Basic Principles and Methodology of Clinical Research in Dentistry

Basic Principles and Methodology of Clinical Research in Dentistry

By

Lars Bondemark and Per-Erik Isberg

Cambridge
Scholars
Publishing

Basic Principles and Methodology of Clinical Research in Dentistry

By Lars Bondemark and Per-Erik Isberg

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-0364-0519-2
ISBN (13): 978-1-0364-0519-9

CONTENTS

PREFACE/FOREWORD

We are pleased to have been given the opportunity to write a textbook on basic principles and methodology in clinical research. The book is primarily intended for researchers in dentistry but also for researchers in medicine and nursing. The book is also suitable as support in basic university courses on research methodology as well as an ideal guide for students at both graduate and undergraduate levels. Moreover, clinicians interested in a sound clinical research may benefit from the book.

Each chapter is organized in an analogous way, providing a structured and pedagogical approach for reading and understanding the principles of clinical research.

In a reader-friendly manner the book starts with descriptions of evidence-based care, different study designs, levels of evidence, scientific knowledge gaps, research needs in dentistry along with the entire research process. A central and unique chapter, based on authentic research studies, is to thoroughly review and comment on the pros and cons of the various research methods and statistical applications that the studies have used. Another important chapter describes how a systematic literature review including meta-analysis is carried out. Furthermore, the book contains special sections on research ethics and sample size calculation. Also, principles for literature reference management and reference systems are presented. In the concluding chapter, advice is given on how new research results can be presented in an informative way. The focus is on the scientific article, what applies to making an informative oral presentation, how a poster and a press release can be designed and written in the best way.

Yngsjö and Slagerstad in May 2024

Lars Bondemark and *Per-Erik Isberg*

ABOUT THE AUTHORS

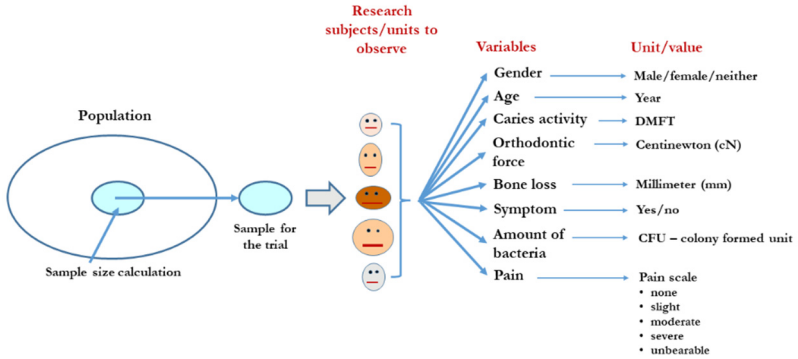
The authors have many years of experience in clinical research production along with teaching research methodology and statistics at university level.

Lars Bondemark, professor emeritus, doctor of odontology and specialist in orthodontics, has for several years been active at the Faculty of Odontology, Malmö University, Sweden, with many years of producing clinical research in dentistry. He has written several informative and international publications on research design and methodology as well as a book in Swedish "Grundläggande ortodonti" published with Gothia kompetens, Sweden, 2021, and have been one of three authors of the book "Essential orthodontics" published with Wiley Blackwell, John Wiley, and Sons, 2018.

Per-Erik Isberg is a statistician and has for many years worked at the Department of Statistics, Lund University, Sweden. He has for a long time participated as a statistician in a very large number of projects mainly in odontology, medicine, psychology, and ecology. He has also several years of experience in teaching biostatistics and medical statistics and has frequently taught courses in statistics for researchers and students in dentistry.

BASIC TERMS AND CONCEPTS

When planning research, data should be analysed and calculated. There are several terms and concepts that are necessary to know and to be aware of. Some common terms and concepts of the research process are presented in the figure below.



In the chapters of this book, we explain most terms and concepts in the running text without becoming complete in that regard. Also, with the help of the index of the book, explanations for many basic terms and concepts can be found and identified. For educational reasons, we want to give the reader another instrument or entry to gain knowledge and understanding of several basic terms and concepts which have been recorded below. Note that the terms and concepts are not listed in alphabetical order but are presented based on how they relate to each other.

Variable A characteristic that we measure for a subject/unit.

Quantitative variable A numerical or quantitative variable such as length, weight, or temperature.

Qualitative variable A non-numerical variable such as patient satisfaction, eye, or hair colour. In these examples, the variable cannot assume a numerical or quantitative value, instead they reflect a characteristic of a non-numerical nature.

