



Institute for Quality and Efficiency in Health Care

Glossary

for the General Methods 3.0¹

Version 1.0 of 27.05.2008

Contact:

Institute for Quality and Efficiency in Health Care (IQWiG)

Dillenburger Str. 27

51105 Cologne

Germany

Tel: +49 (0)221 / 35685-0

Fax: +49 (0)221 / 35685-1

E-Mail: methoden@iqwig.de

¹ This translation is based on the German document “Glossar zu den Allgemeine Methoden 3.0” (Version 1.0 of 27.05.2008).

General comment

This document supplements the document “General Methods” (Version 3.0) published by IQWiG and contains selected statistical, methodological, and epidemiological terms. This glossary primarily explains terms related to IQWiG’s work, and is targeted towards non-professional readers.

To ease comprehension, we have often knowingly chosen wording that, from the point of view of an expert, may not be entirely precise. If precise definitions are required, they should be sought in the relevant literature.

The glossary is also available on the Institute’s website and will be extended when necessary.

References

- 1 Deutsches Netzwerk Evidenzbasierte Medizin (DNEbM) [*German Network for Evidence-based Medicine*]: <http://www.ebm-netzwerk.de/grundlagen/glossar>
- 2 Institut für evidenzbasierte Medizin (DIEM) [*German Institute for Evidence-based Medicine*]: http://www.di-em.de/data/EG_Glossar_200510.pdf
- 3 Glossar der Nationalen Versorgungsleitlinien [*Glossary of the German Clinical Practice Guidelines*]: <http://www.versorgungsleitlinien.de/glossar/glossar/index?key=A>
- 4 Bandolier: <http://www.jr2.ox.ac.uk/bandolier/glossary.html>
- 5 Centre for Evidence-Based Medicine: <http://www.cebm.net/?o=1011>
- 6 American Academy of Family Physicians: <http://www.aafp.org/online/en/home/publications/journals/afp/afpebm/glossary.html>
- 7 Clinical Data Interchange Standards Consortium: <http://www.cdisc.org/glossary/CDISCGlossaryV5.pdf>
- 8 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): <http://www.ich.org/cache/compo/276-254-1.html>
- 9 Cochrane Collaboration: <http://www.cochrane.org/resources/glossary.htm>
- 10 Fletcher RH, Fletcher SW. *Klinische Epidemiologie: Grundlagen und Anwendung*. [*English title: Clinical Epidemiology: The Essentials*]. German version adapted and published by Haerting J, Rink C, and Stang A. Second (completely revised) edition. Hans Huber, Bern, 2007.

Adverse effect

A harmful event where there is at least a justified suspicion of a causal relationship with the use of an intervention.

Adverse event

A harmful event that occurs during or after the use of an intervention (e.g. a drug) without assessment as to whether the event is a causal consequence of the intervention

AGREE instrument

The [AGREE instrument](#) (Appraisal of Guidelines Research and Evaluation) includes a checklist for the quality appraisal of clinical practice guidelines (CPGs). AGREE is a tool for developers and users of CPGs for the appraisal of their methodological quality.

See also [DELBI](#) (*the German instrument*)

Allocation concealment

Allocation concealment is a collective term for measures to ensure that, before the start of a randomised controlled trial, study participants really are randomly allocated to the comparison groups. If patients or researchers know in advance or can predict which participant is allocated to which group next, this could deter certain patients from taking part in the study. This would then prevent the random composition of groups, and would increase the risk of selection bias. The question as to whether allocation was actually concealed is an important criterion in the quality assessment of a randomised controlled trial [6].

Bias

The term “bias” refers to a disposition to produce a result that systematically deviates from the true value.

The aim of scientific studies is to estimate the true difference between two (diagnostic or therapeutic) interventions and hence to exclude the influence of other factors. Bias occurs if this is not achieved and known or unknown factors exist that systematically increase, decrease or even reverse a difference. As a result, the difference measured is not only determined by the different interventions but also by other factors. Bias can be so extreme that it may appear to indicate a benefit of an intervention, even though the intervention actually causes more harm. Study results can be biased by a range of confounding factors (see *Bias, types of*). Study results do not have sufficient certainty without appropriate protection against bias, which is the key quality characteristic of clinical trials and without which there is no certainty of results.

Studies should be planned, conducted, and analysed in such a way that bias is minimised [2,7,10].

