Part 1

Thinking like an epidemiologist

1 Thinking critically, problem-solving and setting priorities in epidemiology

Epi what?

Learning objectives

After studying this chapter, you will be able to:

1. Appreciate the importance of critical thinking and understand how it underpins epidemiology
2. Describe the role of epidemiology in identifying and responding to health priorities
3. Understand the relationship between the scientific question and epidemiological research methods

Introduction

If you are reading this book, you are either particularly interested in health and the field of epidemiology or you are undertaking a program of study in the health sciences. While you may have picked up this text out of interest, it is highly probable that you are undertaking a compulsory component in your study program and are now approaching the subject of epidemiology with some degree of trepidation. Perhaps some of you really enjoy this kind of thing, but others may be worried that this text will be heavy with daunting formulae and complicated calculations. Statistics (commonly referred to as *bio*statistics in the health sciences) is an incredibly important and intensely interesting field of endeavour; however, with apologies to those who may be disappointed, this book is *not* a text on statistics. Epidemiology is the study of patterns and determinants of disease and other health states in populations. It primarily uses quantitative methods (which deal with counting, measuring and comparing things) that definitely *use* statistics and include statistical methods, but in this book we will not be talking about performing any statistical acrobatics more complicated than completing a sudoku puzzle.

Rachael: That all sounds good – I haven't done a lot of maths, but I do like solving puzzles!

Hugo: Well, I hope they mean it. I'm scared of numbers and I'm not even sure I can manage a sudoku.

Author (Emma): Lots of people find numbers a little scary at first, but this is usually just the thought, rather than the reality, of them. I don't mind a number or two myself but remember that I am an epidemiologist and not a biostatistician. I like to solve health puzzles using numbers (and I'm sure you will too after reading this book!) but when serious statistical manoeuvres are called for, my most important solution involves calling the team biostatistician …

It is probably time to point out that you are not alone here. We'd like to introduce Rachael and Hugo, who will be coming along for the ride with you. Their role is to ask the questions you might think about asking and to keep us authors on track. Their contributions will be clearly signposted, so you should feel free to skip those bits if this is not your thing, but their questions (and author responses) might really help to clarify things along the way.

Other things that might be useful are the real-life examples that will be included in every chapter, along with explanation boxes where relevant. The examples will describe epidemiological concepts, often (though not always) in the form of studies or investigations that have been done in the real world. Most of these examples will involve large population trends and events, including the recent **COVID-19** pandemic. If those examples include them, you will get the opportunity to try to calculate some fancy-sounding measures for yourself and the explanation boxes will show you how to do them in clear terms. At the end of each chapter, you can have a go at the review questions and keen readers can access further review questions online.

COVID-19: Coronavirus disease 2019, a disease caused by infection with the SARS-CoV-2 virus, first noted in 2019.

Now, let's set the scene by beginning to address the main question of this chapter – 'Epi what?'

Epidemic: (synonym 'outbreak') An increase in the number of cases, which is beyond that normally expected for a particular region during a particular period of time.

Risk factor: In the context of epidemiology, a factor/variable that increases the chance of having a particular disease or health condition among the people who are exposed to this factor.

Surveillance: Continuous monitoring of diseases or health conditions in a defined population or geographic location. Involves systematic and ongoing data collection, analysis, interpretation and dissemination to detect potential outbreaks and inform timely control measures.

Epidemiology

If asked to describe the principal purpose of the health sciences, a reasonable starting point might be to suggest that it is to prevent illness and maintain good health in the population. One might then extend this to enhancing quality of life, curing those with disease and minimising complications in those with illnesses that can't be cured, even promoting health equity and inhibiting inequality in society. What if I were to tell you that epidemiology could help with all of these goals and more?

'Epidemiology', taken from the Greek words '**epidemic**' (first used by Hippocrates, way back around 400 BC) and '-ology' (science or discipline) is traditionally defined as 'The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems' (1, p 42).

