Study design I

Kate Ann Levin
Dental Health Services Research Unit, Dundee, Scotland, UK

This is the first of a series of articles that will describe the different types of study design; considerations when choosing a study design; and the advantages and disadvantages of each type of study. This first article explains the importance of choosing an appropriate design and the decisions to be made when doing so.

Evidence-Based Dentistry (2005) **6,** 78–79. doi:10.1038/sj.ebd.6400355

Why is study design important?

A study objective determines broadly who and what is to be studied. The study design — the way in which health status and risk factor data are to be measured and collected and a hypothesis tested — is rarely straightforward. Often there is more than one way to carry out a study.

A badly designed study can lead to erroneous results, or it may not answer the question presented. The design of a study also determines the methods used to analyse the data. It is therefore important to consider the design with a view to how the data will later be analysed. The design of a study can rarely be changed once the study has begun.

What are the main types of study design?

Observational vs *experimental* Choosing between an observational and experimental study design (Table 1) depends upon the purpose of the study. Observational studies are used to monitor or describe the health status of a population. They are referred to as observational because the investigators merely observe what is happening rather than attempting to intervene in any way. Health data are collected about one or more groups of subjects, from which an inference is made about a target population.

Often measurements of risk factors or exposure data are also collected. These are used to describe the relationship between exposure and the outcome measure of interest, which is often disease or some measure related to disease.

Experimental studies examine the effect of an intervention on the outcome of interest. Comparisons can be made between an intervention and no intervention or between one intervention and another. Health data are collected after and often prior to the intervention, to assess its relative effectiveness.

Prospective vs *retrospective* The prospective–retrospective dimension (Table 2) describes the way in which data are collected. Prospective studies collect data forwards in time in order to examine the aetiology of disease (observational study) or to assess the effectiveness of an intervention (experimental study). Retrospective studies collect past exposure information on participants through interview or recorded information. Observational studies may be either prospective or retrospective. Experimental studies are always prospective.

Cross-sectional vs *longitudinal* Cross-sectional studies give a snapshot of the outcome of interest in a population at any given time and are often used to look at prevalence. Prospective and retrospective studies are examples of longitudinal

Table 1. Advantages and disadvantages of observational vs experimental designs.

Study type	Advantages	Disadvantages
Observational	Data might already be available; can be used to investigate both harms and benefits	Hard to make causal inferences between exposure and outcome of interest
Experimental	Tailor-made to answer a specific research question	Dropout rates may be high if intervention is unpleasant; limited to interventions that are thought beneficial (eg, not ethical to ask people to start smoking in order to asses its impact on their health)

Table 2. Advantages and disadvantages of prospective–retrospective designs.

Study type	Advantages	Disadvantages
Prospective	Specific risk factor measurements can be included in study	Time-consuming; loss to follow-up
Retrospective	Relatively inexpensive and easy to carry out	Prone to recall bias and/or incomplete data

Table 3. Advantages and disadvantages of cross-sectional and longitudinal designs.

Study type	Advantages	Disadvantages
Cross-sectional	Can encompass a broad scale of information	Prone to confounding
Longitudinal	Powerful study of causal association between exposure and outcome	Time-consuming, expensive and often result in loss to follow-up

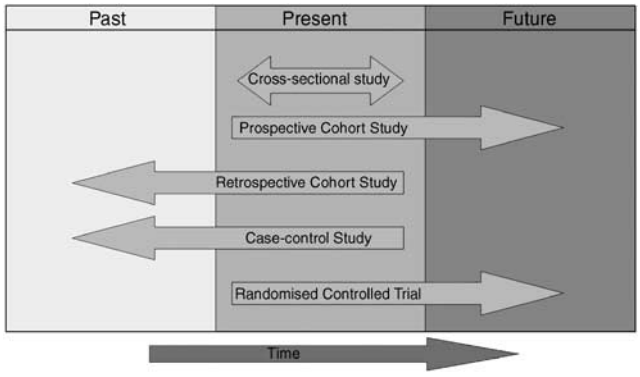

Figure 1. Main study types.

studies (Table 3). These are used to examine changes in health status over time, including incidence rates of disease outcomes. Repeated cross-sectional studies are sometimes used as a pseudo-longitudinal study, with a different group of participants at each timepoint.

The main study types (Figure 1) will be addressed in series of forthcoming publications, addressing cross-sectional, cohort, case–control, ecological and finally randomised controlled trials.

Power, reliability and causality

A number of dimensions indicate the quality of a study. The power of a test or a study is defined as the probability of rejecting the null hypothesis when it is false. The greater the power, the better the study is, therefore, at answering the research question, ''is there an association between exposure/ intervention and outcome of interest?'' when the answer is, ''Yes''. In other words, if an association does exist, the power indicates how likely it is that the study will be able to detect this. A more in-depth discussion of power is provided in the final article of the series on randomised controlled trials.

The reliability of a study is a measure of its ability to reproduce the same results if run second or subsequent times. Optimal study design is the key to both a powerful and reliable study.

It is often difficult to make causal inferences between exposures/interventions and outcomes, particularly if based solely on the results of observational studies. Experimental studies are generally better in the assessment of cause and effect because they have better protection from bias and/or confounding (see later articles in this series). There is a standard set of guidelines for causation, first set out by Sir Austin Bradford-Hill in 1965.[1]

What factors need to be considered when designing a study?

When designing a study, it is important to consider the following points.

First and foremost, what question do you wish to answer?

This should be clearly focussed and ideally written in a form that is amenable to testing (eg, a ''null hypothesis'' if this is applicable).

Availability of data: obtaining secondary-source data, eg, dental records, will save extra time and cost and may provide a more comprehensive set of results for observational studies.

Sampling methods: the aim is to ensure that the sample is representative of the target population. Inappropriate sampling methods may result in an unrepresentative sample, low response rates or incomplete information, large sample size requirements, contamination and/or bias (see below).

