Cambridge Books Online

http://ebooks.cambridge.org/

Essential Epidemiology

An Introduction for Students and Health Professionals

Penny Webb, Chris Bain

Book DOI: http://dx.doi.org/10.1017/CBO9780511997303

Online ISBN: 9780511997303

Paperback ISBN: 9780521177313

Chapter

1 - Epidemiology is… pp. 1-28

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511997303.003

Cambridge University Press

1

Epidemiology is . . .

Box 1.1 Epidemiology is . . .

'The science of epidemics' (*Concise Oxford Dictionary*, 1964)
'The science of the occurrence of illness' (Miettinen, 1978)
'The study of the **distribution** and **determinants** of disease in humans' (MacMahon and Pugh, 1970)
'The study of the distribution and determinants of **health-related states or events** in specified populations, and the **application of this study to control of health problems**' (Porta, 2008)

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

So what is epidemiology anyway? As shown in Box 1.1, the *Concise Oxford Dictionary* (1964) defined it accurately, but not very helpfully, as 'the science of epidemics'. In 1970, MacMahon and Pugh came up with something a bit more concrete: 'the study of the *distribution* and *determinants* of disease'. Their definition succinctly identifies the two core strands of traditional epidemiology: *who* is developing disease (and *where* and *when*), and *why* are they developing it? The final definition, from the *Dictionary of Epidemiology* (Porta, 2008) takes it two steps further by broadening the scope to include health in general, not just disease, as well as highlighting the direct role of epidemiology in disease control.

Epidemiology, therefore, is about measuring health, identifying the causes of ill-health and intervening to improve health; but what do we mean by 'health'? Back in 1948, the World Health Organization (WHO, 1948) defined it as '. . . a state of physical, mental and social well-being'. Now, while this view is clearly what we hope for as individuals, the inclusion of 'mental and social well-being' would until recently have induced despair in epidemiologists. In practice what we usually measure is *ill-health* or disease: more disease equals poorer physical health, and this focus is reflected in the content of most routine reports of health data and in many of the health measures that we will consider here. However, methods that do attempt to capture the more elusive components of mental and social wellbeing are now emerging. Instead of simply measuring 'life expectancy', the WHO introduced the concepts of 'health-adjusted life expectancy' (HALE) and subsequently 'disability-adjusted life years' (DALYs) to allow better international comparisons of the effectiveness of health systems. In doing so they recognised that it is not longevity per se that we seek, but a long and healthy life. We will discuss these and other measures in more detail in Chapter 2.

Perhaps epidemiology's most fundamental role is to provide a logic and structure for the analysis of health problems both great and small. It also emphasises the sound use of numbers – we have to count and we have to think. We have to think about what is worth counting and how best to count it, about what is practical and, importantly, about how well we (or others) finally measured whatever it was we set out to measure, and what it all means. Accurate measurement of health is clearly the cornerstone of the discipline, but we believe the special value of epidemiology flows from a way of thought that is open, alert to the potential for error, willing to consider alternative explanations and, finally, constructively critical and pragmatic.

We offer this book as an aid to such thought. It does not aim to turn you into a practising epidemiologist overnight but will give clear directions if that is where you decide to go. Its primary goal is to help you interpret the mass of epidemiological literature and the various types of health data that you may come across. We hope that you will see, by reading and by doing, that the fundamental

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Table 1.1 Numbers of people who became ill after eating various foods at a youth camp.

	People who ate the food		People who didn't eat the food	
Food	Total	Number ill	Total	Number ill
Friday dinner:				
Hot chicken	343	156	231	74
Peas	390	175	184	55
Potato fries	422	184	152	46
Saturday lunch:				
Cold chicken	202	155	372	75
Salad	385	171	189	59
Saturday dinner:				
Fruit salad	324	146	250	84

(Adapted from Hook *et al.*, 1996, with permission from John Wiley and Sons.)

concepts and tools of epidemiology are relatively simple, although the tasks of integrating, synthesising and interpreting health information are more challenging. But before we go any further, let us do some public health epidemiology.

A case of food poisoning

Epidemiology is a bit like detective work in that we try to find out why and how disease occurs. Our first example illustrates this. After an outbreak of food poisoning at a youth camp, the local public health unit was called in to identify the cause (Hook *et al.*, 1996). They first asked everyone at the camp what they had eaten prior to the outbreak and some results of this investigation are shown in Table 1.1.

Looking at the numbers in Table 1.1, it is difficult to see which of the foods might have been responsible for the outbreak. (Note that everyone is recorded as either having eaten or not eaten each food; and that most people will have eaten more than one of the foods.) More people became ill after eating potato fries than after eating cold chicken (184 versus 155) – but then more people ate the fries (422 versus 202). How then can we best compare the two foods? One simple way to do this is to calculate the *percentage* of people who became ill among those who ate (or did not eat) each type of food. For example, 156 out of 343 people who ate hot chicken became ill and

$$156 \div 343 = 0.45 = 45\%$$

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

So 45% of people who ate hot chicken became sick. This is known as the **attack rate** for hot chicken, i.e. 45% of hot-chicken eaters were 'attacked' by food poisoning.

Calculate the attack rates for the other foods. Which food has the highest attack rate?

Although cold chicken has the highest attack rate (77%), not everyone who ate it (or, more precisely, who *reported* eating it) became ill and 20% or one in five people who did *not* eat cold chicken still became ill. This is to be expected; no matter what the cause of concern, it is rare that everyone who is exposed to it will show the effects (in this case, become ill). What can help here is to work out how much *more likely* people who ate a particular food were to become ill than those who did not eat it. For example, 45% of people who ate hot chicken became ill, compared with 32% of people who did not eat hot chicken. Hot-chicken eaters were therefore 1.4 times ($45\% \div 32\% = 1.4$) more likely to become ill than people who did not eat hot chicken. This measure gives us the risk of sickness in hot-chicken eaters *relative* to non-eaters, hence its name – **relative risk**.

Calculate the relative risk of developing food poisoning associated with each of the other food items. Which food is associated with the highest relative risk of sickness?

We can now conclude that the food item most likely to have been responsible for the outbreak was the cold chicken – people who ate this were almost four times more likely to become ill than those who did not. This is quite a strong relative risk; in comparison, eating any of the other foods was associated with no more than one and a half times the risk of disease. The relevant data, including the attack rates and relative risks, are summarised in Table 1.2, which is much more informative than the raw numbers of Table 1.1.

In identifying the cause of the outbreak you have just solved an epidemiological problem. The 'attack rates' and 'relative risks' that you used are fairly simple to calculate and are two very useful epidemiological measures. We will discuss them further in Chapters 2 and 5 and they will appear throughout the book.

Subdisciplines of epidemiology

The outbreak investigation above is an example of what might be called *public health epidemiology*, or *infectious disease epidemiology*, with the first name reflecting the broad field of application and the second the nature of both the aetiological (causal) agent and the disease. It is quite common now to specify such sub-fields of epidemiology, which range on the one hand from *nutritional* through *social* to *environmental epidemiology*, and on the other from *cancer* to

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Table 1.2 Numbers of people who became ill after eating various foods at a youth camp and attack rates and relative risks for each food.

	People who ate the food			People who didn't eat the food			
Food	Total	Number ill	Attack rate	Total	Number ill	Attack rate	Relative risk[a]
Friday dinner:							
Hot chicken	343	156	45%	231	74	32%	1.4
Peas	390	175	45%	184	**55**	30%	1.5
Potato fries	422	184	44%	152	46	30%	1.4
Saturday lunch:							
Cold chicken	202	155	77%	372	75	20%	3.8
Salad	385	171	44%	189	59	31%	1.4
Saturday dinner:							
Fruit salad	324	146	45%	250	84	34%	1.3

[a] Note, RR are calculated using the exact percentages and not the rounded values shown.
(Adapted from Hook *et al.*, 1996, with permission from John Wiley and Sons.)

injury or *perinatal epidemiology*: the former grouping being exposure-oriented and the latter focused on the particular disease or outcome. Nonetheless, the core methods and techniques of epidemiology remain common to all subdisciplines, so the contents of this book are relevant to all. Setting sub-speciality boundaries largely reflects the explosion of knowledge in these areas, although some areas do present special challenges. For example, capturing a person's usual diet is remarkably challenging and the subsequent data analysis equally so; epidemiologists coming fresh to the field of nutritional epidemiology will need to develop experience and expertise in that specific area. As you read on you will meet examples from a wide cross-section of health research and the common threads of logic, study design and interpretation will, we trust, become apparent.

