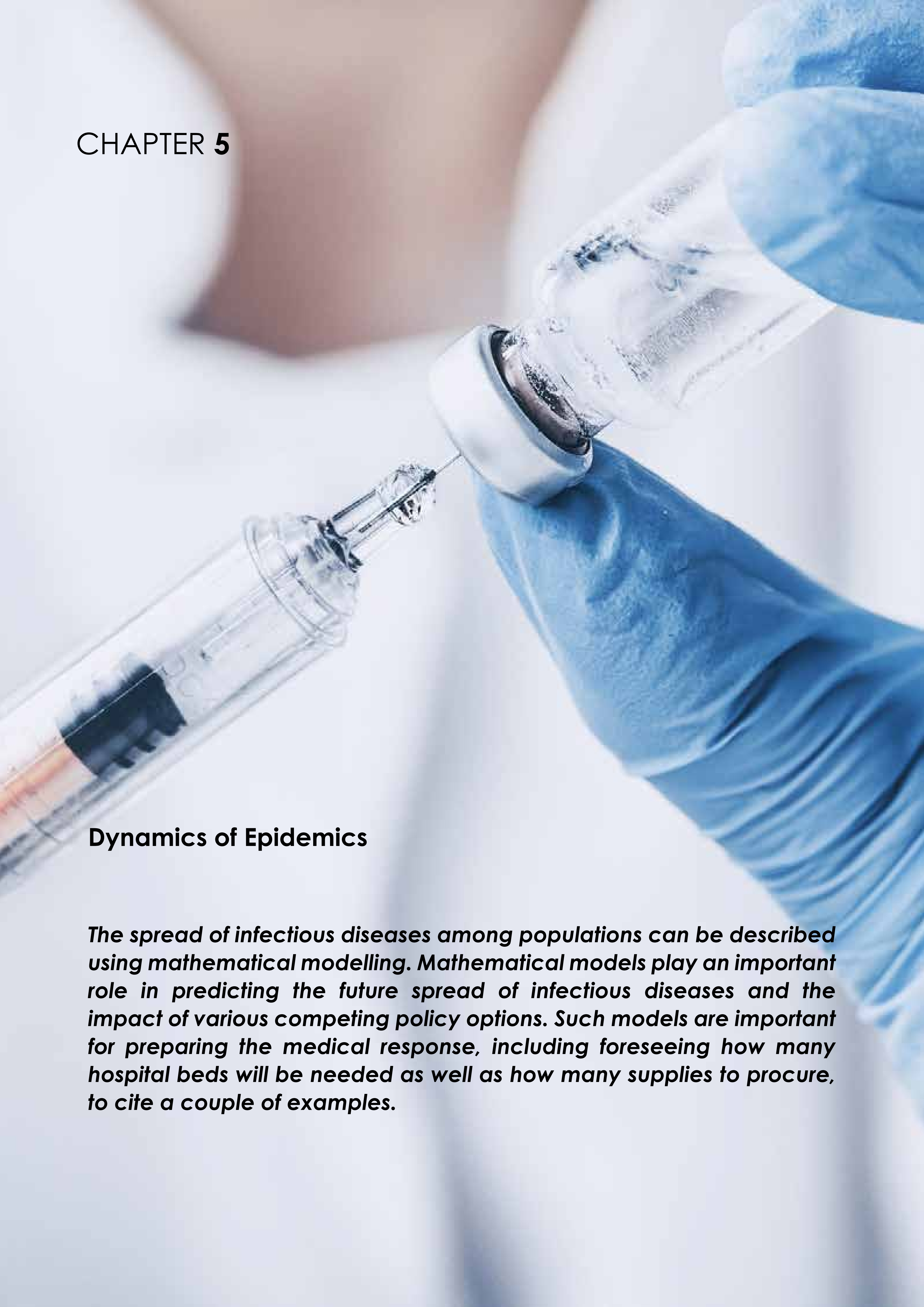
Essential facts about

Covid-19

The disease, the responses and an uncertain future

For South African Learners, Teachers and the General Public

Commissioned by the Academy of Science of South Africa (ASSAf)



CHAPTER **5**

Dynamics of Epidemics

The spread of infectious diseases among populations can be described using mathematical modelling. Mathematical models play an important role in predicting the future spread of infectious diseases and the impact of various competing policy options. Such models are important for preparing the medical response, including foreseeing how many hospital beds will be needed as well as how many supplies to procure, to cite a couple of examples.