Data collection: the assumption here is that data can be feasibly collected. Data may be unobtainable due to the sensitivity of the study question or the rareness of disease outcome. Where data can be collected, methods include survey by telephone, mail and by face-to-face interview, as well as methods of collection of medical data by health professionals.

Cost of the design, time implications and loss to follow-up: these must be considered, especially when designing a longitudinal study.

Controls: conclusions can be drawn only about the relationship between exposure/intervention and outcome when studies include a control group (ie, they are comparative). Controls may be subjects who have no intervention in a randomised controlled trial or subjects who do not have the outcome of interest in a case–control study. The choice of controls, particularly in an observational study, may not be straightforward (see fifth article in series on case–control studies).

Ethical issues: these arise in response to both the content of the data collected and the methods used in obtaining that data. Particularly in the case of experimental studies, ethical approval of the study will be required. There are potentially considerable ethical implications, where positive interventions exist for one arm of the study and not another.

1. Hill AB. The environment and disease: association or causation? Proc R Soc Med 1965; 58: 295–300.

Recommended reading

2. Altman DG. Practical Statistics for Medical Research, London: Chapman & Hall; 1991.
3. Campbell MJ, Machin D. Medical Statistics: A Common Sense Approach. 3rd Edn. Chichester: Wiley; 1999.
4. Coggon D, Rose G, Barker DJP. Epidemiology for the Uninitiated. 4th Edn. BMJ Publishing; 1997; [accessible at bmj.bmjjournals.com/epidem/epid.html].
5. Pocock SJ. Clinical Trials: A Practical Approach, Chichester: Wiley; 1983.
6. Rochon PA, Gurwitz JH, Sykora K, *et al.* Reader's guide to critical appraisal of cohort studies. 1. Role and design. Br Med J 2005; 330:895–897.

Study design II. Issues of chance, bias, confounding and contamination

Kate Ann Levin
Dental Health Services Research Unit, Dundee, Scotland, UK

In the first article in the series I explained the importance of study design and gave an overview of the main types of design. Here, I describe the ways in which the results of a study may deviate from the truth and the measures that can be taken to help minimise this when designing a study.

Evidence-Based Dentistry (2005) **6,** 102–103. doi:10.1038/sj.ebd.6400356

Examination of the association between outcome of interest and exposure or intervention requires accurate measurements that are representative of the target population. Accurate measurements are also required if the purpose of the study is merely to monitor a population's health so that prevalence or incidence rates can be determined. Spurious associations and inaccurate estimates mainly arise due to chance, bias, confounding and/ or contamination. We must endeavour to minimise these at the design phase, as often they cannot be adjusted for later when the data are being analysed.

Findings of a study sample will often be extrapolated to a larger population. Deviations from the true population measure may be due to chance — also called variation — which is measured as random error. Chance can never be eliminated entirely, but it can be minimised through replication of measurements, and by increasing the size of the study.

Bias occurs when there is a systematic difference between study measurements and the true population values. Types of bias include:

Selection bias: systematic difference between those selected into the sample and those not selected. The sample is therefore not representative of the population.

Observer or measurement bias: systematic difference in measurement of health status or risk factor between observers.

Recall bias: differences in reporting experiences between those who have and those who do not the outcome of interest; occurs particularly in retrospective studies.

Publication bias: tendency to report studies that make strong statements about outcomes of interest.

Bias results from poor study design and cannot be corrected for at the analysis stage. Confounding occurs when a spurious association is made at the analysis stage between outcome and exposure which, in reality, results from a secondary exposure that was not included in the analysis. For example, in a study it may be found that people of one town have higher rates of oral cancer than another. Interpreting this to mean that oral cancer is dependent on area of residence is incorrect if it is known that people of the first town are, in general, heavier smokers than the populace of the second. Confounding can be minimised at the design stage by:

Randomisation: In an experimental study, subjects are assigned to control and intervention groups at random. This ensures that members of the same group are less likely to have higher than usual rates of other potentially confounding characteristics in common.

Matching: By matching pairs of subjects according to potential confounding variables, for example, sex and age, the impact of confounding is kept to a minimum (to be covered in more detail in Article V in this series).

Crossover design: Where two or more interventions are being compared, the same subject is assigned to both interventions, and relative effectiveness of the two interventions is assessed within subjects. The idea is similar to that of matching, but within-subject variation is smaller than between-subject variation. By using the same subject, potentially confounding characteristics are standardised (to be covered in the final article of this series).

Restriction (or blocking): Subjects are grouped together according to characteristics that are potential confounders and a specified, identical proportion of each group is randomly assigned to an intervention or control group. This maintains the balance of subjects with potential confounding characteristics assigned to each arm of the study.

Stratification: Like restriction, this ensures that characteristics possibly influencing the health outcome measurement are optimally balanced between intervention and control groups. Stratification can be used at the analysis stage, but the procedure has implications for sample size and, if it is not considered at the design stage, the power of an experimental study test will be greatly reduced.

Confounding can be controlled for at the analysis stage by using statistical models that adjust for more than one variable at once. This can only be done, however, if the confounder is known and the appropriate data have been collected.

Contamination occurs when an intervention administered to an intervention group of an experimental study filters into the control group. An example might be where oral health education is given to the intervention arm and this is repeated informally by a subject in the experimental group to a subject in the control group. This may either dull or entirely mask an existing association between intervention and outcome.

Contamination can be avoided by carrying out a clustered study design, where clusters defined by geography or dental practice, for example, are the experimental unit assigned to one or more interventions.

Recommended readings

1. Altman DG, Bland JM. Treatment allocation in controlled trials: why randomise. Br Med J 1999; 318:1209.
2. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002; 359:248–252.
3. Mamdani M, Sykora K, Li P, *et al.* Reader's guide to critical appraisal of cohort studies. 2. Assessing potential for confounding. Br Med J 2005; 330:960–962.