Bias, types of

1. Attrition bias = systematic difference caused by study discontinuations. Participants who discontinue a study often do so because of side effects, dissatisfaction, or poor outcomes. If these participants are excluded from the analysis, this could result in overestimation of the effectiveness of a medical intervention. Countermeasure: conduct of an intention-to-treat analysis (an analysis that also includes participants who discontinue a study).
2. Detection bias / information bias = systematic difference caused by the fact that different techniques are applied to determine outcomes in comparison groups. For example, to determine whether a tumour is still present after chemotherapy, different techniques may be applied (e.g. computer tomography, ultrasound, or clinical examination), which may lead to different results. Countermeasure: the same investigation techniques should be used in all study groups. The best way to achieve this in practice is the blinded assessment of outcomes.
3. Performance bias = a systematic difference between groups of patients, e.g. because one group is offered additional treatment that is not being investigated in a study. Information on concomitant treatment should therefore always be provided in order to enable the evaluation of any potential differences between comparison groups. Countermeasure: blinding of patients and treating staff to prevent differences between comparison groups that are caused by concomitant interventions.
4. Publication bias = a systematic difference caused by the fact that studies showing a negative effect (or no statistically significant effect) between the intervention and control group are less frequently published (or published later) than studies with positive and significant effects. Systematic reviews or meta-analyses based only on published studies may overestimate the true effect of an intervention. Countermeasure: search for and inclusion of previously unpublished studies.
5. Selection bias = a systematic difference caused by the unequal composition of comparison groups (e.g. if one group is older or sicker than the other). Counter measure: Random allocation of participants to groups and adequate concealment of allocation.
6. Lead time bias = a systematic difference in the assessment of screening tests purely caused by the earlier diagnosis of a disease in screened patients compared with patients diagnosed after clinical symptoms occurred. The period between the time between diagnosis and health deterioration/death appears to be longer in screened patients, even if the screening test has no effect on the prolongation of life. Countermeasure: conduct of controlled studies in which all participants are assessed from a defined time point, and not from the time point of diagnosis.
7. Length bias = a systematic difference in the assessment of screening tests caused by the tendency of screening to detect diseases that have a slowly progressing and less aggressive course. A benefit of screening is implied if diseases not covered by the screening programme are not considered in the assessment. Countermeasure: conduct of controlled studies in which all participants are assessed from a defined time point, and not from the time point of diagnosis.

Blinding (double, single)

Blinding is a measure taken in a clinical trial to conceal until the end of the study from patients, but also from physicians, nurses, and/or researchers which patients received which medical intervention. The aim of blinding is to minimise bias, which can be caused by the fact that the assessment of the effects of a treatment can be affected by knowledge of the allocation to treatment.

Patients or physicians often show prejudice towards a treatment; this can lead to an overestimation of the effects of one of the treatment alternatives. It is also possible that if physicians have knowledge of the allocation to treatment groups, they may consider one group of patients to be disadvantaged and apply additional measures that could also bias the results of the study.

Blinding can be maintained in therapy studies by administering a type of sham or placebo intervention that appears to be identical to the test intervention (e.g. tablets that look identical). Blinding can also be achieved if the staff who analyse the study results are not informed about which results belong to which patients.

In single-blinded studies, only the patients are not informed about their allocation to treatment groups. In double-blinded studies, this is concealed from patients, treating staff, and outcome assessors. However, the terminology used is not consistent, so that in a blinded study it is better to clearly describe exactly who is blinded [5].

Case report

A report published in a medical journal on an individual patient with a specific characteristic.

Case series

A report published in a medical journal on a series of individual patients with a specific characteristic.

Case-control study

Case-control studies are usually based on people with a specific disease (cases); people without that disease are chosen as controls. Cases and controls are then questioned or their medical history is analysed in order to identify differences in past exposure to factors that may act as risk factors for a disease [9].

Certainty of results

Certainty of results is a characteristic of an individual study or a systematic review/meta-analysis. It refers to the certainty that a result found in a study (or several studies) is close to the true result. Certainty of results arises from the assessment of the bias potential of a study and the size of statistical uncertainty.

Clinical practice guideline (CPG)

Clinical practice guidelines (CPGs) are (ideally) systematically developed, scientifically founded and practice-orientated decision-making aids that present appropriate approaches to specific health problems. They provide orientation by means of decision and action paths. In justified cases, guideline users may (or even must) deviate from these paths [2].

Cluster-randomised trial

Trials in which clusters of individuals (e.g. doctors' practices or hospitals), rather than individuals themselves, are randomised to different groups. This type of study design is selected if it is difficult to treat different patients in a very different way in a doctor's practice or hospital. Specific statistical methods need to be applied to analyse cluster-randomised trials.

Cochrane Collaboration

The [Cochrane Collaboration \(CC\)](#) is an international non-profit organisation that aims to make up-to-date information and evidence on health care topics readily available worldwide in order to support decision-making in health care. This is mainly done by the production, updating, and dissemination of systematic reviews of healthcare interventions. The Cochrane Collaboration was founded in 1993 and named after the British epidemiologist Sir Archibald Leman Cochrane.

Cohort study

A cohort is a group of people followed over a defined period of time, in order, for example, to determine the incidence of a specific disease. Cohort studies may be conducted in a prospective or retrospective manner [1,9].

Confidence interval (confidence range, confidence limit)

The confidence interval is the range in which the "true" value (e.g. the effect of an intervention) can be expected with a certain probability. A confidence interval of 95% is usually applied; this means that the confidence limits will include the "true" value with a probability of 95% [1].

Confounder/confounding

A confounder is a factor that is associated with both an intervention (or exposure) and the outcome investigated in a study. For example, in a clinical trial, if participants in the group using Drug A are younger than participants in the Drug B group, it will be difficult to decide whether benefits shown in the first group are due to treatment or to younger age (which is then the confounder).

Countermeasure: randomisation is used to minimise these imbalances. Known confounders recorded in trials can be considered by applying suitable statistical methods (adjusted analyses) [9].

Consensus techniques

Consensus techniques are informal and formal methods that are applied to achieve consensus within groups that initially had different opinions. The most important formal consensus techniques include the Delphi technique and the nominal group process [3].

CONSORT statement

The CONSORT (Consolidated Standards of Reporting Trials) statement describes which standard information on the results of randomised controlled trials should be included in publications. CONSORT comprises a checklist together with a flow diagram presenting the progress of all participants throughout the trial. The aim is to ensure that the publication of a trial includes the information relevant for the assessment of the certainty of results.

