Chapter 11
Methods and Concepts of Epidemiology

Rafael Mikolajczyk

11.1 Introduction

The purpose of this chapter is to review the basic concepts of epidemiology, including definitions of measures of disease occurrence and measures of association, brief descriptions of study designs and ethical principles of epidemiological research. Additionally, the theory and criteria of causation, systematic and random errors in epidemiological studies and methodological issues related to diagnostic tests are discussed. The concepts are outlined and some examples are given.

11.2 Definitions of Epidemiological Terms

11.2.1 Measures of Disease Occurrence

Terms used to quantify the occurrence of a disease in a population are listed in Table 11.1.

Table 11.1 Measures of disease occurrence

Measure	Description
Prevalence	Number or proportion of persons with a specific disease at a specific time point in the population
Incidence	Number or proportion of persons developing a specific disease during a time period
Morbidity	Ambiguously used: prevalence or incidence
Mortality	Number or proportion of persons dying during a time period
Fatality rate	Proportion of persons dying from a specific disease among all persons with the disease
Attack rate	Proportion of cases developing the disease among all persons who were exposed to the disease

R. Mikolajczyk (✉)
Department of Public Health Medicine, School of Public Health, University of Bielefeld, Germany
e-mail: rafael.mikolajczyk@uni-bielefeld.de

A. Krämer et al. (eds.), *Modern Infectious Disease Epidemiology*,
Statistics for Biology and Health, DOI 10.1007/978-0-387-93835-6_11,

For frequent diseases, *prevalence* can be provided in percentages; for rare diseases, prevalence can be presented as the number of cases per 1,000, 100,000, or even 1 million within the population. For example, 5.6 persons with chronic hepatitis C per 1,000 persons of the population on January 1, 2006. The prevalence can be age or gender specific; in such cases, additional information such as "per 1,000 women aged 20–30" should be included. Prevalence can be defined at a given time point (*point prevalence)*, for example, when all participants in a telephone survey are asked about having a specific disease. *Period prevalence* is defined as the cumulative number of cases observed during a given time period.

In contrast, *incidence* is the number of new cases with a specific illness occurring during a given time period (for example, 10 new HIV infections per 10,000 inhabitants during the year 2005). The time period can be flexibly specified as a daily, weekly, monthly, or yearly incidence, depending on how much variability is present in the number of cases observed over time. For example, providing only yearly incidence for a disease with a strong seasonality like influenza will not capture the information about the temporally overcrowded doctor offices during influenza season. While prevalence is important for assessing the need of treatment in the population, incidence provides information about the dynamics of the spread of infection and is the most important measure during outbreaks or epidemics. *Mortality* is the incidence of deaths.

The term *morbidity* is used inconsistently: most often to indicate prevalence but sometimes also to indicate incidence.

Case fatality describes how many of infected persons die because of the disease. It does not provide any information about the duration of the disease before death, so in chronic infections the deaths can spread over long period of time, while in the case of acute infections like influenza or SARS they are concentrated in short period of time – despite possibly the same case fatality.

Attack rate is the fraction of persons acquiring an infection among all persons who were exposed to the infection. The attack rate depends on the composition of the population, especially with regard to immunity.

11.2.2 Measures of Association

Measures of association commonly used in epidemiology to assess the relationship between exposure and outcome are presented in Table 11.2 and Fig. 11.1.

Table 11.2 Measures of association

Measure	Description
Absolute risk	Probability of a specific outcome
Relative risk (risk ratio)	Ratio of absolute risks in two distinct, mutually exclusive subgroups with different exposure status
Absolute odds	Chance of a specific outcome (if 1 person gets the disease and 9 do not, then the odds is 1:9)
Odds ratio (relative odds)	Ratio of two absolute odds in two distinct, mutually exclusive subgroups with different exposure status

Exposure / Outcome	1	2	Sum
Yes	a	b	All sick cases
No	c	d	All healthy cases
Sum	All cases with exposure 1	All cases with exposure 2	All tests / all cases

Outcome can be a disease or death, exposure can be any risk behaviour or any characteristic of the person (like gender or age group). The letters "a" to "d" indicate numbers of cases in each cell – for example "a" - the number of persons with the disease and exposure status 1. The information to fill the table for a concrete research question can be obtained from routine statistics or from epidemiological studies.

Absolute risk for persons with exposure 1: a/(a+c)

Relative risk for exposure 1 compared to exposure 2: $\frac{a/(a+c)}{b/(b+d)}$

Odds ratio for exposure 1 compared to exposure 2: $\frac{a/c}{b/d} = \frac{a \times d}{b \times c}$

Fig. 11.1 Figure for illustration of relative risk and odds ratio

The *absolute risk* of a disease can be different among persons with a specific exposure (risk factor) than among persons without this exposure, for example, the risk of acquiring HIV can be different for males and females. *Relative risk* provides information regarding how many times higher or lower the risk is among individuals with one exposure status compared to individuals with another exposure status. For example, in a given population the risk of having HIV infection can be 3.2 times higher for males than females. In the simplest scenario, the *relative risk* can be calculated from a two-by-two table displayed in Fig. 11.1.