Study design III: Cross-sectional studies

Kate Ann Levin
Dental Health Services Research Unit, University of Dundee, Dundee, Scotland, UK

In this series, I previously gave an overview of the main types of study design and the techniques used to minimise biased results. Here, I describe cross-sectional studies, their uses, advantages and limitations.

Evidence-Based Dentistry (2006) **7,** 24–25. doi:10.1038/sj.ebd.6400375

Cross-sectional studies are carried out at one time point or over a short period (Figure 1). They are usually conducted to estimate the prevalence of the outcome of interest for a given population, commonly for the purposes of public health planning. Data can also be collected on individual characteristics, including exposure to risk factors, alongside information about the outcome. In this way cross-sectional studies provide a 'snapshot' of the outcome and the characteristics associated with it, at a specific point in time.

Why carry out a cross-sectional study?

A cross-sectional study design is used when

- The purpose of the study is descriptive, often in the form of a survey. Usually there is no hypothesis as such, but the aim is to describe a population or a subgroup within the population with respect to an outcome and a set of risk factors.
- The purpose of the study is to find the prevalence of the outcome of interest, for the population or subgroups within the population at a given timepoint.

Cross-sectional studies are sometimes carried out to investigate associations between risk factors and the outcome of interest. They are limited, however, by the fact that they are carried out at one time point and give no indication of the sequence of events — whether exposure occurred before, after or during the onset of the disease outcome. This being so, it is impossible to infer causality.

The next four publications of *Evidence-based Dentistry* describe other study designs that may be more appropriate for the purposes of understanding associations between exposure to risk factors and the outcome of interest. Nevertheless, cross-sectional studies indicate associations that may exist and are therefore useful in generating hypotheses for future research.

Repeated cross-sectional studies may be carried out to give a pseudolongitudinal study, where the individuals included in the study are either chosen from the same sampling frame or from a different one. An example might be the British Association for the Study of Community Dentistry Survey in which 5-year-old children are examined annually and prevalence of caries is recorded. The prevalence of caries for this age group is monitored over time and this information is used in public health policy planning and in the development of targeting strategies.

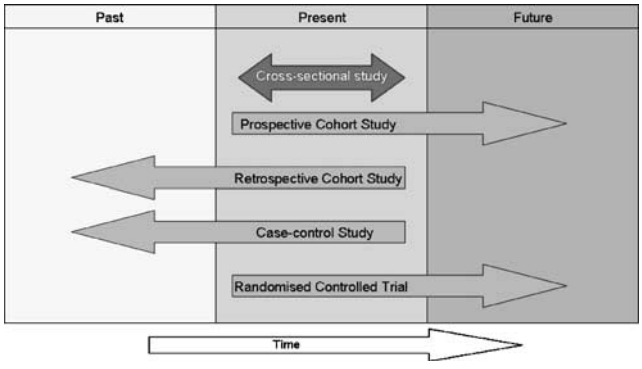

Figure 1. Cross-sectional studies.

Sample selection and response rates

The sample frame used to select a sample and the response rate determine how well results can be generalised to the population as a whole. The sample used in a large cross-sectional study is often taken from the whole population. This is the optimum situation: if the sample is selected using a random technique it is likely that it will be highly representative. In order for the results to be representative of the population, however, not only must the selected sample be representative but so must the responders. Nonresponse is a common problem in wide-scale surveys; techniques to minimise nonresponse include telephone and mail prompting, second and third mailing of surveys, letters outlining the importance of replying and a range of incentives.

The level of nonresponse is one concern, but a greater one still is that of biased response, where a person is more likely to respond when they have a particular characteristic or set of characteristics. Bias will occur when the characteristic in question is in some way related to the probability of having the outcome. The response rate of a survey conducted by door-to-door interview looking at a particular disease, for example, may be highest in the elderly and unemployed because these groups are

more likely to be in their home during the day. These two groups are also more likely to experience higher levels of disease, therefore biasing the results.

Measures of outcome and exposure

A lot of information can be collected about potential risk factors in a cross-sectional study. Loss to follow-up is a common concern in longitudinal studies and one of the strategies used to overcome this is to minimise the amount of information collected. This is not a problem in cross-sectional study design.

It is advisable to think carefully about what might be relevant because this is a good opportunity to gain a broad base of knowledge about subjects who have/do not have the outcome of interest, but it is also important to maintain optimum response levels. Associations between outcomes and exposures of long duration are particularly difficult to establish using cross-sectional studies.

Advantages of cross-sectional studies

- Relatively inexpensive and takes up little time to conduct;
- Can estimate prevalence of outcome of interest because sample is usually taken from the whole population;
- Many outcomes and risk factors can be assessed;
- Useful for public health planning, understanding disease aetiology and for the generation of hypotheses;
- There is no loss to follow-up.

Disadvantages of cross-sectional studies

- Difficult to make causal inference;
- Only a snapshot: the situation may provide differing results if another time-frame had been chosen;
- Prevalence-incidence bias (also called Neyman bias). Especially in the case of longer-lasting diseases, any risk factor that results in death will be under-represented among those with the disease.

Recommended reading

1. Bland M. An Introduction to Medical Statistics. 3rd Edn. Oxford: Oxford University Press; 2001.
2. Pine CM, Pitts NB, Nugent ZJ. British Association for the Study of Community Dentistry (BASCD) guidance on sampling for surveys of child dental health. A BASCD coordinated dental epidemiology programme quality standard. Commun Dent Health 1997; 14(Suppl 1):S10–S17.