Dichotomous variable A variable that can only have two values, for example absence (zero) or presence (one) of a phenomenon or a characteristic. Common examples are illness versus no illness or symptoms versus no symptoms.

Continuous variable A numerical variable that can assume any value (in a certain range), for example number of bacteria in saliva or blood sugar level.

Discrete variable A numerical variable that can only assume certain values, usually integers (whole numbers), for example number of children in a family.

Categorical variable A variable that can have two or more groups or categories usually without any order, and thus, cannot be ordered from highest to lowest. A typical categorical variable is nationality.

Nominal scale For this scale the variable values imply division into groups with no order between them. Examples are gender, diagnoses, type of disease, professions, and thus, the value of the variable must be explained with words such as man/woman, caries/periodontitis, dentist/dental hygienist/dental nurse/dental technician.

Ordinal scale In this scale the variable values can be ranked, for example high school education/university education/doctoral education, whereby the educations are arranged based on increased level of education, however, it is not possible to specify a numerical value.

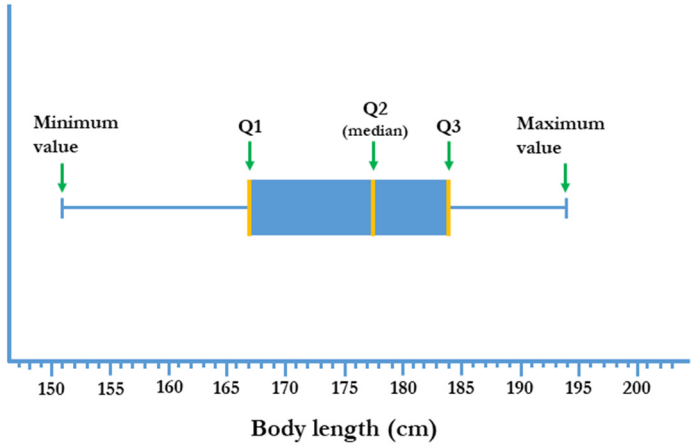
Interval scale What is measured are numerical values and there is a constant distance when we move one unit on the scale (equidistance). An example is temperature measured in degrees Celsius: 15 degrees Celsius is 5 degrees warmer than 10 degrees Celsius and 12 degrees Celsius is 3 degrees colder than 15 degrees Celsius. The scale can have negative values.

Ratio scale A scale with equidistance (equality in distance) and an absolute zero point. This makes it possible to form ratios between two values, and this ratio has a reasonable interpretation. An example is body length measured in centimetres whereby one individual may be twice as tall as another. Most numerical variables that we measure have a ratio scale, for example age, weight, height, or number of teeth.

Mean value The sum of all values/observations divided by the number of values/observations. To be able to calculate the mean, the data must be on an interval or ratio scale.

Median The median is obtained by sorting the numbers in size of magnitude and where the number of values is greater than the median is equal to the number of values less than the median. For 2, 6, 7, 11, 16 (odd number of numbers) the median is 7, while for 2, 6, 7, 9, 11, 16 (even number of numbers) the median is the average of the middle two numbers, i.e., 7 + 9 divided by 2 which becomes 8. The median, just like quartiles and percentiles, requires that the data can be ranked, i.e., ordinal, interval, or ratio scale.

Quartiles Three points that divide a dataset into four parts where 25% of the observations are smaller than the first quartile (Q1), also called the lower quartile. The second quartile (Q2) is identical to the median (Md) since two quarters correspond to the half. The third quartile (Q3), also called the upper quartile, where 75% of the observations are less than Q3. A graph called a boxplot, see below, can be constructed using the quartiles, whereby the data set is graphically illustrated.



Percentile Means hundredths and the observations or values are sorted in size of magnitude and then divides the data set into hundredths. A percentile is a value that divides the data set so that a certain percentage of the values/observations come either below or above. For example, the 20th percentile (P20) is the value that divides the observations so that 20 percent are smaller than and 80 percent are greater than the value of P20. In addition, the median is always the 50th percentile and the twenty-fifth percentile is identical to the lower quartile (Q1), see the figure above.

Mode The **v**alue that appears most often in a set of data. Among the values 2, 3, 3, 4, 4, 4, 6, 8 the mode is equal to 4. A mode can be calculated for all types of variables. A problem with the mode is that there can be several modes in a material or set of data.