While the two-by-two table can be constructed whenever there is a dichotomous outcome and exposure, the measures of absolute and relative risks are not always meaningful. For example, when the study population consists of persons who were selected because they have a given disease and compared to another group who does not have the disease (see case–control study below), absolute risk of having the disease cannot be obtained from these data. In such a study, the relative risk as a ratio of two absolute risks cannot be calculated either. However, even if the absolute risk is not known, the two-by-two table contains useful information. The information can be used to calculate *absolute odds* and *odds ratio*. Both measures are calculated from the inner cells of the table (see Fig. 11.1). *Odds* is the chance of disease, i.e. the ratio of sick to healthy persons – for example, if in a population there are 100 HIV-infected individuals and 900 noninfected individuals, the odds of being infected is 1/9. Risk of being infected is 100/(100+900)=1/10, which is close to the odds, but not identical. (The lower is the prevalence of the disease in the population, the closer are odds and risks to each other.)

Both these measures – relative risk and odds ratio – apply to a disease status at one given time point. Two further measures of association are absolute hazard and corresponding hazard ratio, which measure the occurrence of disease over time. If a

given exposure not only causes a disease but also accelerates its development, then these findings would be better conveyed by hazard ratio than risk ratio or odds ratio.

A different concept – not directly used for analysis but for interpretation of the relative risk – is the *population attributable risk*.

> *Population attributable risk* (PAR) is the fraction of cases with a disease which can be avoided if the risk factor is totally removed from the population.

PAR can be calculated according to the formula PAR = PE(RR–1)/(1+PE (RR–1)), where PE is the prevalence of the risk factor and RR is the relative risk associated with the risk factor (Greenland and Robins 1988). The measure is very appealing for presenting the public health impact of a specific risk behaviour or intervention, for example, PAR of 30% for intravenous drug use for hepatitis B means that when this behaviour can be removed as a risk factor, 30% of cases of hepatitis B infection can be prevented. Similarly, increasing vaccination coverage would lead to a decrease in the incidence of the specific disease and this can be reported as PAR (see Chapters 12 and 14). While the calculation of PAR for non-communicable diseases follows the above formula, additional considerations are required for communicable infectious diseases. For communicable diseases, the removal of a given risk factor often results not only in directly avoided infected cases (primary cases) but also in a reduction in the risk of infection for other persons in the population (secondary cases) who would be otherwise infected by primary cases. These effects may not be proportional, and removing a specific risk factor can result in smaller or larger effect on the incidence than in the above formula.

11.3 Populations, Study Samples and Random Error

All the above definitions are based on the assumption that there is a "true" value of the incidence, the prevalence and the association between risk and outcome in a given population. Population might be defined widely as all inhabitants of a country or more specifically as a group of individuals with special characteristics, for example, intravenous drug users. In some simple cases we might know the incidence of cases in the whole population based on registries, like, for example, the death certificate registry; however, in the case of most diseases and risk factors, it is not feasible to assess them in the whole population. In such cases, for the purpose of specific study, smaller groups might be sampled from the population and assessed more in depth. Since subjects of the study are only a subgroup of the population, by chance the study might have recruited more individuals with a given disease than the prevalence in the population. This is based on the same principle as obtaining four times heads in a row while flipping the coin although the "true" proportion is 50%. This is called *sampling error*, a subtype of *random error* encountered in epidemiological studies. Another potential source of random error is non-sampling error, which can be introduced by the uncertainty in measurement. Non-sampling error would

lead to different results when the study is repeated in the same sample; sampling error would provide different results if a new study sample is recruited from the population. Historically, the first solution to deal with sampling error was testing of significance, with results from the test reported as *p* values. More recently, there has been a trend towards reporting the confidence intervals for the results (Armstrong 1998). While both results are often reported together, confidence intervals provide more information about the studied relationship. Confidence intervals not including a "zero effect" (which is 1 for odds ratio or relative risk) can have the same interpretation as a significance test; they allow the rejection of the tested hypothesis. But additionally, a narrow confidence interval demonstrates that the study is big enough to provide precise information and therefore the size of the observed effect is relatively certain. Modern statistical analysis provided solutions to many specific problems and while sampling error is not avoidable, it can be usually quantified. Other types of error will be discussed later.

11.4 Common Types of Epidemiological Studies

The purpose of epidemiological studies is to provide estimates for measures of disease occurrence in the population and information about the relationship between risk factors and disease using specifically selected samples of the population. There are three main study designs: survey, case–control study and cohort study (including randomized clinical trial) (Table 11.3) (von Elm et al. 2007).

