No cover image available

Critical Appraisal of Epidemiological Studies and Clinical Trials (3rd edn)

Mark Elwood

https://doi.org/10.1093/acprof:oso/9780198529552.001.0001

Published online: 01 September 2009 **Published in print:** 22 February 2007 **Online ISBN:** 9780191723865 **Print ISBN:** 9780198529552

CHAPTER

3 The results obtained from studies of causation

J. Mark Elwood

https://doi.org/10.1093/acprof:oso/9780198529552.003.03 Pages 53–74

Published: February 2007

Abstract

The previous chapter presented a classification of study design which separated the criteria used to select the groups of individuals to be compared from the time relationships of the study. The logic of this system will appear clear in this chapter, which discusses the results obtained from these studies. The format of the results and therefore the appropriate methods of interpreting them depend on the groups being compared. The extent to which particular problems and biases occur depends largely on the time relationships. Self-test questions are provided at the end of the chapter.

Keywords: study designs, cohort studies, intervention studies, surveys, cause and effect, bias

Subject: Public Health, Epidemiology

Collection: Oxford Scholarship Online

> *When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science.*
>
> —William Thomson, Lord Kelvin: Popular lectures and addresses; 1891–1894

In the previous chapter we presented a classification of study design which separated the criteria used to select the groups of individuals to be compared from the time relationships of the study. The logic of this system will appear clear when we discuss the results obtained from these studies. The format of the results and therefore the appropriate methods of interpreting them depend on the groups being compared. The extent to which particular problems and biases occur depends largely on the time relationships.

Cohort and intervention studies

We shall first consider the cohort design, which applies to both observational cohort studies and intervention trials. In this design we are comparing groups of individuals who are classified by their exposure to the putative causal factor. The simplest situation is with two groups, which we can regard as 'exposed' and 'non-exposed'. This will often be the real situation; for example, we might compare women using oral contraceptives with women who do not use oral contraceptives, or patients on one therapy with patients on another. In other situations, the cohorts may be more numerous and may be ordered, such as groups of individuals characterized by the amount they smoke, or groups of patients with different severities of disease. The format of the results in any cohort study is the same. The same method of analysis is used for aetiological studies looking at individuals exposed to external agents in the environment, observational studies comparing the outcome of groups of patients given different therapies, and randomized trials. It makes no difference to the analysis
p. 54 whether the cohort study is prospective (where subjects were identified at the ↳ start of the study and followed forward in time), or retrospective, where the subjects are identified using records and their experience up to the current time provides the outcome data.

The simplest situation is illustrated by **Ex.** 3.1 which shows the results of a study in expectant mothers who had epilepsy during pregnancy, which was treated with one anti-epileptic drug, i.e. monotherapy. The comparison is between those treated with one specific drug, valproic acid, and those treated with other single agents. Mothers with epilepsy who required no drug treatment and those who received two or more drugs are excluded, as they are likely to have less or more severe epilepsy. The data come from a prospective cohort study in the UK based on enrolments of women during pregnancy from both general practitioners and specialists [1]. Since this is a study of pregnancy outcome, the follow-up period is fixed and relatively short, and the outcome of interest is the frequency of a major congenital abnormality in the offspring. Therefore the results have a simple format. For each group of mothers, exposed and unexposed, we have the number of infants delivered, and the number of these infants who has an abnormality.

RESULTS OF A COHORT STUDY (1)

Exposure	Number of infants			Prevalence of malformation (%)
	Malformed	Not malformed	Total	
Valproic acid	44	647	715	6.15
No valproic acid	47	1697	1710	2.75

Ex. 3.1. Results of a cohort study comparing among mothers with epilepsy, those treated with valproic acid with those receiving other drugs (each as single agents) with regard to the occurrence of major malformations in their offspring. From Morrow *et al*. [1]

The results show the two proportions, or prevalence rates at birth. If there were no difference in outcome between the two groups i.e. if the 'null hypothesis' were true, the prevalence rates in the two groups would be the same, apart from random variations. It is logical to compare either the difference in the rates, or the ratio of the rates, and these considerations lead to important and widely used epidemiological measures. **Ex.** 3.2 shows the measures which can be derived.

RESULTS OF A COHORT STUDY (2)

Exposure	Number of infants: Malformed	Not malformed	Total	Prevalence of malformation (%)
Valproic acid	44	647	715	6.15
No valproic acid	47	1697	1710	2.75
Whole population (assuming sample is representative)	91	2344	2425	3.75

RR, Relative risk or risk ratio = 6.15/2.75 = 2.24

OR, Odds ratio = (44/647)/(47/1697) = 2.46

AR, Attributable risk or risk difference = 6.15 – 2.75 = 3.41%

AP_{exp}, *Attributable proportion in exposed subjects* = attributable risk / total risk in exposed = 3.41/6.15 = 55.3%

Also, $AP_{exp} = (RR-1)/RR = (2.24 - 1)/2.24 = 55.3\%$

AP_{pop}, *Attributable proportion in the population*:

For the whole population, risk = 3.75%
attributable proportion in the whole population = (3.75 – 2.75)/3.75 = 26.8%
(using full data to avoid rounding errors)

More generally, where p = the proportion of the population exposed, estimated from the study or from other sources, the attributable proportion in the population is given by

$$AP_{pop} = \frac{p(RR-1)}{p(RR-1)+1}$$

Here p = 715/2425 = 29.4%
So AP_{pop} = 26.8%

Ex. 3.2. Derivation of measures of association from the results of the cohort study shown in Ex. 3.1. Note: in this and subsequent exhibits, the calculations are shown with limited decimal places for clarity, but the calculations themselves are carried out from the basic data to avoid rounding error.