It is of some interest to know a bit more about a few of the special epidemiologies. *Occupational epidemiology* has the longest history of all, with influential early observations of diseases linked to occupations such as mining appearing in the sixteenth century, and a systematic treatise on occupational diseases was published by Ramazzini back in 1700 (Rosen, 1958). Occupational health research in general, and epidemiology in particular, continue to contribute to enhancing workplace health today. Seminal contributions in the field include identification of the pulmonary (lung) hazards of asbestos for miners and construction workers (Selikoff *et al.*, 1965) and the work practices that led to an epidemic of a rare fatal cancer in workers in the polyvinyl chloride industry (Makk *et al.*, 1974). Company records of job tasks can provide measures of past exposure among employees, allowing researchers to look back in time and link, for

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

example, past asbestos exposure to subsequent deaths in the workforce. (This type of study is a *historical cohort design* – see Chapter 4. It is only possible when there are good records of both exposure and outcome, usually death, and for this reason has proved particularly useful in occupational studies where such records often do exist.)

Far more modern are the subdisciplines of *molecular epidemiology* and *clinical epidemiology*. The former aims to weld the population perspective of epidemiology with our rapidly increasing understanding of how variations in genes and their products affect the growth, form and function of cells and tissues. It thus has the potential for defining genetic contributions to disease risk and can also provide biological markers of some exposures (e.g. changes to DNA following exposure to tobacco smoke). In contrast, clinical epidemiology does not rely on advanced technology but differs from other branches of epidemiology in its focus on enhancing clinical decisions to benefit *individual patients*, rather than improving the health of *populations*. For this reason, clinical epidemiology is sometimes regarded as a completely separate discipline, a view which is encouraged by the fact that it has developed its own names for many standard epidemiological measures. The foundations are, however, identical to those of public health epidemiology and when appropriate we will discuss the two in parallel, highlighting any differences in language or approach along the way.

On epidemics

If we take the word 'epidemiology' itself, its origins from '*epidemic*' are clear. If we talk about an epidemic we immediately conjure up pictures of an acute outbreak of infectious disease but, both for practical and for etymological reasons, it seems reasonable to use the term to describe a notable excess of any disease over time. Many developed countries could, for example, be described as undergoing an epidemic of lung cancer over the last few decades (Figure 1.1). Notably the pattern of lung cancer over time differs for men and women; rates in men rose sharply between 1950 and 1980 but have been falling for some years now, while those in women rose later and are only just starting to fall – a consequence of the fact, that as a group, women took up smoking more recently than men. To describe this excessive occurrence of disease (or death) as an 'epidemic' captures some of the urgency the numbers demand.

The derivation of the word 'epidemiology' itself is from the Greek *epi*, upon, *demos*, the people, and *logia*, study. Literally, therefore, it means the 'study (of what is) upon the people'. Such study suggests a simple set of questions that have long lain at the heart of epidemiology.

- **What** disease/condition is present in excess?
- **Who** is ill?

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

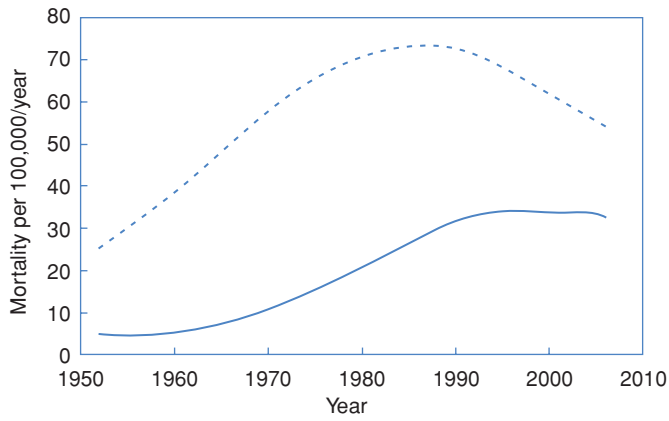

Figure 1.1 Time trends in lung cancer mortality rates (age-standardised to the 1970 US population) for white males (- - - - -) and females (——). (Drawn from: Devesa *et al.*, 1999 and CDC Wonder Database (CDC), accessed 8 January 2010.)

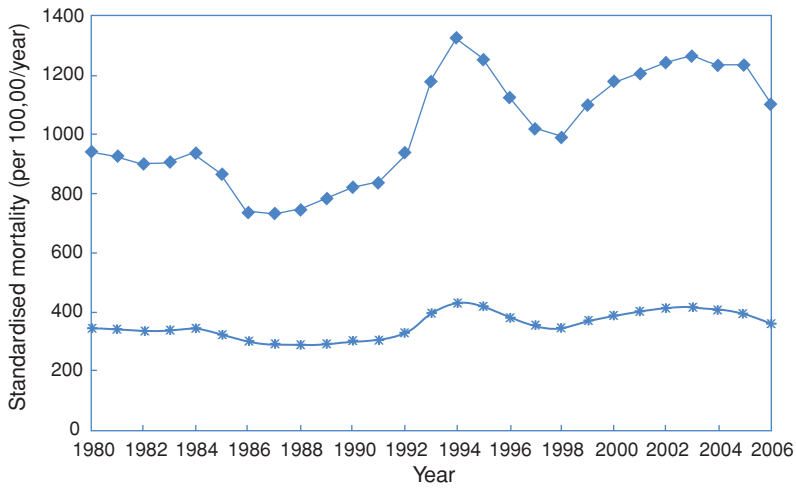

Figure 1.2 Changes in all-cause mortality at ages 0–64 years in the Russian Federation from 1980–2006 (♦, men; ∗, women). (Data from the European Health for All Database. WHO Regional Office for Europe, Copenhagen, Denmark, accessed via http://data.euro.who.int/hfadb/, 30 December 2009.)

- **Where** do they live?
- **When** did they become ill?
- **Why** did they become ill?

The first question reflects the need for a sound, common definition of a disease so that like is compared with like. Epidemiology is all about comparison – without some reference to what is usual, how can we identify excess? The next three questions form the mantra of **descriptive epidemiology**: '*person, place and time*'. As Figure 1.2 shows, an 'epidemic of premature mortality' occurred in the mid-1990s among young and middle-aged men in Russia and again in the early 2000s. This description captures the essence of the problem and prompts the next question: what caused these epidemics? What changed in the circumstances of younger Russian men to reverse the pattern of falling mortality in the

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Table 1.3 An historical event.

	Adult males		Adult females		Children (both sexes)		Total population	
SES[a]	Total	% Dead	Total	% Dead	Total	% Dead	Total	% Dead
High	175	67.4	144	2.8	6	–	325	37.5
Medium	168	91.7	93	14.0	24	–	285	58.6
Low	462	83.8	165	53.9	79	65.8	706	74.8
Other	885	78.3	23	13.0	0	–	908	76.7
Total	1690	80.0	425	25.6	109	47.7	2224	68.0

[a] SES, socioeconomic status.
(*Source:* http://www.anesi.com; for full details see reference list.)

early 1980s and then cause it to almost double in less than 10 years? And why did this happen again in the late 1990s? Other data show that there were no such mortality changes in Western Europe, nor among older Russian men or infants, nor (to the same extent) in Russian women. This simple graph captures a public health disaster for Russia and prompts urgent causal speculation: *why did this happen?* Solving and responding to this final question is critical for public health progress, but there is clearly no simple solution. In this case, a high proportion of the deaths were linked to excess consumption of alcohol: increases in mortality coincided with periods of economic and societal crisis, and rates fell when the economic situation improved (Zaridze *et al.*, 2009). This example highlights the central importance of paying close attention to descriptive data that provide a 'community diagnosis' or take the public health 'pulse' of a nation. Much can be gleaned from apparently simple data to give a quite precise description of the health event or state of interest, as the following exercise shows.