Table 11.3 Selected types of epidemiological studies

Type	Description
Survey	Collection of data at a single time point in a sample of the population
Case–control study	Selection of the sample revolves around the disease status, i.e. participants are included if they have a disease (cases), while a comparison group without this disease is selected (controls)
Cohort study	Collection of data performed over time in the same defined group of participants with different exposure status; information about different time points is either obtained from existing records or reported by participants at different occasions
Randomized clinical trial	Persons meeting specified inclusion criteria are enrolled in the study and randomized into subgroups receiving different treatments
Meta-analysis	Joint analysis of previous studies based typically on published results only, but sometimes also on original data (individual data meta-analysis)

A *survey*: a study in which information concerning each participant is collected only once and the participants are selected to reflect the studied population. It can be used to estimate the prevalence of disease and of risk factors. The estimate of incidence can be obtained only based on retrospective information about when the disease started. Similarly, information about changes in the risk behaviour can be obtained only retrospectively. Therefore, for the current risk behaviour and disease

status, it is not clear whether disease preceded the risk behaviour or not (Hernan and Robins 2006; Martin 2008). Some persons might adapt a healthy lifestyle after developing the disease. Although retrospective information about behaviour before the onset of a disease can be obtained in a survey, its accuracy depends on the memory of the participant and the quality of reporting which can be subject to reporting bias (Hartman et al. 2002). Surveys can be conducted through personal interviews, mailed questionnaires or on the phone. In particular, the two latter forms are relatively cheap and easy to implement, making the survey design a very attractive research method when information needs to be quickly obtained and should be representative for the population. In some cases, the researcher might not be interested in the whole population, but in a specific subgroup like intravenous drug users or prostitutes, he or she would still attempt to obtain a good representation of this subgroup in the study sample. On the other hand, survey design can also include the collection of biological samples. In the epidemiology of infectious diseases, biological samples are collected in serological surveys assessing for HIV status (Montana et al. 2008) or vaccination against childhood diseases in the population (Nardone et al. 2003). Surveys can be used to link seroprevalence of antibodies and information about risk behaviours, but again the possibility has to be considered that the current behaviours might not be related to the disease at all or that the behaviours might have been influenced by knowledge of the disease status.

Case–control studies can provide neither prevalence or incidence estimates nor information about absolute or relative risk. However, case–control studies allow the comparison of exposure status in persons with a disease (cases) and healthy participants (controls) using odds ratios. The case–control design is especially advantageous when risk factors for a very rare disease have to be investigated. Instead of recruiting a very large population for a longitudinal study with the expectation that some cases of the rare disease will occur over time in the study population, for the case–control study, persons who developed the disease are selected. For example, rare complications of vaccinations can be studied in case–control studies. While the identification of cases may sometimes pose diagnostic difficulties, the main challenge in a case–control study is the selection of controls. The aim in selecting controls is choosing individuals representative of the population from which the cases originated (Wacholder et al. 1992a; Wacholder et al. 1992b). Many approaches have been proposed to the selection of controls: hospital controls, controls having a different disease with similar symptoms, best friend controls (Lopes et al. 1996). All these approaches can be seen as convenience samples. A better approach, but more demanding and expensive, is to use population-based controls. In order to make the cases and controls more similar to each other, for example, in terms of gender or age, the researcher might perform *matching* – either as frequency matching (the number of controls in the same age group matched to the number of controls) or as individual-based matching (one or more controls are matched to each individual case for all requested characteristics) (Wacholder et al. 1992c).

Cohort study: a distinction should be made between an open cohort which is defined by a specific framework (for example, all inhabitants of a specific town) and allows new entries (younger inhabitants who reach the age for inclusion in the study

or people who move into the studied area) and a more traditional closed cohort for which a defined group of participants is selected and only loss to follow-up but no new entries are allowed (Philippe 2001). Cohort studies can provide information about the incidence occurring over time as well as the prevalence of the disease at a given time point. To address the problem of changing cohort size over time (due to deaths or loss to follow-up), incidence can be measured as *incidence density*, where the sum of the times each participant contributed to the study during the specific period is in the denominator. Apart from the above distinction, the cohort study itself can be either retrospective or prospective. A prospective cohort study is by far the most demanding and the most expensive study design. Retrospective cohort studies have the advantage of using existing records for a defined population. These studies are often used to identify occupational hazards due to biochemical exposures. However, while more convenient, retrospective studies are restricted to routinely collected information, which may not include many questions of interest. Through the longitudinal character, cohort studies can provide clear information about the sequence of exposure and disease. A drawback is that only very common diseases can be studied with moderate cohort size and the less frequently a disease occurs, the larger the cohort has to be. Another practical problem can be the very long duration of a cohort study when development of the disease occurs over a long process. In infectious disease epidemiology, cohort studies are commonly used to assess the transmission or the progression of the disease. Prospective cohort studies were used to study stages of AIDS (Badri et al. 2006) or seroconversions among HIV serodiscordant couples (Bunnell et al. 2006).