Relative risk and relative odds

The *relative risk* or *risk ratio* is the ratio of the rate of disease among those exposed (mothers using valproic acid)
p. 55 to the rate in those not exposed ↳ (mothers taking other drugs), and as it is a ratio it has no units. In this example the relative risk is 2.24.

We can also consider a further measure, the *odds ratio*. The odds of a baby in the exposed group being malformed are 44:647 (approximately 1 in 15), compared with odds of 47:1697 (about 1 in 36) in the unexposed group. The ratio of these is the odds ratio, which is 2.46, similar but not identical to the relative risk. The odds ratio and the relative risk are similar when the frequency of the outcome is low; they diverge as the outcome
p. 56 becomes more frequent. ↳ For example, if the prevalence in the exposed group is 50 per cent (odds 1:1), and that in the unexposed group is 25 per cent (odds 1:3), the risk ratio is 2, but the odds ratio is 3. For both relative risk and odds ratio, the value of 1.0 corresponds to the situation of no association between exposure and outcome.

Risk difference (attributable risk)

The difference between the two rates is the *risk difference*, or *attributable risk*. The latter term is less satisfactory as it implies that there is a causal relationship. The risk difference gives, in absolute terms, the frequency of the outcome which is associated with the exposure. In this example the attributable risk is 3.41 per cent. If there were no association present, the risk difference would be zero.

If we divide the attributable risk by the total risk in the exposed group, the result is the proportion of disease in those exposed to the factor which is associated with the exposure; this can be called the 'attributable proportion in exposed subjects'. Here it is 3.41/6.15 = 55 per cent.

If, for the population under study, the proportion of subjects who are exposed is known, the proportion of total disease in the population which is associated with the causal factor can be calculated; this is the 'attributable proportion in the population'. In this example it is 27 per cent.

The uses of risk difference and relative risk

Unlike relative risk, the risk difference, or attributable risk, describes the absolute quantity of the outcome which is associated with the exposure. Therefore it is useful in considering the practical implications of studies once a decision has been reached that the association represents causation. Thus a comparison of the attributable risks in terms of total mortality conferred by a number of environmental exposures gives an indication of how much mortality will be prevented by successful action on each one of the exposures, and such an approach has been useful in setting priorities for public health and health promotion campaigns. A similar consideration in terms of the long-term outcome of a complex disease such as diabetes, whose course may be influenced by a large number of factors, may be used to indicate what aspects of care, if dealt with successfully, will result in the greatest improvement in outcome. Therefore attributable risk is particularly useful in well-researched situations where the implications of soundly supported results are being considered. It is of less value in the preliminary stages of assessment of a possible causal relationship.

The great advantage of the relative risk estimate is the empirical finding that in many human disease situations, the relative risk of disease incidence or mortality associated with a particular exposure is fairly
p. 57 constant over a wide ↳ range of populations and groups of subjects, even where other factors differ. If the relative risk for a particular exposure is the same in two different populations, but because of differences in other factors the total risks of the outcome are different, then the attributable risks will be different. In comparing evidence from a number of sources concerning a particular causal relationship, it is more reasonable to expect consistency in terms of relative risk than of attributable risk. This property of consistency makes relative risk more valuable in evaluating whether a particular relationship is or is not likely to be causal. It is also rather easier to predict the effects of non-causal factors on the observed relative risk than on the observed attributable risk.

There is another important reason for concentrating on relative risk, and that is because this measure, unlike attributable risk, can be derived from any of the main study designs. As will be shown, case–control studies provide estimates of relative risk (by the odds ratio), but do not provide direct measures of attributable risk. For these reasons the relative risk is the more useful index to summarize the results of a study in order to discuss the interpretation of the association. Therefore in the rest of this book we will concentrate on the interpretation of relative risk.

Preventive factors and application to intervention studies; number needed to treat

If the factor under consideration is protective, the rate of outcome in the exposed group will be less than in the unexposed group, and therefore the relative risk will be less than 1, and the risk difference will be negative. There is no difficulty with this apart from terminology.

Ex. 3.3 shows the result of a large-scale randomized trial in which patients with myocardial infarction (a heart attack) were treated with a new drug (captopril), and compared with a placebo group [2]. This was a very large trial with about 29 000 subjects in each group. As shown, there was a protective effect of the new drug, with a

relative risk of 0.94 and a risk difference of −4.9 deaths per 1000 subjects. Rather than calling this a risk difference or, if we assume causality, an attributable risk of −4.9 deaths per 1000, a better terminology is to call it an attributable benefit of 4.9 deaths averted per 1000 subjects. Such terminology is logical, although not widely used. We can also simply refer to this absolute benefit as the 'number of deaths averted per 1000 treated', which will again be 4.9. We can also invert that figure, and express the results of the trial in terms of the 'number of subjects treated with the new drug to prevent one death', often called the 'number needed to treat' (NNT). This is a 1000/4.9, or approximately 200. This is an easily understandable and readily applicable figure
p. 58 [3]. Thus the result of this study (assuming we interpret ↳ it as a causal relationship) is that treatment of these patients with this new drug resulted in a reduction in death rate, with one death averted for every 200 subjects treated. Whether this benefit is worthwhile will depend on how it is balanced by the disadvantages of the new drug, in terms of side effects, costs, and so on. (More correctly, one death is averted for every 200 offered treatment, as not all those randomized to captopril would have necessarily accepted it or taken the full dosages.)