- Mathematical modelling is imperfect, given limitations in the available data and its reliability as well as uncertainty as to the precise details of how the disease spreads. Human behaviour and psychology as well as messaging by public health officials, governmental authorities, and the media also play an important role. Nevertheless, such modelling provides insight through presenting a range of possibilities for future evolution of the epidemic.

- We first explain the notion of exponential growth. When the number of people infected is growing exponentially, the number of new infections increases by a constant factor after a same time period–say a week. This means that the number of people infected during the previous week always gets multiplied by the same number. When there is exponential growth, a tiny number of initial infections can lead to a huge number of infections rather quickly. This happens when each infected person infects more than one other person. This number is known as R or R_0 when the entire population is susceptible.

- Public health measures, such as social distancing, can decrease R, and if R becomes less than one, the number of infected people will decrease rather than grow, as we have seen at various times.

- However, as people tire of the sacrifices required to keep R below one, second, third, and higher waves of the epidemic may follow. We have seen this already in South Africa and abroad.

- Exponential growth can also be stemmed through vaccination or herd immunity. When enough people have already been infected, contacts that in a fully susceptible population would have led to growth are insufficient for each infected person to infect at least one other person on average. The same protection can be given by vaccines, as we have recently seen.

- A serious worry now is the impact of new mutated strains of the virus, some of which have been seen to spread more vigorously and also to evade, at least partially, some of the present vaccines. It is of great importance to put an end to this pandemic before more dangerous new strains emerge.

In this chapter we talk about how a contagious disease like Covid-19 spreads through a population. Whereas earlier we focused on the details of transmission–that is, exactly how the disease spreads from person to person, here we shall look mainly at the big picture: how the numbers of people infected change with time. In doing so, we will use some simple mathematics. This mathematics is also applicable to other contexts, such as finance, banking, and engineering. Of course, researchers have found other more accurate ways to model diseases than the one we consider here. Understanding how the disease spreads through a population is very important for planning and for figuring out what steps can best stop or slow down the spread of the disease. It can also help us better understand how certain actions, such as wearing masks and social distancing, beyond protecting ourselves, also serve to protect others and to promote the common good.

Daily new confirmed COVID-19 cases

Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases, the main reason for that is limited testing.

