same period). This proportion is by definition equal to 1 (100%) at the most, and can be extremely low in certain cases, for example, when the adverse effect is well-known and not serious.

The inverse of this proportion is the *under-reporting coefficient*, the factor by which the number of reported cases should be multiplied in order to estimate the number of cases which have actually occurred.

Example: Twelve cases of liver injury were reported in the second year of the availability of a drug on the market. It is later revealed that 72 cases actually occurred during this period. The proportion of reported cases is 12 / 72 = 0.17 (17%) and the under-reporting coefficient is 72 / 12 = 6.

Unfortunately, although these concepts are useful in theoretical calculations and modelling, in practice it is rare

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

to know the precise number of cases actually occurring, in order to be able to calculate the proportion of reported cases or the under-reporting coefficient. It is only possible to make estimates of varying validity.

See also: **capture-recapture, spontaneous reporting, under-reporting coefficient.**

Unexpected adverse drug reaction, unexpected adverse effect Harmful and undesirable manifestation attributed to a drug but not apparently connected with one of its known pharmacological properties.

It is generally acknowledged that an unexpected adverse effect reveals one or more risk factor(s) inherent to certain patients (e.g. prior immunoallergic sensitivity, metabolic anomaly, deficit of a given amino acid, a specific receptor population, etc.). The incidence of the effect can be low or very low, according to the distribution of these factors in the exposed population.

The prior ignorance of these risk factors justifies the use of the term '*unexpected*' even if the mechanism can subsequently be established. An unexpected adverse effect is sometimes called a 'Type B reaction'.

By extension and somewhat abusively, current international definitions tend to denote an *unlabelled effect* as 'unexpected', that is, an effect not clearly mentioned in the most recent version of the summary of product characteristics, or in the investigator's brochure in a clinical trial situation.

See also: **Summary of Product Characteristics.**

Unlabelled *See* **labelled.**

Untoward effect *See* **adverse drug reaction.**

V

Variable Object of a measurement, which may take any value within previously defined limits (*continuous variable*) or several states according to the individuals or circumstances (*qualitative variable*).

If the value of a variable is not modified by the presence of a factor or the value taken by another variable, we speak of an *independent variable*; in the contrary case, the variable is said to be *dependent*.

See also: **continuous variable, qualitative variable, regression analysis.**

Volunteer Subject willingly participating in biomedical research after having freely consented, with full knowledge of the goals, constraints and possible risks inherent in this research.

It should be emphasized that a person cannot legally be made to participate in research without this consent having first been obtained.

According to the International Covenant on Civil and Political Rights (*see* References): 'No one shall be subjected to torture or to cruel, inhuman of degrading treatment or punishment. In particular, no one shall be subjected without his free consent to medical or scientific experimentation'.

A subject is called a *healthy* or *diseased volunteer* according to whether or not he or she presents the disease that the drug of interest is supposed to treat. For ethical and methodological reasons (in the latter case, to avoid having

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

W

Warning *See* **alert.**

Washout More or less complete elimination of the active principle of a drug at the conclusion of a period without treatment.

A sufficiently long washout period is indispensable for bringing the study subjects back to their baseline status, that is, to assume that the active principle is no longer able to interfere with the proposed measurement. Such free intervals can be planned before inclusion of previously treated subjects, or, in the case of *crossover studies*, in between treatment sequences being compared.

