

Argy Polychronopoulou

7.1 Introduction

This chapter addresses key issues related to the design of epidemiologic studies as they apply to research in orthodontics. First, the fundamental measures of epidemiology and measures derived from them to quantify causal effects are presented. In addition, basic epidemiologic study design strategies as well as their strengths and limitations are explored. Lastly, sources of error both in epidemiologic study design and methods of error control and evaluation are described. Emphasis is placed on epidemiologic principles and concepts without resorting to mathematical notation.

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study in controlling health problems. The application of epidemiological principles and methods to the practice of clinical medicine/dentistry is defined as *clinical epidemiology* [1, 2].

According to the above definition, two regions of epidemiologic research are determined: *descriptive epidemiology*, which is related to the study of the distribution and development of diseases, and *etiologic or analytic epidemiology*, which is related to the investigation of likely factors that form these distributions. Consequently, the study and the calculation of measures of occurrence of illness or any outcome of interest as well as effect indicators of illnesses constitute the central activity in epidemiology [3].

Before the basic indicators that are used in epidemiology are described, some information should be provided on the *meaning of population*. Populations are at

A. Polychronopoulou
Department of Preventive and Community Dentistry,
School of Dentistry, University of Athens,
Thivon 2, Athens 11527, Greece
e-mail: argypoly@dent.uoa.gr

the center of epidemiologic research because epidemiologists are concerned about disease occurrence in groups of people rather than individuals. A population can be defined as a group of people with a common characteristic such as age, gender, residence, life event, and use of medical/dental services. Moreover, the population can be fixed or dynamic/open. In the fixed population, no new individuals are added. Membership is based on an event and is permanent. On the contrary, in the open population, new individuals are added and removed. Membership is based on a condition and is transitory.

7.2 Measures of Disease Occurrence

The need for the use of indicators arose from the fact that absolute numbers in epidemiology do not provide useful information. The basic indicators that are used for the measurement of frequency of the appearance of illness are: (a) *prevalence* and (b) *incidence* [1].

Prevalence measures the probability of having a disease or any outcome of interest, whereas *incidence* measures the probability of getting a disease. Incidence indicators always concern new cases of illness (newly appeared or newly diagnosed). There are two incidence indicators: *incidence proportion* and *incidence rate* [4, 5].

Epidemiologic indicators can be calculated for the entire population but also for parts of the population on the basis of any characteristic. In the first case, they are called “crude” and in the second case, “characteristic specific.”

7.2.1 Prevalence

Prevalence describes the proportion of people in a population who have a particular disease or attribute. There are two types of prevalence measures, *point* and *period prevalence*, that relate prevalence to different time amounts [3].

The calculation formula of point prevalence is:

$$\text{Point prevalence} = \frac{\text{number of existing cases of disease at a point in time}}{\text{total population at risk at that point in time}}$$

while the calculation formula of period prevalence is:

$$\text{Period prevalence} = \frac{\text{number of existing cases of disease in a specific period}}{\text{total population at risk during that period}}$$

Prevalence is a proportion. It is dimensionless and can only take a numeric value in the range of zero to one (Table 7.1). *Point prevalence* is used to get a “snapshot” look at the population with regard to a disease. *Period prevalence* describes how much of a particular disease is present over a longer period that can be a week, month, or any other specified time interval.

Table 7.1 Comparison of measures of disease occurrence

Measure	Interpretation	Range	Unit	Cases	Source population	Uses
Prevalence	Probability	0–1	None	Existing	At risk of becoming a case	Planning
Incidence proportion	Probability	0–1	None	New	At risk of becoming a case	Etiologic studies
Incidence rate	Inverse of waiting time	0 to ∞	1/Time	New	At risk of becoming a case	Etiologic studies

Thus, a study group with an orthodontic treatment needing prevalence of 0.45 shows that 45 % of the subjects require some type of treatment at the time of the examination.

7.2.2 Incidence Proportion

Incidence proportion is defined as the proportion of a population that becomes diseased or experiences an event over a period of time [5].

The calculation formula of incidence proportion is:

$$\text{Incidence proportion} = \frac{\begin{array}{l} \text{the number of people who got the illness in the duration} \\ \text{of a specific time period} \end{array}}{\text{number of people in the population who are in danger of becoming ill at the beginning of the period}}$$

Incidence proportion is dimensionless and can only take numeric value in the range of zero to one (Table 7.1). Additionally, it is always referred to in the specific time period of being observed. Incidence proportion is also called cumulative incidence, average risk, or risk. Both the numerator and the denominator include only those individuals who, in the beginning of the time period, were free of illness and were thus susceptible to developing it. Therefore, cumulative incidence refers to those individuals who went from being “free of illness” at the beginning of the time period to being “sick” during that particular time period. Consequently, cumulative incidence can evaluate the average danger for individuals of the population to develop the illness during this time period. Cumulative incidence is mainly used for fixed populations when there are small or no losses to follow up. The length of time of monitoring observation directly influences the cumulative incidence: the longer the time period, the bigger the cumulative incidence. Thus, a 2-year incidence proportion of 0.20 indicates that an individual at risk has a 20 % chance of developing the outcome over 2 years.

A useful complementary measure to cumulative incidence is the *survival proportion* [5]. Survival is described as the proportion of a closed population at risk that

Person	Time						Person Years accrued
	2000	2001	2002	2003	2004	2005	
A	◆	→					2
					(Diagnosed with outcome/disease)		
B		◆	→				4
						(Moves away/lost)	
C		◆	→				5
						(End of follow-up)	
D	◆	→					5
					(Diagnosed with outcome/disease)		
E	◆	→					4
						(Moves away/lost)	
Total							20

Fig. 7.1 Person-time measurement in a hypothetical population, individual follow-up times are known

does not become diseased within a given period of time and is the inverse of incidence proportion. Incidence and survival proportion are related using the following equation:

$$\text{Survival proportion} = 1 - \text{incidence proportion or}$$

$$\text{Incidence proportion} = 1 - \text{survival proportion}$$

7.2.3 Incidence Rate

Incidence rate is defined as the relative rate at which new cases or outcomes develop in a population [3]. The mathematic formula of incidence rate is:

$$\text{Incidence rate} = \frac{\text{number of illnesses that are expressed in a population for the duration of a period of time}}{\text{sum, for each individual of the population, the time observed in which he / she is at risk of developing the illness}}$$

The sum of the time periods in the denominator is often measured in years and is referred to as “person-years,” “person-time,” or “time at risk.” For each individual of the population, the time at risk is the time during which this individual is found in danger of developing the outcome under investigation. These individual time periods are added up for all the individuals (Fig. 7.1).

