CHAPTER 1
Understanding basic numbers

Numbers are not necessarily easy to understand and, notwithstanding stories of grandmothers and teaching them to suck eggs, this chapter will try to cover some of the basics for understanding numbers. The chapter warns about 'orphan' numbers and how percentage changes are difficult.

When is a number large?

Consider the following examples:

1. On 6 May 2020 there were 30,000 deaths due to COVID-19 in the UK, 75,000 in the USA and 265,000 in the world.[1]
2. There were about 634,000 deaths in the UK population, 2,909,000 in the USA and 58 million deaths in the world in 2018.[2]
3. The UK Government stated in 2018: 'We have invested an extra £1 billion in the NHS [National Health Service] this year.'[3]
4. The UK sent £350 million to the European Union every week.[4]
5. The Global Burden of Disease Report (GBDR) on sepsis estimated that there were 48.9 million cases in 2017, and 11 million deaths, across 195 countries and territories.[5]

Are these large amounts? They certainly sound like large amounts, but how do we come to terms with what they mean? Large numbers are often quoted on their own by people in authority, to try to impress the public with how big the numbers are. (A useful term might be 'orphan' numbers because they are not related to other numbers.) However, there is an old joke that if you ask a statistician how well they are, they will reply 'Compared with whom?' Likewise, numbers on their own are by and large meaningless; it is only with comparisons

Statistics at Square One, Twelfth Edition. Edited by Michael J. Campbell.

Companion website: www.wiley.com/go/Campbell12e

that we can extract a meaning. In example 1, the COVID-19 deaths are, on their own, just large numbers. However, we can employ an analogy to give them some meaning. The first number of deaths roughly equates to the same number of people at an average Premier League football club in the UK, whereas the second is closer in number to a capacity crowd at Old Trafford, home ground of Manchester United. The third is the size of an average town in the UK (e.g. Southampton). These analogies put the number of deaths into a very human perspective. However, to get a better understanding of these numbers we need more specific comparisons.

A helpful basis for comparison is knowing the approximate size of the populations to which each statistic is referring. In 2019, the population of the UK was 67 million, that of the USA was 330 million and that of the world 7.7 billion (7700 million).[6] We can then calculate the ratio of the number of deaths to the size of the population. A ratio is simply one number (numerator) divided by another (denominator). In this case, since the numerator is a subset of the denominator, we have proportions. The deaths per head of population are 0.044%, 0.023% and 0.003% for the UK, USA and the world, respectively. These percentages lead to another comparison: that between countries. The UK appears to be doing worse than the USA, which is doing worse than the rest of the world. Is this a reasonable conclusion? Cause of death is often very unreliably reported. Completion of a death certificate is often assigned to a junior doctor with little training. In an elderly patient with multiple diseases, it can be especially difficult to ascribe one main cause. So in this example we should consider how we know the person died of COVID-19. Presumably the patient was tested before they died or they had symptoms similar to COVID-19. However, testing rates have varied widely between countries and diagnosing symptoms of COVID-19 is very subjective. Thus, these numbers for death rates due to COVID-19 are not at all reliable and a reliable comparison is therefore difficult.

In contrast, deaths (from any cause) are reliably reported in the UK and the USA and probably well reported for the rest of the world. In example 2, again the numbers by themselves are not meaningful, but compared to the size of the relevant populations we can extract some meaning. A quick calculation reveals that 0.95% of the UK population dies every year, compared to 0.88% in the USA and 0.76% in the world. These numbers on their own are interesting. In the UK about 1 person in 100 dies each year.

This brings the numbers down to something we can appreciate. Again, we can compare the proportions dying by country, and once more it appears that the UK is so much less healthy than the USA, and both countries are less healthy than the rest of the world. This may lead to further investigations.