See also: **carry-over effect, crossover trial.**

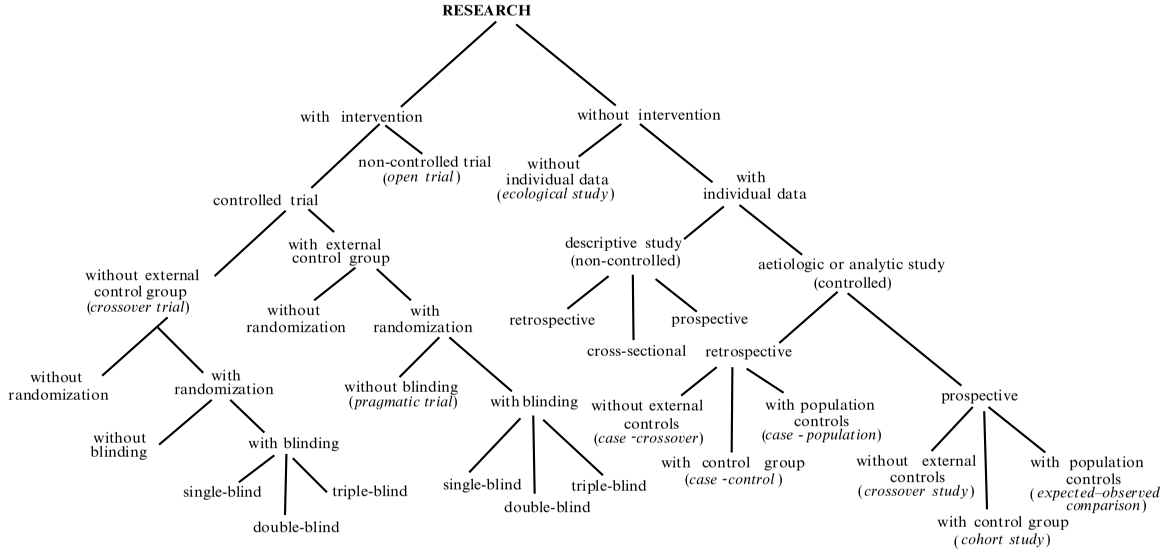
Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Appendix 1
Different Types of Studies in Pharmacoepidemiology

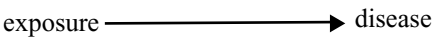
Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

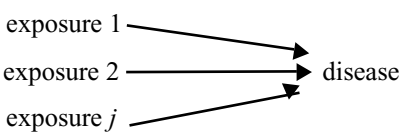

RESEARCH
with intervention
without intervention
non-controlled trial
(open trial)
without
individual data
(ecological study)
with
individual data
controlled trial
with external
control group
descriptive study
(non-controlled)
aetiologic or analytic study
(controlled)
without external
control group
(crossover trial)
without
randomization
with
randomization
retrospective
prospective
cross-sectional
retrospective
without
randomization
with
randomization
without blinding
(pragmatic trial)
with blinding
without external
controls
(case -crossover)
with population
controls
(case - population)
prospective
without
blinding
with blinding
with control group
(case -control)
without external
controls
(crossover study)
with population
controls
(expected–observed
comparison)
single-blind
triple-blind
double-blind
single-blind
double-blind
triple-blind
with control group
(cohort study)

Appendix 2
Different types of relationships between exposure and disease

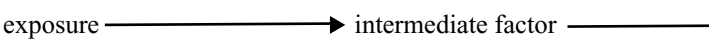
Direct causality (single factor):



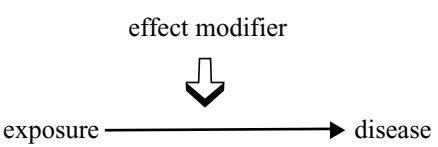
Multifactorial causality:



Indirect causality:



Effect modification (interaction):

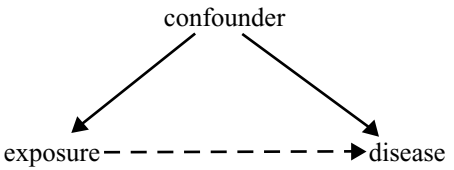


The obtained effect is not the same in the presence and absence of the effect modifier, or according to the values it takes.

Dictionary of Pharmacoepidemiology
Author: Bernard Bégaud

ISBNs: 0-471-80361-8 (Hardback); 0-470-84254-7 (Electronic)

Confounding:



The confounder, linked to both exposure and disease, explains—partly or wholly—the association found between exposure and disease