Similar guidelines have been published for systematic reviews and meta-analyses (QUOROM), observational studies (MOOSE), and diagnostic studies (STARD) [4].

Cross-over trial

In this type of trial, participants first receive one treatment and then, upon completion of the first study phase, are switched to the alternative treatment. For example, in the first study phase, participants in the first study arm initially receive Drug A and participants in the second study arm receive Drug B. After a defined treatment period, patients switch treatments and the second study phase commences: the first study arm now receives Drug B and the second study arm receives Drug A. In cross-over trials, the sequence of treatments (not the treatment itself) is randomly assigned to patients.

Under certain conditions, the effectiveness of the treatments can be determined at the end of study, both by intra- and inter-group comparisons. In cross-over trials, treatment phases can be switched more than once and more than two treatments can be compared with each other [1,4].

Cross-sectional study

In a cross-sectional study, a population is investigated at a certain point in time, in order, for example, to enable statements on the frequency (prevalence) of a particular disease.

Cross-sectional survey

See [Cross-sectional study](#)

DELBI

DELBI (Deutsches-Leitlinien-Bewertungsinstrument, www.delbi.de) is a commented checklist for assessing the methodological quality of clinical practice guidelines. DELBI is the German adaptation of the international AGREE instrument, and largely follows the AGREE structure. DELBI is published by the German Association of the Scientific Medical Professional Societies and the German Agency for Quality Assurance in Medicine.

Delphi technique

The Delphi technique is a method in which a selected panel of experts is repeatedly asked to give their opinion on a topic. This technique comprises several rounds in which the experts complete a questionnaire. After each round, the answers are summarised and sent back to the participants. The aim is to achieve a consensus within the group. Delphi techniques are used as formal consensus procedures in the development of clinical practice guidelines [4].

Disease management programme (DMP)

Disease management programmes (DMPs) are structured programmes for chronically ill patients. The main aim of DMPs is to improve the quality of health care (including treatment across health care sectors) in chronically ill patients.

Dropout

A participant in a clinical trial who fails to continue until the planned end of the trial.

Effect measure

A measure that describes the size of an effect of an intervention. For example, effects of treatments that are designed to prevent certain events (e.g. heart attacks) can be quantified by reporting the risk difference or relative risk (RR).

Equivalence hypothesis

A definition formulated in the planning of studies that specifies under which conditions the results of two medical interventions can be seen as equivalent.

Equivalence range

Predefined range of values in which the results of different medical interventions are seen as equivalent.

Equivalence trial

Clinical trials are mostly designed to show the superiority of one medical intervention over another. Equivalence trials investigate whether the response to two or more medical interventions differs by an amount that is clinically irrelevant. If the observed difference (including the statistical uncertainty) lies within a predefined range of values (equivalence range), the interventions can be seen as equivalent [9].

See also [non-inferiority trial](#)

Error probability, alpha, beta

Alpha is the specified maximum probability of observing a difference in a study by chance, a difference that does not actually exist in reality (= type 1 error; see also p value).

Beta is the probability of not detecting an effect that actually exists because of a sample size that is too small (= type 2 error, see also power) [2].

Evidence level, evidence hierarchy, evidence grade

Scales for a graded classification of the certainty of results of the available evidence. Different scales and definitions (which are not standardised) are used on an international level. In general, studies with a high susceptibility to bias have a lower evidence level than studies with a low risk of bias. For example, high-quality randomised controlled trials have a higher evidence level than observational studies or case series [2].

Exposure

A term used in observational studies to describe the factor whose effects are to be investigated. For example, in studies investigating the health effects of vitamin products, the exposure recorded is the intake of vitamins.

Extraction sheet

A predefined form used for documenting study characteristics and results.

Federal Joint Committee

The [Federal Joint Committee](#) (Gemeinsamer Bundesausschuss) is the supreme decision-making body of the conjoint self-administration of physicians, dentists, psychotherapists, hospitals, and health care funds in Germany. For over 70 million insured members, it specifies the benefits catalogue of the statutory health insurance (SHI) funds by means of directives, and thus determines which health care services are reimbursed by the SHI funds. In addition, this Committee decides on quality assurance measures for inpatient and outpatient health care sectors.

Final report, report

Term for IQWiG scientific assessments that are prepared on the basis of a commission awarded by the German Federal Joint Committee or Federal Ministry of Health. Reports are produced in a defined process and are freely accessible on the IQWiG website www.iqwig.de. Reports fulfil the preconditions to serve as a basis for decisions on directives by the Federal Joint Committee.

Fixed effects model

In meta-analysis: a statistical model that calculates the effect estimate and its uncertainty. The model is based on the assumption that all differences between results of the studies investigated are caused by chance. An alternative model is the random effects model [9].

Focus group

A group usually comprising 8 to 12 individuals who are asked to comment on predefined questions in a moderated discussion. Because of the small sample size, the results of these rounds of questions can never be representative. However, basic arguments and underlying reasoning and motives may be inferred. Focus groups are used, for example, in the preparation of information leaflets and questionnaires [3].

Follow-up

An observation period in a study during which the occurrence of events in participants is documented.

Forest plot

The graphical representation of (a) the results of individual studies included in a meta-analysis and (b) the combined meta-analysis result of these studies [9].

Funnel plot

A graphical representation in meta-analyses for the investigation of publication bias. If certain patterns are found, this may indicate the existence of unpublished data [9].