Logic follows that the total number of individuals who change from being healthy to being sick, in the duration of each time period, is the product of three factors: the size of the population, the length of the time period, and the “strength of the

unhealthiness” that acts upon the population. Dynamic incidence measures this specific strength of unhealthiness. The entry and exit of individuals from the population during the study period for reasons such as immigration, fatality from other causes, or any other competing risks are automatically taken into account. Therefore, including the time of danger in the definition, the incidence compensates for the main disadvantages that come into question from the calculation of the incidence proportion. The dynamic incidence is not a percentage, like the two previous indicators, since the numerator is the number of incidents and the denominator is the number of person-time units. The size of incidence rate is always bigger or equal to zero and can go to infinity (Table 7.1).

Thus, using the data presented in Fig. 7.1, we can estimate the following incidence rate: 2 cases/20 person-years=0.1 cases/person-year, that indicates that for every 10 person-years of follow-up, 1 new case will develop.

7.2.4 Relationship Between Incidence and Prevalence

Among the indicators, various mathematic relations have been formulated, taking, however, certain acknowledgements into consideration [2, 3]. The equation that connects prevalence with incidence rate is:

$$\frac{\text{Prevalence}}{1 - \text{Prevalence}} = \text{Incidence rate} \times \text{Average duration of the disease}$$

That is to say, prevalence depends on incidence as much as the duration of the illness. This is in effect when concerning a steady state, where the incidence of illness and the duration of illness remain constant with the passage of time. If the frequency of disease is rare, that is, less than 10 %, then the equation simplifies to:

$$\text{Prevalence} \approx \text{Incidence rate} \times \text{Average duration of the disease}$$

7.3 Measures of Causal Effects

A central point in epidemiologic research is the study of the cause of illnesses. For this reason, in epidemiologic studies, the frequency of becoming ill, among individuals that have a certain characteristic, is generally compared to the corresponding frequency among those that do not have this characteristic. The compared teams are often referred to as “exposed” and “not exposed.” Exposure refers to the explanatory variable, that is, any agent, host, or environment that could have an effect on health. The *effect indicators* are useful in order for us to determine if exposure to one factor becomes the cause of illness, to determine the relation between a factor and an illness, and to measure the effect of the exposure on the factor [3, 5, 6].

Table 7.2 Organization of prevalence and incidence proportion data

Exposure		Disease		
		Yes	No	Total
Yes		a	b	a + b
No		c	d	c + d
Total		a + c	b + d	a + b + c + d

Where: $a + b + c + d$ total number in the study, $a + b$ total number exposed, $c + d$ total number not exposed, $a + c$ total number diseased, $b + d$ total number not diseased, a number exposed and diseased, b number exposed but not diseased, c number not exposed but diseased, d number neither exposed nor diseased

Table 7.3 Organization of incidence rate data

Exposure		Disease		
		Yes	No	Person-time
Yes		a	–	Person-time exposed
No		c	–	Person-time unexposed
Total		a + c	–	Total person-time

Comparing measures of disease occurrence in either absolute or relative terms is facilitated by organizing the data into a *fourfold* Table [7]. For example, the recorded data for the estimation of the prevalence and incidence proportion can be seen in Table 7.2.

While the presentation of facts for the estimation of the dynamic incidence can be seen in Table 7.3.

7.3.1 Absolute Measures of Effect

Absolute effects are differences in prevalences, incidence proportions, or incidence rates, whereas relative effects involve ratios of these measures. The absolute comparisons are based on the difference of frequency of illness between the two teams, those exposed and not exposed [7]. This difference of frequency of illness between the exposed and not exposed individuals is called risk or rate difference, but also attributable risk or excess risk and is calculated using the following mathematical formula:

$$\text{Absolute effect} = [\text{measure of disease occurrence among exposed}] \\ - [\text{measure of disease occurrence among nonexposed}]$$

The difference measure gives us information on the absolute effect of exposure on the measure of disease occurrence, the difference in the risk to the exposed population, compared to those not exposed, and the public incidence of the exposure. The risk or rate difference has the same dimensions and units as the indicator which

Table 7.4 Comparison of effect measures

Measure	Type	Range	Unit	Interpretation
Risk difference	Absolute	-1 to +1	None	Excess risk of disease among exposed population
Risk ratio	Relative	0 to +∞	None	Strength of relationship between exposure and disease, the number of times the risk of exposed subjects is higher or lower than the risk of nonexposed
Incidence rate difference	Absolute	-∞ to +∞	1/time	Excess rate of disease among exposed population
Incidence rate ratio	Relative	0 to +∞	None	Strength of relationship between exposure and disease, the number of times the rate of exposed subjects is higher or lower than the rate of nonexposed

is used for its calculation (Table 7.4). The difference refers to the part of the illness in the exposed individuals that can be attributed to their exposure to the factor that is being examined, with the condition that the relation that was determined has an etiologic status. Thus, a positive difference indicates excess risk due to the exposure, whereas a negative difference indicates a protective effect of the exposure.

7.3.2 Relative Measures of Effect

On the contrary, the relative measures of effect are based on the ratio of the measure of disease occurrence among exposed and not exposed [1]. This measure is generally called the risk ratio, rate ratio, relative risk, or relative rate. That is to say, the relative effect is the quotient of the measure of disease occurrence among exposed by the measure of disease occurrence among nonexposed.

$$\text{Relative effect} = \frac{[\text{measure of disease occurrence among exposed}]}{[\text{measure of disease occurrence among nonexposed}]}$$

The relative comparisons are more suitable for scientific intentions [2, 3]. The relative risk gives us information on how many times higher or lower is the risk of somebody becoming ill. It also presents the strength of association and is used in the investigation of etiologic relations. The main reason is that the importance of difference in the measure of disease occurrence among two populations cannot be interpreted comprehensibly except in relation to a basic level of disease occurrence. The relative risk is a clean number.