Although historically thought of as concerned primarily with disease and disease **risk factors**, in more recent times there has been far greater recognition of epidemiology's capacity to explore the social, political and health experience of populations (2). Yet epidemiologists tend to be thought of by many as 'disease detectives' – complete with metaphorical trenchcoats and deer hunter hats, sniffing out causes of disease and identifying the culprits and their associates. Although there is some truth in this stereotype, in fact there are multiple 'flavours' (or fields) of epidemiologists who work across the gamut of health-related scientific endeavour. These flavours include, but are not limited to: communicable and non-communicable disease **surveillance** and control; immunology and vaccine development; medical, clinical and pharmaceutical; veterinary science; environmental health; and the analysis of '**big data**'. In this book, we will be discussing examples of research in many of these fields of epidemiology and more.

Hugo: I was thinking epidemiology had something to do with skin disorders – I guess I got confused because the word 'epidermis' is about skin.

Rachael: You're not the only one. When she heard I was learning about epidemiology, my Aunt Mary started talking to me about her rash!

Big data: Very large population-level datasets that are analysed for epidemiological trends and patterns in health, risk and human behaviour.

Outcomes: The different health states a person might experience.

What all fields of epidemiology have in common is the application of critical thinking to understand and describe phenomena, identify causative or contributing factors to **outcomes**, provide and evaluate evidence from which to develop actions to modify health outcomes, and even evaluate the effect of those actions once they are taken. Of course, critical thinking is not just the preserve of epidemiology; it has wide and important application in just about every field of human endeavour. Developing your critical thinking skills can help you with all parts of your current studies and future work. In fact, the ability to question, analyse, interpret and reach an informed judgement could be seen as vital to making sense of just about everything that happens in the modern world (3).

As well as raising the quandary of the best place to put your chair, Example 1.1 describes a hypothetical case of everyday critical thinking: questioning the source of information, seeking opposing arguments and synthesising the evidence before reaching an informed judgement or conclusion. In this chapter, we will learn a bit more about critical thinking in epidemiology and how this is used to identify, prioritise and address health priorities in society.

EXAMPLE 1.1

You hear you could catch a cold from the draught of an open window. Your first instinct might be to move your chair as far away from the window as possible, but then you might start thinking about this a bit more:

- I wonder how many of the colds I have caught in the past might have happened after I sat next to an open window?
- What about those people having their lunch outside in the cold – will they get more or fewer colds if they eat their lunch indoors?
- What about the person who told me about the window thing – would I normally trust their advice?
- Do other people have different theories about the risk of sitting near open windows?

Based on these ruminations, you could either conclude that the original advice is likely to be correct or begin to wonder whether you should be looking into this further.

An approach to critical thinking

The concept of critical thinking is not new but, in what is starting to be a repeating theme here, goes back at least as far as ancient Greece. The idea of critical thinking was born from the work of Socrates, around 2500 years ago (4). His research indicated that the majority of people in his own society, even those who were supposed to be in charge, were prone to irrational and often contradictory beliefs that were usually based on little or no evidence. Socrates stressed the importance of gathering good evidence and developed a method of challenging underlying assumptions and beliefs, and their implications, through an approach known as 'Socratic questioning', a method that is still used in psychotherapy and teaching (5) – although, in true critical thinking style, there is some debate about the consistency of understandings and application of Socratic questioning in practice (6).

While Socrates' observations were made about the illogical thought processes prevailing at the time, it could be argued that the tendency to build irrational systems of beliefs (what I like to term '**feel-osophy**' – based on a mixture of emotional reaction, peer influence, resentment, misunderstandings and personal agenda), is ever present when it comes to human beings. Illogical and irrational assumptions arising from the absence of critical thinking can have devasting consequences on personal, societal and global levels. These consequences can range from the breakdown of personal relationships to the type of shared societal hysteria that brought about the cruel and deadly witch trials that occurred periodically across Europe and North America between the fourteenth and eighteenth centuries (7, 8). More recently, a lack of critical thinking in relation to COVID-19 vaccinations could well have had global health impacts. In one example, the owner of a multinational technology corporation stands accused of a plan to microchip the whole world population using vaccines – a surprisingly widespread belief that seriously impacted vaccination rates in the United States and elsewhere (9). You may have heard about this and other COVID-19 related 'feel-osophies' and the conspiracy theories they spawn (10). Reduced protective behaviour – such as decreased mask-wearing, social distancing and handwashing – combined with low vaccination coverage and high