Study design IV. Cohort studies

Kate Ann Levin
Dental Health Services Research Unit, University of Dundee, Dundee, Scotland, UK

Previously in this series I have given an overview of the main types of study design and the techniques used to minimise biased results. In this article I describe more fully cohort studies, their uses, advantages and limitations.
Evidence-Based Dentistry (2003) **7**, 51-52. doi:10.1038/sj.ebd.6400407

A cohort study is one in which a group of subjects, selected to represent the population of interest, is studied over time. Much like a cross-sectional study, information is collected about the outcome of interest and exposure to risk factors, but in cohort studies subjects are followed over time. Subjects are disease-free at the outset of the study and at distinct points in time, data are collected relating to health outcomes and exposure to risk factors. This type of study is observational and used to examine causal factors. Cohort studies may be either fixed, where the study subjects do not vary over time and dropouts are not replaced, or dynamic, where new subjects enter the study in accordance with eligibility criteria.

Prospective versus retrospective cohort studies

Prospective studies, those studies carried out from the present time into the future, can be tailored to collect specific exposure data, but there may be a long wait for events to occur, particularly where the outcome of interest is associated with old age. Studies can therefore be expensive to carry out and are prone to high dropout rates, although these can be overcome by incorporating a dynamic study design.

In contrast (Figure 1), retrospective or historical cohort studies look at medical events from some timepoint in the past up to the present time. The advantage of historical cohort studies is that the information is available immediately. There may be difficulty in tracing subjects for such studies, however. A further disadvantage is the reliance on the memory of subjects and/ or the quality of recorded information.

Prospective cohort study design is more commonly used because accurate and complete data, necessary for historical cohort studies, are rarely available. The rest of this article refers to prospective cohort studies unless otherwise stated.

Selection and follow-up of subjects

The probability of having the outcome of interest will be affected by the selection of subjects into a cohort study.

The 'healthy entrant effect' occurs because of the necessity of a disease-free status on entry to the study. Initially, subjects are seen to have lower levels of disease than might be true of the population in general, with an acceleration of disease rate over time. Following cohorts from birth may overcome healthy entrant effects of this sort, but may result in a lengthy and costly study, depending on the average age at onset of disease.

Follow-up of subjects is carried out to monitor changes in health status over time. It is essential to have a mechanism in place that achieves the lowest possible dropout rate from the study. Loss to follow-up will increase with the length of study. A greater concern than number of dropouts are any systematic differences related to the outcome or exposure to risk factors, between those who drop out and those who stay in the study. Analysis of data must include a comparison of risk factors between individuals who remain in the study and those who have dropped out. If loss to follow-up is ignored, the reliability of study conclusions may be called into question.

Inferring causality

To infer causality with any degree of certainty, an experimental study design is required. The longitudinal nature of cohort studies, however, enables the assessment of causal hypotheses, as it is known if exposure occurred prior to outcome. Furthermore, measuring changes in levels of exposure over time alongside changes in outcome measure gives an insight into the dose–response relationship between exposure and outcome. Higher levels of exposure, associated with higher levels of outcome provide further argument for causality.

Hypotheses of cohort studies

Analysis of cohort study data takes one of two forms. The first is a straightforward

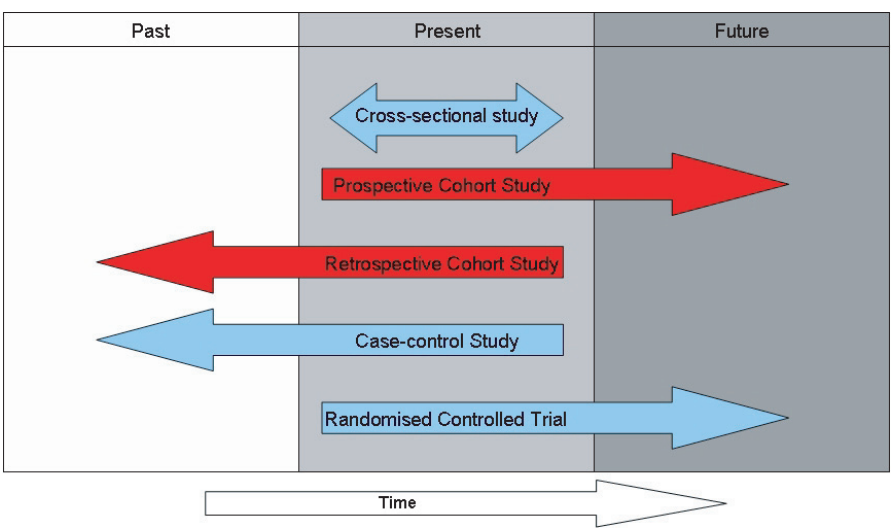

Figure 1. Prospective and retrospective studies

Table 1.

	Exposed to risk factor:		
Outcome	**Yes**	**No**	**Total**
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	N

comparison of two groups, those with exposure and those without. The hypothesis (H_o) is then: on average, the disease in the exposed group is no different to that of the unexposed group. An example of this might be comparing the dental health of those who smoke with those who do not.

The second type of analysis occurs when there are more than two groups. An example might be where smokers are split into those who do not smoke, those who smoke fewer than 20 cigarettes a day and those who smoke over 20 cigarettes. Where there is this type of grouping in strata, hypotheses fall into three broad categories:

- H_o: on average, the outcome rates in groups of subjects are no different to that of the population as a whole.
- H_o: on average, the outcome rates in groups of subjects are no different from each other.
- H_o: on average, there is a linear trend in outcome rates across groups.

Analysis of data

Using the data collected in a cohort study, the following statistics may be obtained:

Crude rates of outcome: this is the number of individuals with the outcome out of the total cohort study size. In Table 1 this is given by:

$$\frac{(a+b)}{n}$$

Standardised rates and ratios of the outcome can also be calculated, using demographic information so that rates of the outcome are adjusted for other potential risk factors such as age and sex.