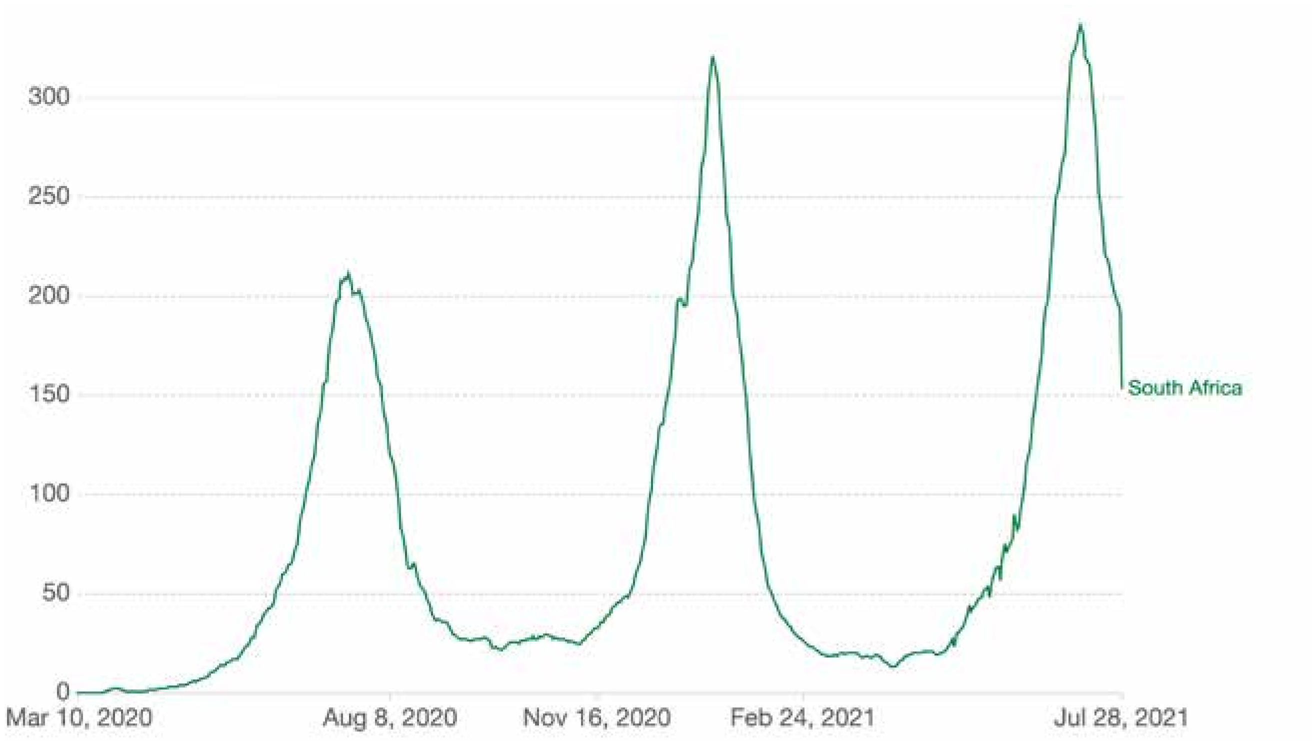


Figure 5.1: Recorded daily number of new Covid-19 cases in South Africa

Source: https://ourworldindata.org/covid-cases. Johns Hopkins University CSSE COVID-19 Data.

We begin by looking at exponential growth, which serves as an accurate model of how the numbers change over the short term, and possibly also the medium term. Exponential growth, however, cannot continue forever. If the numbers get too large, eventually other factors come into play and take over, as we later explain.

What is exponential growth?

Sometimes a gossip story is said to have 'gone viral' on social media. This means that such a story is spreading very fast, like a virus. We read in the newspapers that the virus is "growing exponentially." What exactly does this mean?

Problem: Let us begin with a simple exercise. Let us examine the three sequences A, B and C of numbers and try to figure out the rule.

Sequence A:
1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1 024, 2 048, 4 096, 8 192,16 384, 32 768, 65 536, 131 072, 262 144, 524 288, 1 048 576, 2 097 152, 4 194 304, 8 388 608, 16 777 216, 33 554 432, 67 108 864, 134 217 728, 268 435 456, 536 870 912, 1 073 741 824, 2 147 483 648, 4 294 967 296, 8 589 934 592, …

Sequence B:
1, 3, 9, 27, 81, 243, 729, 2 187, 6 561, 19 683, 59 049, 177 147, 531 441,1 594 323, 4 782 969, 14 348 907, 43 046 721, 129 140 163, 387 420 489, 1 162 261 467, 3 486 784 401, 10 460 353 203, …

Sequence C:
1, 4, 16, 64, 256, 1 024, 4 096, 16 384, 65 536, 262 144, 1 048 576, 4 194 304, 16 777 216, 67 108 864, 268 435 456, 1 073 741 824, 4 294 967 296, …

We have continued each sequence up to 10 billion, but these sequences could go on forever. The trailing three dots … mean 'and so forth'.

Answer: For Sequence A, each number is equal to two times the number before itself. For sequence B, each number is equal to three times the number before itself. For sequence C, each number is equal to four times the number before itself.

These are special cases of a geometric sequence. *A geometric sequence follows the pattern that every number is multiplied by the same fixed number to get the next number.* We may also say that the ratio between every number and the number before it stays the same or is constant.