In example 3, we could compare the extra sum invested in the NHS to the annual budget for the NHS, which is about £130 billion, so this extra £1 billion is less than 1% of the total. Another way to look at this is to consider that we now know there are about 67 million people in the UK, so £1 billion equates to about £15 for every person in a year, roughly the cost of five pints of beer (at current UK prices outside of London). It doesn't sound so big now, does it?

In example 4, it is worth knowing that the UK economy was worth £8.8 trillion a year in 2016 (a trillion is 1000 billion).[7] The £350 million a week given to the EU is £18.2 billion a year, so the amount the UK sends to Europe is $\frac{18.2}{8800} \times 100 = 0.2\%$ of the UK economy. Again, it doesn't sound so big now, does it?

If we combine the information from example 5 with the worldwide death data in example 2, we would deduce that approximately 1 in 5 deaths worldwide is due to sepsis. This certainly is a large number! However, all unusual numbers should be subjected to a little scrutiny. As a quick reality check, you might start by asking yourself whether of the people you know who died recently, did 1 in 5 die of sepsis? One would expect the answer to be no. Thus, we might query whether the GBDR is right. One issue is that sepsis can be difficult to diagnose and the rate of diagnosis varies hugely from one country to another, so local experience may be misleading in that in another country sepsis might be more readily diagnosed.

When you hear a number given that you believe the presenter wants to sound big, it is always worth applying reality checks such as those described in Box 1.1. A light-hearted example has been provided in a video from the Sheffield Methods Research Institute[8] concerning a news report that stated that floods in New Zealand had caused 30,000 pigs to be washed down a river. This was then reported uncritically by other news outlets, until someone thought: '30,000? That's an enormous number, is it believable? How many pigs are likely to fall into a river at any one time?' Going back to the original broadcast, it turned out that the reporter had in fact said 30 *sows and* pigs, but owing to their New Zealand accent, this got 'misheard' and repeated uncritically to the wider media.

A further question about a large number is to ask what period of time the number refers to. By expanding or contracting the time scale, a presenter can make a number look big or small. When a large sum of money is promised, one should ask: How much does this equate to per year? In example 2 above, 634,000 deaths sounds large, whereas 0.95% sounds small. However, if one stated that approximately 1650 people die every day, or about 1 every minute, it may sound even bigger, since in our everyday experience people are not dying every minute! It is worse when reports state the 'risk of death' or 'lives saved' without stating a time period. The risk of death in the long run is one!

Definitions of the quantities discussed in this chapter are given in the Glossary. Ways of questioning numbers are given in Box 1.1.

Ratios

As we have shown, a number on its own is difficult to comprehend, but when compared to another number it can be given meaning. The simplest way to compare two numbers is to divide one by the other. A ratio is simply one number (numerator) divided by another (denominator). Ratios of continuous variables are often used to 'adjust' the numerator by the denominator. Possibly the most commonly used ratio in medicine is the Body Mass Index (BMI), which is a person's weight in kilograms divided by their height in metres squared (kg/m^2). The idea is that tall people are naturally heavier than small people because they are bigger, but that doesn't make all

Box 1.1 Things to think about to help understand numbers

Where did the number come from?
Why is this number being given and what is it supposed to show?
Can one trust the source?
Is there a useful comparator?
If there is a comparator, why was it chosen?
Is it the best one?
What period of time is the number covering?
If the number is a proportion, is the numerator relevant to the denominator?
If the number is the death rate of people with a disease, ask: How do we know that the people who died (numerator) had the disease? How do we know whether people who did not know had the disease?
What is the size of the population from which the number is coming?

tall people overweight! To decide whether someone is overweight, you can't just compare weight, you need to make some allowance for height. The idea of dividing by the square of height is credited to Adolphe Quetelet (1796–1874), who observed in a cross-sectional study that weight increased as the square of height. However, it may seem simplistic to think that a simple ratio can 'adjust' for the denominator. For example, the BMI has received much criticism in that it doesn't properly account for height, thus is more likely to classify short people as overweight, and also because it doesn't account for muscle mass, which is more dense than fat. Consider that Arnold Schwarzenegger and Tom Cruise are both estimated to have a BMI over 30, which classifies them as obese![9]