Feel-osophy: A made-up term denoting systems of belief relying on inaccurate information selected on the basis of factors such as peer influence, resentment, misunderstandings and personal agenda.

infection rates generates the perfect environment not only for more virus transmission, but also for the development of new variants of the virus. As the continuing emergence of new COVID-19 variants also makes the task of developing specific vaccines more difficult (11), pandemics are probably not a great time to be forgoing critical thinking.

Recognition of the importance of critical thinking today is reflected in the explicit goals of most universities in striving to produce graduates with critical thinking skills as a core attribute. This is because critical thinking is strongly associated with academic success and employment opportunity (12). Critical thinkers have the capacity to be innovative, to solve problems, are often creative and reach evidence-based conclusions. In fact, one study found that critical thinking was a better predictor of positive life outcomes than intelligence (13).

Hugo: That might explain how really smart people can make some really silly decisions.

Rachael: Maybe it also explains how my big, dumb brother always wins at chess – actually, he always wins at everything …

Fortunately, unlike inherent traits such as intelligence, critical thinking is something people can learn and develop over time. As touched on, critical thinking is a cognitive process that involves questioning, analysing, evaluating and reaching a judgement about information you might hear or read. It can be applied to information coming from social media, television, news sources, family and friends, your teacher or lecturer, and all published works (even this book!). Transferability across contexts and tasks is in fact an important and useful hallmark of critical thinking skills (14). Critical thinking provides a way of sorting out, on a preponderance of the evidence, what is most likely to be the most accurate version of reality. Being able to reach reliable judgements and then apply this to problem-solving is important across all social, academic and professional activities in life, but has particular relevance to the health sciences. People working in the health sciences may be providing clinical services, developing clinical guidelines or health policy, or researching and providing the evidence to inform all of these activities. Without trying to sound melodramatic here, the well-being of the population or even people's lives might well depend on those judgements (15).

And how do we 'do' critical thinking? Well, there is much written on the topic (and you should check out the 'Further reading' section at the end of the chapter to find out more), but most methods include the following components: analysing, evaluating, reasoning, problem-solving and decision-making.

Analysing

Analysis may sound like something that could be complicated, involving numbers and possibly the wearing of lab coats, but you are probably doing it all time without thinking too much about it. When first presented with new information, unless you are the type to automatically let your eyes glaze over, you will make an assessment of how relevant those data are to you. As in the window-opening example (Example 1.1), you could think about how much you trust the source of the information and whether there are alternative facts to consider. This natural ability (not the eye glazing-over thing) can be harnessed and improved in a logical way.

At the core of analytical thought is questioning the argument being proposed. In this context, the argument is a proposition or claim. The claim may be in the form of a statement of belief or may be more extensive, such as a theory (a system of ideas aimed at explaining a phenomenon) or even a feel-osophy (just had to get that term in there one more time).

In epidemiology, a claim might be about what causes disease. In Example 1.1, the claim was that the draught coming in through the window could cause the body to become more susceptible to respiratory illnesses. This is actually a common belief with its roots in folklore, where illness is thought to emanate from body temperature changes caused by exposure to cold or damp in the environment (e.g. 'catching a chill') (16). This theory may have developed to explain the objective observation that most respiratory illnesses are associated with the colder months of the year. This is a kind of theory involving reasoning to connect conditions – it is colder outside than inside, and colds are associated with colder seasons. The competing argument, particularly reinforced by recent observations involving COVID-19, is that air movement from open windows might dilute virus particles that could be in the indoor environment, usually secreted from people huddling inside, therefore decreasing the likelihood of being infected (17). This is also based on a theory: when it is cold, people tend to congregate inside for longer, heightening the risk of virus transmission between people who may be infected. But how do we sort out which claim is likely to be the most accurate? The first claim comes along with years of tradition, handed down from trusted personal authority to trusted personal authority. The competing claim comes from people we have never met, who tend to use complicated methods and language that is barely comprehensible at times.