Risk ratio of outcome: the risk of the outcome in exposed subjects relative to those not exposed is given in Table 1 by:

$$\frac{a/(a+c)}{b/(b+d)}$$

Where one or more of the second set of above hypotheses is being examined, more complex data analysis is used. Regression analysis allows investigation of two or more groups of subjects whilst adjusting for characteristics that might act as confounding risk factors.

Advantages of cohort studies

- The temporal dimension, whereby exposure is seen to occur before outcome, gives some indication of causality
- Can be used to study more than one outcome
- Good for the study of rare exposures
- Can measure the change in exposure and outcome over time
- Incidence of outcome can be measured

Disadvantages of cohort studies

- Costly (less so for retrospective) and may take a long time, particularly where onset of the outcome measure can occur both early and late on in life
- Require accurate records for retrospective studies
- When studying rare outcomes, a very large sample size is required
- Prone to dropout
- Changes in aetiology of disease over time may be hard to disentangle from changes observed as age increases
- Selection bias: a difference in incidence of the outcome of interest, between those who participated and those who did not, would give biased results

Recommended reading

1. Breslow NE, Day NE. Statistical Methods in Cancer Research. Volume II. The Design and Analysis of Cohort Studies. IARC Scientific Publications no. 82. Town: Oxford University Press; 1987.
2. Grimes DA, Schultz KF. Cohort studies: marching towards outcomes. Lancet 2002; 359:341–345.
3. Squires BP, Elmslie TJ. Cohort studies: what editors want from authors and peer reviewers. CMAJ 1990; 143:179–180.
4. Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. Am J Epidemiol 2005; 161:280–288.

Study design V. Case–control studies

Kate Ann Levin
Dental Health Services Research Unit, University of Dundee, Dundee, Scotland, UK

Previously in this series I have given an overview of the main types of study design and the techniques used to minimise biased results. In this article I describe more fully case control studies, their uses, advantages and limitations.
Evidence-Based Dentistry (2003) **7**, 83-84. doi:10.1038/sj.ebd.6400436

Like cohort studies, the purpose of case-control studies is to establish association between exposure to risk factors and disease. Unlike cohort studies, however, members of the population with the disease are selected into the study at the outset and risk factor information is collected retrospectively (figure 1). These are known as the cases. A second group of individuals who do not have the disease, the controls, is also included in the study (figure 2). The case-control study design is often used in the study of rare diseases or as a preliminary study where little is known about the association between the risk factor and disease of interest.

Case control studies are prone to bias and confounding. In order to minimise bias, care must be taken in the selection of both cases and controls, in establishing definitions of disease, risk factors and in ensuring there are no confounding associations between detection of disease and risk factor exposure.

Choice of cases

Care must be taken when choosing cases for the study. In particular it is important to distinguish between stages or subtypes of disease and to define a measure of health status. For example, when studying oral cleanliness it is important to define what is meant by oral cleanliness in terms of cause, nature and degree. It is also important to establish whether interest is in incident cases, subjects entered into study on detection of disease, or prevalent cases, those who have been diagnosed as having the disease prior to the study. Views, behaviours and reports of exposure to risk factor amongst these two groups will tend to differ, as those who have been diagnosed previously are likely to be more informed about the disease and may have altered their behaviour and attitudes since diagnosis. Incident case design is preferred as it reduces recall bias and over-representation of cases with long standing disease

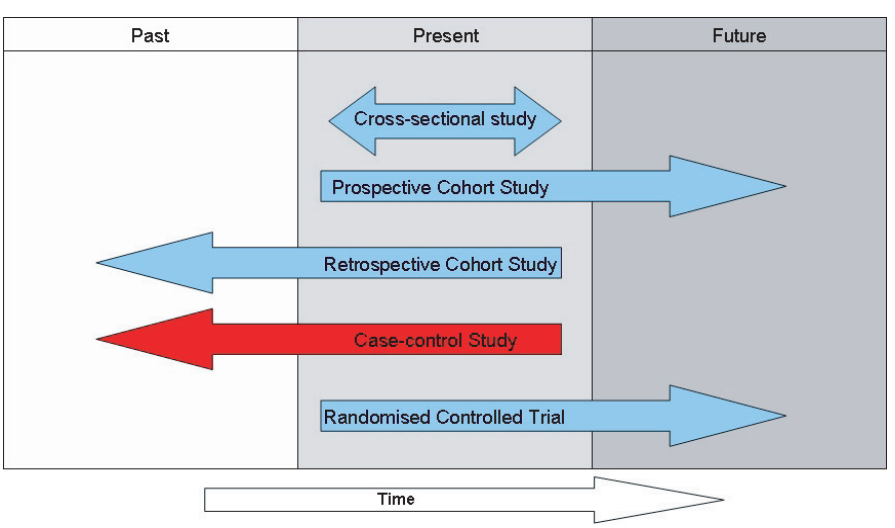

Figure 1. Prospective and retrospective studies

Choice of controls

Controls should come from the same population at risk of disease, should not have the disease and should be representative of the target population. Selecting controls often proves harder than cases and requires great care in the prevention of bias. A sampling frame of hospital patients is often used to select controls, however risk factors such as diet and smoking are commonly linked to many diseases. Selecting controls in this way might therefore over-estimate population exposure to such risk factors, resulting in an underestimation of association between disease and exposure. Using more than one control group helps to overcome this type of issue.

Multiple controls can be used for each case, giving the study greater power, particularly where the number of cases is small, due for example, to the disease being rare.

Exposure to risk factors and matching

Exposure measurements are reliant either on memory where cases and controls are interviewed retrospectively, and/or medical records. Exposure estimates are therefore vulnerable to recall bias; commonly those with the disease are more likely to remember exposure than those without, interview or measurement bias; where the interviewer interviews or reports findings systematically differently between cases and controls and confounding factors. Interview and measurement bias can be overcome by including blinding in the design so that they do not know who is a case and who is a control at the time of interview.