Another way to say this is *that the numbers obey the rule:*

(Next number) = (Constant factor) x (Current number)

The factor is exactly the same for each number to follow; hence we call it a constant or the common ratio.

The growth factor does not have to be a whole number. It could be any real number greater than zero. Also, the first element does not have to be one but could be any positive number. For example, the following is a geometric sequence, even though it is not so obvious.

2.73, 3.00, 3.30, 3.63, 3.99, 4.39, 4.83, 5.31, 5.85, 6.43,…

To confirm that it is geometric, or not, we would have to calculate the ratios of successive numbers to show that they are all equal, or perhaps not.

Exercise: Which of the following are geometric sequences?

Sequence X:
11.63, 23.63, 29.03, 28.71, 33.07, 51.04, 77.81, 96.33, 93.90

Sequence Y:
11.96, 23.93, 35.89, 47.86, 59.83, 71.79, 83.76, 95.72, 107.69

You may wish to use a calculator to solve this problem.

We can use this type of mathematical method to understand the rates of infection or growth of the Covid-19 pandemic. Let us look at the table of data below of the number of people infected with Covid-19 in South Africa, in which each row is separated by one week, and see how well the data are described by a geometric pattern:

Table 5.1 Growth in Covid-19 cases in South Africa in June and July 2020

Week	Date	Cumulative number of cases	Current week/Previous week ratio
0	6 June	45 973	-----------
1	13 June	65 736	1.43
2	20 June	92 681	1.41
3	27 June	131 800	1.42
4	4 July	187 977	1.43
5	11 July	264 184	1.41
6	18 July	350 879	1.33
7	25 July	434 200	1.24

We observe that from Week 0 through Week 5 the growth from one week to the next is approximately 0.42, or in other words about 42% (on average). We say that the rate behaves like a plateau because it is almost constant in time. However, from Week 5 through Week 7 the average weekly growth rate is 28%, with evidence of a gradual drop in the rate.

How long might the epidemic last based on this data? We do not have enough information to make an accurate estimate. There is no reason to believe that the present growth rates will continue into the future unchanged, and the data does not indicate a constant growth rate but a dropping trend from a plateau. Nevertheless, there is great value in making rough estimates as long as one keeps in mind that these estimates are only approximate, and, moreover, that future events might possibly suggest that a new trend is forming. In other words, the predictions could be completely wrong.

Nevertheless, let us proceed with a certain amount of courage. Obviously, the total number of people infected cannot exceed the total population of South Africa, which is presently ~ 58 million. Consequently, let us assume very roughly that the epidemic continues to grow at the present rate until everyone has become infected. (We shall see that the endgame will be less abrupt, but let us for the moment ignore this complication.)

Consequently, from 6 June 2020 (with 4.6×10^4 infections) to when everyone has become infected (5.8×10^7 infections) requires growth in the number of infected people by some factor. How do we determine this? We estimate this by dividing the total possible number of infections (58 million people in the population) by the number of people infected at 6 June

$$\frac{5.8 \times 10^7}{4.6 \times 10^4} = 1{,}260869 \times 10^3$$

rounded off to 1261.

The monthly growth measured above by a factor of 1.42 per week amounts to 1.42^4, which is approximately equal to 4 per month. To find the number of months needed for everyone in South Africa to get infected, we must solve the equation

$$4^x = 1261$$

where x is the number of months needed for everyone to be infected and the epidemic ends. 4^x means 4 multiplied by itself x times, with fractional power defined using nth roots.

One way to solve this equation is by guessing with the help of a calculator.

We find that $4^5 = 1024$ which is slightly too small and $4^6 = 4096$ which is way too big. Consequently, x lies between 5 and 6. The correct estimate should be close to 5, so after trying values like 5.25, 5.2, and then 5.15, we could settle on x = 5.15.