General commission

The Federal Joint Committee awarded a general commission to IQWiG in December 2004 in order to strengthen the Institute's scientific independence. This commission enables IQWiG to select and work on topics on its own initiative. Scientific reports produced within the framework of the general commission are referred to as working papers.

Gold standard

This term refers to a method, procedure, or measurement that is widely accepted as being the most accurate diagnostic or therapeutic tool available, and which should serve as a benchmark for new developments [9].

Good clinical practice (GCP)

A written guideline that specifies standards for the design, conduct, monitoring, analysis, and publication of clinical trials. These standards provide assurance that the data and results of a trial are accurate and credible, and that the rights of participants are protected [8].

Health literacy

“Health literacy represents the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health.” (http://whqlibdoc.who.int/hq/1998/WHO_HPR_HEP_98.1.pdf)

Health literacy is defined “by a person’s competence to make decisions in his or her daily life that have a positive effect on health – at home, at work, in the health care system and in society in general. Health literacy promotes the freedom of making one’s own decisions and taking one’s own actions with regard to health issues, and improves skills to find, understand, and implement health information.”

<http://www.gesundheitskompetenz.ch/request.php?site=definitionen&siteID=112&lang=de&futurepatient=74f815c8c9e06aaef5d6b07adae76ca4>

Health technology assessment (HTA)

Health technology assessment (HTA) is a comprehensive and systematic assessment of new or existing health care technologies (e.g. drugs, medicinal product, procedures, organisation systems), primarily regarding their medical, social, and financial effects. The main aim of HTA reports is to provide information for decision-making in the health care system. The German HTA Agency (DAHTA) is located in the [German Institute for Medical Documentation and Information](#) (DIMDI) [2].

Hearing, debate, comments

For specific interim steps in the production of an IQWiG report, interested persons or institutions have the possibility to submit written comments (written hearing).

If aspects presented in the comments remain unclear, an additional oral debate may be held. The hearing procedure takes place after publication of the preliminary versions of report plans and preliminary reports.

Heterogeneity / Homogeneity

Heterogeneity describes the variability of results or the size of differences between individual studies in systematic reviews that goes beyond statistical uncertainty. The causes of heterogeneity may include differences in study design or in the selection of participants. If heterogeneity exceeds a certain level, it is not usually meaningful to pool studies in a meta-analysis.

Heterogeneity between studies can be documented with statistical methods (e.g. the I^2 measure). It can be assessed by using appropriate test methods whether the deviations between studies are so large that

occurrence by chance is improbable. This then indicates that undetected differences exist between various studies that have a relevant influence on results [4].

Hypothesis

A statement or theory whose correctness or non-correctness is tested by means of a study or an experiment.

Hypothesis, one-sided

A specification before the start of a study that “Intervention A” will differ from “Intervention B” in a certain direction (e.g.: “A is superior to B”).

Hypothesis, two-sided

A specification before the start of a study that “Intervention A” will differ from “Intervention B” in any direction.

I² measure

A measure used in meta-analyses to assess the extent of heterogeneity of study results. It describes what proportion of the variability of the results of different studies is caused by heterogeneity and not by random variations [9].

Incidence

The term “incidence” describes the number of new cases of a disease in a defined population over a particular period of time.

Intention-to-treat principle (intention-to-treat analysis, ITT)

The intention to treat (ITT) principle is a strategy for analysing results of controlled trials. Even if participants drop out during a trial or switch therapies, in the final analysis they should still be included in the group to which they were originally allocated, as was intended at the start of the study. This is to minimise the risk of allowing a group to have an unfair advantage. For example, if participants drop out because of a change of therapy, adverse events, or death, this could be associated with the study intervention. The ITT principle improves the reliability of study results. If a difference is shown in an ITT analysis, this increases the certainty that this difference was actually caused by the therapy investigated. The alternative to an ITT analysis is a per-protocol analysis [6].

Interaction

A situation in which the effect of one factor on an outcome is strengthened or weakened by a second factor (e.g. if a treatment has a stronger effect on an outcome in men than in women) [9].

Intervention

In medicine: a collective term for measures that can or should change the course of a disease. This includes therapies, but also preventive measures or diagnostic tests that may lead to changes in behaviour.

Intervention trial

A trial in which the allocation of participants to different interventions is planned. This type of study is distinguished from an observational study, where it is not planned in advance how interventions are allocated, but where participants are followed without any external intervention [2].

Likelihood ratio

The likelihood ratio (LR) describes the accuracy of a diagnostic test. The LR describes the ratio of the probability that a positive (or negative) test result will occur in individuals with the target disorder to the probability of the same test result occurring in individuals free of the disorder. The LR enables a statement on how strongly the test result changes the probability of having or not having a disease [1].

Literature, grey

The term “grey literature” refers to materials that are not published in journals or databases easily accessible through a computerised search. Grey literature includes documents such as research abstracts presented at conferences [9].

Mean

The mean is the sum of all values divided by the number of values.

Measurement error

In an experiment, a measurement error is the deviation of a measurement from the true result. Depending on the cause of the error, one distinguishes between systematic and random measurement errors. Random errors are caused by variations in circumstances and by inaccuracies, and cause a result to sometimes deviate in one and sometimes in the other direction. Systematic errors are (often unrecognised) characteristics of a measurement method that consistently lead to the deviation of a result in one direction.

Median

The median is the middle value in a data set in which the values are ranked in order.