Standard deviation Is a measure that is approximately equal to the average deviation from the mean in a series of values. If the standard deviation is large, the spread between the values is large. The standard deviation is calculated by first calculating the difference between all measured values (x) and the mean value (m). Square each of these differences and sum these squares together. Divide by the number of values minus one, and finally, calculate the square root of the quotient, see figure below. The standard deviation of a sample is often denoted s or SD. It is also common to use small sigma (σ) to symbolize the standard deviation in a population.

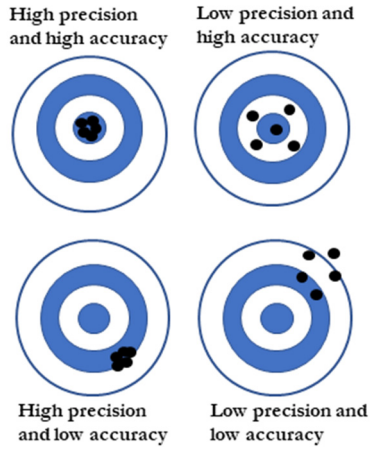
$$s = \sqrt{\frac{\sum(x-m)^2}{n-1}}$$

Range (maximum–minimum) The difference between the largest and smallest value in a numeric series of values; for 2, 6, 7, 9, 11, 16, the range becomes 2 to 16 (or 16 - 2 = 14). With a large width or a large interval, there is a risk of a misleading average value, see median above.

Confidence interval An interval that describes the likely values for a certain quantity. Common ranges are about the mean value or coefficients in some statistical model. The interval always has a certain degree of confidence, usually 95 percent. Note that a 95 percent confidence interval for the mean value does not describe the individual values, but rather where the population mean value is likely to lie.

Precision How close the measurements are in relation to each other, i.e., a type of measurement or observation error, see the figure below.

Accuracy How close a value is to the true value, i.e., compared to precision, is another type of measurement or observation error, see the figure below.



Reliability High reliability means that the result must be the same during repeated measurements and regardless of who performs the test, or if it concerns a study, the result will be the same if the study is performed again (also called repeatability).

Validity To measure what it was intended to measure or to what extent the measurement results are valid and reflect reality, for example: Does a bacteria sample really measure disease activity?

Internal validity for a study Is based on the sample of research subjects included in the study, that is, whether the results of the study are reliable or not. It is also about estimating the risk of systematic deviations from the actual relationship between treatment and effect.

External validity of a study To what extent the results of a study have a general validity (degree of generalisability), i.e., to what extent the results of the study can apply to other groups or in general clinical situations.

Efficacy Does the treatment/intervention work during a clinical trial/study? (Can it work?). The ability of a treatment/intervention to produce beneficial effects.

Effectiveness Does the treatment work in widespread use in healthcare? (Does it work?).

Efficiency Is the cost reasonable in relation to the treatment effect? (Is it worth it?).

Population A group of subjects present in a certain area at a certain time. A clinical study is usually carried out on a sample of a population, i.e., a study population.

Census The entire population is involved and examined (also called total survey).

Sample (survey) Part of the population is involved and examined.

Prevalence An epidemiological term indicating the proportion of subjects in a population (usually as a percentage) who have a certain disease or condition and usually at a certain time.

Incidence The number of new cases of a disease or condition that develop in a population during a certain time interval.

Frequency Denotes occurrence in greater or lesser numbers (degree of regularity), how often something occurs or how many subjects belong to a specific category.

Sensitivity The probability that a diagnostic test or an assessment method will give a positive response in subjects who have the disease or condition (true positive rate). A highly sensitive test means that there are few false negative results, and thereby, few cases with the disease are missed.

Specificity The probability that a diagnostic test or an assessment method will give negative answers in subjects who do not have the disease or condition (true negative rate). A test with high specificity means that there are few false positive results.

Risk factor Is used to determine the relationship between an exposure and a specific outcome, for example the relationship between frequent meals and caries. Usually, the risk factor is expressed in terms of absolute risk or relative risk. If the risk for exposed subjects is 45 percent and 15 percent for unexposed control subjects, the absolute risk difference is 30 percent units, and the relative risk is 3.0. Consequently, the risk is three times higher for those exposed. If the risk difference is 0 (zero), for example, 20 percent for both exposed and unexposed, the relative risk would be 1.0, i.e., no connection between exposure and outcome.

Outcome A specific result or effect that can be measured, and to document the impact that a given intervention or exposure has on a certain population. Examples of outcomes include decreased pain, reduced periodontitis, or improvement of oral health related quality of life.