If the risk ratio is equal to 1, the risk in exposed persons equals the risk in the nonexposed. If the risk ratio is greater than 1, there is evidence of a positive association, and the risk in exposed persons is greater than the risk in nonexposed persons.

Table 7.5 Strength of association

Relative risk/rate	Relationship interpretation
1.1–1.3	Weak
1.4–1.7	Modest
1.8–3.0	Moderate
3.0–8.0	Strong
8.0–16	Very strong
16–40	Dramatic
40+	Overwhelming

If the risk ratio is less than 1, there is evidence for a negative association and possibly protective effect; the risk in exposed persons is less than the risk in nonexposed. A relative risk of 2.0 indicates a doubling of risk among the exposed compared to the nonexposed. The power of the relative risk or relative rate can be accredited according to Table 7.5.

We have examined the most commonly used measures in epidemiologic research that serve as tools to quantify exposure-disease relationships. The following section presents the different designs that can utilize these measures to formulate and test hypothesis.

7.4 Study Design

Epidemiologic studies can be characterized as measurement exercises undertaken to get estimates of epidemiologic measures. Simple study designs intend to estimate risk, whereas more complicated study designs intend to compare measures of disease occurrence and specify disease causality or preventive/therapeutic measures' effectiveness. Epidemiologic studies can be divided into two categories: (a) descriptive and (b) etiologic or analytic studies [2, 7, 8].

7.4.1 Descriptive Studies

Descriptive studies have several roles in medical/dental research. They are used for the description of occurrence, distribution, and diachronic development of diseases. In other words, they are based on the study of the characteristics of individuals who are infected by some disease and the particularities of their time-place distribution.

Descriptive studies consist of two major groups: those involving individuals and those that deal with populations [7]. Studies that involve individual-level data are mainly *cross-sectional studies*. A cross-sectional study that estimates prevalence is called a prevalence study. Usually, exposure and outcome are ascertained at the same time, so that different exposure groups can be compared with respect to their disease prevalence or any other outcome of interest prevalence. Because associations are examined at one point in time, the temporal sequence is often impossible to work out. Another disadvantage is that in cross-sectional studies, cases with long

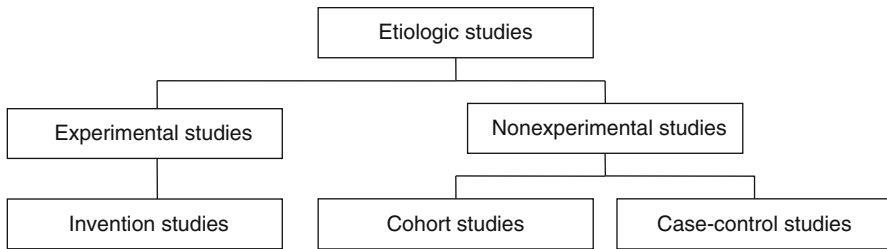


Fig. 7.2 Types of etiologic studies

duration of illness are overrepresented whereas cases with short duration of illness are underrepresented.

Studies where the unit of observation is a group of people rather than an individual are *ecological studies*. Because the data are measurements averaged over individuals, the degree of association between exposure and outcome might not correspond to individual-level associations.

Descriptive studies are often the first tentative approach to a new condition. Common problems with these studies include the absence of a clear, specific, and reproducible case definition and interpretation that oversteps the data. Descriptive studies are inexpensive and efficient to use. Trend analysis, health-care planning, and hypothesis generating are the main uses of descriptive design.

7.4.2 Etiologic Studies

The basic objective in epidemiologic science is the search for etiologic relations between various factors of exposure and various diseases. For this reason, etiologic studies are used (Fig. 7.2). It is about studies, experimental and not, which aim to investigate the etiology of a disease or to evaluate a preventive/therapeutic measure, through documentation of the association of a disease and a likely etiologic or preventive/therapeutic factor on an individual basis. Etiologic studies are distinguished in experimental (intervention studies) and in nonexperimental (observational studies). Nonexperimental studies are distinguished in cohort studies and in case-control studies [2, 3, 9].

7.5 Nonexperimental Studies

7.5.1 Cohort Studies

Cohort studies are also called follow-up studies, longitudinal studies, or incidence studies. The choice of individuals on which the study is based is made with the exposure or nonexposure on that factor as a criterion, for which its etiologic contribution to the illness is being investigated. These groups are defined as study cohorts.

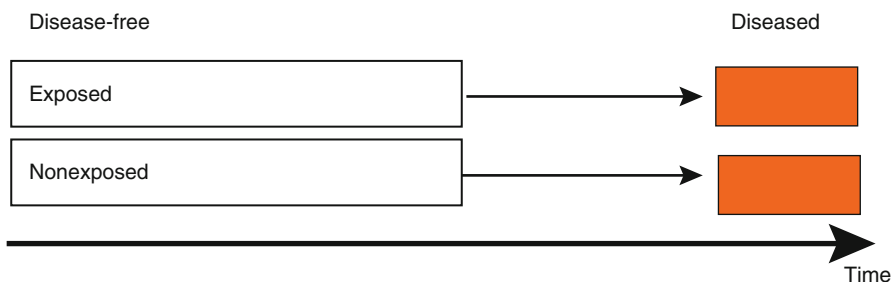


Fig. 7.3 Design of a cohort study

All participants must be at risk of developing the outcome. The individuals of the study are followed for a set period of observation, which is usually long, and all the new cases of the illness being studied are identified. Comparisons of disease experience are made within the study cohorts. It should be specified that the population studied is not free of all diseases but is free of the disease being studied at the beginning of follow-up (Fig. 7.3) [1, 3, 10].

There are two types of cohort that can be defined according to the characteristics of the population from which the cohort members are derived. The *closed cohort* is one with fixed membership. Groups are followed from a defined starting point to a defined ending point. Once follow-up begins, no one is added to the closed cohorts. The *open cohort*, or dynamic cohort, is one that can include new members as time passes; members come and go, and losses may occur.

The choice of the *exposed group* depends on the etiologic case, the exposure frequency, and the practical difficulties of the study, such as record availability or ease to follow up. The *nonexposed group* intends to provide us with information on the incidence of illness that would be expected in the exposed group if the exposure being studied did not influence the frequency of illness. Therefore, the nonexposed group is chosen in such a way as to be similar to the exposed group, in regard to the other risk factors of the illness being studied. It would be ideal for the exposure factor to constitute the only difference among the compared populations.