BENEFICIAL EFFECTS AND NUMBER NEEDED TO TREAT

Actual results

	Deaths	Survivors	Total	Death rate per 1000
Treated with captopril	2088	26940	29028	71.9
Placebo group	2231	26791	29022	76.9

Relative risk	= 71.9/76.9	=	0.94
Risk difference per 1000	= 71.9 – 76.9	=	– 4.9
Deaths averted per 1000 treated	= 76.9 – 71.9	=	4.9
Number treated to prevent one death	= 1000/4.9	=	202

Hypothetical results if relative risk remains the same, but baseline mortality rate doubles

	Deaths	Survivors	Total	Death rate per 1000
Treated with captopril	4176	24852	29028	143.9
Placebo group	4462	24560	29022	153.7

Relative risk	= 143.9/153.7	=	0.94
Risk difference per 1000	= 143.9 – 153.7	=	– 9.9
Deaths averted per 1000 treated	= 153.7 – 143.9	=	9.9
Number treated to prevent one death	= 1000/9.9	=	101

Ex. 3.3. Results of a clinical trial: number needed to treat. Results of randomized trial of a drug treatment (captopril) in patients with suspected myocardial infarction; deaths in first 35 days. From ISIS-4 study [2]. As noted in Ex. 3.2, calculations avoid rounding errors, which is why the risk difference is −4.9, not −5.0 from the simple calculation shown.

Consider what would happen if the relative effect of the drug were the same, i.e. a relative risk of 0.94, but the death rate of both the treated and the comparison groups were higher, as would happen if the treatment were offered to patients with particularly severe disease. If the relative risk is constant and the mortality rate in the non-treated group is doubled, the number of deaths averted per 1000 treated (the attributable benefit) is doubled, and the number who are treated to prevent one death falls to about 100. The difference between 100 and 200 patients treated to avert one death could well be important in terms of the comparison with ill effects and costs. Although in this example only the data on early mortality are shown, the overall results of this major
p. 59 study looking ↳ at longer-term mortality were essentially similar, and from consideration of this and other literature, the authors indeed concluded that the use of the captopril therapy would save about one life per 200 treated in average-risk patients, and about one life per 100 treated in higher-risk groups.

Perceptions given by relative and absolute risks: framing

Therefore the association shown by studies can be expressed as a ratio mea surement, or a measurement on an absolute scale. Results which are actually the same can produce different impressions when presented in different ways, an effect referred to as ‘framing’, and the question of whether relative or absolute measurement of benefits or risks should be used can be controversial. Effects expressed as absolute benefits are usually numerically smaller than when expressed as relative risks, and often have a lesser impact [4], for example on influencing doctors to prescribe cholesterol-lowering drugs [5]. In a population survey, 80 per cent of people said that they would accept a screening test when presented with data on relative risk reduction, while only 45 per cent would accept it when given equivalent data expressed as number needed to screen [6]. Both absolute and relative measures have value. The trial just described shows a very useful property of relative risk, in that this was fairly constant over the range of underlying absolute mortality rates. The absolute benefit measure, which can be expressed as the number needed to treat to avert one outcome, is a very useful interpretation in terms of clinical and cost implications. For example, in recent trials of the monoclonal antibody trastuzumab (Herceptin), use of this drug in addition to standard chemotherapy after primary surgery in suitable patients with breast cancer (adjuvant treatment) gave a relative risk of 0.67, i.e. a one-third reduction in the average death rate [7]. This is quite dramatic, but these are early results. At 3 years follow-up, the absolute mortality was reduced from 8.3 per cent to 5.7 per cent, which is a risk difference of 2.6 per cent, corresponding to a number needed to treat of 38 (1/0.026). Thus treating 38 eligible women with the new drug would result in one less death. Given that the cost of a course of the drug may be around US $50 000, the cost of averting one death would be about US $2 million. Understandably, these important new results are causing considerable discussion in terms of appropriate drug costs. Longer follow-up in the trial may change the results; if the relative risk persists, the absolute mortality reduction will increase and the number needed to treat will fall. Indeed, the estimated 4-year follow-up results show this; the mortality is reduced from 13.4 per cent to 8.6 per cent, giving a risk difference of 4.8 per cent and a number needed to treat of 21.

p. 60

Person-time as the denominator of rates

In the data for pregnancies in Ex. 3.2, the outcome measure is the proportion of malformed births. In cohort studies of other outcomes, the follow-up time is not fixed and may not be the same for each subject. Consider a cohort study in which men are classified by their level of exercise, and subsequently the occurrence of deaths from heart disease is recorded. Each subject contributes information from the time he enters the study until his death, until the end of the follow-up period, or until some specified time at which the data collection ends. Some subjects may leave the study before any of these endpoints; they are ‘censored’, or become ‘lost to follow-up’, but still contribute to the study up to the time at which their outcome status was last known. To assess the incidence rate, the number of heart disease deaths is divided by the total follow-up period, i.e. the sum of the follow-up times for all individuals, expressed as person-time (person-years, man-months, etc.).

One of the classic studies showing a relationship between physical exercise and coronary heart disease mortality was the prospective cohort study of longshoremen (dock workers) in San Francisco [8]. Between 1951 and 1961, men who were assessed at a health screening clinic and were aged 35–74 were entered into the study; men first assessed at a younger age entered the study when they turned 35. The follow-up continued until death, attaining the age of 75, the end of the follow-up period in 1972, or the date of loss to follow-up. Less than 1 per cent were lost to follow-up. Thus the follow-up varied from very short, if an early death occurred, to 22 years. In total, 6351 men entered the study, contributing 92 645 man-years of experience; there were 598 deaths from coronary heart disease, giving a crude death rate of 598/92 645 = 64.5 deaths per 10 000 man-years. The total incidence rate in this example is the average for the whole follow-up period. If the rate of outcome varies greatly over the follow-up period, this simple average will not be adequate and more complex methods of analysis are necessary. As the issues of interpretation can be adequately dealt with using simple data, such methods are not discussed here but are presented in Chapter 7 as survival and life-table methods (p. 264).