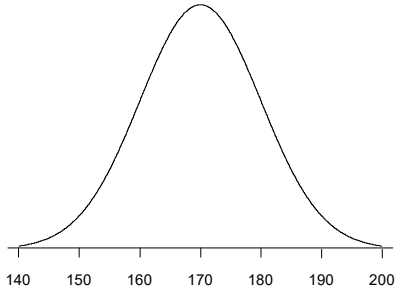
Primary outcome The outcome that the researchers consider to be the most important among the many outcomes that are to be examined in the trial. The primary outcome needs to be pre-defined at the time when the trial is designed. It is preferable to use participant/patient-centred outcomes, i.e., an outcome that matters to the participant/patient, for example quality of life or pain relief.

Secondary outcome A pre-specified and additional outcome that is considered as being relevant, but less important than the primary outcome. The second outcome can help to interpret the results of the primary outcome, for example a laboratory test may constitute a secondary outcome.

Patient-reported outcome measures (PROMs) Normally when a trial is performed, the researchers' collect data by observations, clinical registrations, or by results from different tests. However, data can be directly reported by the subject via interviews or questionnaires, and that type of data is labelled patient-reported outcome measures, i.e., PROMs. Usually, PROMs consider functional status of a subject, well-being, symptoms, health-related quality of life, treatment compliance, and satisfaction. The patient-reported outcome instruments (different questionnaires or interviews) can be used to collect the PROMs and be designed to produce generic, disease-specific, dimension-specific, or site-specific data.

Surrogate outcome A surrogate outcome is used as a substitute for a clinically useful outcome that for example measures directly how a subject feels or functions. The purpose is that changes of the surrogate outcomes should reflect what is desirable to evaluate. Common surrogate outcomes are blood pressure, blood values, number of bacteria in dental plaque or probing depth instead of the attachment level of a tooth. To map and analyse surrogate outcomes, a shorter observation period is often required compared to outcome measures with significance for health. Therefore, studies based on surrogate outcomes can be shortened, and thus, carried out faster and with lower costs. There are many examples of interventions/treatments that present beneficial effects regarding the surrogate outcomes, but parallel or simultaneously, outcomes that are important for the patient imply weak or even unhealthy effects for the patients' health.

Normal distribution Theoretical distribution for a continuous variable. A normal distribution is determined by the mean and standard deviation. Many statistical methods assume that the sample being analysed comes from a population that is normally distributed, see the figure below. In the case of large samples, this assumption can sometimes be of less importance for the validity of the analysis.



A normal distribution with a mean of 170 cm and a standard deviation of 10.

Binomial distribution Discrete distribution that describes the number of times a certain event occurs when we perform independent trials with a constant probability in each trial in which the event will occur. For example, we assume that 20 percent of the individuals in a very large population have a certain disease. Calculate the probability that exactly 30 percent of the subjects (fifteen subjects) in a random sample of 50 people have the disease. This probability can be calculated using a binomial distribution.

Chi-square distribution A statistical distribution that is often used to evaluate, for example, percentages and variances (standard deviations squared). The Chi-square distribution is based on the normal distribution.

T-distribution A statistical distribution that is often used to evaluate, for example, average values, differences between average values and various measures in regression models. The T-distribution is based on the normal distribution.

Placebo or placebo effect This is a favourable medical treatment effect that occurs despite the treatment being carried out with a placebo, that is with a drug preparation that lacks an active substance or with a treatment method that is designed to resemble an active one but has been changed to be inactive. The placebo effect, which often consists of alleviated symptoms in the event of illness or injury, can primarily be attributed to positive

expectancy effects, i.e., that the person being treated has faith that the treatment will be beneficial. Placebo is Latin and implies "I will please".

Nocebo or nocebo effect This is the opposite of the placebo effect and means that negative expectations in a patient worsen the relevant symptoms. In placebo-controlled studies, side or adverse effects may occur even in the group receiving ineffective medication; this is then attributed to the nocebo effect. Nocebo is Latin and means "I shall harm".