It is important to note that one can make a ratio smaller by either reducing the numerator or increasing the denominator. For example, cholesterol is a dense, fatty substance found in every cell of your body. The two main types in the blood are high-density lipoprotein (HDL, the 'good' cholesterol) and low-density lipoprotein (LDL, the 'bad' cholesterol). The cholesterol ratio is the ratio of the total cholesterol in the blood to the HDL, and its main advantage is that it is a single number that is easy to remember. A low cholesterol ratio is good. A person with a low HDL ratio could still have dangerously high levels of LDL if their total cholesterol was high.

A further example shows that care is needed to determine the denominator. The Summary Hospital Mortality Index (SHMI) is a ratio of the observed number of deaths in a hospital in a year with the number expected, based on the demography of the hospital population using a prediction equation.[10] Thus a hospital can lower its SHMI by *reducing* the observed number of deaths or by *increasing* the expected number. One way of doing this is to ensure that all comorbidities are included in the prediction equation. (A comorbidity is an underlying health issue that is not the reason for coming into hospital this time.) For example, a person may have heart disease, but come into hospital with a broken leg. Including the comorbidity in the coding will increase the expectation of death and so reduce a hospital's SHMI. This is because anyone admitted with heart disease will increase the expected number of deaths, even if they don't change the observed number of deaths. Suppose that for people with heart disease their risk of death was 2 in 100. Each person coded with heart disease will increase the expected number of deaths by 0.02, and with large numbers of patients admitted to a hospital these increases add up! Thus, hospitals might strive to increase their expected values as much as possible. Not including the comorbidity might penalise a hospital because its SHMI would appear too high if the patient happened to have a heart attack in hospital. (There are other issues with SHMI that will be discussed later.)

When two numbers are made up of the same units, then the ratio can be easier to understand if we multiply it by 100. Thus, the NHS increased investment in example 3 above by 0.007 of the total. This is easier to understand as 0.7%, which is why we stated earlier that the amount was less than 1%. Similarly, the deaths in the UK and the world in example 2 are given as percentages of total populations.

Using ratios to adjust for other variables

The BMI is an example of using a ratio to 'adjust' for another variable. This is a common technique in medicine, but it needs to be done with care. In respiratory medicine, a commonly used index is the FEV1/FVC ratio. The forced expiratory volume (FEV1) is the volume of air one can blow out in one second and the forced vital capacity (FVC) is the total volume of air one can blow out. This ratio, in theory, allows for the fact that larger people will have larger lungs, since an FEV1 reading on its own will not have much meaning. The ratio is expressed as a percentage, and a normal value is considered to be greater than 70%. The advantage of using the FVC is that we would expect it to change with height and age in the same way as the FEV1, so it in some way 'adjusts' for these quantities. Similarly, the HbA1c% is the percentage of haemoglobin molecules that have a glucose molecule attached to them. and this adjusts for whether a person has a high haemoglobin level or not. The assumption underlying these ratios is that the adjustments are proportional. This means that a healthy person with an FVC of 4 litres would be expected to have an FEV1 of more than $0.7 \times 4 = 2.8$ litres. A person with an FVC of 5 litres would be expected to have an FEV1 of greater than $5 \times 0.7 = 3.5$ litres. Another way of looking at this is that if a person has an FVC 25% higher than another person, we would expect their FEV1 to be 25% higher if they are healthy.