Meta-analysis

In a systematic review, a meta-analysis is a statistical technique used to summarise quantitatively the results of several studies on the same question to an overall result. This increases the evidential value (certainty of results) compared with an individual study [1].

Meta-regression

In a systematic review, meta-regression is a statistical technique used to investigate the relationship between characteristics of a study or its participants (e.g. concealment of allocation; participants' baseline characteristics) and study results [9].

Morbidity

Rate of non-fatal disease events.

Mortality

Rate of fatal disease events based on the total population.

Nominal group process (NGP)

The nominal group process (NGP) is a technique that is applied to achieve a consensus. Its main elements are:

- Presentation of preprepared texts/items;
- Comment by each group member on a specific aspect (e.g. recommendation of a guideline, explanatory text etc.);
- Collection of all the comments by the moderator;
- Summarisation of similar comments;
- Voting on items of discussion/priority setting;
- Discussion of all comments, and (if necessary) revision of the draft;
- Subsequently, another discussion of the draft, and (if necessary) conduct of a new discussion round.

Non-inferiority hypothesis

see [Non-inferiority trial](#)

Non-inferiority trial

Clinical trials are usually designed to demonstrate the superiority of one medical intervention over another. In contrast, a non-inferiority trial is designed to demonstrate that a medical test intervention is inferior to an alternative intervention by not more than a clinically irrelevant difference, or that the test intervention is even superior to the alternative intervention. For this purpose, a non-inferiority limit must be specified in the planning of the study; any intervention above this is assessed as being at least equivalent.

See also [Equivalence trial](#)

Null hypothesis

The precondition for performing a statistical significance test in a study is that two statements (null hypothesis, alternative hypothesis) that exclude each other are formulated in advance. For example, the null hypothesis states that there is no difference between study groups regarding the outcome of interest. After completion of the study, the null hypothesis is tested by means of suitable statistical

techniques. If the p-value falls short of the predefined alpha error, the null hypothesis is rejected and the alternative hypothesis accepted [1].

Number needed to treat (NNT)

The number needed to treat (NNT) is an estimate of how many people on average need to receive a specific treatment before a harmful outcome in one person is prevented. An NNT of 20 means that on average, 1 in 20 treated patients has an advantage with regard to a specific outcome (e.g. prevention of a heart attack). An NNT usually refers to a defined treatment period (e.g. 5 years), which should also be reported [9].

Observational study

In observational studies, in contrast to experimental studies, investigators do not intervene in the application of medical interventions. On the one hand, characteristics and behaviour (exposure) of participants are recorded, and on the other, relevant health care events. Observational studies often aim to record and describe the natural course of a disease, as well as to describe associations between exposure factors and specific events.

Participants with a specific characteristic usually differ from other participants with regard to this characteristic (and also other characteristics), so that the relevance of individual characteristics cannot be clearly distinguished. Observational studies are therefore susceptible to bias (e.g. confounding and selection bias) and therefore usually cannot prove causality (cause-consequence-effect) [2,9].

Odds ratio

In colloquial English, the term “odds” often refers to the odds of a bet (e.g. “The odds are 9 to 1”). As an analogue in medicine, this term refers to the ratio of persons in a group with a specific outcome to persons without this outcome. If 30 of 100 persons experience this outcome (and 70 do not) the odds ratio (OR) is “30 to 70” or 0.43.

The OR is defined as the ratio of the odds of an event occurring in a test group to the odds of it occurring in a control group. For example: the odds in the control group are the same as above (30 to 70 = 0.43), but in the test group, only 20 of 100 patients experienced the outcome. This results in an odds of “20 to 80” = 0.25 in the test group. The odds ratio is then calculated as $0.25 / 0.43 = 0.58$.

The term “odds” is not the same as “probability”. Probability is calculated from the ratio of persons in a group with a specific outcome to all persons in the group. If 30 of 100 persons experienced the outcome, the probability is “30 of 100” or 0.30. Analogously to the odds ratio, in the comparison of groups the relative risk is calculated as the ratio of the probabilities in both groups: $0.20 / 0.30 = 0.66$.

The example shows that odds ratio and relative risk convey similar information, but should not be confused. However, differences between the OR and the RR become negligible for small risks (from 1:100 onwards) [3].

Outcome, continuous

An outcome that is measured on a continuous scale. Blood pressure is an example of a continuous outcome.

Outcome, dichotomous (binary)

An event that either occurs or does not occur in patients. For example, participants of a study may either experience a heart attack or not, or they may survive or not. This term is also called “binary outcome” and is distinguished from the term “continuous outcome”.

Peer review

A peer review is a refereeing process for scientific documents where experts in a specific area review the quality and importance of each other’s work (e.g. of articles submitted for publication) [9].

Per-protocol analysis

An analysis including only patients that completed a study as planned in the study protocol. This type of analysis may lead to an overestimation of differences.

See [*Intention-to-treat analysis*](#)

Placebo

A placebo is a therapeutic intervention that cannot be distinguished from the active treatment by the type of administration or other characteristics such as appearance, colour, taste, and smell. However, the placebo is not an active substance with a specific known mechanism of action. The term “placebo” is usually used in the context of drug trials. Placebos are administered in trials so that study participants and treating staff do not know who receives which treatment (blinding). Sham interventions serve the same purpose, and can be used to achieve blinding in studies investigating operations and other procedures [10].

Placebo effect

A collective term for influences (in part caused by psychological effects) that are based on the circumstances of the administration of a treatment and not on the specific effect of the treatment.