CHAPTER 1

EVIDENCE-BASED CARE, DIFFERENT TYPES OF STUDIES, RANKING OF STUDIES AND EVIDENCE LEVELS, SCIENTIFIC KNOWLEDGE GAPS AND NEED FOR RESEARCH

Learning objectives

- Be able to explain the concept evidence-based care
- Have knowledge of different types of studies and be able to describe their distinguishing characteristics
- Be able to describe how studies are ranked in relation to achieved evidence
- Be able to explain the concept of scientific knowledge gaps
- Have knowledge of the needs for dental care research

Evidence-based care

In general, all care should be based on “science and proven experience”, which means that science and research as well as clinical skill and experience are important for well-executed care. The care must also be easily accessible and build on respect for the patient's self-determination, integrity, and safety, as well as promote effective communication between patients and caregivers. For the standard of care to be as high as possible, the care must be evidence-based, which means a conscious systematic use of the best available scientific evidence for decisions about care and to ensure the most effective and least risky treatments for the patients (Sackett et al., 1996). The philosophy behind evidence-based care consists of three parts that must always be weighed together when making care decisions. The three parts are the clinical skill of the caregiver, evidence from research, and the patients' needs and wishes (Figure 1.1).

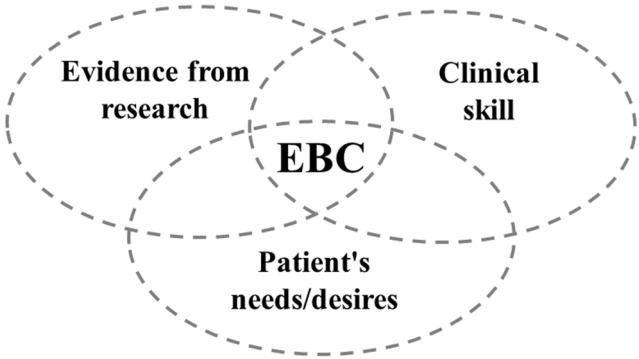


Figure 1.1. The concept of evidence-based care (EBC) consists of three components that must always be considered when individual care is to be carried out. The figure is modified from Sackett et al., 1996.

It is important to be aware that continuous evaluation of both established methods and new medical innovations is part of the concept of evidence-based care. For example, if we hope or believe that a specific treatment is useful or reliable without asking for proof, it is possible that the treatment is ineffective or even harmful. It can also be the opposite, that we reject a method as insufficient and unsafe, but after a scientific investigation has been carried out of the method, we can conclude that it is reliable and safe. Therefore, it is important to scientifically analyse all care interventions with carefully designed investigations so that the evidence or level of evidence for the care intervention can continuously be updated.

The systematic literature review is an important part of the concept evidence-based care. The review starts with a clinically relevant question such as: What caries prevention effect does fluoride varnish have on root caries? or: Does premature birth cause an increased prevalence of malocclusions? Then, searches are carried out in various databases to find relevant literature and publications that are relevant to the question. To be able to determine the quality of the studies, these are systematically analysed based on a strict set of regulations, and finally, a statement of the level of evidence that exists for the requested treatment or diagnostic method is created. In chapter 4, more and detailed information is provided on how to carry out a systematic literature review.

If the conclusions from the systematic literature review is that an insufficient level of evidence or a scientific knowledge gap exists, it is of

course recommended and urgent to carry out new studies. These studies must be well-designed, and the new research results must have the potential to change clinical practice, and thereby, help healthcare providers to deliver new high-quality evidence-based care.

Different types of studies

There are various types of clinical studies, and one type of study classification can be: experimental studies, observational studies (also called descriptive studies) and registry-based studies. Another classification is: prospective (forward-looking) or retrospective (backward-looking) studies. In addition, the studies can be in the form of a cross-sectional study (study at one point in time) or longitudinal study (study over time).

Experimental study

An experimental study can also be called an intervention study or a clinical trial. It can be carried out to study various biomedical, odontological or other health-related issues. The study implies that the research participants are actively exposed to a planned intervention such as a new treatment method, a new drug, a new type of diet/a dietary supplement or a medical technology product. The main objectives are to investigate and determine the effectiveness of the intervention including safety and side effects.

For evaluations of treatment effects, an experimental study in which subjects are randomised to treatment or control is considered to be the most solid or robust study design. The ideal is thus to be able to randomly allocate the participants to different groups, but when this cannot be done, the experimental study may be carried out without randomisation to different groups, which is then referred to as a quasi-experimental study.

Usually, the new intervention/treatment is compared to a control consisting of an established intervention/treatment, alternatively with a placebo treatment or no intervention at all. Since there are established treatments for most conditions and diseases, placebo treatment or no treatment at all can be ethically questionable for the subjects in the control group because they are deprived of an effective treatment. Therefore, if possible, comparisons against placebo should be avoided and instead compared against an established intervention/treatment. In this regard it is important that the established intervention/treatment in the control group is the best available because comparison against a suboptimal treatment may lead to a positive false or overestimated treatment effect for the new intervention.