Randomized clinical trials (RCTs) recruit participants who meet predefined inclusion and exclusion criteria and then randomly allocate them into groups which receive different interventions (Moher et al. 2001). At minimum, there are two groups of participants, one in which the intervention is conducted and one placebo group, but RCTs can also test several interventions against each other or against standard therapy or placebo treatment. The control group is necessary because of the subjectivity of human perception which might result in a mistaken impression that the treatment is more effective than in the reality. Two other mechanisms to ensure objectivity of RCTs are randomization and blinding. Randomization should ensure that the studied groups are comparable and the selection of patients is not influenced by the expectations of the researcher (Kang et al. 2008). Different techniques can be used for randomization and it was shown that less rigorous randomization, possibly including incomplete concealment of the randomization sequence, can affect results of the study (Schulz et al. 1994; Schulz et al. 1995; Schulz and Grimes 2002a). To further improve the objectivity, RCTs apply blinding, which means, at minimum, that patients do not know whether they receive active medication or placebo, or that neither patients nor their doctors (double blinding) know the group of the patient (Schulz and Grimes 2002b). Sometimes, triple blinding is also used, which means that during the analysis of the data, anonymous codes are used and only after the results are obtained the meaning of the codes is revealed. While there is a strong agreement about the necessity of blinding in RCTs, the blinding is not

always correctly performed and reported in publications (Haahr and Hrobjartsson 2006).

In the control group, placebo therapy is used, but some problems arise if the application of placebo has its own effects (Hrobjartsson and Gotzsche 2001). Another difficulty is when the intervention involves manipulations which are obvious to the patient. In such case, application of placebo infusions, needle insertion outside of areas used in standard acupuncture or even surgical scars are sometimes used to blind the patient (called *sham* procedures (Sutherland 2007)). While testing against placebo is usually the most desirable way to obtain information about the true effectiveness of the treatment, this option is unethical when effective treatment for the studied question already exists because a treatment which is considered to be effective cannot be withheld from the patient. In some cases, when the new medication is not expected to be more effective, but, for example, have fewer side effects or simpler administration, a new medication can be tested in *non-inferiority trials* in comparison with the standard medication (Piaggio et al. 2006). Non-inferiority trials require substantially larger sample sizes than trials oriented towards demonstrating a difference in treatments, but are increasingly common in modern pharmaco-epidemiology.

A subtype of randomized clinical trials is the *cross-over trial*, which is especially oriented towards the assessment of physiological processes. In these trials, participants are randomized for a given sequence of intervention or placebo (Putt and Ravina 2002). In the simplest case, a participant, after being in the placebo group, switches to the intervention group and vice versa. In more complex studies, several different medications can be applied sequentially. The advantage of cross-over trials is that patient's values under treatment can be compared with his/her own values under placebo; however, their use is limited if the patient can be cured or his condition permanently changes.

While originally designed for simple one-step interventions, clinical trials can also be used for the assessment of *complex interventions* (Campbell et al. 2000). Complex interventions are defined as a specified sequence of treatments which might be applied in different combinations to individual patients. Studies assessing the effectiveness of complex interventions usually require a large sample size.

All the above types of trials are person based in which the unit of evaluation is a single person. A specifically epidemiological type of an intervention trial is a *community-based intervention* (Doyle et al. 2008). In this type of studies, groups of participants rather than individual patients are randomized; these groups can be based on geographic or other criteria (for example, minorities can be studied in specific settings which do not follow a geographic distribution). When the intervention is applied, the results are evaluated as rates on the level of randomization units.

Another type of epidemiological studies is systematic review, which is a joint evaluation of several, possibly all existing studies addressing a specific question. It is called *meta-analysis* when statistical methods are also used to obtain a combined estimate from all included studies. Meta-analysis can be based either on published

estimates or on originally collected data from several studies (more on systematic reviews and meta-analyses in Blettner et al. 1999).