Well, let's break it all down …

Collecting and synthesising data

To begin, we need to analyse these claims about windows in the context of the events to which they pertain. They are both about contracting a respiratory illness in cold weather and the risk associated with airflow (increased or decreased). The first task might be to find out how many respiratory viruses arise in a given population and when they occur. For that we will need to collect some reliable data on how many people are getting respiratory illnesses and when. Friends and family members would probably not make the best source of information, as there are several potential factors that would impact the quality (or reliability) of the data. Personally, I have increasing trouble remembering accurately all that transpires the further that time moves on from the event. Your friends and family might not always be able to give you accurate information about whether or when they contracted a respiratory virus last year, for instance. Relying on this information could introduce a type of unintentional systematic error in your data known by epidemiologists as '**recall bias**'.

A **bias** can occur due to the existence of any factor or trend in the collection or analysis of data that can lead people to arrive at erroneous conclusions. Epidemiologists know about many different types of bias, some of which could arise from asking your friends and family for information. Another type that springs to mind here is the potential for **selection bias** – the type of bias that can occur when we select people to provide data who may have characteristics that differ from those about whom we want our conclusions to make a claim (such as the population in general). Those closest to you might share many genetic and/or behavioural characteristics that are related in some way to infection transmission but may differ from other people in society. Further, respiratory infections are transmitted from person to person so we might not be able to tell whether the open window is implicated in transmission or the

Recall bias: Occurs when there is differential recall of exposures or experiences between comparison groups, which is most likely to occur in studies using self-reported data (a form of measurement bias).

Bias: Any unintentional systematic factor, or trend in the collection or analysis of data, that leads to erroneous conclusions.

Selection bias: Systematic error introduced by the selection of participants, or inclusion of their information, with characteristics that differ from those not included in the study.

Confounding: Occurs when understanding of the relationship between one potential cause and an outcome is distorted by one or more additional factors associated with both that potential cause and the outcome.

De-identified data: Information about characteristics of a population (or sample from a population) from which all identifying information has been removed; it is not possible to recognise particular people from whom the data have been collected.

virus spreads since friends and families tend to hang out together, regardless of proximity to windows. This last is an example of another potential source of error called **confounding**, when the true relationship between one potential cause and an outcome is distorted by the existence of another factor that is associated with both that potential cause (being near an open window) and the outcome (viral transmission). You will be hearing much more detail about bias and confounding and their various impacts in later chapters, but this is enough to be going on with.

So where do we get the required reliable information? Happily, we are now living in an era of unprecedented access to high-quality information about almost every part of human experience. This is a time I like to call 'paradise for nerds'. All this information is collected across the population, turned into **de-identified data** (all identifying information is removed) and regularly presented in downloadable number files (e.g. in Microsoft Excel and similar) or summarised in reports. Assuming one has access to the internet, in many countries it is now possible to access the latest population data on health, welfare, workforce, housing and more. International health data are available online from institutions such as the World Health Organization (18). I have included a link to this website in the 'Further reading' section near the end of this chapter for you to go and see what is there, but you could and should do some exploring online to see what is available from government and non-government agencies in your country.

Accessing data from one such public health organisation in my own country, I learnt that there were 12 083 cases of influenza reported in 2022 in the population of around 1.8 million people of South Australia (19, 20). This was quite a big year for influenza following some quiet years, which was thought to be due to international border restrictions in place in Australia during the earlier years of the COVID-19 pandemic (21). Influenza is not the only virus that causes respiratory infections like colds, but it does follow a similar pattern of occurrence to most other respiratory viruses and, handily for us, is subject to surveillance in many jurisdictions. Looking at the information available, and loosely estimated in Figure 1.1, it is clear that most of the cases of influenza in South Australia occurred in late autumn and winter – the same seasonal pattern that has been note around the world (22).

Rachael: I knew there would be graphs!

Hugo: Do you think you could explain it a bit, Emma?