When it comes to evaluating, for example a new drug or vaccine, the evaluation or clinical trial is divided into four phases:

Phase I is usually conducted on 20–100 voluntary healthy participants, and analyses are made of safety, immune response, and side effects. It may also test how to administer the new treatment (by mouth, infusion into a vein, or injection) and how the treatment affects the body.

Phase II is carried out on a larger participant group than in phase I – there can be up to several hundred participants in phase II. The primary objective is to investigate how effective the drug/vaccine is in treating the disease or condition. In phase II, it is also common to analyse which dose of the drug/vaccine is the most optimal and it will be this dose that will then be used for the test subjects in phase III.

Phase III is conducted on a very large participant group, up to several thousand participants, to finally determine how successful and effective the drug/vaccine is for treating the disease or condition. Efforts are made for the participant group to mimic or match as far as possible, for example with age and gender, the population for which the final drug/vaccine will be used in the future. Either the drug/vaccine is compared with placebo or existing standard treatment. After phase III, an application is made to have the drug/vaccine approved by the medicine authorities in each country. In Europe approval may be given by the European Medicines Agency (https://www.ema.europa.eu/en/human-regulatory-overview/research-development).

Phase IV is carried out with extensive studies once the drug/vaccine has come on the market. Treatment effectiveness and safety are monitored as well as mapping of usual and unusual side effects.

Observational study

An observational study can also be called a descriptive study, non-interventional study, or non-experimental study. It is a study that may include thousands of participants and does not involve any intervention or experiment. The researchers do not try to influence the course of events but observe what develops or happens to individuals or groups under certain conditions over time (longitudinal study) or at a single point in time (cross-sectional study).

Observational studies can include epidemiological investigations with the aim of identifying and evaluating risk factors as well as finding causes of different diseases and premature death. It may also be a question of mapping influencing factors such as smoking, alcohol consumption, dietary habits, exercise habits, or risky substances in different work environments.

Register-based study

In this type of study, data from various registers established by authorities or organisations are used. Many countries have several national care registers that can be used for register-based research in medicine, public health, and social sciences. Usually, registered data of medical and biological characteristics, lifestyle and social factors from large population groups are used to analyse whether there are links to various diseases or whether general health has been affected.

An advantage of register-based research is that researchers do not need to collect data from the study participants directly but can instead receive data from the registers. In this way, the results can be implemented more quickly, and the research can become relatively more cost-effective.

Prospective study

In a prospective study, the participants are evaluated or watched during the study period, and thus, treatments, diseases, conditions, or various risk factors for different diseases/conditions are analysed over time. New participants must be recruited into the study and registrations of data are carried out at the start of the study, while the study is in progress and at the end of the study. It is common for this to be done as a cohort study and an experimental study is always prospective.

Retrospective study

In a retrospective study, participants are sampled, and information is collected about their past. Usually, the analyses are carried out using medical records or other register documents which means that all the data is already available when the study starts. The study may apply to the evaluation of several types of completed treatments or to analyse how diseases or conditions have developed. The study is less resource-intensive and costly because all the data are available and there is no need to recruit new participants into the study. The disadvantages are that important data

may be missing from records/register documents. In terms of treatments, there is also a risk that only successful treatments have been fully registered, and that randomisation is not possible. The disadvantages also implies that retrospective studies receive a lower level of scientific evidence than prospective studies. A case-control study in which a group of cases is compared with a group of controls with respect to a previous exposure to risk factors is a typical example of a retrospective study.

Cross-sectional study

In this type of study, different variables for a participant or group/population are recorded on a single occasion. The study is often used to evaluate prevalence of various conditions and diseases. One disadvantage is that it can be difficult to define the causal relationship between the intervention/exposure and outcome. Another disadvantage is that chance or unintentional coincidences can influence the single measurement occasion.

Longitudinal study

In a longitudinal study, various variables are often evaluated with repeated measurements for a participant or group/population over a longer time-period. The study design makes it possible to analyse changes over time, and thus, obtain reliable information about various diseases, individual behaviour, and relationships. Compared to a cross-sectional study, the longitudinal study is safer for defining causal relationships because multiple registrations can be made over time. The role of chance also becomes smaller compared to cross-sectional studies since several registrations take place over time.