Population

A population is a group of people. Populations may be defined by characteristics such as geographical borders, age, gender, or certain diseases. In order to answer research questions, a sample taken from a population is studied, which is preferably representative for the overall population [9].

Power

In a clinical trial, power is the probability that a trial will actually detect an existing difference as being statistically significant (e.g. between two treatment groups) [9]. Among other things, the power of a trial depends on the size of the difference and on the frequency of the event, by means of which the treatments are compared [1].

Pragmatic trial

A trial where experimental treatments are tested in real life situations. For this purpose, few restrictions are applied concerning the selection of participants or the use of concomitant treatments. This contrasts with trials that are conducted under ideal conditions in order to determine whether a therapy has the ability to achieve a benefit under such conditions [9].

Preliminary report

Interim step in the production of an IQWiG report. The preliminary report outlines the preliminary results of an IQWiG report. After the publication of a preliminary report, a hearing is conducted.

Prevalence

The proportion of persons with a particular characteristic (e.g. a disease) at a particular point in time in a particular population (see also cross-sectional study) [2].

Prevention

Prevention of diseases and their complications. Prevention aims to avoid or at least delay the onset of a disease. Depending on the time of application of preventive measures in the course of a disease, one can distinguish between primary, secondary, and tertiary prevention.

Prevention, primary

Measures that prevent or delay the onset of a disease by eliminating or reducing its causes. Primary prevention includes measures such as vaccinations, the use of condoms, hygienic measures for drinking water, or prevention of obesity. Primary prevention usually takes place as a part of daily life outside the framework of the health care system [10].

Prevention, secondary

Measures that are employed to detect early stages (before the onset of clinical symptoms) of a disease that has already started. These measures are to prevent or delay the progression of the disease. Examples include screening for cancer [10].

Prevention, tertiary

Measures that are employed after the onset of a disease to prevent or delay further deterioration and reduce the rate of complications. Examples include the use of acetyl salicylate, beta blockers, or statins after a heart attack [10].

Prospective study

In a prospective study, the event of interest to researchers (e.g. a particular disease) has not yet occurred at the start of the study. In this type of study, researchers have the possibility of precisely defining in advance the events to be assessed and the influencing variables of interest [2].

p-value

The p-value describes the probability that the effects observed in a study could have occurred by chance alone (the p-value lies between 0 and 1). The smaller the p-value, the lower the probability that the result can be explained by chance. For most research questions it has been agreed upon to regard a p-value of (or smaller than) 0.05 as “statistically significant” (this corresponds to the probability of a result caused by chance of at most 5%).

QUOROM statement

The QUOROM statement (Quality of Reporting of Meta-analyses) outlines what information publications on meta-analyses of clinical trials should include as a standard. The main elements include a checklist and a flow chart that present the handling of identified studies.

Random effects model

In meta-analysis: a statistical model that calculates the effect estimate and its uncertainty. Systematic differences between individual studies are also included in the calculation. An alternative model is the fixed effects model [9].

Randomisation

Randomisation is the process of allocating participants into the arms (e.g. an intervention and control arm) of a controlled trial without allowing allocation to be affected by any subjective influences. This is to ensure that both groups are as equally composed as possible, i.e. do not differ with regard to characteristics such as age of participants or disease severity. If a difference is shown between groups during the course of the trial, then a causal association with the test intervention can be made.

Randomisation, stratified

A type of randomisation where participants are initially allocated to subgroups on the basis of important characteristics. Subsequently, the participants in each subgroup are randomly allocated to the study groups. This process aims to ensure that factors of particular importance in a disease are actually distributed equally to the study groups.

Randomised controlled trial (RCT)

An experimental trial in which participants are allocated to intervention and control groups following a clearly defined random procedure (in particular to observe the occurrence of predefined events).

Rapid report

“Rapid report” refers to a type of document published by IQWiG. Rapid reports are primarily produced to provide information at short notice on relevant health care developments, including new technologies. It is usually required to complete this type of document in a short period of time. The production procedure for a rapid report differs from that of a full report in two main points:

1. No report plan or preliminary report is published.
2. No hearing is conducted.

Rapid reports are commissioned by the Federal Joint Committee or the Ministry of Health.

Receiver operating characteristic (ROC)

A presentation (often done graphically) of the association between sensitivity (proportion of true-positive results) and the proportion of false-positive results of a diagnostic test.

Regression analysis

A statistical technique used to describe the type and strength of the association between two or more factors (e.g. the effect of age on the prevalence of a disease) [9].

Regression model, multifactorial (multiple regression)

A statistical technique used to describe the type and strength of an association between more than two factors (e.g. the effect of age, weight, and smoking on the prevalence of a disease).

Report plan, (preliminary version)

An interim step in the production of an IQWiG report. The main components of the report are defined in advance (e.g. how the benefit of different medical interventions will be assessed). For this purpose, it is specified which group of patients is to be studied, which medical interventions are to be compared, and which treatment outcomes are to be considered in the assessment. It is also specified how studies are to be searched for, selected, and analysed.

IQWiG first publishes a preliminary report plan. After a written commenting procedure (written hearing), the preliminary version is revised if necessary, and then the report plan is published.

Responder

1. Patients who respond to a specific treatment in the expected way (e.g. patients with high blood pressure whose blood pressure decreases noticeably after taking an antihypertensive drug).
2. Persons who return a questionnaire sent to them within the framework of a survey.