Some additional study designs are also used in specific fields. For example, *nested case–control studies* can be used to address more specific questions based on cases and controls recruited from an existing cohort study (Ernster 1994; Cologne et al. 2004). To study the effects of transient exposures (stress events, jogging in the morning or vaccination) as triggers of certain diseases, *case cross-over design* was proposed (Maclure 1991; Maclure and Mittleman 2000; Park et al. 2004; Reintjes et al. 2005). In this study the exposure of the participant at some earlier time point is compared with the exposure during the time window in which the exposure could potentially trigger the disease. A methodological innovation in recruiting survey participants from the so-called hidden populations [for example, populations with illegal (like intravenous drug users) or socially stigmatized behaviours (prostitutes, homosexual men)] is *respondent-driven sampling* (Heckathorn 1997; Heckathorn 2002). This method is based on the idea of "snowball" sampling: the participant is asked to recruit further persons from the same population, but advanced methodologies are used to convert the derived estimates into representative estimates for the whole studied populations (Heckathorn 2007). Developmental psychologists invented the *accelerated longitudinal design* also called cohort sequential or mixed longitudinal design (Prinzie and Onghena 2005). This design uses multiple cohorts which are overlapping with respect to age at the end of the study. The information from the separate cohorts is combined to evaluate the trajectory over the whole age spectrum of the participants, which can be several times larger than the duration of the observation period (Miyazaki and Raudenbush 2000).

11.5 Ethic in Epidemiological Research

The most crucial ethical requirement for clinical or epidemiological studies is the *informed consent*, which assures that study participants are informed about possible consequences of the study, understand them and agree to participate in the study (World Medical Association 1964). Historically, the debate about ethical aspects of research and informed consent was triggered by a study of syphilis infection in the African-American population in Alabama from 1932 to 1972 (Jones 1993). In order to assess the natural course of the syphilis infection in untreated patients, no treatment was provided to the participants, even after an effective treatment was developed and introduced in the general population. Over the course of the study, 28 participants died directly of syphilis, 100 died of complications, 40 infected their wives and 19 of their children had been born with congenital syphilis. While this drastic example is surely an issue of the past, epidemiological studies bear the potential of abuse, especially when conducted in disadvantaged populations. Recent reviews of informed consent practices and whether patients are able to understand them indicate a further need of improvement (Flory and Emanuel 2004; Sugarman et al. 2005; Hill et al. 2008).

To assist the researchers in the planning of studies, *ethical approval* by *institutional review boards* is provided. This practice is standard in many developed countries, but in some developing countries these control mechanisms might be

deficient (Bhutta 2002; Creed-Kanashiro et al. 2005; Marshall 2005; Sumathipala et al. 2008).

11.6 Causality in Epidemiological Research

Seminal work on the criteria of causality for associations observed in epidemiological studies was conducted by a British medical statistician Austin Bradford Hill (1897–1991). He proposed a list of nine criteria, of which only one was logically necessary but none was sufficient. More recently, a theoretical work on the scientific foundations of epidemiology, including the understanding of causation, was requested (Susser and Susser 1996a; Krieger, Zierler 1997). In a recent review, Parascandola and Weed (Parascandola and Weed 2001) identified five definitions of causation used by epidemiologists:

- *Production*: a given variable is understood as a cause when it creates or produces a specific outcome; this definition was repeatedly criticized because it applies an equally unclear concept of production to define the causation.
- *Necessary cause*: only the logically necessary conditions are understood as causes. This definition is strongly influenced by the notion of scientific determinism. If in some cases the expected outcome does not occur in the exposed participant, this has to be attributed to other, not yet known deterministic causes.
- *Sufficient-component cause*, in which several variables play a role as component causes and jointly form a sufficient cause. This definition was proposed by Rothman and Saunders (Rothman and Greenland 2002) and follows older philosophical traditions. Parascandola and Weed (Parascandola and Weed 2001) argue that this definition is also based on the concept of scientific determinism, since for a specific patient it implies that he/she either has the sufficient cause or not.
- *Probabilistic causation* is opposed to the deterministic understanding of science. According to this definition, cause is a variable or a characteristic which increases the risk of disease. This definition can be used to accommodate the concepts of sufficient cause (which raises the probability of outcome to 1) and necessary cause (which raises the probability of outcome above 0) (Parascandola and Weed 2001).
- *Counterfactuals*: the definition is based on the idea of comparing the outcomes under two contrasting scenarios, one including a specific effect and one excluding it. The definition is not directly opposed to the concept of necessary and sufficient cause and can also be combined with both the scientific determinism and the probabilistic causation. It is also not opposed by the other definitions but rather adds a specific aspect of understanding the causal relationship.

While Parascandola and Weed (Parascandola and Weed 2001) argue in favour of the probabilistic causation, they also stress that distinction should be made

between defining causation and recognizing a cause. For the latter the counterfactual definition can be very helpful.

Another issue which received recent attention is the understanding of proximate and distal or direct and indirect causes (Susser and Susser 1996a). Susser and Susser (Susser and Susser 1996b) proposed that after the era of risk behaviours assessed in the individual patient, social structures on the one side and genetic epidemiology on the other side will receive more attention in the 21st century. The effects of social structures were studied in social epidemiology but now they will be combined with molecular knowledge in multilevel models. The different levels of these models are also associated with a different understanding of causation (Parascandola and Weed 2001; Susser 2001).