An historical epidemic

Table 1.3 shows some data that relate to an actual human experience. It tells you how many people there were in various age, sex and socioeconomic groups and what percentage of these people died during the 'epidemic'. The challenge is to use these data to describe the event systematically in terms of **whom** this happened to (we have no data on place or time) and then to think about the sort of event that might have induced such a pattern.

The following questions are designed to help you identify key features of the data.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

1. What is distinctive about this isolated population with regard to
   - the numbers of men and women (sex distribution),
   - the numbers of adults and children (age distribution) and
   - the numbers in each socioeconomic group (socioeconomic distribution)?
2. What strikes you about the percentage of people who died (the 'death rate')? Is this different for (a) adults and children, (b) men and women, (c) high and low socioeconomic status (SES) and (d) any particular combinations of the above?
3. How many times more likely were
   - men to die than women and
   - those of low SES to die than those of high SES?
4. To what historical event might these data refer?

Table 1.3 displays more complicated data than Table 1.2, since you had to consider the joint effects of three factors (sex, socioeconomic status and age) on mortality. The sequence of questions aims to underline a general principle in describing such tables, i.e. to look at overall patterns first, then move on to more detail (see Box 1.2 on the next page). We all see things in different ways, but until you develop your own style this approach is one that can help you avoid becoming lost in the array of possible relationships. You need first to grasp the size of the *whole group* under study and how many died; then check the overall patterns (numbers and mortality rates) across each 'exposure' separately (sex, SES, age). These are sometimes called the 'marginal' rates based on row and column totals; e.g. first look at the rates for all adult males, ignoring their SES, or for all people of high SES, ignoring their age and sex. Only then consider the more complex 'inner' set of joint effects such as the influence of SES on mortality among women.

In tackling this and the previous problem you have already done some serious epidemiology: you have *described* data, *interpreted* the patterns you observed and used *epidemiological measures* to help do this. We will build on this throughout the book, but first let's step back a little and see what other lessons we can learn from the past.

The beginnings[1]

The 'great man' approach has fallen out of favour in modern historical practice; however, linking historical events to people adds character so we will focus on some of the main players in this brief overview of the development of population health and epidemiology.

[1] The material in this section is drawn from a mix of primary and secondary sources, with the latter including a number of texts, most helpful being those of Stolley and Lasky (1995) and Lilienfeld and Lilienfeld (1980).

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.2 An historical event

Things to note about the population include

- the predominance of adult males (1,690 ÷ 2,224 = 76%), the much smaller proportion of adult females (19%) and the very few children;
- the substantial excess of persons of low SES (men and children in particular); and
- the total population (2,224) is quite large – a village, small town, an army barracks . . . ?

Things to note about the 'death rates' include the following.

- The overall death rate is very high – more than two-thirds died. (Note: these death rates are essentially identical in form to the attack rates in Table 1.2.)
- Overall, death rates increased with decreasing SES.
- The death rate in men (80.0%) was much higher than that in women (25.6%); the death rate in children was between these two.
- In men, the death rate was high in all socioeconomic classes, although those of high SES fared better than the rest; in women, the death rate was always less than that for males of equivalent SES, but it increased strikingly from high to medium to low SES.
- The only children to die were of low SES.

Overall, the relative risk (RR) for men versus women is	80.0 ÷ 25.6 = 3.1
The RR for low versus high SES is	74.8 ÷ 37.5 = 2.0
The RR for women of low SES versus women of high SES is	53.9 ÷ 2.8 = 19.3
The RR for men of low SES versus women of high SES is	83.8 ÷ 2.8 = 29.9

A disaster has occurred, causing a high death rate that predominantly affected men (of all social classes) and, to a lesser extent, women and children of low social class. Overall there is a modest benefit of belonging to a higher social stratum, and among women this protection was exceptionally strong (a 19-fold higher risk of dying for low versus high SES).

Such substantial differences in risk reflect powerful preventive effects and in this instance it was a mix of social custom and the physical consequences of social stratification. The event was the sinking of the Titanic, where those of higher SES (the first class passengers) were situated on the upper decks and were therefore closer to the lifeboats than those of medium and low SES (those travelling second and third class, respectively). The males gallantly helped the females and children into the lifeboats first. Those of 'other' SES were the crew.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.3 On airs, waters and places

Whoever wishes to investigate medicine properly, should proceed thus: in the first place to consider the seasons of the year, and what effects each of them produces . . . Then the winds, the hot and the cold, especially such as are common to all countries, and then such as are peculiar to each locality. We must also consider the qualities of the waters . . . In the same manner, when one comes into a city to which he is a stranger, he ought to consider its situation, how it lies as to the winds and the rising of the sun; for its influence is not the same whether it lies to the north or the south, to the rising or to the setting sun. These things one ought to consider most attentively, and concerning the waters which the inhabitants use, whether they be marshy and soft, or hard, and running from elevated and rocky situations, and then if saltish and unfit for cooking; and the ground, whether it be naked and deficient in water, or wooded and well watered, and whether it lies in a hollow, confined situation, or is elevated and cold; and the mode in which the inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labour . . .

(Extracted from Hippocrates of Cos, 400 BC.)

Good epidemiological practice and reasoning started long ago. Perhaps the first proto-epidemiologist (*proto* because he did not actually count anything) was Hippocrates of Cos (460–375 BC), who recognised that both environmental and behavioural factors could affect health (see Box 1.3).

The Dark Ages and Middle Ages (AD 500–1500) have little to say to us, other than in the development of causal reasoning, which we will set aside until later in the book (Chapter 10). The introduction of more quantitative methods into epidemiology, and, in fact, into biology and medicine in general, has been attributed to John Graunt (1620–1674), a haberdasher and early Fellow of the Royal Society in London, although his friend William Petty may well have been a seminal influence too. In 1662 Graunt published his *Natural and Political Observations Mentioned in a Following Index and Made Upon the Bills of Mortality*. He studied parish christening registers and the 'Bills of Mortality', and noted many features of birth and death data, including the higher numbers of both male births and deaths in comparison with females, the high rates of infant mortality and seasonal variations in mortality. He also provided a numerical account of the impact of the plague in London and made the first attempts to estimate the size of the population. In an attempt to define a 'law of mortality' he constructed the first life-table (Table 1.4). This summarised the health of a population in terms of the chance of an individual surviving to a particular age. Notice that at this time only

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Table 1.4 An historical example of a life-table.

Exact age (years)	Deaths	Survivors	Chance of living to that age (%)
0	–	100	
6	36	64	64
16	24	40	40
26	15	25	25
36	9	16	16
46	6	10	10
56	4	6	6
66	3	3	3
76	2	1	1
86	1	0	

(Adapted from Graunt, 1662.)

three out of every hundred people reached the age of 66, and the majority of deaths occurred in early life. This technique was a forerunner of that used by life insurance companies for calculating insurance premiums today, as well as a fundamental approach to measuring a population's health. As you will see when we come back to consider life-tables in more detail in Chapter 2 (see also Appendix 5 for details of how to construct a life-table), things have improved considerably since Graunt's time, with about 85 of every 100 men and 90 of every 100 women now making it to the age of 66 in developed countries such as Australia.