Blinded study

The word "blinding" means masking, and in a blinded study it is hidden which participant receives one or the other treatments/interventions to be evaluated. It is preferred that all participants in the study – subjects, researchers/doctors/dental staff and those evaluating the research data – are blinded. Consequently, in a blinded study, an attempt is made to reduce the risks of researchers' and participants' expectations affecting the study results (performance bias). It has been shown that unblinded studies tend to produce larger treatment effects than blinded.

In a single-blind study, only one part of the participants is unaware of which treatment has been administered. Most often it is the participant/patient that is blinded while the researchers are not.

In a double-blind study, neither participants/patients nor researchers/ doctors/dental staff know which treatments have been administered. A traditional double-blind study is when a new drug is to be tested against an inactive control preparation (placebo). One group of participants receives the new drug/tablets, and the other group receives placebo tablets, whereby it is required that both the active tablets and the placebo tablets look identical and are dosed in the same way.

In a triple-blind study, the blinding includes participants/patients, researchers/doctors/dental staff and persons who evaluate, process, and compile the data (result compilers, statisticians etcetera).

A study is called an open label study/trial when no blinding has been carried out or has been used. Usually, blinded studies have lower bias, and thereby, produce higher level of scientific evidence than open label studies.

Ranking of studies and evidence levels

Regardless of the study design and for the studies to be as reliable as possible, the researchers must prevent sources of errors and bias of the results considering measurement methods, analyses, reporting of results and interpretations of the results. It is also fundamental to recruit enough research subjects and limit the dropout rate (Bondemark, 2019). To find out how many participants are needed in a study, a sample size calculation must be performed together with a statistician. Read more in Chapter 5 about conducting a sample size calculation.

There exists a hierarchy or ranking of study types based on the potential of the study to produce reliable results and evidence. In this context, the risk of bias of the study is assessed, whereby also an assessment is made of the study's internal and external validity. An assessment of internal validity involves estimating the risk of systematic deviations from the actual relationship between treatment and effect. External validity is about to what extent the results of a study have a general validity (degree of generalisability), i.e., to what extent the results of the study can apply to other groups or in general clinical situations.

Most often, the study hierarchy has been illustrated as a pyramid. The type of study that creates the highest level of evidence is placed at the top. A randomised controlled trial (RCT) is considered to create the highest level of scientific evidence and is often referred to as the »gold standard«. However, it can be noted that the systematic literature reviews (SRs) or meta-analyses have been placed at the top of the pyramid for several good reasons. A SR or meta-analysis is considered even more reliable than an RCT because the SR compiles research from several RCTs and thereby contributes to even greater evidence than a single study, and in addition includes a critical review of the combined knowledge. Cohort and/or case-controlled studies can create a relatively high level of evidence, while case series, case reports and expert opinions, editorials or blog posts generate weak or no evidence (Figure 1.2).

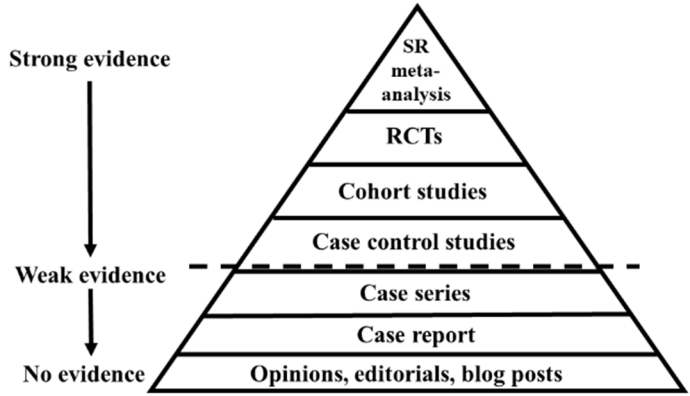


Figure 1.2. Diverse types of studies and hierarchy of evidence. It has been claimed that the strongest or highest evidence is created by a systematic literature review (SR) where research from several randomised controlled trials (RCTs) has been compiled. As a single study, an RCT ranks highest and is considered the »gold standard« in treatment research. Study variants or reports below the dashed line provide weak or no evidence (Bondemark & Ruf, 2015; Pandis, 2011).