Proportions, percentages and odds

The proportion is a special kind of ratio, where the numerator is a count and is a subset of the denominator. As before, we often multiply by 100 to give a percentage. Thus it is easier to state that the percentage of people with diabetes (Types 1 and 2) in the UK is 6% rather than that the proportion is p = 0.06.[11] We can express this as a count by saying that for every 100 people in the UK, 6 have diabetes. These whole numbers are termed natural frequencies and convert the percentage to a count. Gigerenzer[12] showed that people (indeed even doctors!) understand counts, or natural frequencies, better than proportions and so it is better to use frequencies in communications. If the proportions are small, the denominator can be increased to ensure that the proportion is a whole number. For a percentage the denominator is 100, but for smaller proportions denominators of 1000 or 10,000 are used. In another example from a Lullaby Trust report,[13] there were 200 deaths from sudden infant death syndrome (SIDS) in the UK in 2017. There were 755,000 live births in the UK that year, so the proportion of SIDS to live births is 200/755,000 = 0.00026. This is expressed in the report as per 1000 live births, which makes it 0.26 SIDS per 1000 live births. However, one might argue that 26 SIDS per 100,000 live births is easier to understand, especially when comparing different time periods or countries.

Note that the denominator should be a *relevant* population. Thus for SIDS the denominator is the number of live births, not the number of people in the population. When considering deaths from a particular disease, it is tempting to use as the population the proportion of people with the disease. This can be a particular problem, as the recent COVID-19 pandemic illustrated. We can only know if people have the disease by testing for it. If one country carries out a larger number of tests compared to another, it will make the disease appear more prevalent.

Another way to express a proportion (p) is to use the reciprocal (1/p). Thus, we can say that approximately 1 person out of every 17 (1/0.06) in the UK has diabetes or that (1/0.00026) or about 1 in 4000 babies born every year will die of SIDS. According to the Office of National Statistics in January 2021, approximately 1 in 7 people in private households has had COVID-19. Such numbers are often easier to understand than proportions.

Percentage difference and percentage change: importance of baseline

We use percentages to express differences as a fraction of the whole. However, it is crucial to define the baseline. The mean height of British adults aged 20 years is 177.3 cm for men and 163.6 cm for women, a difference of 13.7 cm.[14] Women are 100 × (13.7/177.3) = 7.7% shorter than men, whereas men are 100 × (13.7/163.6) = 8.4% taller than women. There are two percentage differences, depending on which sex is used as the baseline! The same problem does not arise with the absolute difference: women are, on average, 13.7 cm shorter than men, and men are, on average, 13.7 cm taller than women.

Percentages are often used to show change. However, there are a number of issues to be aware of. In particular, percentages critically depend on the baseline. Suppose a headline was: 'The number of cases of knife crime has risen by 20% this year'. One should automatically ask: 'from what to what and when?' The baseline could have been 10, and it has gone up to 12, or it could have been 1000 and gone up to 1200. The interpretation of the same percentage is very different in the two scenarios! The ratio of baseline to final figure is 1.2, thus a simple way to find the final figure is to multiply the baseline by 1.2. If knife crime continues to rise as a fixed proportion of the previous year (which is termed an exponential rise), then from a baseline of 100 the next year we would expect 1.2 × 1000 and the following year 1.2 × 1.2 × 1000 = 1440 (in the same manner as compound interest). Note that the percentage change is not additive. If the rate goes up 20% each year, then after two years the rate is 44% above baseline, not 40%.

This problem of understanding percentages is even worse for percentages over 100. Suppose there were 100 deaths from some cause in a year. If you were told that deaths had risen 200% the following year, how many deaths would you expect? Since 200% of 100 is 200, the increase in the number of deaths is 200 and one would expect 300 = 200 + 100 deaths to have occurred this year. However, some people would think the deaths the following year would rise to 200 (a rise of 100%)! It is better to state the beginning and end numbers to avoid confusion. (I am always bemused by people who make '110% effort'.)

A problem with baselines is illustrated in the recent statement: 'Between 15 March and 27 March (2020) the UK government's numbers on death from COVID-19 have been more than 100% less than the actual figures on three occasions.'[15] If something reduces 100% from baseline then it is zero! Here, the authors have used the UK government's figures as the baseline and so what they really mean is that the actual deaths are 100% *greater* than the UK government figures.