During the nineteenth century, the collection and use of health statistics for what we now call 'descriptive epidemiology' continued to develop in England and also, briefly, in France. Of particular influence as a teacher was Pierre Charles-Alexandre Louis (1787–1872), who conducted some of the earliest epidemiological studies of treatment effectiveness when he demonstrated that bloodletting did not aid recovery from disease. Among his students was William Farr (1807–1883), physician, statistician and director of the Office of the Registrar General for England and Wales from 1837, its second year of operation. Farr studied levels of mortality in different occupations and institutions and in married and single persons, as well as other facets of the distribution of disease. He published these and other findings in the *Annual Reports of the Registrar General*, and the present UK system of vital statistics stems directly from his work.

John Snow (1813–1858), a physician and contemporary of Farr, was better known at the time for giving chloroform to Queen Victoria during childbirth, but is now remembered for his pioneering work in elucidating the mode of transmission of cholera (Snow, 1855). This remains a classic and exciting example of

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

epidemiological detection and some of Snow's personal account of it is given below and again later in the chapter. His initial observations were based on a series of reports of individual cases of cholera and, in every instance, he was able to link the case to contact with another infected person (or their goods), thereby demonstrating that the disease could spread from person to person. He then surmised, contrary to popular belief at the time, that cholera could be transmitted through polluted water, a view that was strengthened by his observations linking a terrible outbreak of cholera around Broad Street, London, in 1854, to the local water pump (Box 1.4).

Box 1.4 John Snow and the Broad Street Pump (1854)

Within two hundred and fifty yards of the spot where Cambridge Street joins Broad Street, there were upwards of five hundred fatal attacks of cholera in ten days . . . The mortality would undoubtedly have been much greater had it not been for the flight of the population . . . so that in less than six days from the commencement of the outbreak, the most afflicted streets were deserted by more than three-quarters of their inhabitants.

There were a few cases of cholera in the neighbourhood of Broad Street, Golden Square, in the latter part of August; and the so-called outbreak, which commenced in the night between the 31st of August and the 1st of September, was, as in all similar instances, only a violent increase of the malady. As soon as I became acquainted with the situation and extent of this eruption of cholera, I suspected some contamination of the water of the much-frequented street-pump in Broad Street . . . but on examining the water . . . I found so little impurity in it of an organic nature, that I hesitated to come to a conclusion. Further inquiry, however, showed me that there was no other circumstance or agent common to the circumscribed locality in which this sudden increase of cholera occurred, and not extending beyond it, except the water of the above mentioned pump.

On proceeding to the spot, I found that nearly all the deaths had taken place within a short distance of the pump. There were only ten deaths in houses situated decidedly nearer to another street pump. In five of these cases the families of the deceased persons informed me that they always sent to the pump in Broad Street, as they preferred the water to that of the pump which was nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street. Two of them were known to drink the water; and the parents of third think it probable that it did so. The other two deaths, beyond the district which this pump supplies, represent only the amount of mortality from cholera that was occurring before the irruption

(*continued*)

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.4 (*continued*)

took place . . . (*Snow used a spot map to show the spread of cases in relation to this and other pumps.*) I had an interview with the Board of Guardians of St James's parish, on the evening of Thursday, 7th September, and represented the above circumstances to them. In consequence of what I said, the handle of the pump was removed on the following day.

Snow was also able to explain why some groups of people within the area did not develop cholera:

The Workhouse in Poland Street is more than three-fourths surrounded by houses in which deaths from cholera occurred, yet out of five hundred and thirty-five inmates, only five died of cholera, . . . The workhouse has a pump well on the premises, . . . and the inmates never sent to Broad Street for water. If the mortality in the workhouse had been equal to that in the streets immediately surrounding it on three sides, upwards of one hundred persons would have died. (*Note Snow's comparison of the 'observed' number of cases with the number 'expected'.*)

There is a Brewery in Broad Street, near to the pump, and on perceiving that no brewery men were registered as having died of cholera, I called on Mr Huggins, the proprietor. He informed me that there were above seventy workmen employed in the brewery, and that none of them had suffered from cholera . . . The men are allowed a certain quantity of malt liquor, and Mr Huggins believes they do not drink water at all . . .

The limited district in which this outbreak of cholera occurred, contains a great variety in the quality of the streets and houses; Poland Street and Great Pulteney Street consisting in a great measure of private houses occupied by one family, whilst Husband Street and Peter Street are occupied by the poor Irish. The remaining streets are intermediate in point of respectability. The mortality appears to have fallen pretty equally amongst all classes, in proportion to their number.

(Extracted from Snow, 1855.)

Snow went to a lot of trouble to explain why some people developed cholera when they were believed *not* to have drunk the water from the Broad Street pump. He attributed these cases to the use of water from the pump in the local public houses, dining rooms and coffee shops. He was also able to explain why some groups of people within the area did not develop cholera when they lived in the affected area. If these low-risk groups (brewery workers, workhouse dwellers) had been users of the nearby Broad Street pump, Snow's hypothesis would have been in tatters. His findings among the 'exceptions' of both sorts thus bolster his arguments considerably: for the most part he found convincing explanations

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

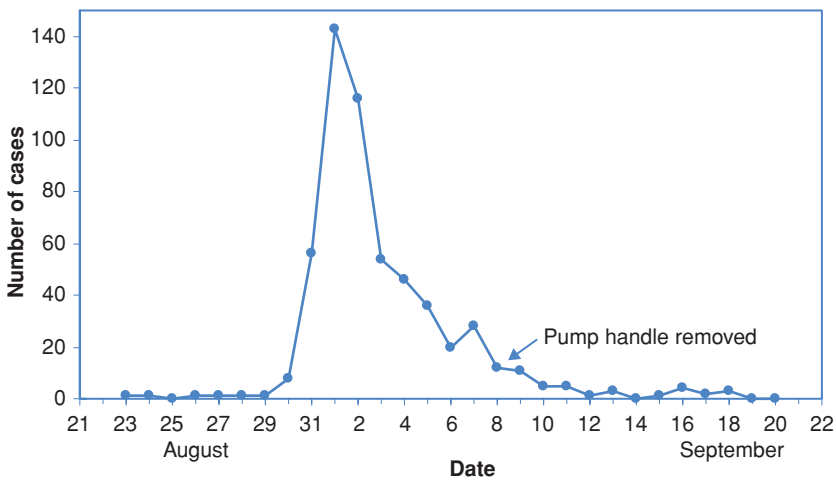

Figure 1.3 The Broad Street cholera epidemic, 1854. (Drawn from: Snow, 1855.)

for why some people apparently at risk did not fall ill, and so too for the small group not living near the pump who did contract cholera. His openness to collecting all the facts, not just those that obviously supported his contention, is a salutary reminder of what constitutes good science – and that effective public health action requires realistic information about the problem at hand.

In addition to mapping the distribution of cases by place, Snow tabulated the numbers of cases and deaths over time. His time data are displayed graphically, showing what is called an 'epidemic curve', in Figure 1.3.

When did the epidemic start? When did it end? What role did Snow's dramatic removal of the pump handle on 8 September play in interrupting its course?

The epidemic curve shows that the rise above the preceding baseline began on 30 August, with a dramatic increase over the next two days. And although the fall from the peak starts shortly thereafter, case numbers are high for quite some days later, not getting close to the preceding baseline until two weeks from the commencement. The epidemic had waned substantially before Snow's intervention on 8 September, probably largely due to the flight of much of the populace. However, since the graph shows the total *number* of cases occurring and does not take into account the size of the population, the *rate* of disease (the number of new cases occurring among the smaller number of people remaining in the area) could still have been fairly high. Snow's action may therefore truly have contributed to containment of the outbreak.