Nevertheless, there are counter arguments to placing a systematic review or meta-analysis at the top of the pyramid, and it is also argued that it is a simplistic way to design the pyramid with horizontal lines to separate or illustrate the differences in strength of evidence between the diverse study designs. Therefore, two new modifications have been introduced into the pyramid (Murad et al., 2016). Based on the quality of the primary studies compiled, the reliability or credibility of the systematic review can vary

considerably. In other words, a systematic review or meta-analysis of well-conducted RCTs with low risk of bias cannot be equated with a systematic review or meta-analysis of RCTs or observational studies with low quality and high risk of bias. Clearly, a systematic review or meta-analysis that compiles evidence from studies with poor quality and high bias should not be at the top of the pyramid. Therefore, in the first modification, the systematic reviews or meta-analyses have been removed from the top of the pyramid and are instead used as a lens through which all types of studies are compiled and assessed for quality (Figure 1.3). In the second modification of the pyramid, the straight horizontal lines that separate the different study designs have been replaced with wavy lines because different studies can be of varying quality (Figure 1.3). Accordingly, the wavy lines go up and down to reflect the GRADE method which grades studies up or down based on the different domains of evidence quality.

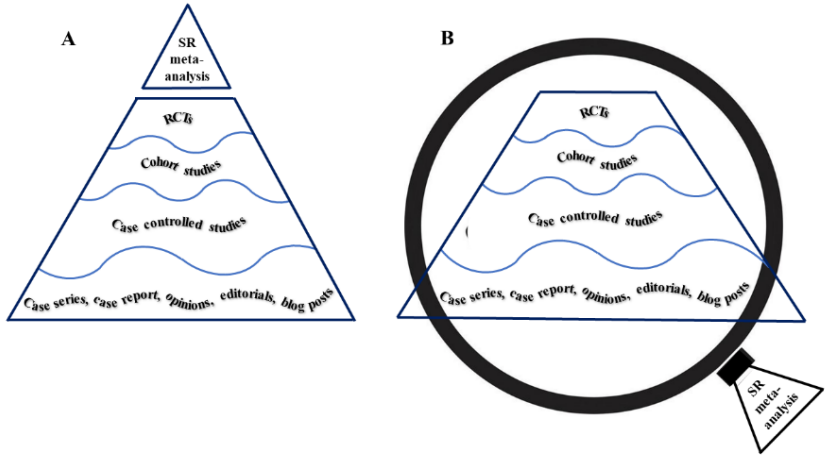


Figure 1.3. The new pyramid. In A, the top (systematic literature review or meta-analysis) of the traditional pyramid has been removed, and the lines separating the study designs have become wavy. In B, the new modified pyramid is illustrated whereby the systematic reviews/meta-analyses form a lens through which all types of studies are compiled and assessed for quality.

In general, it can be stated that prospective studies create stronger evidence than retrospective studies, since in retrospective studies there is often a lack of complete registration of data or that only successful cases are registered. Correspondingly, controlled studies always produce higher evidence than non-controlled ones, because no comparisons have been made in uncontrolled

studies. Instead, only the results of the treated subjects/patients can be reported.

Another question is whether it is enough to produce only one well-conducted study with reliable results. The short answer is almost never since a single well-conducted study will rarely provide enough evidence to definitively decide which treatment to choose. The rule is that several different studies regarding the treatment method are required, and then, the results from each study should be evaluated in relation to each other. If the results then coincide, the combined study results can support a reliable and evidence-based decision about the treatment method.

Systematic Literature Review

Overall, a very large amount of clinical research and scientific publications are produced annually in the world. For dentistry, over 30 000 new scientific publications are produced per year. In the concept of evidence-based care, it is important that new research results are established and can be transferred into daily clinical practice to help caregivers deliver evidence-based care of high quality. However, it can be difficult for the hard-working healthcare provider to find time to read and analyse each new research publication and then evaluate the quality of the research results. In this context, the systematic literature review can be a useful tool to contribute to updating the healthcare providers.

Thus, the systematic literature review is an important part of the concept of evidence-based care and starts by asking a clinically relevant question regarding the treatment that is desired to be evaluated. To collect relevant publications, a systematic literature search is then carried out in one or more different databases. MEDLINE is the largest and most dominant medical research database, indexing abstracts from over 4000 scientific journals worldwide in medicine, dentistry, health sciences, and preclinical sciences. The PubMed version of MEDLINE is free to the user and article titles and abstracts are available to all and can be easily accessed via the National Library of Medicine (https://pubmed.ncbi.nlm.nih.gov). Other databases that may be useful are Embase, Cochrane Library, CINAHL, Google Scholar, Web of Science and PsycINFO. After the relevant publications have been collected, then based on a specific protocol a systematic quality evaluation of the collected publications is carried out to finally determine the grade of evidence for the current treatment. Read more in Chapter 4 about the process and method steps for conducting a systematic literature review.