11.7 Systematic Error in Epidemiological Studies

While the issue of causation is related to the more general understanding of epidemiological research, at the technical level the issue of a correct measurement of associations in epidemiological studied receives a lot of attention.

> *Systematic error (Bias)* is a phenomenon which causes a distortion of the result – the measured effects are either too small or too large.

Bias can occur in the study at different stages: during planning, conducting the study, analysing the data or publishing the results. Extensive lists of possible bias were published in the literature *(Good* 1962*; Sackett* 1979*; Hartman et al.* 2002*)*. Some forms of bias are rather obvious and easily understood, but others require technical knowledge and reflection. Probably the most commonly discussed bias refers to the selection of participants for the study: in any study related to health, persons with a special interest in this area are more likely to participate. This might not affect the ill cases in a case–control study, but it may affect the selection of controls. If controls are the most healthy individuals from the population, they might have fewer risk factors – and in such a case the association between specific risk factors and the disease might appear too strong. While the selection bias may not be avoidable, the researcher should be aware of the direction in which such bias would affect the results. If the bias would affect the results towards a zero effect (odds ratio or relative risk of 1), and the effect still can be observed, then the researcher can propose that the effect is "at least" that large. When the direction in which the bias is most likely to affect the results is towards a smaller effect and no effect is observed in the study, the study is inconclusive, but such "negative" study will also receive less attention. The truly difficult case is when the bias is towards a larger effect and a positive effect is found in the study. Given this relationship, epidemiologists try to keep unavoidable errors in the direction which does not support the study hypothesis, and in such a case at least a conservative estimate of the effect can be

obtained. Finally, it has to be kept in mind that each study has its own potential in providing systematically distorted results.

11.8 Methodological Issues of Diagnostic Tests

Assessment of any disease is based on diagnostic tests, which is a specific area of measurement error. There are four basic terms which refer to diagnostic tests: sensitivity, specificity, positive and negative predictive value (Fig. 11.2).

Test / Disease	Positive	Negative	Sum
Yes	True positive	False negative	All sick cases
No	False positive	True negative	All healthy cases
Sum	All positive tests	All negative tests	All tests / all cases

Sensitivity = True positive/all sick cases
Specificity = True negative/all healthy cases
Positive predictive value (PPV) = True positive/all positive tests
Negative predictive value (NPV) = True negative/all negative tests

Fig. 11.2 Figure for illustration of sensitivity, specificity, positive and negative predictive value

Sensitivity is the ability of the test to detect sick cases. Some of the detected cases have the disease in reality, but others might be incorrectly classified as sick. *Specificity* is the ability of the test to distinguish between false-positive and truly negative cases.

A method which diagnoses all sick cases as sick and does not include any healthy cases in this group would be a perfect test and has a 100% sensitivity and a 100% specificity. However, in many cases either due to technical limitations of the test or due to overlapping values in both healthy and ill subjects, the sensitivity and the specificity are interconnected. Increasing sensitivity (i.e. detecting more of the diseased cases) results in a decreased specificity of the test (along with the higher detection of true cases, the number of false-positive tests also increases). Balancing sensitivity and specificity by selecting an appropriate cut-off for positive test is therefore based on the rationale of the test. For example, a screening test should have a high enough sensitivity to detect a high fraction of the sick cases. But since many of the cases might be incorrectly classified as positive, a second test can be used to confirm the result. This second test should have high specificity in order to discard the truly negative cases, after which only the truly positive cases should remain. Since the second test is used only in the subgroup which had positive results in the first test, the requirements on its sensitivity are not as high. This relationship points towards the fact that the performance of the test depends on the prevalence of

the disease in the sample: when the disease is very rare, along with the truly positive cases, there will be many false-positive cases.

Two other measures were proposed to provide information which incorporates the prevalence of the disease in the population:

> *Positive predictive value* is the probability of being ill when the test is positive.
> *Negative predictive value* is the probability of being healthy when the test is negative.

Tests with the same sensitivity and specificity will have a lower positive predictive value in a population where the disease is less frequent as compared to a population with a higher prevalence. This relationship is the opposite for the negative predictive value.

Test characteristics also affect the estimates of prevalence or incidence of a disease: when a real-world test is used, neither specificity nor sensitivity can be 100%. Therefore some cases will be falsely classified as sick while not having the disease and some others will be classified as healthy although they have the disease. When the sensitivity and the specificity are known, they can be used to correct the prevalence estimate. (For further reading on sensitivity and specificity, see Loong 2003.)