The second half of the nineteenth century saw the expansion of epidemiology in the direct service of public health in the UK, with a similar trend in the USA starting early the next century. Infectious diseases remained the core

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

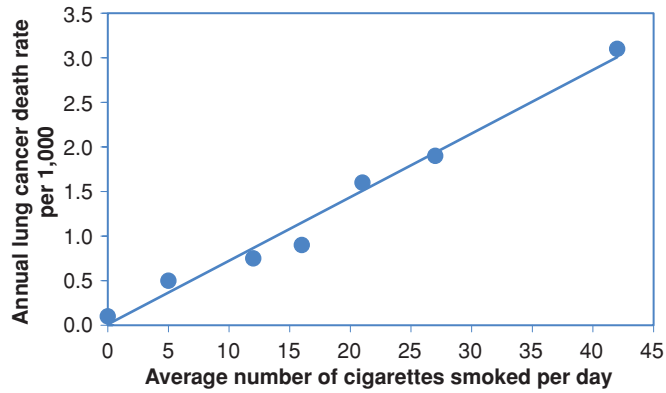

Figure 1.4 Age-standardised death rates from lung cancer in relation to the number of cigarettes smoked per day, British Doctors Study, 1951–1961. (Reproduced from: Doll and Hill, *BMJ*, 1964;1: 1399–1410, with permission from BMJ Publishing Group Ltd.)

interest until the early 1900s when Joseph Goldberger, a Hungarian physician working in the US Public Health Service, showed that pellagra[2] was not infectious but of dietary origin and Wade Hampton Frost, another pioneer in the field, articulated the value of non-experimental epidemiology in discovering disease origins. Then in 1950, the publication of two case–control studies of lung cancer, by Richard Doll (epidemiologist) and Austin Bradford Hill (statistician) in the UK and Ernest Wynder (medical student) and Evart Graham (surgeon) in the USA, publicly marked the start of modern epidemiology.

Both papers (Doll and Hill, 1950; Wynder and Graham, 1950) showed that patients with lung cancer (*cases*) tended to smoke much more than people without lung cancer (*controls*). Doll and Hill then set out to confirm their findings using a different, prospective design (a *cohort study*). They wrote to a large number of British doctors to find out how much they smoked and then 'followed' them (by mail and death records) over subsequent years to see what they died from. They again showed quite clearly that those who smoked cigarettes were much more likely to die of lung cancer than those who did not smoke, and the more they smoked the higher their risk (Figure 1.4). What is now known as the 'British Doctors Study' ran for more than 50 years (Doll *et al.*, 2004). Unfortunately, in spite of this and other clear evidence of the harmful effects of smoking, it was many years before attempts to discourage people from smoking were made, and it is only recently that tobacco companies have begun to admit that their products cause disease.

Twenty years after those key case–control studies came the publication of the first comprehensive and widely influential disciplinary text: *Epidemiology: Principles and Methods* by Brian MacMahon and Thomas Pugh (1970). Highly readable and erudite, it remains a benchmark for successors.

[2] A disease common in poorer areas, characterised by diarrhoea, dermatitis, dementia and ultimately death.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

What does epidemiology offer?

You will have discerned parts of the answer to this question from what you have already read and done in reaching this point. Here we recap and expand to bring the elements together more directly and fully, and thereby effectively map the content of the rest of the book. This section sketches the broad purposes of epidemiology and the next aims to illuminate these through some concrete examples.

A large part of public health is about identifying health problems within a community (who is becoming ill, where and when?), identifying what is causing the problems and then testing possible solutions to try to resolve or reduce the problem. Epidemiology is fundamental in providing the data needed to make public health judgements in each of these areas and the data come from studies of 'populations' (groups of people) of all sorts and sizes. Epidemiology largely deals with descriptions and comparisons of groups of people who vary widely in their genetic make-up, behaviour and environments. The great challenge for epidemiologists is to deal with such multiple influential health-modifiers in a systematic and logical way that produces information of practical value (to improving a community's health). How this challenge is met is what this book is all about.

Description of health status of populations

The observation and recording of health status makes it possible to identify sudden (and not-so-sudden) changes in the level of disease over time that might point to a need for action or further investigation. Similarly, differences between groups of people in one area, or between different geographical areas, can also give clues regarding the causes of disease (or health) in those groups. Such *descriptive statistics* are also important for health authorities and planners who need to know the nature and size of the health challenges faced by their communities.

Causation

Once a problem has been identified, we need to know what causes it, and probably the best-recognised use of epidemiology is in the search for the causes of disease. In some cases strong genetic factors have been identified, as for example with cystic fibrosis, a lung disease that occurs because of specific genetic defects. In other instances major environmental factors are crucial, such as asbestos in the development of lung mesothelioma (a rare form of lung cancer). In general, though, there is almost always some interaction between genetic and environmental factors in the causation of disease. Epidemiological tools are central to

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

the identification of modifiable factors that will allow preventive interventions. (Note that in epidemiology and public health there remains some confusion over what is meant by *environmental factors*. We, and most others, take this to mean the sum of all non-genetic factors, including psychological, behavioural, social and cultural traits.)

Evaluation of interventions

Once we have identified a factor that causes disease, we then want to know whether we can reduce a population's exposure to this factor and so prevent the occurrence of disease – a 'primary' prevention programme (we will discuss prevention further in Chapter 14). Epidemiology has a core role to play in this process and is also key to the evaluation of different treatments for a particular disease (an aspect of both mainstream and clinical epidemiology) and assessments of the effectiveness of health services.

Natural history and prognosis

Epidemiologists are also concerned with the *natural history*, or the course and outcome, of disease, both in individuals and in groups. *Prognosis* often implies the course of disease after treatment, but the terms tend to be used rather interchangeably. Such knowledge has obvious value for discussing treatment options with individual patients, as well as for planning and evaluating interventions. Of particular interest is whether early disease is present for long before symptoms drive someone to seek medical attention. If this 'sub-clinical' disease can be detected and if, as a result, treatment is more effective, this opens the way for screening programmes that aim to improve treatment outcomes. (We will discuss screening further in Chapter 15.)

What do epidemiologists do?

How then are these objectives of epidemiological research attained? Let us look briefly at some more examples of what the practice of epidemiology can yield across some of its main dimensions.

Descriptive studies: person, place and time

By 'person'

In some countries there is concern over health differences between indigenous people and the rest of the population. Figure 1.5 shows Australian mortality data comparing Indigenous with non-Indigenous people. The bars show how many times higher mortality from circulatory, respiratory and infectious

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

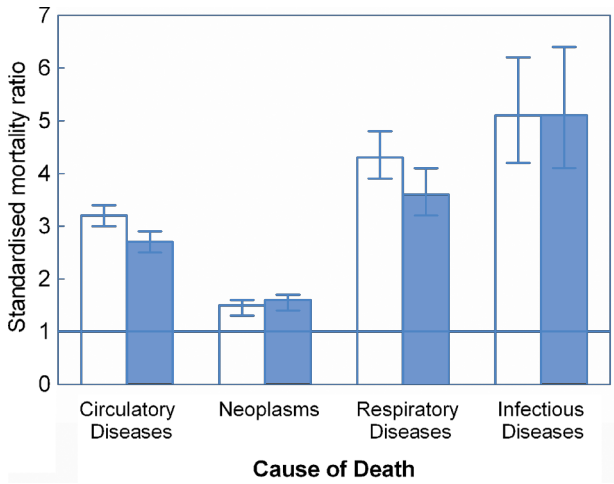

Figure 1.5 Age-standardised mortality ratios for selected diseases in the Indigenous compared to the non-Indigenous population in Australia, 2001 to 2005. (Drawn from: Pink and Allbon, 2008.) The bars indicate how much higher mortality was among Indigenous men (open bars) and women (solid bars) compared to non-Indigenous people. The vertical lines indicate 95% confidence intervals for these estimates (a measure of how certain the figures are).

diseases and cancer is in Indigenous men and women in Australia compared to non-Indigenous Australians (the horizontal line at the level '1' indicates the point where mortality rates in Indigenous and non-Indigenous people would be equal).

How many times higher is mortality from circulatory diseases in Indigenous males than in non-Indigenous males?

What is the obvious striking fact about relative mortality in Indigenous people in general?