Confounding factors must be identified prior to the start of the study. Individual cases can be matched to controls where is it thought that other factors, aside from those risk factors of interest, might contribute to the development of disease and confound the causal association under investigation. Cases and controls are commonly matched by age and sex. The factor upon which cases and controls are matched cannot be studied as a risk factor. An alternative method of overcoming confounding is to collect relevant information on the confounding factor during the course of the study and adjust for this at the analysis stage.

Analysis of Data

The table used to analyse the data, Table 1, looks much like that used in a cohort study

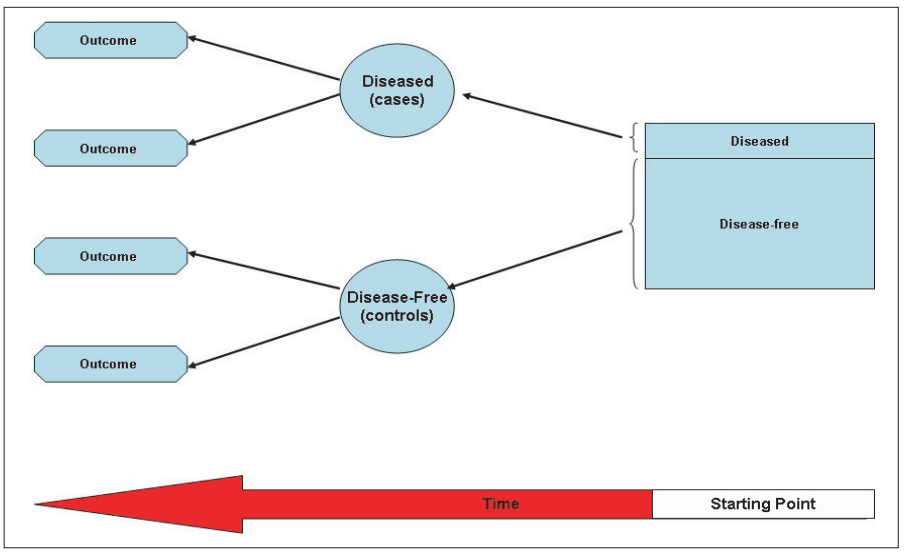

Figure 2. Diagrammatic representation of case-controlled study

Table 1

	Exposed to risk factor		
Status	Yes	No	Total
Case	a	b	a+b
Control	c	d	c+d
Total	a+c	b+d	n

design (see Pub 4), however, the case-control study is retrospective so that instead of measuring relative risk of disease based on exposure, we measure the odds of exposure based on disease.

The odds ratio is made up of two components,

- the odds of exposure for cases; the number of cases exposed/ the number of cases unexposed given by Odds/cases = a/b
- the odds of exposure for controls; the number of controls exposed/ the number of controls unexposed given by Odds/ controls = c/d

The estimated Odds Ratio is then Odds = ad/bc

The Odds Ratio is interpreted as

- OR<1 : Odds of exposure for cases are less than those for control. Exposure appears to reduce risk of disease.
- OR=1 : Odds of exposure for cases are the same as those for control. Exposure does not appear to be a risk factor.
- OR>1 : Odds of exposure for cases are more than those for control. Exposure appears to increase risk of disease.

A 95% confidence interval gives an indication of the confidence we have in the estimated Odds Ratio, eg. if the entire 95% confidence interval is above 1, it is concluded that exposure significantly increases the risk of disease at the 95% level.

Advantages of case-control studies

Case-control studies are quick and cheap and are particularly suited to the study of rare diseases as the diseased are selected at the outset of the study.

Disadvantages of case-control studies

The disadvantages of case-control studies include

- difficulties in overcoming potential bias and confounding
- the successful selection of both cases and controls who are representative of their respective populations is often difficult.
- an inability to infer causality and no information on the chronology of disease and exposure.
- Inefficient in studying risk factors which are rare
- Studies are often not population based, therefore it is impossible to calculate incidence of disease.

Recommended Reading

1. Breslow NE, Day NE. Statistical methods in cancer research. Volume I- the analysis of case-control studies. IARC, Lyon: IARC, 1980.
2. Schlesselman JJ. Case control studies. Design, conduct, analysis. Oxford: University Press, 1982.
3. Schultz KF. Case- control studies: research in reverse. Lancet, 2002; 359: 431-434.
4. Squires BP and Elmslie TJ. Reports of case-control studies: what editors want from authors and peer reviewers. CMAJ, 1990; 143: 17-18.

Study Design VI - Ecological Studies

Kate Ann Levin
Child and Adolescent Health Research Unit, University of Edinburgh, Edinburgh, Scotland, UK

Previously in this series I have given an overview of the main types of study design and the techniques used to minimise biased results. In this article I describe more fully ecological studies, their uses, advantages and limitations.
Evidence-Based Dentistry (2003) **7**, 60-61. doi:10.1038/sj.ebd.6400454

An ecological study is an observational study defined by the level at which data are analysed, namely at the population or group level, rather than individual level. Ecological studies are often used to measure prevalence and incidence of disease, particularly when disease is rare. They are inexpensive and easy to carry out, using routinely collected data, but they are prone to bias and confounding. Also, because they are area-level studies, care must be taken when extrapolating either to individuals within the area level of measurement, or to a higher population level. Although other study designs are generally considered more reliable, particularly in the inference of causation, the population context of individual characteristics has been shown to be a stronger determinant of disease at population level than individual level risk factors.

Why carry out an ecological study?

An ecological study design is used when;

- the purpose of the study is to monitor population health so that public health strategies may be developed and directed;
- the purpose of the study is to make large-scale comparisons, eg, comparisons between countries;
- the purpose of the study is to study the relationship between population-level exposure to risk factors and disease, or in order to look at the contextual effect of risk factors on the population;
- measurements at individual level are not available, eg. confidentiality might require that individuals are anonymised by aggregation of data to small area level; or
- the disease under investigation is rare, requiring aggregation of data for any analysis to be carried out.