11.9 Outlook on the Use of Mathematical Modelling in Infectious Disease Epidemiology

For many research questions, infectious disease epidemiology uses the same methods as general epidemiology. For example, the estimates of prevalence can be obtained from surveys and that of incidence from cohort studies. However, the situation is more complicated for analytical studies, i.e. studies assessing risk factors of being infected. For communicable diseases, the probability of getting an infection depends not only on behavioural or genetic factors but also on the prevalence of the disease. This dependency requires the use of special statistical methods or mathematical models to obtain accurate results. Especially models designed to predict future spread of the disease should consider the relationship between risk of getting an infection and prevalence of this infection. When a fast investigation is necessary (as in outbreak investigation, see Chapter 9), conducting the analysis without taking into account the potential dependency between infected persons could be an option providing initial results. However, when not an immediate intervention, but estimation of the impact of specific risk factors is necessary, it should be performed taking into account the potential dependency between infected persons. To address this issue, statistical models incorporating dependency and dynamic transmission models were developed. Solutions to the dependency issue based on dynamic transmission models are presented in Chapter 12.

References

Armstrong BG (1998) Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med 55(10): 651–6

Badri M, Lawn SD and Wood R (2006) Short-term risk of aids or death in people infected with HIV-1 before antiretroviral therapy in South Africa: A longitudinal study. Lancet 368(9543): 1254–9

Bhutta ZA (2002) Ethics in international health research: A perspective from the developing world. Bull World Health Organ 80(2): 114–20

Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T and Friedenreich C (1999) Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol 28(1): 1–9

Bunnell R, Ekwaru JP, Solberg P, Wamai N, Bikaako-Kajura W, Were W, Coutinho A, Liechty C, Madraa E, Rutherford G and Mermin J (2006) Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda. Aids 20(1): 85–92

Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D and Tyrer P (2000) Framework for design and evaluation of complex interventions to improve health. Bmj 321(7262): 694–6

Cologne JB, Sharp GB, Neriishi K, Verkasalo PK, Land CE and Nakachi K (2004) Improving the efficiency of nested case-control studies of interaction by selecting controls using counter matching on exposure. Int J Epidemiol 33(3): 485–92

Creed-Kanashiro H, Ore B, Scurrah M, Gil A and Penny M (2005) Conducting research in developing countries: Experiences of the informed consent process from community studies in Peru. J Nutr 135(4): 925–8

Doyle J, Armstrong R and Waters E (2008) Issues raised in systematic reviews of complex multisectoral and community based interventions. J Public Health (Oxf) 30(2): 213–5

Ernster VL (1994) Nested case-control studies. Prev Med 23(5): 587–90.

Flory J and Emanuel E (2004) Interventions to improve research participants' understanding in informed consent for research: A systematic review. Jama 292(13): 1593–601

Good IJ (1962) A classification of fallacious arguments and interpretations. Technometrics 4(1): 125–32

Greenland S and Robins JM (1988) Conceptual problems in the definition and interpretation of attributable fractions. Am J Epidemiol 128(6): 1185–97.

Haahr MT and Hrobjartsson A (2006) Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. Clin Trials 3(4): 360–5

Hartman JM, Forsen JW, Jr., Wallace MS and Neely JG (2002) Tutorials in clinical research: Part iv: Recognizing and controlling bias. Laryngoscope 112(1): 23–31

Heckathorn DD (1997) Respondent-driven sampling: A new approach to the study of hidden populations. Social Problems 44(2)

Heckathorn DD (2002) Respondent-driven sampling ii: deriving valid population estimates from chain.Referral samples of hidden populations. Soc Probl 49(1): 11–34

Heckathorn DD (2007) Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. Sociol Methodol 37(1): 151–207

Hernan MA and Robins JM (2006) Estimating causal effects from epidemiological data. J Epidemiol Community Health 60(7): 578–86

Hill Z, Tawiah-Agyemang C, Odei-Danso S and Kirkwood B (2008)Informed consent in Ghana: What do participants really understand? J Med Ethics 34(1): 48–53

Hrobjartsson A and Gotzsche PC (2001) Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 344(21): 1594–602

Jones JH (1993) Bad blood: The Tuskegee syphilis experiment. New York, Free Press

Kang M, Ragan BG and Park JH (2008) Issues in outcomes research: An overview of randomization techniques for clinical trials. J Athl Train 43(2): 215–21