Mortality for circulatory diseases in Indigenous men is just over three times that in non-Indigenous men and the difference for women is almost as great. The data presented indicate a much worse health situation for Indigenous Australians than for the non-Indigenous population. (Note: these **standardised mortality ratios** are similar to the relative risk in the food poisoning example earlier. They show how many times more likely it was for an Indigenous Australian to die compared with a non-Indigenous Australian in 2001–2005. The process of standardisation also takes account of the fact that Indigenous Australians are, on average, younger than non-Indigenous people. We will discuss these measures further in Chapter 2.)

By 'place'

How 'healthy' is any given country in relation to the rest of the world – are things better or worse there compared with other countries? Figure 1.6 shows cardiovascular disease mortality rates in males in different countries. You can see that men in the Netherlands, for example, are considerably better off than those in the UK, New Zealand and particularly Poland and Hungary; but things could be

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

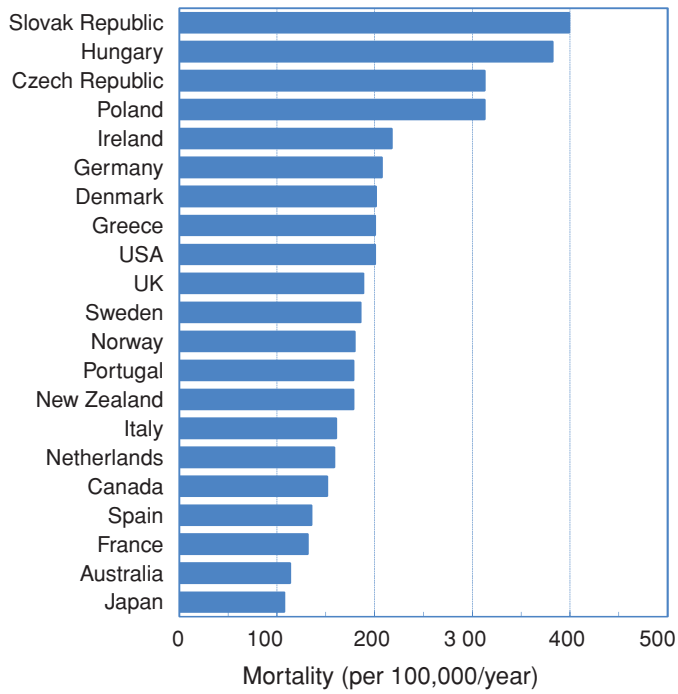

Figure 1.6 Circulatory disease mortality for males, about 2001. (Drawn from: AIHW, 2008.)

better – as shown by the lower rates in Spain, France, Australia and Japan. What is it about Japanese men that makes them less likely to die of cardiovascular disease than Dutch men? If we can work this out then perhaps we could reduce cardiovascular mortality in the Netherlands to the level seen in Japan (provided that the differences are not purely genetic). By studying patterns of disease and relating them to variations in risk factors for the disease we can come up with possible reasons why some people or places have higher rates of disease than others or why disease rates have changed over time.

By 'time'

What emerges if we look at the changing patterns of mortality in a country over time? The graph in Figure 1.7 shows mortality trends for selected conditions and groups over almost three decades (1979–2006) in the USA.

What are the most notable features of Figure 1.7?

The picture we see is mixed, some good news, some concerning. The most obvious health success story is the consistent downward trend in deaths from heart attacks, with more than 100 fewer people in every 100,000 (half as many) dying from them at the end of the period. A less dramatic decline is seen for motor vehicle accidents. Deaths from AIDS rose until 1995 and have fallen since

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

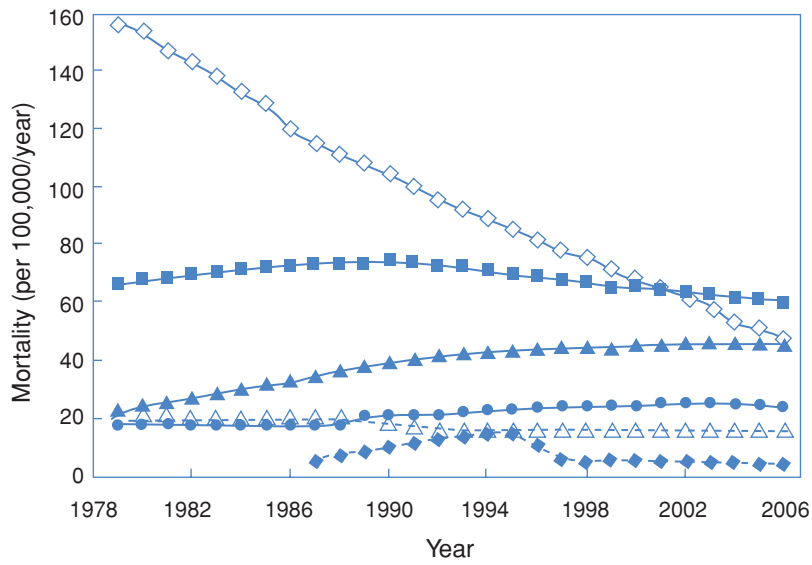

Figure 1.7 US mortality rates, 1979–2006, for heart attack (◇), lung cancer in males (■), lung cancer in females (▲), motor vehicle accidents (△), diabetes (●) and AIDS (♦). (Drawn from CDC Wonder (CDC), accessed 8 January 2010.)

(an epidemic where perhaps the worst is past, at least for the USA). The same is true for lung cancer in men, although on this scale it is not striking. Most worrying is the steady rise in lung cancer deaths among women.

However, these details don't give us the big picture. Some up, some down, some changing direction: what was happening to overall mortality in the USA during the period? Total mortality rates fell from about 1,000 to 810 per 100,000 per year, but we would not be able to fit this information onto the same graph without losing almost all the details we noted above. We could, of course, draw a separate graph showing the total death rate, but we can do both by changing the scale of the vertical axis, as in Figure 1.8.

Instead of a linear scale (1, 2, 3, 4, . . .), we have now used a 'log' (logarithmic) scale (1, 10, 100, 1,000, . . .) where the distance between 1 and 10 (a 10-fold difference) is the same as the distance between 10 and 100 (also a 10-fold difference) and so on. Now we can fit mortality rates as different as 4.0/100,000 (AIDS mortality in 2006) and 1,000/100,000 (all-cause mortality in 1979) on the same page. It also allows us to compare *relative* changes in mortality rates directly, with parallel slopes reflecting equal rates of change. The fall in heart attacks looks much less dramatic now: the drop is only about 2%–3% per year but, as Figure 1.7 showed, this led to a large absolute benefit, because the death rate was so high to start with. The rate of change for AIDS looks much steeper on a log scale because the percentage change is greater, but the absolute benefits are clearly much less. In public health we need to think on both relative and absolute scales: they tell us different things that are useful for different purposes. We will take this further later in the book.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

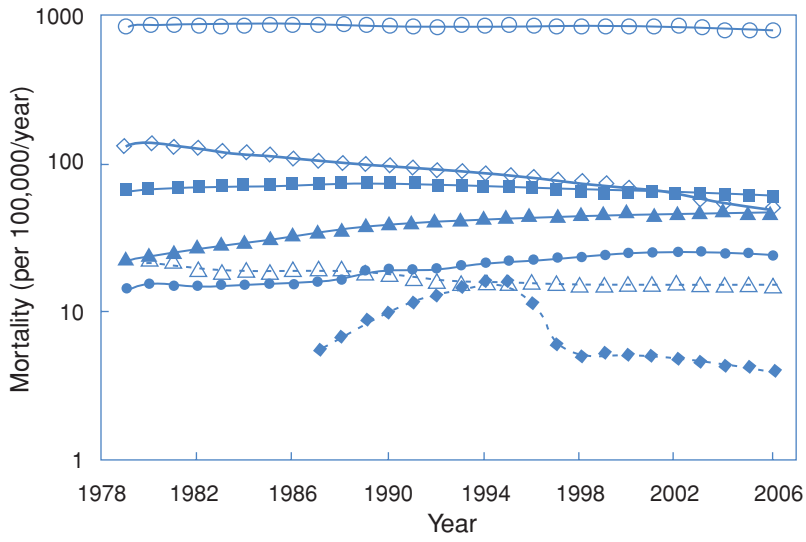

Figure 1.8 US mortality rates (log scale), 1979–2006 for all causes (◯), heart attack (◇), lung cancer in males (■), lung cancer in females (▲), motor vehicle accidents (△), diabetes (•) and AIDS (♦) (Drawn from: CDC Wonder (CDC), accessed 8 January 2010.)