Author (Emma): Reading graphs can be a bit daunting, but the best thing to do is to always take it step by step, as even seasoned graph readers are well advised to do. The title of this graph (Figure 1.1) indicates that it shows the number of cases of influenza reported each week over the seasons in the year 2022. Along the upright axis (known as the y-axis) is depicted the range of the potential number of cases from zero to 1400. Along the horizontal axis (known as the x-axis) are the seasons flowing from summer to summer – left to right. Because the graph concerns influenza in Australia, you can see that winter occurs in the middle of the year and the year starts and ends with summer seasons. The columns (or bars, as this type of graph is called a 'bar chart') across the x-axis represent the number of cases reported each week (refer back to the title). The height of the bars represents the number of cases reported in that week measured against the y-axis.

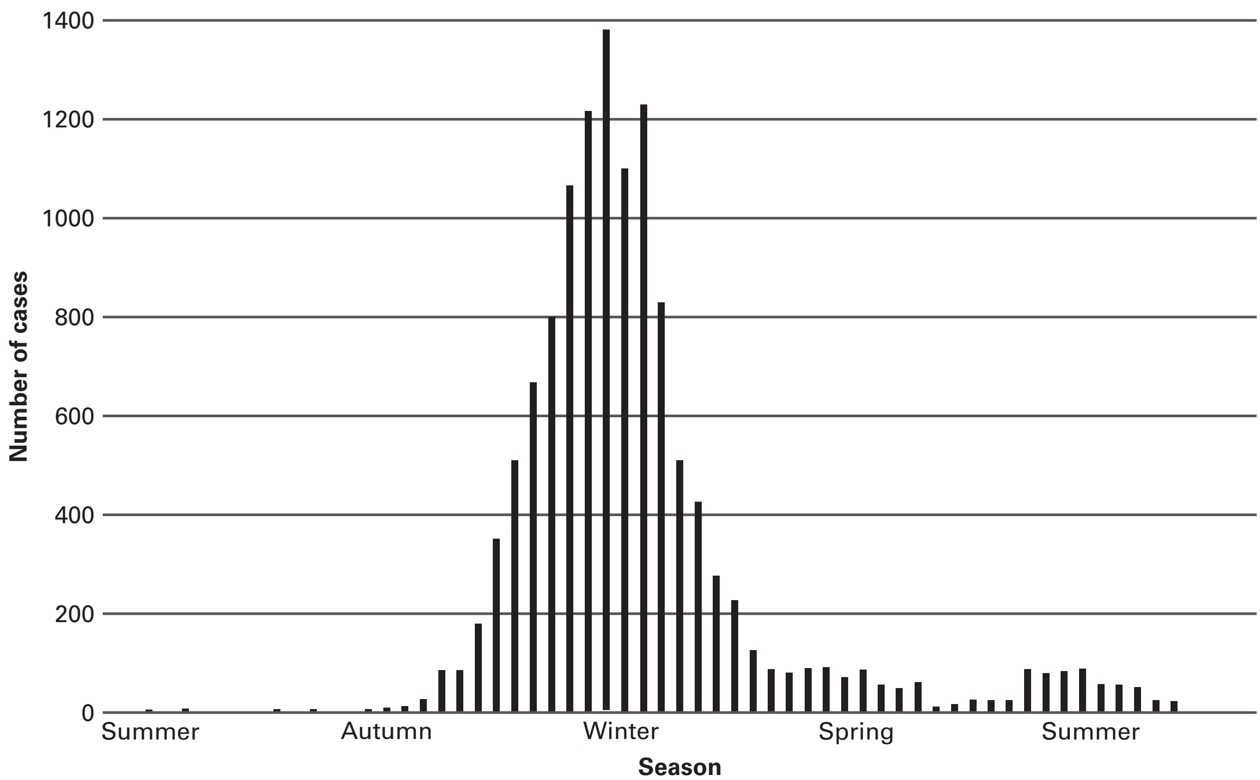

Figure 1.1 Laboratory-confirmed influenza cases per week by season in 2022 in South Australia. Adapted from (20).