The exposure is divided into two types: common and infrequent (increased frequency in certain populations). If the factor being studied is common enough in the general population, then it is possible to select a sample from the general population and then to separate the individuals in teams exposed to a varied degree and/or to nonexposed in the factor that is being examined. These studies are defined as *general-population cohort studies*. Cohort studies that focus on people that share a particular uncommon exposure are defined as *special-exposure cohort studies*. The researcher identifies a specific cohort that has the exposure of interest and compares their outcome experience to that of a cohort of people without the exposure.

The cohort studies that are based on information on the exposure and the illness that has been collected from preexisting sources in the past are called *retrospective cohort studies* (Fig. 7.4). The authenticity, however, of such a study depends on the thoroughness of the certification of the illness in the files on the population and for

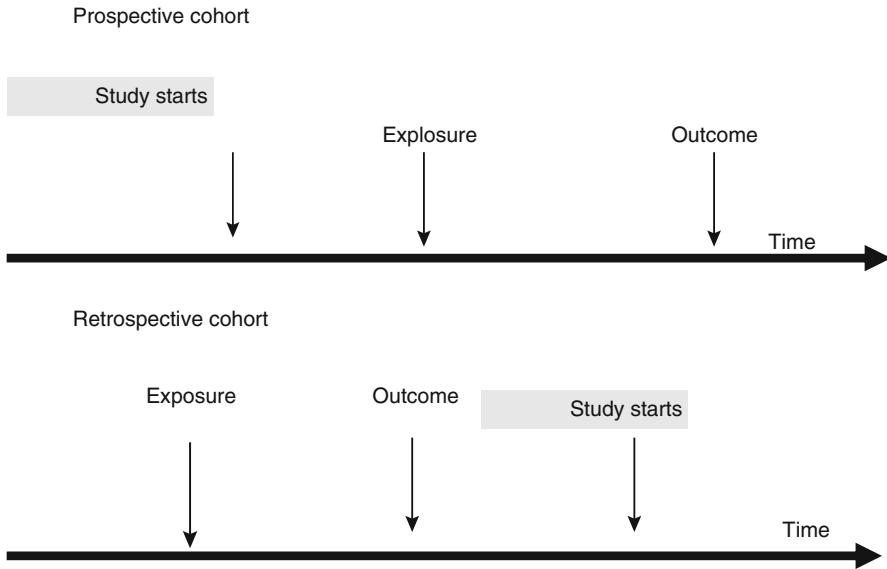


Fig. 7.4 Prospective and retrospective cohort study design

the time period being studied. Moreover, information on the relative confounding factors may not be available from such sources.

During the analysis of data, the frequency of illness is calculated (cumulative or dynamic incidence) in exposed and nonexposed individuals in accordance with the available data. With the use of a fourfold table, presentation and treatment of data take place. The relative risk or rate is calculated so as to determine the strength of the association and so that there exists the possibility of calculating the risk or rate difference.

Advantages and Disadvantages of Cohort Studies: Cohort studies are optimal for the investigation of rare exposures and can examine multiple effects of a single exposure. Since study subjects are disease-free at the time of exposure ascertainment, the temporal relationship between exposure and disease can be more easily elucidated. Cohort studies allow direct measurement of incidence rates or risks and their differences and ratios. However, prospective cohort studies can be expensive and time consuming, whereas retrospective cohort studies require the availability of adequate records. The validity of this design can be threatened by losses to follow-up.

7.5.2 Case-Control Studies

In prospective studies, a large number of individuals must be examined for their exposure conditions and must be followed up for a long time period so that a satisfactory number of outcomes are acquired. Such a study is not often practical or feasible.

Fig. 7.5 Design of a case-control study

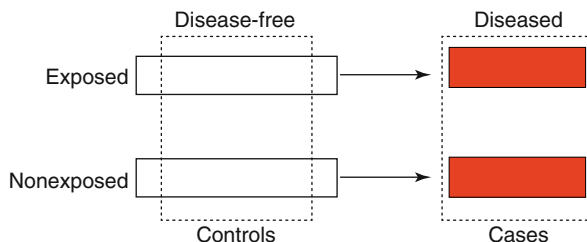


Table 7.6 Organization of case-control study data

Exposure	Disease			Total
	Yes	No		
Yes	a	b		?
No	c	d		?

This problem can be dealt with by using a study plan such as *case-control* (patient-witness). This design aims at achieving the same goal as the cohort study more efficiently using sampling. These studies have a characteristic methodology, during which a sample of the population being studied, featuring the common characteristic that they are cases, is used. Simultaneously, controls free from that particular disease are chosen as a representative sample of the population being studied. Ideally, the control group represents the exposure distribution in the source population that produced the cases. Then, exposure information is collected both for the patients and for the controls. Data are analyzed to determine whether exposure patterns are different between cases and controls (Fig. 7.5) [1, 3, 11].

The basic characteristics of case-control type studies are that the selection of individuals is made based on the criterion that they have or have not been infected by the illness being examined. There are many potential sources of *cases*, such as those derived from hospitals or clinics and those identified in disease registries or through screening programs. For causal research, incident disease cases rather than prevalent disease cases are preferable. Basic criteria in the choice of the group of cases are that they must constitute a relatively homogeneous group from an etiologic point of view and the facts about the illness must come from reliable sources.

Controls are a sample of the population that represent the cases and provide the background exposure expected in the case group. In many cases, the use of more than one witness group is necessary. Several sources are available for identifying controls including the general population (using random sampling), hospital- or clinic-based controls, and relatives and friends of cases. Persons with disease known or suspected to be related to the study exposure should be excluded from being used as controls. In Table 7.6, the organization of facts in a case-control study are described.

In these studies, it is not possible to calculate the frequency and effect indicators that we have known up to this point. This is owed to the fact that the size of the group of controls is arbitrary and is determined by the researcher. In these studies, we can use the odds ratio, which constitutes a very good estimate of the strength of association

of exposure-illness (relative risk) [12]. The odds ratio is defined as the ratio of the odds of being a case among the exposed (a/b) divided by the odds of being a case among the unexposed (c/d). Thus, the odds ratio is calculated as follows:

$$\text{OR} = \frac{a/b}{c/d} = ad/bc$$

The odds ratio is interpreted in the same way as the relative risk. An odds ratio of 1 indicates no association; an odds ratio greater than 1 indicates a positive association. Thus, exposure is positively related to outcome. If exposure is negatively related to the disease in a protective association, the odds ratio will be less than 1.