Box 1.5 Smallpox

The elimination of smallpox had a major impact on the health of millions of people, especially in many of the poorest countries. Descriptive epidemiology played a major role by providing information about the distribution of cases (*jointly by person, place and time*) and levels of transmission, by mapping outbreaks and by evaluating control measures. In 1967, there were 10–15 million new cases and 2 million deaths from smallpox in 31 countries. By 1976, smallpox was being reported in only two countries and the last naturally occurring case was recorded in 1977. Elimination of this scourge was helped by simple but painstaking case-finding and counting.

See Box 1.5 for a practical example of how simple descriptive epidemiology can help to solve a major global health problem.

Analytic studies

Ideas generated by such *descriptive* work can then be tested further in *analytic* studies, looking for associations between potential causal agents and diseases. This research is based on facts collected from groups of individuals, not large-scale population statistics. Are people with higher blood pressure more likely to develop coronary heart disease than those with normal blood pressure? Are people who smoke more likely to develop lung cancer than those who do not? Even

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.6 The Nurses' Health Study

This cohort study of 120,000 US nurses was started in 1976 by Frank Speizer of the Channing Laboratory, Harvard Medical School. The study was initially funded for five years to study whether the oral contraceptive pill caused breast cancer, but the nurses are still being followed over 30 years later. Hundreds of scientific papers have been published, covering scores of diseases and exposures and investigating their inter-relationships. The study has been particularly influential in the field of diet and disease (nutritional epidemiology), owing to diet questionnaires that the nurses have been completing since 1980. As with other long-term follow-ups of cohorts, such as the British Doctors Study of Doll and Hill, its success is jointly dependent on the enthusiasm and commitment of researchers and participants. For the latter this has extended to providing blood, toenail clippings (for measurement of trace metals) and samples of tapwater over the years! This human side to epidemiology does not feature much in textbooks but is fundamental to successful fieldwork.

more usefully, how *much* more likely is a smoker to develop lung cancer than a non-smoker? Does risk depend on the number of cigarettes smoked? That is, how *strong* is the effect of the exposure; and does it increase with higher levels of exposure? In the British Doctors Study mentioned earlier, Doll and Hill found that the risk of lung cancer increased steadily as people smoked more cigarettes (Figure 1.4). This adds weight to the idea that smoking cigarettes really does affect the chance that an individual will develop lung cancer. In Box 1.6 you will find a brief account of another cohort study which has studied many exposures and diseases over the past three decades.

Once we have found an association, the challenge is then to evaluate this in order to determine whether something really *causes* disease or is linked to it only secondarily. If we find that people with a peptic ulcer drink a lot of milk, does this mean that drinking milk causes ulcers or simply that people with an ulcer drink milk to ease their pain? In Chapter 10 we will look more deeply at this challenge.

Intervention studies

Finally, epidemiologists evaluate new preventive measures, programmes or treatments that are designed to reduce ill health or promote good health. They also monitor the effectiveness of these 'intervention' programmes after they have been implemented: do they actually achieve the good they set out to do? These programmes can include evaluations of different health promotion strategies targeted at individuals or whole communities, or clinical trials of new drugs designed to cure disease. Does taking aspirin reduce your chance of having a

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

heart attack? Which of several strategies is better at helping people give up smoking? Is one drug better than another for treating a heart attack?

A natural experiment

We will end this chapter with another example from John Snow's *On the Mode of Communication of Cholera* (1855) because, although this text is more than 150 years old, the methods he used and his combination of flair, skill, logic and dogged persistence remain the cornerstones of modern epidemiology. His work also exemplifies, in more detail than modern papers, the logical dissection of evidence about disease patterns to identify practical preventive strategies – which is still the key function of epidemiology – and it gives an excellent sense of the role and utility of epidemiology in practical public health.

In the early 1850s, London was cholera-free for a number of years and during that period one of the major water supply companies (the Lambeth Company) moved their waterworks out of London, thereby obtaining water free of the sewage of the city. During the next major cholera outbreak in 1853–1854 Snow was able to obtain information about the number of deaths occurring in the different sub-districts of London and he found that cholera mortality was lower in areas supplied by water from the Lambeth Company than in those supplied by the Southwark and Vauxhall water company which continued to take water from Battersea in the city. He did not stop there, but went on to conduct his 'Grand Experiment' (see Box 1.7).

Conclusions

Again we have a vivid picture of a master epidemiologist at work. Not satisfied that his hypothesis had been adequately tested, Snow identified the opportunity to conduct an even more rigorous test – his '*Grand Experiment*' – and in doing so he addressed the major epidemiological issues that still concern us today.

- He identified a situation in which people were unknowingly divided into two groups differing only in the source of their water, thereby creating what was effectively a **randomised trial** (we will look at the different types of epidemiological study in Chapter 4).
- In doing so, he realised the importance of ruling out other differences between the groups (e.g. sex, age, occupation, socioeconomic status) that could explain any mortality differences (a problem known as **confounding** that we will come back to in Chapter 8).
- He worked long and hard to acquire *accurate information* about both the water supply and the number of cholera deaths in each house – we will

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.7 A grand experiment

Although the facts . . . afford very strong evidence of the powerful influence which the drinking water containing the sewage of a town exerts over the spread of cholera, when that disease is present, yet the question does not end here; for the intermixing of the water supply of the Southwark and Vauxhall Company with that of the Lambeth Company, over an extensive part of London, admitted of the subject being sifted in such a way as to yield the most incontrovertible proof on one side or the other . . . A few houses are supplied by one Company and a few by the other, according to the decision of the owner or occupier at that time when the Water Companies were in active competition . . . Each Company supplies both rich and poor, both large houses and small; there is no difference either in the condition or occupation of the persons receiving the water of the different Companies. Now it must be evident that, if the diminution of cholera, in the districts partly supplied with the improved water, depended on this supply, the houses receiving it would be the houses enjoying the whole benefit of the diminution of the malady, whilst the houses supplied with the water from Battersea Fields would suffer the same mortality as they would if the improved supply did not exist at all. As there is no difference whatever, either in the houses or the people receiving the supply of the two water Companies, or in any of the physical conditions with which they are surrounded, it is obvious that no experiment could have been devised which would more thoroughly test the effect of water supply on the progress of cholera than this which circumstances placed ready made before the observer.

The experiment, too, was on the grandest scale. No fewer than three hundred thousand people of both sexes, of every age and occupation, and of every rank and station, from gentlefolk down to the very poor, were divided into two groups without their choice, and, in most cases, without their knowledge; one group being supplied with water containing the sewage of London, and amongst it, whatever might have come from the cholera patients, the other group having water quite free from such impurity.

To turn this grand experiment to account, all that was required was to learn the supply of water to each individual house where a fatal attack of cholera might occur . . .

The epidemic of 1854

When the cholera returned to London in July of the present year . . . I resolved to spare no exertion . . . to ascertain the exact effect of the water supply on the progress of the epidemic, in the places where all the circumstances were so happily adapted for the inquiry . . . I accordingly asked permission at the

(*continued*)

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Box 1.7 (*continued*)

General Register Office to be supplied with the addresses of persons dying of cholera, in those districts where the supply of the two Companies is intermingled in the manner I have stated above . . . I commenced my inquiry about the middle of August with two sub-districts of Lambeth . . . There were forty-four deaths in these sub-districts down to 12th August, and I found that thirty-eight of the houses in which these deaths occurred were supplied with water by the Southwark and Vauxhall Company, four houses were supplied by the Lambeth Company, and two had pump-wells on the premises and no supply from either of the Companies.