Retrospective study

In a retrospective study the disease (event) has already occurred before the start of the study, and risk factors for the disease are searched for retrospectively [2].

Review, non-systematic (review, narrative)

A review that is not based on the same type of methods used in the preparation of a systematic review. Consequences include a subjective selection of a subset of studies.

Risk reduction, absolute

The absolute risk reduction (ARR) is used to quantify a treatment effect in outcomes that can only be present in two forms (e.g. an event either occurs or does not occur; also referred to as dichotomous outcomes). If a treatment has a beneficial effect, the ARR describes the difference between event rates in the control group and the treatment group. For example, if 30 of 100 participants die in the control group the risk is “30 of 100” or 0.30. If 20 of 100 participants die in the treatment group, the risk is 0.20. The ARR is $0.3 - 0.2 = 0.1$. The ARR is often reported in percent: “The ARR is 10 percent.” The reciprocal of the ARR is the number needed to treat ($1/ARR = NNT$).

Risk, absolute

Absolute risk is the probability that a particular event will occur within a particular period of time in an individual. The range is between 0 (no event will occur at all) and 1 (the event will occur in any case). For example, an absolute risk of 0.6 means that the probability of the occurrence of an event is 60%: of 100 persons, 60 will be affected.

Risk, relative

The relative risk (RR) is used to quantify a treatment effect in outcomes that can only be present in two forms (e.g. an event either occurs or does not occur; also referred to as dichotomous outcomes). The RR is based on a comparison of two absolute risks. For example, if 30 of 100 participants in the control group of a study die, the risk is “30 of 100” or 0.30. If 20 of 100 participants die in the treatment group, the risk is 0.20. The RR is the ratio between these two risks: $0.20 / 0.30 = 0.66$. The RR is often reported in percent: “The RR is 66 percent.” A RR of 1 means that there is no difference between the comparison groups. If the RR is < 1 , the risk is reduced; if the RR is > 1 , the risk is increased.

Sample

A sample is a subset of a population. Samples are investigated as the representatives of a population if an investigation of the overall population is either not possible or would involve too much effort. In order to be representative for the overall population, the sample must be sufficiently large and be selected without bias. Ideally this is achieved through random selection.

Sample size planning

An estimation before the start of a study to determine the number of participants and the study duration needed to achieve a good likelihood of actually detecting an existing effect of an intervention.

Screening

Screening is the examination of (mainly healthy) persons who have no clinical symptoms in order to detect diseases at an early stage.

Screening study

A study in which the advantages and disadvantages of a screening test are investigated.

Sensitivity

Sensitivity is the probability that a diagnostic test will correctly detect people with the disease. Sensitivity is also referred to as the true positive rate of a test.

Sensitivity analysis

A sensitivity analysis is a technique used to determine how sensitively a model calculation or a meta-analysis reacts to changes in methodology (e.g. when individual studies are excluded from the analysis) [1].

Standard deviation

The standard deviation is a measure of the spread of measured values, and is calculated as the square root of the variance.

Standard error

The standard error is a measure of the precision of estimates. For example, it is used to determine confidence intervals.

Subgroup analysis

An analysis to determine whether the effect of an intervention in a study differs between various subgroups (e.g. subgroups formed on the basis of age, gender, etc.). The results of subgroup analyses are usually not reliable unless the analyses were specified in advance when planning the study.

Superiority trial

A clinical trial with the objective of showing that one medical intervention is superior to another.

Surrogate parameter (intermediate outcome)

Surrogate outcomes are outcomes that are not of immediate relevance for patients but are associated with patient-relevant outcomes (e.g. reduction in blood pressure as a surrogate outcome for the prevention of a stroke). Surrogate outcomes are often physiological and biochemical parameters that can be measured relatively quickly and easily. Surrogate parameters are often used if patient-relevant outcomes occur rather rarely or not until after a longer period of time.

Even if a surrogate outcome is associated with a patient-relevant outcome, it does not mean that a causal relationship necessarily exists between the two. As long as a causal relationship has not been explicitly demonstrated, changes in a patient-relevant outcome cannot be inferred from changes in a surrogate outcome.

Survey

A poll, review, or study.

Survival time analysis

A technique used to analyse data that describe the time to an event (e.g. death or the next episode of disease) [9].

Systematic review

A systematic review is the scientific summary of a clearly formulated question based on defined methods and a systematic, reproducible approach. Available studies on a research question are searched for, assessed with regard to their relevance, and subjected to a critical evaluation. Relevant studies based on predefined criteria are identified, and the results of these studies are extracted and (if appropriate) summarised applying statistical methods (meta-analysis).

Two by two table

A table used for the comparative presentation of the results of two therapeutic or diagnostic procedures. The main element consists of two columns and two rows that form four cells in the table. For example, the sensitivity and specificity of a diagnostic test can be inferred from a two by two table.

Unblinding

Disclosure of who received which study intervention, at the end of (or during) a study that was originally blinded.

See also [Blinding](#)

Validity internal, external

One distinguishes between the internal and external validity of a study. In the assessment of internal validity it must be reviewed to what extent the influence of confounding factors was minimised by the study design, i.e. whether the results of a study were actually due to the treatment investigated.

External validity describes the transferability (generalisability) of study results to groups of patients or everyday health care conditions. Generalisability depends on factors such as the selection of study participants and the qualifications of the responsible physicians [2].

Value, predictive

Positive predictive value: the proportion of individuals with positive test results who do actually have the target disorder.