Types of measurement in ecological studies

In ecological studies health outcomes are aggregates of individual health data, eg: prevalence, incidence, rate of disease. Ecological risk or exposure data takes the form of one or more of the following:

- Aggregate measures; the data are summaries of individual level data eg, mean dmft, percentage of children with no caries, area-level deprivation indices
- Environmental measures; equivalent individual level data are conceivable eg, mean annual exposure to fluoridation
- Global measures; there are no equivalent individual level data eg, number of dental practices, population density.

Types of ecological studies

Geographical; This type of study compares one geography with another by assessing the health of the population of each. Exposures for geographies may also be measured and included in analysis as well as other potential confounding variables such as demographic and socioeconomic information.

Longitudinal; A population is monitored to assess changes in disease over time. Again, confounding factors are often included in analysis.

Migration; Data of migrant populations are collected and analysed. The unit of interest is neither time nor place, but population type.

The ecological fallacy

The ecological fallacy is a type of confounding specific to ecological studies. It occurs when relationships which exist for groups are assumed to also be true for individuals. For example, an area with a majority of girls may have higher mean dmft than that of a second area with a majority of boys. The conjecture might therefore be that girls have higher dmft than boys. There are two types of potential confounding in this inference. The first is confounding which is common to all observational studies, i.e. because of the unknown characteristics of the areas being studied - for example, the first area might have higher deprivation than the second. The conjecture might therefore be that children from deprived areas have higher dmft. The second type of confounding is what is known as the 'ecological fallacy'. Even if all confounders are adjusted for, the aggregate nature of such a study tells us nothing about who in the population has high dmft. Although there are only a few boys in the first population, perhaps these are the only children with dmft, albeit extremely high dmft, bringing the mean above that of the second population. If this were true then our conjecture would be incorrect and may even be reversed at the individual level.

Analysis of ecological data

As with all observational studies, in order to overcome confounding, regression analysis is advisable. Multilevel modelling techniques have been developed, where analysis includes both individual and population level data, thus overcoming the ecological fallacy and enabling examination of contextual effects. Dental data are suited to multilevel modelling due to the clustered nature of disease, within surfaces, teeth, individuals, as well as group clusters such as schools, dental practices and geographical areas. Where individual level data are not known, for example where data are only available in an aggregated form, care must be taken in making causal inferences due to aggregation bias and the ecological fallacy.

Recommended reading

1. Pearce N. The ecological fallacy strikes back. J Epidemiol Community Health, 2000; 54: 326-327.
2. Piantadosi S, Byar DP, Green SB. The ecological fallacy. Am J Public Health, 1988; 127: 893-904.
3. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health, 1994: 84; 825-829.
4. Susser M. The logic in ecological: I. The logic of design. Am J Public Health, 1994: 84; 830-835.

Study design VII. Randomised controlled trials

Kate Ann Levin
Child and Adolescent Health Research Unit, University of Edinburgh , Edinburgh, Scotland, UK

Previously in this series, I have given an overview of the main types of study design and the techniques used to minimise the likelihood of obtaining biased results. In this article I describe more fully randomised controlled trials, their uses, advantages and limitations.

Evidence-Based Dentistry (2007) **8**, 22-23. doi:10.1038/sj.ebd.6400473

The study designs described previously in this series have focused on observational studies, where outcome measures have been recorded either at one specific timepoint (cross-sectional) or repeatedly over time (longitudinal). Information may be collected at an aggregate level (ecological) and may be gathered prospectively (cohort study) or retrospectively (case–control study).

Randomised controlled trials (RCT), or randomised clinical trials, are experimental studies where the effect of an intervention is assessed by collecting data before and after an intervention has taken place. RCT are used to compare an intervention with one or more other interventions or with no intervention. Interventions are often clinical treatments but may also be educational interventions such as health promotion leaflets.

What are the two main features of the RCT?

1. They are comparative

In RCT, an intervention is investigated by comparing one group of people who receive the intervention with a control group or control arm who do not (Figure 1, below). The control group receives the usual or no treatment and their outcome measure, or the change in measure from the starting point or baseline, is compared with that of the intervention group.

2. They are designed to minimise bias

Allocation bias

Allocation bias occurs when the measured treatment effect differs from the true treatment effect because of how participants were selected into the intervention or control groups. In RCT, once the participants are entered into the study, they are randomised to either an intervention group or the control group. Randomisation ensures that characteristics that might affect the relationship between intervention and outcome measures will be roughly equal across all arms of the study, minimising potential bias.

Performance bias

Even after randomised allocation, bias can occur. Performance bias occurs when participants' response to the treatment is affected by knowledge of the group to which they are assigned, or when health professionals administer treatments differently between treatment arms.