Advantages and Disadvantages of Case-control Studies: Case-control studies are cheaper and easier to conduct than cohort or experimental studies and are the method of choice for investigating rare diseases. In addition, case-control studies offer the opportunity to investigate multiple etiologic factors, simultaneously. However, case-control studies are not efficient designs for the evaluation of a rare exposure unless the study is very large or the exposure is common among those with the disease. In some situations, the temporal relationship between exposure and disease may be difficult to establish. In addition, incidence rates of disease in exposed and nonexposed individuals cannot be estimated in most instances. Case-control studies are very prone to selection and recall bias.

7.6 Experimental Studies

7.6.1 Intervention Studies

Intervention studies, commonly known as trials, are experimental investigations. They are follow-up studies where the researcher assigns the exposure study subjects. They differ from the nonexperimental studies in that the condition under which the study takes place is controlled.

Intervention studies are used for the evaluation of the effectiveness of preventive and therapeutic measures and services. In the first case, they are called preventive trials and, in the second case, therapeutic trials. Preventive trials are conducted among disease-free individuals, whereas therapeutic trials involve testing treatment modalities among diseased individuals. Additionally, two types of trials are determined: individual trials, in which treatment is allocated to individual persons, and community trials, in which treatment is allocated to an entire community. Epidemiologic studies of different treatments for patients who have some type of disease establish a broad subcategory, namely, *clinical trials*. The aim of clinical trials is to investigate a potential cure for disease or the prevention of a sequel [5, 13].

The individuals who participate in intervention studies come from a more general group, for which the results of the research should be in effect. This group is called a *reference population or a target population*. Once those who cannot

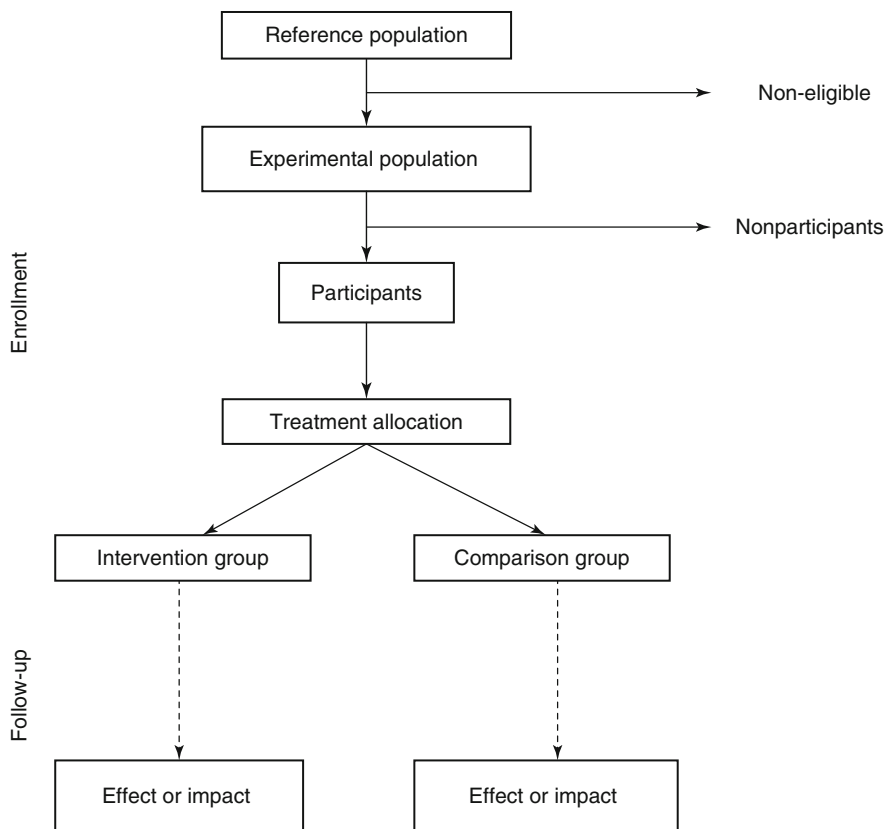


Fig. 7.6 Flowchart of an intervention study

participate in the study are excluded, those remaining, who are all likely candidates for the study, make up the *experimental population*. Once the potential group of subjects is determined, it is essential to get informed consent from the participants. Subjects unwilling to give consent should be removed. The eligible and willing subjects are then allocated into two main groups: (a) *basic-intervention/experiment group* and (b) *comparison group*. The study subjects are followed for a specified period of time under strict conditions, and the effects/outcomes are carefully documented and compared between the groups. Below, there is a flowchart of an intervention study (Fig. 7.6).

In the nonexperimental research (cohort and case-control), the known confounding factors can be monitored either as a choice of the compared groups or during the analysis of the data, but there is no sufficient ability to control the unknown confounding factors. On the contrary, in intervention research, it is possible to methodologically check both the known and the unknown confounding factors with the application of randomization. With *randomization*, each individual has the same

probability of receiving or not receiving the preventive or therapeutic measure being studied. Randomization of treatment allocation, if done properly, reduces the risk that unknown confounders will be seriously unbalanced in the study groups [14].

A particular problem in therapeutic intervention studies is *confounding by indication*, a bias that results from differences in prognosis between patients given different therapeutic schemes. Random treatment assignment ensures that prognostic factors are balanced between groups receiving the different treatments under investigation [15].

Techniques that are used for the achievement of randomization are simple, for example, flip of a coin, use of random number tables, or computer random number generator or complex, such as stratified randomization. But perhaps in these studies, a nonrandomized trial may not be possible and, consequently, the adjustment toward all the factors that become confounded. This can be corrected in part since we can eliminate differences in the phase of the analysis of the data. Thus, the way a nonrandomized intervention study is done and analyzed resembles that of a cohort study.