We now begin to suspect there may be some truth in at least the first part of both claims: increases in respiratory infections are associated with colder weather. But what about the second part, the risks associated with cold air flow? Although there are unlikely to be handy population data sets to access on this issue, there may well be some other evidence in the form of published studies. A lot of research has been done in this area in recent times, particularly in relation to COVID-19, which is caused by a particular respiratory virus type known as a *coronavirus*, **SARS-CoV-2** (Severe Acute Respiratory Syndrome Coronavirus 2). Non COVID-19 type coronaviruses are thought to cause up to 15 per cent of cases of what most people call the 'common cold' (23). So, the real-life study summarised in the Example 1.2 could be highly relevant to our current critical thinking activity.

SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2, the virus that causes COVID-19.

Systematic review: A review that aims to identify, evaluate and synthesise all available evidence from quality-assessed studies, using predetermined explicit and systematic methods to address specific research questions (may or may not involve a meta-analysis).

Rate: The occurrence of an event (e.g. counts of disease or death) divided by the number of a specified population at risk over a specified period of time.

EXAMPLE 1.2

In a special kind of study called a **systematic review**, in which the researchers locate and review as many quality published studies as they can find to fully understand a particular topic, Thornton and colleagues (24) synthesised the findings of 32 different studies that investigated the effectiveness of ventilation – for example, window opening and their placement in the building, exhaust systems, air conditioners – on the transmission of coronaviruses. They found that increasing of ventilation decreased both the **rate** of virus

EXAMPLE 1.2 Continued

transmission and the risk (or probability) of getting infected. Increased ventilation was also associated with decreased persistence of droplets (infected watery particles released into the air from sneezing or coughing) and less viral concentration in the air. Although better results were associated with different types of ventilation (such as exhaust systems, window placements and fans), the researchers found that any ventilation was better at reducing viral transmission than none at all and increasing ventilation and introducing fresh air were among their recommendations.

The study summarised in Example 1.2 seems to support the claim that airflow from open windows might reduce the risk of catching a cold. What about the original claim: that being exposed to the draught of open windows could increase your risk of catching a cold? Although not on draughts from windows specifically, I was able to find some published studies of the effect of cold air in people with chronic respiratory illness, such as exposure to cold air causing asthma attacks (25). I couldn't find evidence that cold air alone might increase the likelihood of getting a respiratory infection; however, there appears to be mounting evidence that respiratory viruses such as influenza are more stable in colder drier weather (26). Thus, viral 'staying power' in the environment (once transmitted by an infected person) might be why we see more infections in winter (27). It is also possible that sitting in a cold draught can dry out the mucosal lining of the respiratory system, making people more susceptible to infection should they be exposed to a virus (28).

Evaluating

It is here that we return to considering the quality of the information underlying the competing claims. This is part of the process of evaluation: assessing the strength, authority, credibility or value of the argument. The quality of information is greatly influenced by the authority and agenda of its source. First, I always like to ask: who is saying this and why they are saying it? The 'who' is the person, author or organisation presenting the information. Finding out the 'why' can be a relatively simple exercise, but sometimes can take a bit more investigation.

From time to time, you may hear about methods to preserve health and prevent illness as you navigate your way through cyberspace. Many theories of health and illness are espoused on social media by 'friends', or presented on the internet via 'influencers' (people who appeal to large online audiences, and who often are paid in money or goods to promote commercial products). Developing, or ascribing to, alternative (and sometimes irrational) explanations for health phenomena is extremely common and has been studied for many years (29). The emergence of social media has allowed for the rapid and unprecedented propagation of information and misinformation across the world (30). In the social media environment, it is not always easy to answer the 'who' question and can be even more difficult to uncover what the motivation for presenting any information might be. While some posters might genuinely wish to pass on what they believe is good health advice, misinformation may also be promoted deliberately. One study suggests that the intentional posting of misinformation may be due to people having a high degree of trust in the online environment and therefore being willing to take risks in passing on unverified information, the fear of missing out (FoMO) and a sort of indifference to caution arising from social media fatigue (31). Overall, the inability to assess