More than one causal agent

Ex. 3.4 shows death rates from cardiovascular disease in cohorts of women defined by two exposures, smoking and contraceptive usage, from a large prospective cohort study [9]. Women exposed to neither oral contraceptives nor smoking had a cardiovascular death rate of 3.0 per 100 000 woman-years. Those who smoked but did not use oral contraceptives had a death rate of 8.9/100 000, while those who used oral contraceptives but did not smoke had a death rate of 13.8/100 000. In this joint exposure situation, it is
p. 61 appropriate ↳ to regard the women who were exposed to neither factor as the baseline or 'referent' group, and to consider attributable risks and relative risks as compared with this group. What rate would we expect in women who both used oral contraceptives and smoked?

COHORT STUDY WITH TWO CAUSAL FACTORS

Exposure	Mortality from circulatory system diseases, per 100 000 woman-years	Relative risk	Attributable risk/10^5 woman-years
Non-smoker, no OC use	3.0	1.0	0 (referent)
Smoker, no OC use	8.9	3.0	5.9
Non-smoker, OC user	13.8	4.6	10.8
Smoker, OC user	?	?	?

Ex. 3.4. Two causal factors. Results of a prospective cohort study comparing women classified by smoking habit and oral contraceptive (OC) use in terms of deaths from circulatory system diseases. For effects of joint exposure, see text. From Royal College of General Practitioners [9]

There are two simple methods by which we could derive such a rate. The first is to assume that women who are exposed to both agents have the baseline risk of the unexposed group, plus the attributable risk associated with smoking, plus the attributable risk due to oral contraceptive use, and so end up with a cardiovascular mortality rate of 3.0 + 5.9 + 10.8 = 19.7 deaths per 100 000 woman-years. We are assuming that the two effects work in an *additive* fashion. Therefore the excess risks produced by each exposure add together, and add to the baseline risk, which is due to the effects of other factors, to give the total risk. The relative risk for the group exposed to both factors, compared with the unexposed group, is 19.7/3.0 = 6.6. We can derive the same figure using relative risk estimates by combining what is called the *excess relative risk*, which is the relative risk minus 1. Therefore the additive calculation is the excess relative risk from smoking (3.0 − 1 = 2.0), added to the excess relative risk from oral contraceptive use (4.6 − 1 = 3.6) plus the baseline relative risk (1.0), which sums to give 6.6 as the relative risk for subjects with both exposures compared with those with neither.

Another simple argument is that if smoking increases an individual's risk by a factor of 3.0, and oral contraceptive use increases it by a factor of 4.6, the joint effect of smoking and oral contraceptive use may increase the risk by a factor of 3.0 × 4.6 = 13.8. If we multiply this relative risk of 13.8 by the baseline absolute risk of 3.0 deaths per 100 000 woman-years, we obtain an expected death rate of 41.4 deaths per 100 000
p. 62 woman-years. We are assuming here ↳ a *multiplicative* model, i.e. the effect of two exposures is the multiple of the effects of each. In few studies of human disease have we enough information on biological mechanisms to predict confidently which of these models, or indeed which of a large range of other models, fit the real situation. We have to observe what happens in practice. In this oral contraceptive study the cardiovascular death rate per 100 000 woman-years in women who both smoked and were exposed to oral contraceptives was in fact 39.5 per 100 000 woman-years. This is close to the expected result on a multiplicative model, and so this is the more appropriate model for these data.

However, we should be cautious, as the numbers of deaths in each group shown in Ex. 3.4 are small (2, 3, 5, and 19 in the four groups) and more information would be needed to be confident that the multiplicative model is a better fit to the data than the additive model. This example is based on a early report of this cohort study with

follow-up from the start of the study in 1968 to 1976. A later report from this study, based on follow-up to 1993, showed that the increased mortality from circulatory diseases is largely restricted to current and recent users of oral contraceptives [10], and a further report from another cohort study of the topic showed that the increase in deaths from ischaemic heart disease associated with oral contraceptive use was confined to heavy smokers [11].

Case–control studies

The other major design is the case–control study, comparing a group of individuals who have experienced the outcome under study with a group who have not. The exposure of each subject in the study to the factor under consideration is ascertained retrospectively. The results appear as a 2 × 2 table as shown in **Ex.** 3.5 In the case–control design, a sample of all available cases is taken ($a + c$), and a sample of unaffected subjects (controls) is drawn independently ($b + d$). As these two groups are sampled separately, rates of disease in the exposed or unexposed groups cannot be calculated, nor can relative risk be measured directly. However, the *odds ratio* can be obtained, and so it is the primary measure of association in case–control studies.

CASE-CONTROL STUDY RESULTS

	Cases	Controls
Exposed	a	b
Unexposed	c	d
	$a + c$	$b + d$

Odds ratio = $(a/b)/(c/d)$
= ad/bc the 'cross-products' ratio

Ex. 3.5. Simplest form of results for a case-control study. The odds ratio, or relative odds, is the key measure of association. Because of the sampling used, the total number of exposed subjects is *not* $a + b$, and the risk in exposed subjects is *not* $a/(a + b)$; see text for explanation

Note the simple algebra of the calculation of odds ratio; the number a of exposed cases is multiplied by the number opposite it on the diagonal of the table, and this result is then divided by the two other numbers multiplied together; because of this simple arithmetic, the odds ratio is sometimes referred to as the *cross-products ratio*.