Another way to make numbers look bigger is to choose the smallest baseline. Given a slowly changing rate, the baseline may be chosen some way in the past. Thus the rate per 1000 live births of SIDS in the UK declined from 0.48 to 0.26 from 2004 to 2017 (or from 48 to 26 per 100,000 live births), a reduction of 84%.[13] While this is a noticeable achievement, one might ask: Why choose 2004 as the baseline year? In fact, this is when this particular definition of SIDS started to be used (but one has to dig to find this). A much easier statistic to recall is that the death rate almost halved from 2004 to 2017.

Even if the choice of baseline is clear, such as the previous year, there is another problem with percentage change. If you go up x% and then down x%, you don't end up where you started! Thus we may find out that that the number of cases of knife crime last year was 1000, and we are told it has risen by 20% this year, so one would expect the number this year to be 1200. Suppose there was then a successful campaign to reduce cases of knife crime by 20% the next year. Since 20% of 1200 is 240, one would expect the actual number of cases to drop to 960, 40 less than the baseline! Similarly, if deaths from some cause dropped from 300 to 100, what is the percentage drop? In fact it is (300 – 100)/300 = 67%. Thus, although the rise from 100 to 300 is 200%, the drop from 300 to 100 is 67%. The issue is that the baseline had changed and so the meaning of the percentage change is different. Always look carefully at numbers purporting to be percentage changes.

When a measure is itself a percentage, it is even more important to avoid confusion about the percentage change. If an HbA1c has gone down from 7.5% to 7%, then this a 0.5% drop, but it can also be expressed as a 6.7% (0.5/7.5) drop. It is better to use the term 'absolute percentage points' for differences in percentages, so we would be better to say that HbA1c% has dropped by 0.5 absolute percentage points.

Rounding proportions and percentages

If we had three patients with similar heart disease risk and one was on statins, we might say that 33% of the group was on statins. However, while 1/3 on a calculator is 0.333 ...., it would be a mistake to state that 33.3% of people are on statins, because the numbers don't justify the apparent accuracy of the statistic. Similarly, we would write 10/30 as 33%. When we get to 100/300 it might be worth stating this as 33.3%, but it would rarely be worth going beyond three significant figures, so 1000/3000 would also be 33.3% (although good practice is *always* to give the numerator and denominator as well as the percentage, so as to distinguish between 1/3 and 1000/3000). To avoid appearing too accurate, statistics should be rounded. Thus 100/300 is 33.33… (where … means recurring), which would be rounded to 33.3%. The general rule is that if a digit is 0, 1, 2, 3 or 4, then to round to the previous digit just drop that and subsequent digits. If the digit is 5, 6, 7, 8 or 9, then add one to the previous digit and drop the rest. Thus 2/3 to two significant figures is 0.67 or 67%. 200/300 would be 66.7% to three significant figures.

Probabilities and risks

One particular type of proportion is a probability. The probability of choosing a 'diamond' from a shuffled pack of cards is 0.25, because there are 13 diamonds in a pack of 52 cards and so the proportion of diamonds is 13/52 = 0.25. One could think of this as the probability that a randomly chosen card is a diamond. Similarly, we could say that a baby born in 2017 in the UK has a probability of 26 out of 100,000 of dying from SIDS. Here we are implicitly assuming that the probability applies to a randomly chosen baby, or a 'typical' one. If we knew more about the baby, such as whether the family had suffered a SIDS death before, then the probability would change (this is known as conditional probability, since the probability is conditional on other factors).

The proportion of events occurring over a particular time is called a *rate* (a term we used without definition earlier). Thus, we can talk of the rate of SIDS in 2017 as being 26 per 100,000 live births. This can also be described as the *risk* of dying from SIDS in 2017. Note that when the event is bad, many people prefer the term 'risk' to the term 'probability' in such contexts.