Krieger N and Zierler S (1997) The Need for epidemiologic theory. Epidemiology 8(2): 212–4
Loong TW (2003) Understanding sensitivity and specificity with the right side of the brain. Bmj 327(7417): 716–9
Lopes CS, Rodrigues LC and Sichieri R (1996) The lack of selection bias in a snowball sampled case-control study on drug abuse. Int J Epidemiol 25(6): 1267–70
Maclure M (1991) The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epidemiol 133(2): 144–53
Maclure M and Mittleman MA (2000) Should we use a case-crossover design? Annu Rev Public Health 21: 193–221
Marshall PA (2005) Human rights, cultural pluralism, and international health research. Theor Med Bioeth 26(6): 529–57
Martin W (2008) Linking causal concepts, study design, analysis and inference in support of one epidemiology for population health. Prev Vet Med 86(3-4): 270–88
Miyazaki Y and Raudenbush SW (2000) Tests for linkage of multiple cohorts in an accelerated longitudinal design. Psychol Methods 5(1): 44–63
Moher D, Schulz KF and Altman DG (2001) The consort statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 357(9263): 1191–4
Montana LS, Mishra V and Hong R (2008) Comparison of HIV prevalence estimates from antenatal care surveillance and population-based surveys in sub-Saharan Africa. Sex Transm Infect 84 Suppl 1 : i78–i84
Nardone A, Pebody RG, van den Hof S, Levy-Bruhl D, Plesner AM, Rota MC, Tischer A, Andrews N, Berbers G, Crovari P, Edmunds WJ, Gabutti G, Saliou P and Miller E (2003) Sero-epidemiology of mumps in western Europe. Epidemiol Infect 131(1): 691–701
Parascandola M and Weed DL (2001) Causation in epidemiology. J Epidemiol Community Health 55(12): 905–12
Park T, Ki M and Yi SG (2004) Statistical analysis of Mmr vaccine adverse events on aseptic meningitis using the case cross-over design. Stat Med 23(12): 1871–83
Philippe P (2001) Density incidence and cumulative incidence: A fundamental difference. Internet J Intern Med 2(2)
Piaggio G, Elbourne DR, Altman DG, Pocock SJ and Evans SJ (2006) Reporting of noninferiority and equivalence randomized trials: An extension of the consort statement. Jama 295(10): 1152–60
Prinzie P and Onghena P (2005). Cohort Sequential Design. B. Everitt and D. Howell, (eds.),. Encyclopedia of statistics in behavioral science. John Wiley & Sons
Putt ME and Ravina B (2002) Randomized, placebo-controlled, parallel group versus crossover study designs for the study of dementia in Parkinson's disease. Control Clin Trials 23(2): 111–26
Reintjes R, Kajueter H, Ehrhard I, van Treeck U and Ammons A (2005) Applying a case-crossover study design to examine transient exposures in the transmission of *N. Meningitides*. Eur J Epidemiol 20(7): 629–33
Rothman K and Greenland S (2002) Modern epidemiology. Lippincott-Raven, Philadelphia, PA.
Sackett DL (1979) Bias in analytic research. J Chronic Dis 32(1–2): 51–63.
Schulz KF, Chalmers I, Altman DG, Grimes DA and Dore CJ (1995) The methodologic quality of randomization as assessed from reports of trials in specialist and general medical journals. Online J Curr Clin Trials Doc No 197: [81 paragraphs]
Schulz KF, Chalmers I, Grimes DA and Altman DG (1994) Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. Jama 272(2): 125–8
Schulz KF and Grimes DA (2002a) Allocation concealment in randomised trials: Defending against deciphering. Lancet 359(9306): 614–8.
Schulz KF and Grimes DA (2002b) Blinding in randomised trials: Hiding who got what. Lancet 359(9307): 696–700

Sugarman J, Lavori PW, Boeger M, Cain C, Edsond R, Morrison V and Yeh SS (2005) Evaluating the quality of informed consent. Clin Trials 2(1): 34–41

Sumathipala A, Siribaddana S, Hewege S, Lekamwattage M, Athukorale M, Siriwardhana C, Murray J and Prince M (2008) Ethics review committee approval and informed consent: An analysis of biomedical publications originating from Sri Lanka. BMC Med Ethics 9: 3

Susser M (2001) Glossary: Causality in public health science. J Epidemiol Community Health 55: 376–8

Susser M and Susser E (1996a) Choosing a future for epidemiology: i. eras and paradigms. Am J Public Health 86(5): 668–73

Susser M and Susser E (1996b) Choosing a future for epidemiology: Ii. From black box to Chinese boxes and eco-epidemiology. Am J Public Health 86(5): 674–7

Sutherland ER (2007) Sham procedure versus usual care as the control in clinical trials of devices: Which is better? Proc Am Thorac Soc 4(7): 574–6.

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC and Vandenbroucke JP (2007) The strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. PLoS Med 4(10): e296

Wacholder S, McLaughlin JK, Silverman DT and Mandel JS (1992a) Selection of controls in case-control studies. I. Principles. Am J Epidemiol 135(9): 1019–28

Wacholder S, Silverman DT, McLaughlin JK and Mandel JS (1992b) Selection of controls in case-control studies. II. Types of controls. Am J Epidemiol 135(9): 1029–41

Wacholder S, Silverman DT, McLaughlin JK and Mandel JS (1992c) Selection of controls in case-control studies. III. Design options. Am J Epidemiol 135(9): 1042–50

World medical association declaration of Helsinki. ethical principles for medical research involving human subjects [Http://Www.Wma.Net/E/Policy/Pdf/17c.Pdf] Website