To understand how the odds ratio in a case–control study is derived, it is useful to consider an unusual example in which we can compare the results with the situation in the whole population from which the study participants ↳ were selected. **Ex.** 3.6 shows the results of a case–control study in which 1391 births with anencephalus, a severe and fatal abnormality, were compared with 5000 live births; the exposure was the mother's past history of stillbirth [12]. The odds ratio is 4.13, showing a strong association. Ignore the results under (b) at present.

p. 63

CASE-CONTROL STUDY RESULTS

	Cases numbers	Controls numbers
Exposed (one or more previous stillbirths)	141	133
Unexposed (no previous stillbirths)	1250	4867
	1391	5000

(a) Odds ratio = $(141 \times 4867)/(133 \times 1250) = 4.13$

(b) Relative risk undetermined but will be similar to relative odds as other information shows that anencephalus is uncommon (about 1–2 per 1000 births).

Attributable risk undermined.

Attributable proportion in exposed subjects $= \frac{4.13 - 1}{4.13} = 75.8\%$

Attributable proportion in population; if controls are representative of all unaffected births, then prevalence of exposure $p = 133/5000 = 0.0266$
and attributable proportion in the population =

$$\frac{p(RR - 1)}{p(RR - 1) + 1} = \frac{0.0266 \times 3.13}{(0.0266 \times 3.13) + 1} = 7.7\%$$

Ex. 3.6. Format of results from a case-control study assessing the relationship between a history of a previous stillbirth and the occurrence of anencephalus. Mothers of all notified babies with anencephalus in a defined population form the case series (n = 1391); an arbitrary number of 5000 mothers of liveborn babies were chosen as controls. From Elwood *et al*. [12]

p. 64 To see how these results were derived, **Ex.** 3.7 shows the source population, i.e. all births in certain cities in Canada over a 20-year period. There were 1391 births with anencephalus, all of which were included in the case–control study; and 1 193 600 live births, of which only 5000 were selected to form the control series. From Ex. 3.7, all the measures of risk can be calculated as shown. Note that the relative odds and the relative risk are virtually the same; this is because the frequency of the disease is low, and so the numbers of unaffected and of total subjects are not very different.

A RELATIONSHIP IN A POPULATION

	Anencephalus births	Unaffected births	Total births	Prevalence of anencephalus per 1000 births
Previous stillbirth	141	31 750	31 891	4.42
No previous stillbirth	1250	1 161 850	1 163 100	1.07
	1391	1 193 600	1 194 991	1.16

relative odds = (141/31 750)/(1250/1 161 850) = 4.13
relative risk = (141/31 891)/(1250/1 163 100) = 4.11
attributable risk = 4.42 – 1.07 = 3.35 per 1000 births
attributable proportion in exposed subjects = 3.35/4.42 = 75.8%
attributable proportion in population = (1.16 – 1.07)/1.16 = 7.8%

Ex. 3.7. Relationships in a population. The relationships between anencephalus and previous stillbirths in all pregnancies in the population in which the case-control study shown in Ex. 3.6 was conducted. From Elwood *et al*. [12]

Now, return to Ex. 3.6 and see what measures can be calculated. The odds ratio estimate is valid, differing from that in the whole population only by sampling variation; knowing that the outcome being studied is uncommon allows us to use the odds ratio as an estimate of relative risk. What cannot be calculated from Ex. 3.6 is the actual rate of disease in either the exposed or unexposed groups, or the attributable risk. If Ex. 3.6 is misinterpreted by a failure to consider that the data come from a case–control study, it might be thought that the risk of this defect in births to mothers who had had a previous stillbirth was 141/(141 + 133) = 51 per cent, an order of magnitude greater than the true value (4.4 per 1000) given in Ex. 3.7.

Attributable proportion in case–control studies

In Ex. 3.2 it was shown that there is a simple formula linking the attributable proportion in exposed subjects to
p. 65 the relative risk, with the attributable proportion being equal to $(RR - 1)/RR$; thus this can be estimated in a ↳ case–control study, as shown in Ex. 3.6. The attributable proportion in the population can be calculated if there is an estimate of the proportion of the total population which is exposed to the causative factor (Ex. 3.6). In some circumstances the control group in a case–control study can be considered as a representative sample of the total population, and this proportion is given by the proportion of controls exposed. However, often this is not so; for example, if the controls have been matched to the cases on certain characteristics, they will not be representative of the population and an independent source of evidence will be necessary to give the proportion of the population which is exposed.

Use of the odds ratio

As has been already shown, the odds ratio is a very good estimate of relative risk in most situations, the exception being where the outcome is very frequent. Cohort designs for outcomes which are frequent will be efficient, as a reasonable number of subjects will have to be followed to obtain an adequate number who experience the outcome under study. The great disadvantage of cohort studies is in attempting to study outcomes which are uncommon, so that many hundreds or thousands of individuals have to be followed before a reasonable number of them experience the outcome. It is in this situation that case–control studies are most useful. Therefore in most situations for which case–control studies are advantageous, the difference between relative odds and relative risk will be trivial, and certainly much less important than other potential sources of error in the study. Therefore in most literature odds ratios produced from case–control studies are referred to as relative risks.

Odds ratio from case–control studies; a more formal derivation

(For the more curious reader; others may wish to skip to the next section.)

The difference between the results from a case–control study and those from the entire source population arise because the sampling fractions are different for cases and controls. In the study shown in Ex. 3.6 and Ex. 3.7, the sampling fraction for the cases was 1.0, and that for the controls was 5000/1 193 600 = 0.00419. Usually the sampling fractions are unknown in a case–control study, but they will often be of this nature, in that the sampling fraction of the cases will be 1.0 or very high, and the sampling fraction of the controls will be very low.