For readers in grade 12, the method of logarithms can give the answer quicker. We use the natural logarithm with base e, which is the Euler number as opposed to base 10. Taking the natural logarithm of both sides, we obtain

$$\ln(4^x) = x \ln(4) = \ln(1261)$$

or

$$x = \frac{\ln(1261)}{\ln(4)} = 5.15$$

So, according to this estimate the epidemic should end by the beginning of 2021 after about 5 months. We may get an idea of the possible error by using the lowest growth rate where 4 is replaced by $1.24^4 = 2.36$. In this case we would get $x = \ln(1261)/\ln(2.36)$ which is 7.93 or roughly 8 months. This means that the end would be estimated for around mid-2021 instead. But if the growth rate in the future drops further, the epidemic would last longer.

The general rule for a geometric sequence is given by:

$$F_n = Ax^n$$

where n = 0, 1, 2,… is an integer. Such a sequence may be plotted using dots, as follows.

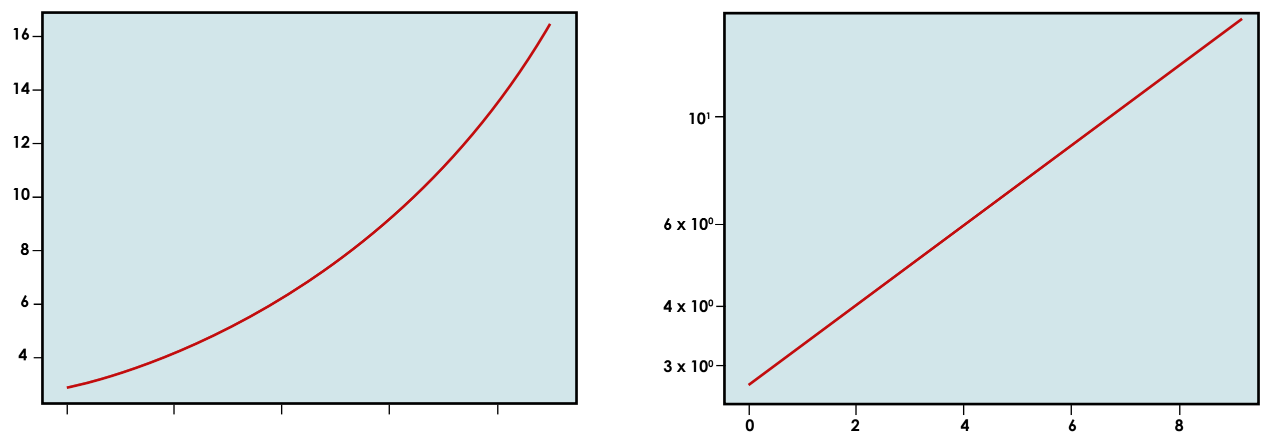


Figure 5.2: Plotting geometric sequences

A geometric sequence is plotted on the left using a linear scale for the vertical axis, and on the right using a logarithmic scale instead. With the logarithmic scale, the dots arrange themselves perfectly along a straight line, showing visually that the growth is geometric.

The mathematics of exponential growth and decay appears in a wide variety of diverse applications such as radioactive decay, the absorption of light, nuclear chain reactions, finance and banking. In the simplest model of inflation (with a constant rate of inflation), prices grow with time according to an exponential curve.

Because inflation growth is an exponential function, we sometimes find some countries have rates of inflation in the thousands of percent whereas other countries, such as South Africa, keep the inflation rate down to single digit numbers through a strict monetary policy. Money in a savings bank account also grows exponentially, provided that the rate of interest does not vary.

Why (and when) do we expect case numbers to follow a geometric progression?

Let us consider a simple model of how the disease spreads. Suppose that each person who gets infected in turn, on average, infects the number R_0 of other people, after a time T, known as the 'serial interval'. As long as not too many people have been infected, with each successive generation, R_0 as many people are infected as had become infected at a time earlier by T. We know that the virus spreads from person to person by respiratory droplets and aerosols. These are exhaled, to some extent by simply breathing, but even more so by speaking, singing, panting as the result of strenuous exercise, or coughing. These droplets may be directly inhaled by someone nearby, or the small aerosols may linger in the air for hours (especially in an enclosed, poorly ventilated area) and then inhaled by someone. The serial interval T is more a property of the infection itself, but R_0 is strongly dependent on people's behaviour. If people stay at home and hardly go out, R_0 will be greatly reduced, and if R_0 falls below 1 for long enough, the disease will eventually die out and become extinct, at least in the geographical area where that is the case. Wearing masks will also reduce R_0, as will avoiding events where many people gather. However, in places like South Africa, where people living