The functionalism of such a study can, to a large degree, be influenced by the fact that the participants and the researcher know the group the members of the study belong to. Knowledge of the treatment might influence the evaluation of the outcome. The solution, then, is *blinding* [3, 5]. It should be noted that blinding is desirable; however, it is not always feasible or necessary; for example, in orthodontic appliance research, because appliances differ in appearance, the researcher or even the participant could be aware of the intervention received.

There exist three types of *blinding*: (a) simple blinding: the evaluator assessing the outcome knows the assigned treatment, the participant does not; (b) double blinding: neither the evaluator nor the participant knows; and (c) triple blinding: neither the evaluator nor the participant knows, and the person who administers the treatment does not know which treatment is being assigned either.

Placebo treatment for the comparison group is often used to facilitate blinding. The placebo is inactive, morphologically similar with the tried therapeutic or preventive measure, medicine that is applied to the comparison group (when no other measure with documented effectiveness exists). When intervention studies involve procedures rather than pills, sham procedures take place to match the experience of the treatment and comparison groups as close as possible. The beneficial effect produced by an inactive pill or sham procedure is reported as the *placebo effect* and is attributed to the power of suggestion.

The results of the experimental studies can be organized in a fourfold table, and measures of disease frequency and association can be estimated. Although experimental study data analysis is straightforward, two issues should be kept in mind. Application of the *intention-to-treat principle* states that all randomized participants should be analyzed in order to preserve the goals of randomization. All subjects assigned to treatment should be analyzed regardless of whether they receive the treatment or complete the treatment. In addition, analysis of nonrandom subgroups threatens study validity and is not universally acceptable [3].

Ethical considerations are intrinsic to the design and conduct of intervention studies. No trial should be conducted without due consideration to ethical issues. These studies should be reviewed and approved by an ethics committee.

Advantages and Disadvantages of Intervention Studies: Well-designed intervention studies often provide the strongest support for a cause-effect relationship. Confounding factors that may have led to the subjects being exposed in cohort studies are not a problem here as researchers decide on who will be exposed. However, many research questions cannot be tested in trials, for example, if the exposure is fixed or if the outcome is rare. Intervention studies can also be more difficult to design than nonexperimental studies due to their unique problems of ethics and cost.

In this section, various design strategies have been considered for epidemiologic and clinical research. Table 7.7 provides key features and examples of different etiologic research methods in orthodontics. Any study is an effort to estimate an epidemiologic measure although this estimate could differ from the correct value. Steps that researchers can take to reduce errors are presented in the next section.

7.7 Sources of Error in Study Design

In order for an epidemiologic study to be considered credible, what is being measured, either frequency of appearance of an illness or the result of some report on the frequency of an illness, has to be authentic (accuracy in the measurement of a parameter). Accuracy is a general term denoting the absence of error. Authenticity depends on two factors: (a) precision and (b) validity [7, 8].

Precision is the repetitiveness of the result of a study, that is to say the degree of similarity between its results if it were to be repeated under similar conditions. Loss of precision is reported as a random error. *Validity* is the extent in which the study measures that which it alleges that it measures. Loss of validity is referred to as systematic error or bias. The terms validity and precision are often explained with the help of an objective (Fig. 7.7).

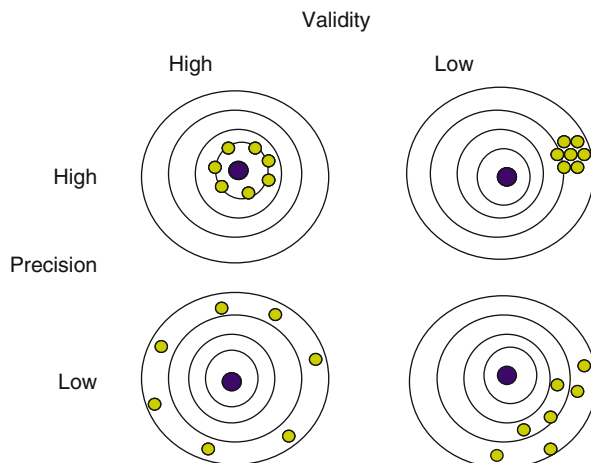
High validity corresponds to the average number of shots hitting near in the center. High precision corresponds to the shots grouping together in a small region.

7.8 Systematic Error or Bias

Bias can occur in all types of epidemiologic studies. Systematic error is usually the result of bad methodology, which leads to the creation of fictitious increases or reductions, differences or effects, the extent of which cannot be easily limited. Systematic error is unaffected by sample size. A study can be biased because of the way study subjects have been selected, the way measurements or classifications are conducted, or some confounding factor resulting from unfair comparisons. Thus, systematic error can be divided into three types: (a) selection bias, (b) information bias, and (c) confounding bias (Table 7.8) [5, 19, 20].

Table 7.7 Etiologic study designs and examples of questions addressed in orthodontic research

Study design	Key feature	Example of study question	Source
Intervention study	Evaluating the effectiveness of an intervention	To evaluate the effectiveness of preoperative ibuprofen in reducing the incidence and the severity of pain after orthodontic separator placement	Steen Law et al. [16]
	Randomization		
	Blinding		
Cohort study	Measuring incidence of disease or outcome	To evaluate the dental and psychological status of individuals who receive and do not receive orthodontic treatment	Kenealy et al. [17]
	Examining causes of disease or outcome		
Case-control study	Longitudinal	To evaluate whether the amount or the structure of mandibular bone affects post-orthodontic mandibular incisor relapse	Rothe et al. [18]
	Examining causes of disease or outcome		

Fig. 7.7 Validity and precision**Table 7.8** Facts about bias

An alternative explanation for an association
Can be of a small, moderate, or large amount
Can overestimate or underestimate the correct value (positive/negative bias)
Is avoided when the study is carefully designed and conducted

7.8.1 Selection Bias

Selection bias results from the processes that are used for the selection of the members of the study and to factors that influence participation in the study [3, 7]. It is caused when the association of exposure-illness differs among those who participate and those who do not participate in the study. This type of bias can more often be found in case-control studies or in cohort studies of retrospective character because exposure to the factor and the illness has already occurred by the time the study has begun.

Selection bias can occur in different ways such as differential surveillance, diagnosis, or referral of study participants according to exposure and disease status; differential unavailability due to illness or migration (selective survival), or refusal (nonresponse bias); and inappropriate control group selection (control selection bias).