Ex. 3.8 summarizes the algebra of case–control studies, showing that in the odds ratio calculations the sampling fractions cancel out, and therefore the odds ratio is a direct estimate of the odds ratio in the source
p. 66 population, even if the sampling fractions are unknown. This simple algebra emphasizes a crucial ↳ point in the design of case–control studies: the sampling system used for cases must be ‘without bias’ with respect to exposure, i.e. the same sampling fraction f must apply to both exposed and unexposed cases. If that is so, the ratio of the numbers of exposed to unexposed cases in the study will be the same as the ratio in the population. The same logic applies to the controls, which must also be sampled without bias with regard to exposure status, so that the (different) sampling fraction g applies to both exposed and unexposed controls.

ODDS RATIO IN A CASE-CONTROL STUDY

	Cases	Controls
Exposed	$a = fA$	$b = gB$
Unexposed	$c = fC$	$d = gD$
	$a + c = f(A + C)$	$b + d = g(B + D)$

$$\text{odds ratio in population} = (A/B)/(C/D) = AD/BC$$
$$\text{odds ratio in study} = (a/b)/(c/d)$$
$$= (fA/gB)/(fC/gD)$$
$$= AD/BC$$

Ex. 3.8. Algebraic justification for the calculation of odds ratio from a case-control study.

Capital letters = numbers of individuals in the population (A, B, C, D).

Lower case letters = numbers of individuals in the study (a, b, c, d).

f = sampling fraction for cases = $(a + c)/(A + C)$

g = sampling fraction for controls = $(b + d)/(B + D)$

The essential design characteristic of case-control studies is that the fractions f and g are the same for both exposed and unexposed subjects

Different sampling schemes in case–control studies: why a case can also be a control

The traditional, and intuitively clear at first glance, process of selecting controls in a case–control study is to select them as a representative sample of the *unaffected* subjects from a particular population. In **Ex.** 3.9 we present a hypothetical set of data concerning an exposure in pregnancy and a malformation, which is based loosely on the data presented in Ex. 3.1, but quite deliberately shows a much greater prevalence of the malformation at birth. Therefore in this hypothetical example, the exposure is associated with a relative risk of 4.23, but an odds ratio of 5.65. The substantial difference between these two measures is because the frequency of the malformation overall is quite high.

SAMPLING OF CONTROLS IN A CASE-CONTROL STUDY

A: Total population

Exposure	Malformed	Not malformed	Total	Prevalence of malformation (%)
Exposed	120	273	393	30.5
Unexposed	140	1799	1939	7.2
Total population	260	2072	2332	11.1

Odds ratio = 5.65
Relative risk (prevalence ratio) = 4.23

B: Case control design, sampling controls from non-affected only

Exposure	Malformed	Controls = not malformed
Exposed	120	53
Unexposed	140	347
Total population	260	400

Sample estimate of odds ratio = $(140 \times 347)/(53 \times 140) = 5.61$

C: Case control design, sampling controls from all pregnancies

Exposure	Malformed	Controls = not total pregnancies
Exposed	120	67
Unexposed	140	333
Total population	260	400

Sample estimate of relative risk = $(120 \times 333)/(67 \times 140) = 4.26$

Ex. 3.9. Different sampling schemes for controls in a case–control study. Subtable A shows the whole population data; both odds ratio and relative risk can be calculated. Subtable B shows controls drawn as a random sample of unaffected pregnancies; the results yield the odds ratio. Subtable C shows controls drawn as a random sample of all pregnancies; the results yield relative risk.

Now suppose we set up a case–control study in this situation. One design of case–control study would take a representative sample of *malformed* infants as cases, and a representative sample of the *not malformed* infants
p. 67 as controls. Let us assume we sample all the malformed infants (260) and an arbitrary sample ↳ of 400 of the unaffected infants (there is nothing magical about having an equal number, or a fixed ratio of controls to cases, and the actual size chosen will depend on practical issues such as cost). The numbers in subtable B are calculated by integer numbers given by a random sample of 400 controls. This sample yields an *odds ratio* of 5.61, which is a good estimate of the actual odds ratio of 5.65 in the population, but is clearly different from the relative risk.

A second type of case–control design would again sample all the 260 malformed infants as cases, but take a representative sample of *all infants* as controls. Therefore the distribution of the controls would reflect the
p. 68 exposure frequency in the whole population, not just the unaffected babies. The cross-products ratio ↳ from this design yields a sample estimate of *relative risk* directly, as the control sample represents the total population rather than only the 'non-case' population. As can be seen, a good estimate of the relative risk is obtained, which is clearly different from the odds ratio. This is '*incidence density*' sampling.

The practical ways in which these studies would be carried out clarify another seemingly complex issue. In the first design, comparing cases with non-cases, malformed infants would be entered into the study as cases as they are recognized; and a sample of non-malformed infants would be chosen, for instance by selecting a number of such infants from the same hospitals each week and interviewing their mothers, as non-diseased controls.

The second design is equally simple. Again, malformed infants would be classified as cases as they occur, but the comparison sample here would be a representative, perhaps random, sample of all births occurring in the

birth population which yields the cases. In this sampling design, an affected 'case' infant is also eligible for selection as a control; if that case infant appears in the representative sample of controls, it should be used as a control. Thus the same individual can appear as both a case and a control. Although this seems obscure, it is simply analogous to the calculation of any measure of occurrence rate, such as prevalence or incidence rate. In these rates (as noted in Chapter 1), the number of affected individuals is the numerator, and it is divided by the total number of individuals at risk which of course includes those who are affected. Therefore affected individuals appear in both the numerator and the denominator of a calculated frequency rate. Thus in a case–control study designed to produce an estimate of the ratio of such rates, i.e. a relative risk, the same individual can appear as a case and a control.