Covid-19 in the United States

The pandemic struck the United States after Europe. But despite this late start, the United States now leads the world both in cases and deaths. Although the US claims only 4.2% of the world's population, as of 22 July 2021, 26% of the recorded Covid-19 cases and 18% of the recorded Covid-19 deaths were from the US. The epidemic first hit the large cities, in particular the New York City metropolitan area, and then spread across the country to less populated states and rural areas. Although the US government shut borders early on, little else was done to prepare, particularly by the federal government. President Donald Trump claimed persistently that the epidemic was a "hoax" and would simply "go away" by itself. When numbers exploded, hospital emergency department and intensive care unit numbers exceeded capacity in many big cities. Moreover, medical personnel lacked needed personal protective equipment and testing supplies.

(continued on page 46)

in close proximity in informal settlements and traveling using public transport increase opportunities for close contact, some of the preventative recommendations may not be implemented.

In our estimates of the future course of the epidemic based on geometric growth, we assumed that R_0 stays constant. But from the above discussion we see that we, as individuals, may control R_0 by adjusting our behaviour, at least to some extent. The government may also affect R_0 by suggesting guidelines for people's behaviour aimed at slowing, and also possibly reversing, the growth of the epidemic, or by imposing regulations. Identifying infected people and isolating them until they are well are other measures that help stop the spread of the disease. People's perception of the seriousness of the disease and of their own probability of getting infected also plays a role in determining behaviour, and thus the media and also the statements of political and community leaders can make a difference.

Consequently, we see that the spread rate following a geometric sequence assumes that people do not change their behaviour. Another factor affecting the spread is how many people have already been infected. It is generally assumed that when a person is infected and recovers, they retain an acquired immunity for some time, possibly for the rest of their life, as is the case for other types of viruses, although the situation for Covid-19 is uncertain. Consequently, when it is no longer the case that almost everyone is susceptible (in other words, they can be infected), the rate of growth slows down, even if people do not change their behaviour. If a fraction f of people are either presently infected or have recovered and, therefore, are immune, with each new generation $R = (1-f)\, R_0$ people will become infected, rather than the larger R_0.

We, thus, see that the geometric progression or exponential growth described in the previous section is not a hard and fast rule, but simply a rough guide. In principle, more accurate mathematical models

Covid-19 in the United States

(continued)

After Northeastern states managed, through the efforts of their governors, to bring the numbers down, numbers shot up in other states, in particular in states in the South and Midwest, largely as a result of a premature reopening promoted by then President Trump. Unfortunately, in the US, mask wearing and social distancing have been turned into a political issue rather than a question of public health and plain common sense. The US subsequently experienced a serious second wave with hospital capacity overwhelmed in many parts of the country, with daily new cases peaking above 300 000 on one day in January 2021.

Presently thanks to a massive vaccination campaign, with 49% of the population fully vaccinated and 56% having received at least one dose (22July 2021), the number of new cases has been brought down and life is starting to return to normal. However, because of the new more contagious delta variant, numbers are starting to rise again.

Mathematical Functions For Epidemiology

Let $a > 0$ be a positive real number and x any real number. Then a^x (read "a to the power x," or more succinctly "a to the x") is standard notation for the power function, which is included on most calculators and as part of most computer programming languages. The power function obeys the following relations:

$a^0=1$, $a^1=a$, $a^2=a \cdot a$, $\sqrt{}=a^{1/2}$,…
$a^x\, a^y = a^{x+y}$
$a^{xy}=(a^x)^y$

The exponential function, denoted as exp(·), is a special case of the power function that pops up again and again