Because the selection bias cannot be corrected, it should be avoided. This is possible with correct and careful planning of the study and its proper conduct.

7.8.2 Information Bias

Information bias is the result of the method in which information is collected which concerns the exposure as well as the illness of the individuals who participate in the

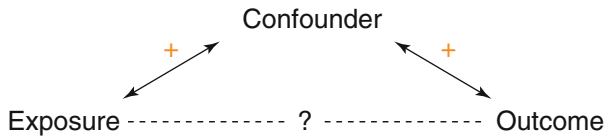


Fig. 7.8 The confounder is associated both with the exposure and outcome

study [7, 19]. It occurs when information that is collected for or by the individuals that participate in the research is erroneous.

Information bias leads to the placement of certain members of the study in the wrong category (misclassification) regarding the exposure, the illness, or both. Misclassification is divided into differential and non-differential. Differential is when misclassification regarding the exposure (or the illness) is different among those that are ill (or are exposed) and those that are not ill (or are not exposed). Non-differential is when misclassification regarding the exposure (illness) is independent from the appearance of the illness (exposure).

There are various types of information bias: recall bias, where the individuals that are ill refer to the exposure differently in relation to those that are not ill; interviewer bias, which is the fault of the researcher and is caused during the recording and interpretation of information of the exposure and illness; and follow-up bias, which results from the fact that those individuals who are not monitored for the duration of the study differ from those that remain until the end of the study.

Information bias can be avoided by carefully designing the study questionnaire, training interviewers, obtaining accurate exposure and disease data, and employing methods to successfully trace study subjects.

7.8.3 Confounding

Confounding is a result of the fact that the relation between the exposure and the illness is influenced by other factors so that it is led to confounding due to a mixture of these effects [2, 3, 20]. As a result, it constitutes an alternative explanation of the relation between exposure-illness, which is different from the actual explanation. The result of the exposure, therefore, is different from what would have resulted if no confounding factors existed, and only the factor being examined was influencing the individual (exposure). That is to say, the relation of the exposure and the illness is disturbed because it is mixed with the effect of some other factor which relates to the illness and exposure being examined. Thus, a change of the real picture is brought on, moving in the direction of either the undervaluation or the overestimation of the situation.

A confounding variable has three features: it is an independent cause or predictor of the disease, it is associated with the exposure, and it is not an effect of the exposure (Fig. 7.8).

There are three methods to prevent confounding in the design phase of a study: randomization, restriction, and matching. Random assignment of study participants

Table 7.9 Commonly used statistical tests for group comparison

Type of data	Number of groups	Independent	Paired
Continuous-normal	2	Student <i>t</i> -test	Paired <i>t</i> -test
	>2	Analysis of variance	repeated measures analysis of variance
Continuous-non-normal or ordinal	2	Mann-Whitney <i>U</i> test	Wilcoxon signed-rank test
	>2	Kruskal-Wallis test	Friedman test
Nominal	2	Chi-square test	McNemar test
	>2	Chi-square test	Repeated measures logistic regression
Survival	≥2	Long rank test	Conditional logistic regression

to experimental groups controls both known and unknown confounders, but can only be used in intervention studies. Restriction involves selection of study subjects that have the same value of a variable that could be a confounder. Restriction cannot control for unknown risk factors. Matching involves constraining the unexposed (cohort studies) or control group (case-control studies) such that the confounders' distribution within these groups is similar to the corresponding distribution of the index group. Once data has been collected, there are two options for confounding control: stratified analysis and multivariate regression models.

7.9 Random Error

The error that remains after systematic error elimination is defined as *random error*. Random error is most easily conceptualized as sampling variability and can be reduced by increasing sample size. *Precision* is the opposite of random error and is a desirable attribute of measurement and estimation. *Hypothesis testing* is commonly used to assess the role of random error and to make statistical inferences [3, 5, 21].

Statistical hypothesis testing focuses on the null and alternative hypotheses. The null hypothesis is the formulation of a non-relation or non-difference among variables that are being investigated (exposure-illness). That is to say, the two compared groups do not differ between themselves as per the size being examined (more than random sampling allows). The control of the null hypothesis takes place with the *statistical test* [21].

Commonly used tests include the student *t*-test and chi-square test depending on the nature of the data under investigation (Table 7.9). The test statistic provides a *p*-value, which expresses the level of statistical importance. The *p*-value is the probability of a result, like the one being observed or a larger one, to be found by chance when the null hypothesis is in effect (it shows the probability of the result that is being observed to occur, if the null situation is real). The researcher wishes for his/her results not be explained by random error, and consequently, he/she wishes the smallest possible *p*-value for the results. If the *p*-value is smaller than 5 % (level of

Table 7.10 Relationship between confidence interval and p -value for measures of effect

Measure of effect	No effect value	Confidence interval	p -value
Risk or rate difference	0	95 % CI includes 0	$p > 0.05$, nonsignificant
Risk or rate difference	0	95 % CI does not include 0	$p < 0.05$, significant
Relative risk, relative ratio or odds ratio	1	95 % CI includes 1	$p > 0.05$, nonsignificant
Relative risk, relative ratio or odds ratio	1	95 % CI does not include 1	$p < 0.05$, significant

statistical importance more often used), then the null situation has the probability of being less than 5 %, and it is rejected. When the null situation is not guaranteed, the alternative situation is adopted according to which the two compared groups differ between them, as to the size being tested (more than a chance sampling allows). The p -value depends on the strength of the association and the size of the sample. As a result, we may have a large sample in which, even a slight increase/reduction in risk, may seem statistically important or a small sample where large increases/reductions do not achieve statistical importance.

Many researchers prefer to use *confidence intervals* to quantify random error [2]. The confidence interval is calculated around a point estimate and quantifies the variability around the point estimate. The narrower the confidence interval, the more precise the estimate. The confidence interval is defined as follows: it is the breadth of values in which the real extent of the effect is found with given probability or specific degree of certainty (usually 95 %). That is to say, it is the breadth of values in which the real value is found with precise certainty. The confidence interval is calculated with the same equations that are used to calculate the p -value and may also be used to determine if results are statistically significant (Table 7.10). For example, if the interval does not include the null value, the results are considered statistically significant. However, the confidence interval conveys more information than the p -value. It provides the magnitude of the effect as well as the variability around the estimate, whereas the p -value provides the extent to which the null hypothesis is compatible with the data and nothing about the magnitude of effect and its variability.