As was shown earlier, if the odds ratio is calculated, affected and unaffected infants appear as the numerator or the denominator respectively, and there is no figure which includes both. The case–control design which produces estimates of the odds ratio keeps the cases and controls separate, and an individual cannot appear in both affected and unaffected samples.

Many case–control studies of major diseases are of the type which directly yields an estimate of the relative risk. A typical design for a case–control study of a chronic disease such as cancer is to enter individuals as cases as soon as they are diagnosed with the disease, and to obtain a control at that time from the entire population from which that case is sampled. Therefore the control series provides a sample based on the population at risk at the time when the case became an incident case. The cross-products ratio, as in subtable B of Ex. 3.9, gives an estimate of the relative risk. In this design, an individual can be sampled as a control, and then shortly afterwards could become a case and be eligible to be sampled as a case. This is quite appropriate theoretically.

p. 69

Surveys

In principle, survey designs yield the same type of results as cohort studies; in cross-sectional surveys, all information on both exposures and outcomes is in terms of prevalence. The terms relative risk, relative odds, and so on can be used, remembering that in this situation these apply to ratios of prevalence rates rather than ratios of incidence rates as in a prospective cohort design. The use of relative odds is also more limited because frequently in a survey the prevalences of conditions being studied are quite substantial, and therefore the relative odds may be substantially different from the relative risk. Thus in the survey comparing Tanzanian villagers with different diets, the prevalence of borderline or frank hypertension was 38.7 per cent in those with a vegetarian diet, and 12.5 per cent in fish-eaters, giving a prevalence rate ratio of 3.1; the odds ratio is 4.4 [13].

Ex. 3.10 shows the results of a cross-sectional survey of the prevalence of smoking in a sample of schoolchildren in England, and explores the association, at that point in time, between the parents' current smoking habit with that of the children [14]. Where one or both parents smoked, the prevalence of 'ever smoking' in the children was 57.0 per cent; where neither parent smoked, it was 51.5 per cent. The relative risk, the prevalence ratio, is 57.0/51.5 = 1.11. The odds ratio can be calculated, as shown, but tends not to be used as it is different from the relative risk as the prevalence of the outcome is high. The attributable risk, and proportions, are calculated as shown and expressed in terms of prevalence.

RESULTS OF A CROSS-SECTIONAL SURVEY

Parent's smoking	Smoking in young children: Ever smoked	Never smoked	Total	Prevalence, ever smoked (%)
One or both smoke	3203	2418	5621	57.0
Neither smokes	1601	1508	3109	51.5
Total	4804	3926	8730	55.0

Prevalence ratio (relative risk) = 57.0/51.5 = 1.11
Odds ratio = (3203/2418)/(1601/1508) = 1.25

The odds ratio is not a very useful measure, as the prevalence ratio can be calculated as easily, and it is substantially different from the prevalence ratio as the outcome is common.

$$\text{Attributable risk or prevalence difference} = 57.0 - 51.5 = 5.5\%$$

$$\text{Attributable proportion in exposed} = \frac{5.5}{57.0} = 9.6\% = \frac{(RR-1)}{RR}$$

$$\text{Attributable proportion in population} = \frac{55.0 - 51.5}{55.0} = 6.4\%$$

$$= \frac{p(RR-1)}{p(RR-1)+1}$$

$$\text{where } p = 5621/8730 = 0.64$$

Ex. 3.10. Results of a cross-sectional survey in which questionnaires were given to 8–9-year-old subjects in industrial areas of northern England; these results show the smoking history of the subjects, compared with their report on their parents' smoking.

Note: for the attributable proportion calculation, rounding errors can be considerable. Thus *RR* = (3203/5621)/(1601/3109) = 1.1066; (*RR* − 1)/*RR* = 0.1066/1.1066 = 9.6% although (1.11 − 1)/1.11 = 9.9%. From Charlton [14]

General applicability of these analytical approaches

We have now presented a format of results for a cohort study and for a case–control study. The cohort format applies to any study in which the comparison is between two or more groups defined by differences in the exposure. Thus, for example, Sloan *et al*. [15] compared the assault and murder rates in Vancouver, Canada, and Seattle, USA, two generally similar cities which have different laws affecting access to firearms. This is a descriptive study, but as the comparison is between the 'exposure' to different legislative systems, the results can be expressed in the format of a cohort study and show the great excess of assault and murder related to firearms in Seattle compared with Vancouver, in contrast with similar rates of assault and murder involving other weapons (**Ex.** 3.11). Relative risks are calculated, and the other measures shown in Ex. 3.2 can be derived.

EXPOSURE-BASED COMPARISON OF DESCRIPTIVE DATA

Crime	Annual rate per 100 000 population: Seattle, USA	Vancouver, Canada	Relative risk
Aggravated assault using:			
Firearms	87.9	11.4	7.7
Knives	78.1	78.9	1.0
Other means	320.6	330.2	1.0
All aggravated assault	486.5	420.5	1.2
Murder (homicide) using:			
Firearms	4.8	1.0	5.1
Knives	3.1	3.5	0.9
Other means	3.4	2.5	1.3
All homicide	11.3	6.9	1.6

Ex. 3.11. Descriptive study with results in a cohort format. These annual rates of assault and murder in Vancouver (Canada) and Seattle (USA) represent the outcome; the city of residence, and its legal system, is the exposure. From Sloan *et al*. [15]

Other surveys may equate to the case–control design, where the essential comparison is between groups with and without the outcome of interest, in which case the format of results applicable to case–control designs is applied.

p. 70
p. 71

Self-test questions (answers on p. 494)

Q3.1 The trial of tuberculosis treatment shown in Chapter 1 showed four deaths from 55 patients treated by streptomycin, compared with 14 deaths from 52 controls. Calculate:

(a) The mortality rates in each group. What type of rates are these?