Negative predictive value: the proportion of individuals with negative test results who do not actually have the target disorder.

Not only do both values depend on the sensitivity and specificity of the diagnostic test, but also on the prevalence of the disease in the target population [1].

Variance

A measure of the variation of several measured values. Calculation: (a) the difference to the mean is formed for each value, (b) this difference is squared, (c) the sum of squares is then formed, and (d) this sum is then divided by the number of values minus 1.

Working papers

Scientific projects within the framework of IQWiG's general commission are referred to as "working papers". These "working papers in cases of urgent need for advice" are prepared according to the scientific standards specified in IQWiG's methods paper. No consultation with the Federal Joint Committee or Federal Ministry of Health is required concerning these topics. Moreover, there are no deadlines for publication.

In contrast to other IQWiG commissions, discussions with the professional public do not take place until after the publication of the document.

Index

Adverse effect.....	2	Exposure.....	7
Adverse event.....	2	Extraction sheet.....	7
AGREE instrument.....	2	Federal Joint Committee.....	7
Allocation concealment.....	2	Final report.....	7
Bias.....	2	Fixed effects model.....	7
Bias, types of.....	2	Focus group.....	7
Blinding (double, single).....	3	Follow-up.....	8
Case report.....	4	Forest plot.....	8
Case series.....	4	Funnel plot.....	8
Case-control study.....	4	GCP.....	<i>see</i> Good clinical practice
Certainty of results.....	4	General commission.....	8
Clinical practice guideline.....	4	Gold standard.....	8
Cluster-randomised trial.....	4	Good clinical practice.....	8
Cochrane Collaboration.....	4	Health literacy.....	8
Cohort study.....	4	Health technology assessment.....	8
Comments.....	9	Hearing.....	9
Confidence interval.....	5	Heterogeneity.....	9
Confidence limit.....	<i>See</i> Confidence interval	Homogeneity.....	9
Confidence range.....	<i>See</i> Confidence interval	HTA.....	<i>see</i> Health technology assessment
Confounder.....	5	Hypothesis.....	9
Confounding.....	5	Hypothesis, one-sided.....	9
Consensus techniques.....	5	Hypothesis, two-sided.....	9
CONSORT statement.....	5	I ² measure.....	9
CPG.....	<i>see</i> Clinical practice guideline	Incidence.....	9
Cross-over trial.....	5	Intention-to-treat analysis.....	<i>See</i> Intention-to-treat principle
Cross-sectional study.....	5	Intention-to-treat principle.....	9
Cross-sectional survey.....	<i>See</i> Cross-sectional study	Interaction.....	10
Debate.....	9	Intermediate outcome.....	<i>See</i> Surrogate parameter
DELBI.....	6	Intervention.....	10
Delphi technique.....	6	Intervention trial.....	10
Disease management programme.....	6	ITT.....	<i>See</i> Intention-to-treat principle
DMP.....	<i>see</i> Disease management programme	Likelihood ratio.....	10
Dropout.....	6	Literature, grey.....	10
Effect measure.....	6	Mean.....	10
Equivalence hypothesis.....	6	Measurement error.....	10
Equivalence range.....	6	Median.....	10
Equivalence trial.....	6	Meta-analysis.....	11
Error probability, alpha.....	7	Meta-regression.....	11
Error probability, beta.....	7	Morbidity.....	11
Evidence grade.....	<i>see</i> Evidence level	Mortality.....	11
Evidence hierarchy.....	<i>see</i> Evidence level	Multiple regression.....	16
Evidence level.....	7		

NGP	<i>See</i> Nominal group process	Receiver operating characteristic	15
NNT	<i>See</i> Number needed to treat	Regression analysis.....	15
Nominal group process	11	Regression model, multifactorial	16
Non-inferiority hypothesis .	<i>See</i> Non-inferiority trial	Report	7
Non-inferiority trial.....	11	Report plan, (preliminary version).....	16
Null hypothesis	12	Responder	16
Number needed to treat.....	12	Retrospective study.....	16
Observational study	12	Review, narrative.....	<i>See</i> Review non-systematic
Odds ratio.....	12	Review, non-systematic	16
Outcome, binary.....	13	Risk reduction, absolute.....	16
Outcome, continuous	13	Risk, absolute.....	16
Outcome, dichotomous	13	Risk, relative	17
Peer review	13	ROC.....	<i>See</i> Receiver operating characteristic
Per-protocol analysis.....	13	Sample	17
Placebo.....	13	Sample size planning	17
Placebo effect.....	13	Screening	17
Population	13	Screening study.....	17
Power	13	Sensitivity	17
Pragmatic trial.....	13	Sensitivity analysis	17
Preliminary report.....	14	Standard deviation	17
Prevalence.....	14	Standard error	17
Prevention	14	Subgroup analysis.....	17
Prevention, primary	14	Superiority trial.....	18
Prevention, secondary	14	Surrogate parameter.....	18
Prevention, tertiary.....	14	Survey.....	18
Prospective study	14	Survival analysis	18
p-value	14	Systematic review	18
QUOROM statement	14	Two by two table	18
Random effects model	15	Unblinding	18
Randomisation	15	Validity, external	18
Randomisation, stratified	15	Validity, internal.....	18
Randomised controlled trial.....	15	Value, predictive.....	19
Rapid report	15	Variance.....	19
RCT	<i>See</i> Randomised controlled trial	Working papers.....	19