As mentioned earlier, the primary way to increase precision is to enlarge the study size. *Sample size calculations* based on conventional statistical formulas are often used when a study is being planned [5, 21]. These formulas relate study size to the study design, study population, and desired power or precision. However, these formulas fail to account for the value of the information gained from the study, the balance between precision and cost, and many social, political, and biological factors that are almost never quantified. Thus, study size decision in the design phase can be aided by formula use; however, this determination should take into account unquantified practical constraints and implications of various study sizes. Calculation of study size after study completion is controversial and discouraged.

Deductively, there are two broad types of error affecting epidemiologic studies: systematic and random error. In designing a study, an effort should be made to reduce both types of error.

Conclusion

Understanding disease frequency and causal effects measures is a prerequisite to conducting epidemiologic studies. Strategy designs include experimental and nonexperimental studies, whereas causal effects can only be evaluated in studies with comparison groups. All research designs are susceptible to invalid conclusions due to systematic errors. Study outcomes should preferably be reported with confidence intervals.

Disease Occurrence

- Incidence and prevalence are measures of disease frequency.
- Prevalence provides the proportion of a population that has a disease at a particular point in time.
- Incidence measures the transition from health to disease status.
- Cumulative incidence provides the proportion of the population that becomes diseased over a period of time.
- Incidence rate provides the occurrence of new cases of the disease during person-time of observation.
- Measures of disease frequency are compared relatively or absolutely.
- Absolute measures of effect are based in the difference between measures and include the rate or risk difference.
- Relative measures of effect are based on the ratio of two measures and include the rate or risk ratio.
- Comparing measures of disease occurrence is facilitated by organizing data into a fourfold table.

Study Design

- Epidemiologic studies are divided into two categories: (1) descriptive studies and (2) etiologic or analytic studies.
- Descriptive studies consist of two major groups: those involving individuals (cross-sectional studies) and those that deal with populations (ecological studies).
- Etiologic studies are distinguished in experimental (intervention studies) and in nonexperimental (observational studies).
- Nonexperimental studies are distinguished in cohort and in case-control studies.
- In a cohort study, subjects are defined according to exposure status and are followed for disease occurrence.
- In a case-control study, cases of disease and controls are defined, and their exposure history is assessed and compared.
- Experimental studies are follow-up investigations where the researcher assigns exposure to study subjects.

- Well-designed experimental studies (trials) provide the strongest support for a cause-effect relationship.

Bias

- Bias or systematic error can occur in all types of epidemiologic studies and results in an incorrect estimate of the measure of effect.
- Bias is divided into three types: (1) selection bias (2) information bias and (3) confounding bias.
- Selection bias results from systematic differences in selecting the groups of study.
- Information bias results from systematic differences in the way that exposure and disease information are collected from groups of study.
- Confounding is a result of the fact that the relation between the exposure and illness is influenced by other factors so that it is led to confounding due to a mixture of these effects.

Random error

- Random error is the error that remains after systematic error elimination.
- Random error is most easily conceptualized as sampling variability.
- Increasing sample size reduces random error.
- Hypothesis testing is used to assess the role of random error and to make statistical inferences.
- Many researchers prefer to use confidence intervals to quantify random error.

References

1. McMahon B, Trichopoulos D (1996) *Epidemiology: principles and methods*, 2nd edn. Little, Brown and Co, Boston
2. Rothman KJ (2002) *Epidemiology: an introduction*. Oxford University Press, New York
3. Aschengrau A, Seage GR III (2003) *Essentials of epidemiology in public health*. Jones and Barlett Publishers, Sudbury
4. Elandt-Johnson RC (1975) Definition of rates: some remarks on their use and misuse. *Am J Epidemiol* 102:267–271
5. Rothman KJ, Greenland S (1998) *Modern epidemiology*, 2nd edn. Lippincott-Raven, Philadelphia
6. Miettinen OS (1974) Proportion of disease caused or prevented by a given exposure, trait or intervention. *Am J Epidemiol* 99:325–332
7. Hennekens C, Buring J (1987) *Epidemiology in medicine*, 1st edn. Little, Brown and Co, Boston
8. Last JM (2001) *A dictionary of epidemiology*, 4th edn. Oxford University Press, New York
9. Antczak-Bouckoms AA (1998) The anatomy of clinical research. *Clin Orthod Res* 1:75–79
10. Grimes DA, Schulz KF (2002) Cohort studies: marching towards outcomes. *Lancet* 359:341–345
11. Schulz KF, Grimes DA (2002) Case-control studies: research in reverse. *Lancet* 359:431–434
12. Bland JB, Altman DG (2000) The odds ratio. *Br Med J* 320:1468
13. Peto R, Pike MC, Armitage P et al (1976) Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 34:585–612
14. Gore SM (1981) Assessing clinical trials: why randomise? *Br Med J* 282:1958–1960
15. Salas M, Hofman A, Stricker BH (1999) Confounding by indication: an example of variation in the use of epidemiologic terminology. *Am J Epidemiol* 149:981–983

16. Steen Law SL, Southard KA, Law AS, Logan HL, Jakobsen JR (2000) An evaluation of preoperative ibuprofen for treatment of pain associated with orthodontic separator placement. *Am J Orthod Dentofacial Orthop* 118:629–635
17. Kenealy PM, Kingdon A, Richmond S, Shaw WC (2007) The Cardiff dental study: a 20-year critical evaluation of the psychological health gain from orthodontic treatment. *Br J Health Psychol* 12:17–49
18. Rothe LE, Bollen AM, Little RM, Herring SW, Chaison JB, Chen CS, Hollender LG (2006) Trabecular and cortical bone as risk factors for orthodontic relapse. *Am J Orthod Dentofacial Orthop* 130:476–484
19. Greenland S (1977) Response and follow-up bias in cohort studies. *Am J Epidemiol* 106:184–187
20. Greenland S, Robins JM (1985) Confounding and misclassification. *Am J Epidemiol* 122:495–506
21. Rosner B (2000) *Fundamentals of biostatistics*, 5th edn. Duxbury Thompson Learning, Pacific Grove