One way of thinking of the probability of an event is to think of the frequency with which an event occurs in a larger population. However, we often use probability in a different way, to indicate our strength of belief that an event will happen. Thus, we might describe the *risk* of someone dying in the next 10 years. An example of this is the risk predictor QRisk.[16] This uses data from a large population to model the expected risk of developing heart disease or stroke. Again, it uses natural frequencies to express risk. Rather than stating (to a 70-year-old man with no other risk factors) 'your risk of heart disease in the next 10 years is 15%', it states 'In 100 people like you, 15 of them will develop heart disease in the next 10 years'. QRisk also uses a 'smiley face' plot to help convey the meaning of a percentage. This is a diagram with 100 faces on it, with 15 in red and frowning and 85 plain and smiling, to show 15% visually. For some people, this is easier to understand than a number. Further suggestions for understanding risk are given by Gigenzerer and Edwards.[17]

Prevalence and incidence rate

Epidemiologists have particular terms for proportions and rates. The proportion of people in a population with a particular characteristic is termed a *prevalence*. In a population without disease at the start of a period, the number who develop the disease in a given period of time, expressed as a proportion of the population, is termed the *incidence rate*. When a rate is quoted the time period should always be given (in the long run our death rate is one!).

Trusting numbers

When given a statistic, one should always ask: Why are they telling me this and where did this statistic come from? The reason for quoting a particular statistic may be to convey a particular message, even if the numbers are unimpeachable. For example, the UK government was not lying when it stated that it had invested an extra £1 billion in the NHS in 2018, but possibly its reason for stating it in this way was to convey the message that the government was committed to the NHS and was giving it a lot more money. Without knowing what the NHS requires, however, it is a meaningless number. One would like to think that the people who wanted the UK to leave the EU (Leavers) were not lying when they stated that £350 million goes to the EU per week from the UK, when in fact, taking the rebate into account, it was 'only' £276 million.[18] Nevertheless, to the person in the street these are both big, meaningless numbers. Expressing them as a comparison renders them much more meaningful, perhaps something the Leavers wanted to avoid.

With ratios, it is important to consider whether the denominator is suitable. In example 1 earlier the denominator was the whole population. However, older people have a much higher risk of dying than younger people and a more relevant population would be the proportion of people over 65 (say). In fact, 18% of the population of the UK is over 65, compared with only 9% in the world. This might explain the 'anomaly' that the UK appears to be less healthy than the world overall.

In this era of 'fake news', vigilance in the use of numbers is more important than ever before.

Conclusions

Numbers are only interpretable when used in comparison with other numbers. Thus, when presented with a number, it is a worthwhile exercise to ask yourself questions such as: Why am I being told this number and has it been made to look big or small to bolster the presenter's arguments? Furthermore, you should consider how the number might change if the time scale were to be altered and think about what it can be compared to. If a comparison is already given, is the comparison valid? How and why was the baseline number chosen? Do not simply be a trusting consumer of numbers; learn to be an avid critic of them.

Percentage changes are always difficult to understand. Always see if you can work out what was the original figure and what was the final figure used to derive the percentage. Always check with a proportion that the denominator is relevant. There are many ways to display even simple proportions, and people find some ways of display easier to understand than others. When you present numbers, help your audience through these difficulties by always offering a variety of display methods. Good 'numberology' is everyone's responsibility.

Further reading

David Spiegelhalter's *Understanding Uncertainty* website[19] has numerous examples of displaying risk, and his book *The Art of Statistics*[20] is a good introduction to the topic. Anthony Ruben's book[17] describes simple ways to avoid being misled by numbers, some of which are in the medical field. An earlier version of this chapter has been published as a paper.[21]

Exercises

1.1 In a general practice there were 700 patients over the age of 65, and 100 of them had diabetes. Express this as a percentage.

1.2 There were 800 newborn babies in a general practice last year. This went up by 10% this year. How many babies were born this year? If we expect another 10% rise next year, how many more babies will the practice have to help?

1.3 If the number of cases of an infectious disease is increasing at 30% per day, how big will the increase be three days after the baseline was measured, as a proportion of the number at baseline?