As soon as I had ascertained these particulars, I communicated them to Dr Farr, who was much struck with the result, and at his suggestion the Registrars of all the south districts of London were requested to make a return of the water supply of the house in which the attack took place, in all cases of death from cholera. This order was to take place after the 26th of August, and I resolved to carry my inquiry down to that date, so that the facts might be ascertained for the whole course of the epidemic.

The inquiry was necessarily attended with a good deal of trouble. There were very few instances in which I could at once get the information I required. Even when the water rates were paid by the residents, they can seldom remember the name of the Water Company till they have looked for the receipt. In the case of working people who pay weekly rents, the rates are invariably paid by the landlord or his agent, who often lives at a distance, and the residents know nothing about the matter. It would, indeed, have been almost impossible for me to complete the inquiry, if I had not found that I could distinguish the water of the two companies with perfect certainty by a chemical test. The test I employed was founded on the great difference in the quantity of chloride sodium [salt] contained in the two kinds of water, at the time I made the inquiry . . .

According to a return which was made to Parliament, the Southwark and Vauxhall Company supplied 40,046 houses from January 1st to December 31st, 1853, and the Lambeth Company supplied 26,107 houses during the same period; consequently, as 286 fatal attacks of cholera took place, in the first four weeks of the epidemic, in houses supplied by the former Company, and only 14 in houses supplied by the latter, the proportion of fatal attacks to each 10,000 houses was as follows. Southwark and Vauxhall 71, Lambeth 5. **The cholera was therefore fourteen times as fatal at this period amongst persons having the impure water of the Southwark and Vauxhall Company as amongst those having the purer water from Thames Ditton.**

(Excerpted from Snow, 1855.)

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

consider sources of data in Chapter 3 and will discuss the problem of error in Chapter 7.

- He *measured the occurrence of cholera* in the two groups of houses served by the different water companies – we will look further at measures such as these in Chapter 2.
- He calculated *how many times more common* cholera deaths were in those houses receiving the contaminated water – we will come back to this measure (again a **relative risk**) in Chapter 5.
- He then integrated all of his information to reach the conclusion that cholera was indeed *caused* by contaminated water – Chapter 10.

He did not stop there, but went on to make a series of clear practical recommendations to prevent transmission of cholera in future – sensible measures including the need for cleanliness and sterilisation that are still practised today.

Snow's work therefore sets the scene for the chapters to come. Chapters 2–8 cover the basic principles and underlying theory of epidemiology in a very 'hands-on' way, leading to Chapters 9–11, which integrate this information in a practical look at how we read and interpret epidemiological reports, think about assessing causality and finally synthesise a mass of information in a single review. Chapters 12–15 then look at some specific applications of epidemiology and Chapter 16 concludes with a fresh look at what epidemiology is and what it can do to help address the health concerns facing the world today.

But, before you move on, take a minute to stop and think. Imagine that someone asked you what epidemiology was and why it was useful. Could you now give them a satisfactory explanation in a few sentences?

REFERENCES

AIHW (Australian Institute of Health and Welfare). (2008). *Australia's Health 2008*. Cat. *no. AUS 99*. Canberra: AIHW.

CDC (Centers for Disease Control and Prevention), National Center for Health Statistics. Compressed Mortality File 1979–1998. CDC WONDER On-line Database, compiled from Compressed Mortality File CMF 1968–1988, Series 20, No. 2A, 2000 and CMF 1989–1998, Series 20, No. 2E, 2003. Accessed at http://wonder.cdc.gov/cmf-icd9.html on 8 January 2010 7:24:23 p.m. and Compressed Mortality File 1999–2006. CDC WONDER On-line Database, compiled from Compressed Mortality File 1999–2006 Series 20 No. 2L, 2009. Accessed at http://wonder.cdc.gov/cmf-icd10.html on 29 December 2009 10:50:10 p.m..

Concise Oxford Dictionary, 5th edn. (1964). Oxford: Oxford University Press.

Devesa, S. S., Grauman, D. J., Blot, W. J. *et al.* (1999). *Atlas of Cancer Mortality in the United States: 1950–1994*. Washington DC: US Government Print Office.

Doll, R. and Hill, A. B. (1950). Smoking and carcinoma of the lung. *British Medical Journal*, **2**: 739–748.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014

Doll, R. and Hill, A. B. (1964). Mortality in relation to smoking: ten years' observations of British doctors. *British Medical Journal*, **1**: 1399–1410, 1460–1467.

Doll, R., Peto, R., Boreham, J. and Sutherland, I. (2004). Mortality in relation to smoking: 50 years' observations on male British doctors. *British Medical Journal*, doi: 10.1136/bmj.38142.554479.AE (published 22 June 2004).

Graunt, J. (1662). *Natural and Political Observations Mentioned in a Following Index and Made Upon the Bills of Mortality*. London. (http://www.books-on-line.com)

Hippocrates of Cos. (400 BC). *On Airs, Waters, and Places*. Translated by Francis Adams. (http://classics.mit.edu)

Hook, D., Jalaludin, B. and Fitzsimmons, G. (1996). *Clostridium perfringens* food-borne outbreak: an epidemiological investigation. *Australian and New Zealand Journal of Public Health*, **20**: 119–122.

Lilienfeld, A. M. and Lilienfeld, D. E. (1980). *Foundations of Epidemiology*, 2nd edn. New York: Oxford University Press.

MacMahon, B. and Pugh, T. F. (1970). *Epidemiology – Principles and Methods*. Boston: Little Brown.

Makk, L., Creech, J. L., Whelan, J. G. and Johnson, M. D. (1974). Liver damage and angiosarcoma in vinyl chloride workers – a systematic detection program. *Journal of the American Medical Association*, **230**: 64–68.

Miettinen, O. S. (1978). *Course Notes – Principles of Epidemiologic Research*. Harvard School of Public Health.

Pink, B. and Allbon, P. (2008). *The Health and Welfare of Australia's Aboriginal and Torres Strait Islander Peoples*. Canberra: Australian Bureau of Statistics and Australian Institute of Health and Welfare.

Porta, M. (Ed.) (2008). *A Dictionary of Epidemiology*, 5th edn. New York: Oxford University Press.

Rosen, G. (1958). *A History of Public Health*. New York: MD Publications.

Selikoff, I. J., Churg, J. and Hammond, E. C. (1965). Relation between exposure to asbestos and mesothelioma. *New England Journal of Medicine*, **272**: 560–565.

Snow, J. (1855). *On the Mode of Communication of Cholera*, 2nd edn. London: Churchill. (http://www.ph.ucla.edu/epi/snow/snowbook.html)

Stolley, P. D. and Lasky, T. (1995). *Investigating Disease Patterns: the Science of Epidemiology*. New York: Scientific American Library.

WHO (World Health Organization). (1948). Text of the constitution of the World Health Organization. *Official Record World Health Organization*, **2**: 100.

www.anesi.com/titanic.htm, The Titanic casualty figures (and what they mean). Accessed 30 December 2009.

Wynder, E. L. and Graham, E. A. (1950). Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma. A study of six hundred and eighty-four proved cases. *Journal of the American Medical Association*, **143**: 329–336.

Zaridze, D., Maximovitch, D., Lazarev, A. *et al.* (2009). Alcohol poisoning is a main determinant of recent mortality trends in Russia: evidence from a detailed analysis of mortality statistics and autopsies. *International Journal of Epidemiology*, **38**: 143–153.

Downloaded from Cambridge Books Online by IP 137.132.123.69 on Sat Oct 04 17:24:23 BST 2014.
http://dx.doi.org/10.1017/CBO9780511997303.003
Cambridge Books Online © Cambridge University Press, 2014