(b) The relative risk.

(c) The odds ratio.

(c) The attributable risk.

(d) The attributable benefit.

(e) The 'number needed to treat' for one death prevented.

Q3.2 In a cohort study of diet and heart disease in men, the study population is considered in two groups. In the high fat diet group there were 12 500 man-years of observation and 48 deaths; in the lower fat group there were 22 000 man-years of observation and 64 deaths.
Calculate:

(a) The mortality rates in each group.

(b) The relative risk.

(c) The attributable risk of the high fat diet.

(d) What proportion of disease in the high fat diet group is attributable to the high fat diet?

(e) What proportion of disease in the whole population group is attributable to the high fat diet?

Q3.3 If the incidence rate of asthma is 400 per 10 000 person-years in those who smoke, 100 in those who are exposed to air pollution, and 50 in those with neither exposure, what is the expected incidence in those with both exposures, assuming:

(a) an additive model

(b) a multiplicative model.

Q3.4 In a case–control study of arthritis of the knee, there are 500 cases, and 180 have a history of a previous knee injury; of 800 controls free from knee arthritis, 120 had a history of injury. Calculate:

(a) The relative risk.

(b) The odds ratio.

(c) The attributable proportion in those with a previous injury.

p. 72

Q3.5 ↳ If, in Q3.4, the cases were all newly diagnosed cases in a given community and the controls were a random sample of that community, each selected at the time a case occurred, what difference would that make to the answers to Q3.4 (a) and (b)?

Q3.6 If we assume that the effects seen in Ex. 3.11 represent an effect of the difference in gun laws, and if Canadian gun laws could be introduced to Seattle:

(a) What proportion of murders in Seattle, using firearms, would be prevented?

(b) How many people would be affected by the change for each firearm-linked murder prevented?

(c) What proportion of all murders in Seattle would be prevented?

References

1. Morrow J, Russell A, Guthrie E, *et al*. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. *J Neurol Neurosurg Psychiatry* 2006; **77**(2): 193–198. 10.1136/jnnp.2005.074203
WorldCat Crossref

2. ISIS-4 (Fourth International Study of Infarct Survival Collaborative Group). ISIS-4: A randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58 050 patients with suspected acute myocardial infarction. *Lancet* 1995; **345**: 669–685. 10.1016/S0140-6736(95)90865-X
WorldCat Crossref

3. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995; **310**: 452–454. 10.1136/bmj.310.6977.452
WorldCat Crossref

4. Edwards A, Elwyn G, Covey J, Matthews E, Pill R. Presenting risk information: a review of the effects of "framing" and other manipulations on patient outcomes. *J Health Commun* 2001; **6**: 61–82. 10.1080/10810730150501413
WorldCat Crossref

5. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ* 1994; **309**: 761–764. 10.1136/bmj.309.6957.761
WorldCat Crossref

6. Sarfati D, Howden-Chapman P, Woodward A, Salmond C. Does the frame affect the picture? A study into how attitudes to screening for cancer are affected by the way benefits are expressed. *J Med Screen* 1998; **5**: 137–140.
WorldCat

7. Romond EH, Perez EA, Bryant J, *et al*. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005; **353**: 1673–1684. 10.1056/NEJMoa052122
WorldCat Crossref

8. Paffenbarger RS, Hale WE. Work activity and coronary heart mortality. *N Engl J Med* 1975; **292**: 545–550. 10.1056/NEJM197503132921101
WorldCat Crossref

9. Royal College of General Practitioners' Oral Contraception Study. Mortality among oral-contraceptive users. Royal College of General Practitioners' Oral Contraception Study. *Lancet* 1977; **ii**: 727–731.
WorldCat

10. Beral V, Hermon C, Kay C, Hannaford P, Darby S, Reeves G. Mortality associated with oral contraceptive use: 25 year follow up of cohort of 46 000 women from Royal College of General Practitioners' oral contraception study. *BMJ* 1999; **318**: 96–100. 10.1136/bmj.318.7176.96
WorldCat Crossref

p. 73 11. Vessey M, Painter R, Yeates D. Mortality in relation to oral contraceptive use and cigarette smoking. *Lancet* 2003; **362**: 185–191. 10.1016/S0140-6736(03)13907-4
WorldCat Crossref

12. Elwood JM, Raman S, Mousseau G. Reproductive history in the mothers of anencephalics. *J Chronic Dis* 1978; **31**: 473–481. 10.1016/0021-9681(78)90011-5
WorldCat Crossref

13. Pauletto P, Puato M, Caroli MG, *et al*. Blood pressure and atherogenic lipoprotein profiles of fish-diet and vegetarian villagers in Tanzania: the Lugalawa study. *Lancet* 1996; **348**: 784–788. 10.1016/S0140-6736(96)01391-8

WorldCat Crossref

14. Charlton A. Children's coughs related to parental smoking. *BMJ* 1984; **288**: 1647–1649. 10.1136/bmj.288.6431.1647
WorldCat Crossref

15. Sloan JH, Kellermann AL, Reay DT, *et al*. Handgun regulations, crime, assaults, and homicide. *N Engl J Med* 1988; **319**:
p. 74 1256–1262. ↳ 10.1056/NEJM198811103191905
WorldCat Crossref