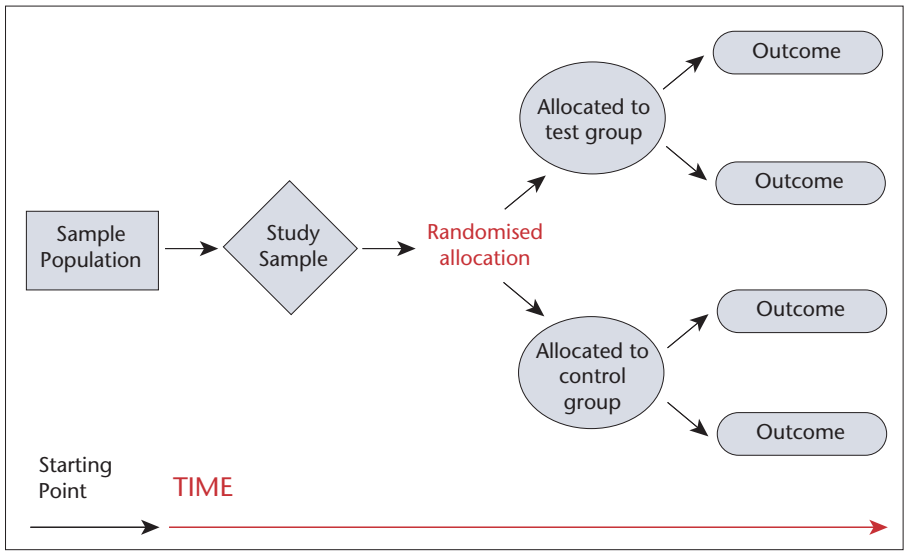

Figure 1. Diagrammatic representation of a randomised controlled trial

Assessment bias

Health professionals assessing the outcome of treatment relative to alternative or placebo interventions may record outcome measures biased by the knowledge of the group to which the participant has been assigned. Over- (or under-) estimation of the effects of an intervention, even if subconscious, is known as assessment bias. Alternatively there might be a systematic difference in measuring the outcomes between the two groups because of the method of recording used. For example, where the control group is assigned to one practitioner and the intervention group to another, or where groups are assessed at different times of the day, there may be a systematic difference between groups in the mean outcome measure recorded. Bias will be minimised where a standardised method of evaluation is used across both groups. Subjective

measures to assess the effectiveness of a treatment will be more prone to bias.

Attrition bias

Attrition bias (also call loss-to-follow-up bias) occurs when patients drop out of the study from one or other of the study groups preferentially. For example, if halfway through a study the treatment has been successful participants may drop out, and information about the success of the treatment is then lost. Conversely, participants in the control group may be unhappy with their lack of progress and may drop out of the study in order to seek alternative help.

Allocation concealment

Bias will be minimised where the allocation schedule is concealed of whom is assigned to which group. Blinding (also known as masking) helps to prevent systematic differences between comparison groups in prognosis or responsiveness to treatments (allocation bias). Blinding of both participants and practitioners prevents performance and assessment bias by ensuring that participants, treatment administrators and those measuring outcomes do not know which treatment was given. Where possible, it is recommended that RCT participants are blind to the treatment they receive. In order to do this it is sometimes appropriate for the control group to receive a placebo. This enables the RCT to be blind and overcomes the placebo effect, where it is the action of taking medication and not the medication itself that results in a positive outcome.

Why carry out RCT?

RCT are prospective longitudinal studies, allowing the investigation of causal associations between interventions and outcomes. The random selection of participants into each arm and the controlled way in which the trial is carried out mean that all factors other than the intervention are considered equal. Although associations may be investigated under observational studies, causality cannot be inferred. If possible, it is always preferable to choose RCT over other study designs when assessing the impact of one or more interventions.

Factors to be considered when carrying out RCT

Sample size

The size of the sample required when carrying out RCT is dependent upon the power of the test and what size of intervention impact is considered meaningful. It also depends on the type of hypothesis the RCT is testing. The smaller the magnitude of difference between groups that is to be detected and the greater the variability in outcomes, the larger the sample size that will be required.

Stratification

Very large trials are likely to have a good balance of patients within each arm. When the samples are small, however, treatment groups may by chance end up with different characteristics, which may affect the outcome of the trial. Stratification is a way of ensuring the treatment groups are balanced on characteristics that are likely to alter the relationship between treatment and outcome.

Crossover design

Crossover trials are another way of overcoming differences in groups by keeping the patients as matched as possible. Instead of having different patients in each treatment group, patients receive first one treatment and then the other, in a random order, with a washout period in between. Within-patient differences are then compared. Thus each patient effectively becomes their own 'test' and 'control'.

Between group contamination

Educational interventions, in particular, are prone to contamination where, for example, a member of the control arm is a friend of a patient receiving the oral health advice-sheet intervention. Information may then pass between the two arms of the trial, and thus alter the results.

One way to overcome this is to use cluster sampling so that natural clusters such as geographic areas or dental practices are randomised rather than individuals.

Ethical issues

RCT are not always possible because of ethical issues when assigning patients to study arms.

If one group of patients receives treatment thought to be effective, while another group does not, the ethics of a trial may be brought into question. Similarly, there are some trials that cannot be carried out because they may actively encourage unhealthy practices such as smoking; people cannot be randomised into smoking and nonsmoking groups.

Analysis of data

RCT are experiments set up to test hypotheses. The null hypothesis (H0) is that the intervention will have no impact on the outcome measure (ie, that the outcome of interest will be similar in both the test and control groups). The alternative to the null hypothesis (H1) is that the intervention will have a meaningful effect and that this effect will be statistically significant. The statistical method required to test this hypothesis will depend on the nature of the outcome of interest. Comparing the proportions of patients achieving a 'successful outcome' between treatment groups (eg, those satisfied with their treatment) will require a different approach from investigating average differences between groups (eg, average probing pocket depths).

Advantages of RCT

- Ability to make causal inferences mean that RCT provide the strongest empirical evidence of a treatment's efficacy.
- Randomisation of participants to the test and control arms and concealment of their allocation ensures that allocation bias and confounding of unknown variables are minimised.
- The study can be tailored to answer a specific research question.

Disadvantages of RCT

- High dropout when the intervention has undesirable side-effects or there is little incentive to stay in the control arm.
- Ethical considerations may mean that a research question cannot be investigated using the RCT design.
- For a descriptive overview it may be cheaper and easier to use an observational design.
- Prior knowledge is required about the level of improvement that is clinically meaningful and the expected variation of improvement in the sample in order to calculate the RCT sample size. These facts are often not known.

Recommended reading

1. Pocock SJ. Clinical Trials: a Practical Approach. Chichester: Wiley; 1983.
2. Altman DG, Bland JM. Treatment allocation in controlled trials: why randomise? Br Med J 1999; 318:1209.
3. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting randomized controlled trials: the CONSORT statement. J Am Med Assoc 1996; 276:637–639.