1.4 If a person is 10% heavier than another person, how much taller must they be to have the same BMI?

1.5 The US spent $360 billion on prescription drugs in 2019.[22] Is that a large amount?

References

1. Office for National Statistics. Dataset: Vital statistics in the UK: births, deaths and marriages. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/vitalstatisticspopulationandhealthreferencetables. Accessed 26 January 2021.
2. Wikipedia. Demography of the United Kingdom. https://en.wikipedia.org/wiki/Demography_of_the_United_Kingdom. Accessed 26 January 2021.
3. Countrymeters. World population. https://countrymeters.info/en/World#population_2019. Accessed 26 January 2021.
4. World Population Review. 2020 world population by country. https://worldpopulationreview.com. Accessed 26 January 2021.
5. Gov.UK. £1 billion of funding to upgrade NHS services in England. 7 December 2018. https://www.gov.uk/government/news/1-billion-of-funding-to-upgrade-nhs-services-in-england. Accessed 26 January 2021.
6. Rudd KE, Johnson SC, Agesa KM, *et al*. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* 2020;**395**(10219):200–211.
7. Office for National Statistics. UK worth £8.8 trillion. 18 August 2016. https://www.ons.gov.uk/news/news/ukworth88trillion. Accessed 26 January 2021.
8. Sheffield Methods Institute. Resources. https://www.sheffield.ac.uk/smi/resources. Accessed 26 January 2021.
9. DocShop. Arnold Schwarzenegger is obese – problems with Body Mass Index (BMI) calculations. 6 April 2015. https://www.docshop.com/2008/04/08/arnold-schwarzenegger-is-obese-problems-with-body-mass-index-bmi-calculations. Accessed 26 January 2021.
10. Campbell MJ, Jacques RM, Fotheringham J, Maheswaran R, Nicholl J. Developing a Summary Hospital Mortality Index: retrospective analysis in English hospitals over five years. *BMJ* 2012;**344**:e1001.

11. Diabetes.co.uk. Diabetes prevalence. 15 January 2019. https://www.diabetes.co.uk/diabetes-prevalence.html. Accessed 26 January 2021.
12. Gigerenzer G. What are natural frequencies? *BMJ* 2011;**343**:d6386.
13. The Lullaby Trust. SIDS & SUDC facts and figures. August 2019. https://www.lullabytrust.org.uk/wp-content/uploads/Facts-and-Figures-for-2017-released-2019-1.pdf. Accessed 26 January 2021.
14. Cole TJ, Altman DG. Statistics notes: what is a percentage difference? *BMJ* 2017;**358**:j3663.
15. Morris J. Coronavirus: actual death toll has been 'more than double' the figures supplied by government. *Yahoo News UK*, 7 April 2020. https://uk.news.yahoo.com/coronavirus-death-toll-reporting-delays-160833341.html/?ncid=emailmarketing_ukcoronavi_a1ylavpigaq. Accessed 8 April 2020.
16. QRisk.org. QRISK®3. https://www.qrisk.org. Accessed 26 January 2021.
17. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. *BMJ* 2003;**327**(7417):741–744.
18. Reuben A. *Statistical: ten easy ways to avoid being misled by numbers*. London: Constable, 2019.
19. Spiegelhalter D. Welcome to Understanding Uncertainty. https://understandinguncertainty.org. Accessed 26 January 2021.
20. Spiegelhalter D. *The art of statistics: learning from data*. Harmondsworth: Penguin, 2019.
21. Campbell MJ, Green D, Barker D. Understanding basic numbers: examples from Brexit, Covid and common medical conditions. *Res Methods Med Health Sci* 2021;**2**(1):31–38.
22. Statista. Prescription drug expenditure in the United States from 1960 to 2020 (in billion U.S. dollars). https://www.statista.com/statistics/184914/prescription-drug-expenditures-in-the-us-since-1960. Accessed 26 January 2021.