$\exp(x)=e^x$ where $e=2.71828\ldots$ known as Euler's number
$\exp(x+y)=\exp(x)\exp(y)$
$\exp(0)=1$, $\exp(1)=e$

When x is small (compared to one), $\exp(x)\approx 1+x$
The inverse of the exponential function is the natural logarithmic function, denoted as ln(·), or sometimes as $\log_e(\cdot)$, which is read logarithm base e, so that

$\exp(\ln(x))=x$
$\ln(\exp(x))=x$

and the natural logarithm has the properties

$\ln(ab)=\ln(a)+\ln(b)$
$\ln(a^x)=x\ln(a)$

can be developed that could account for these complications, and this is precisely what applied mathematicians and epidemiologists do for a living. The success of such modelling is limited because human behaviour is hard to predict. Moreover, more sophisticated models include a large number of input parameters that are hard to determine from the data.

It is also possible that the virus will mutate, possibly in a way that will increase its spread. Mutation means that the virus goes through changes and perhaps becomes a stronger form of the virus. Any vaccine that is developed may not be useful for too long if the virus mutates, and upgraded vaccines may need to be developed. The virus can also change in how it affects those infected. It can become more or less virulent. This means that it becomes more or less severe.

What can we learn from the data?

We already looked at some data showing that a geometric progression is not a bad model to start with in the short term. But let us look deeply at all of the data so far to see what we can learn about how the virus has spread and the impact that different policies have had on the numbers. In the previous section, we saw arguments around why the growth rate may be expected to change with time, at least somewhat. We now look at real data to see what has in fact happened.

The virus started in China and spread throughout the world. With globalisation and frequent air travel, the spread could not be avoided, although many people had hoped otherwise. Many countries closed their borders or tried screening travellers entering, for example by checking their body temperature, as had been done successfully with the SARS virus during the short-lived 2003 epidemic. These measures, however, were not effective enough for Covid-19, because not everyone with Covid-19 gets a fever, and moreover, infected people become infectious–meaning that they could spread the disease to others–before they start showing symptoms. Moreover, later on it was found that a subset of those infected never showed any symptoms but, nevertheless, spread the disease to others. This explains, at least in part, why Covid-19 spread around the world and is now virtually everywhere, whereas SARS was successfully eradicated within a few months.

People's perception of the seriousness of the disease and of their own probability of getting infected also plays a role in determining behaviour.

Different countries adopted different policies, which we will discuss further below. However, the general progression of the epidemic seemed to obey the following general trends:

1. In the beginning, although people had heard about the outbreak in China, they did not actually believe that it would reach their country, or they believed that their country would be spared. During the period (until mid-March 2020), we see a very rapid growth rate.

2. As numbers rose exponentially, and in many places, hospitals began to fill beyond capacity; people and governments reacted by imposing drastic measures such as lockdowns, which reduced–and in many cases also reversed–the growth rate. In some countries such as Australia, China, Israel, Vietnam, and New Zealand, to name a few, these measures succeeded in reversing and virtually eliminating the virus. In other places, a slowdown of the growth or modest reduction of the virus spread was achieved.

3. In many places, however, because of the great economic cost of maintaining a lockdown and opposition from business interests, measures were loosened, accompanied by an increase in the growth rate.

The Covid-19 pandemic is an on-going story of which the ending cannot be predicted with any confidence. There are many plausible scenarios. Yet it is useful to look at the available data from different countries, looking in particular at how the growth rate has evolved over time and is correlated with the various policies implemented.

Probably the most useful way to plot the data is using a logarithmic scale for the daily number of cases. On a logarithmic scale, the instantaneous slope indicates the instantaneous growth rate, and multiplying by a constant factor is equivalent to sliding the curve upward (if $f > 1$) or downward (if $f < 1$). Exponential growth corresponds to a straight line of positive slope, and exponential decay corresponds to a negative slope.

It is important to keep in mind that the data available to us is not perfect. Ideally, we would like to know the exact number of new infections each day in a given country or province, but only a certain fraction of these new infections is reported

Daily new confirmed COVID-19 cases

Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases, the main reason for that is limited testing.

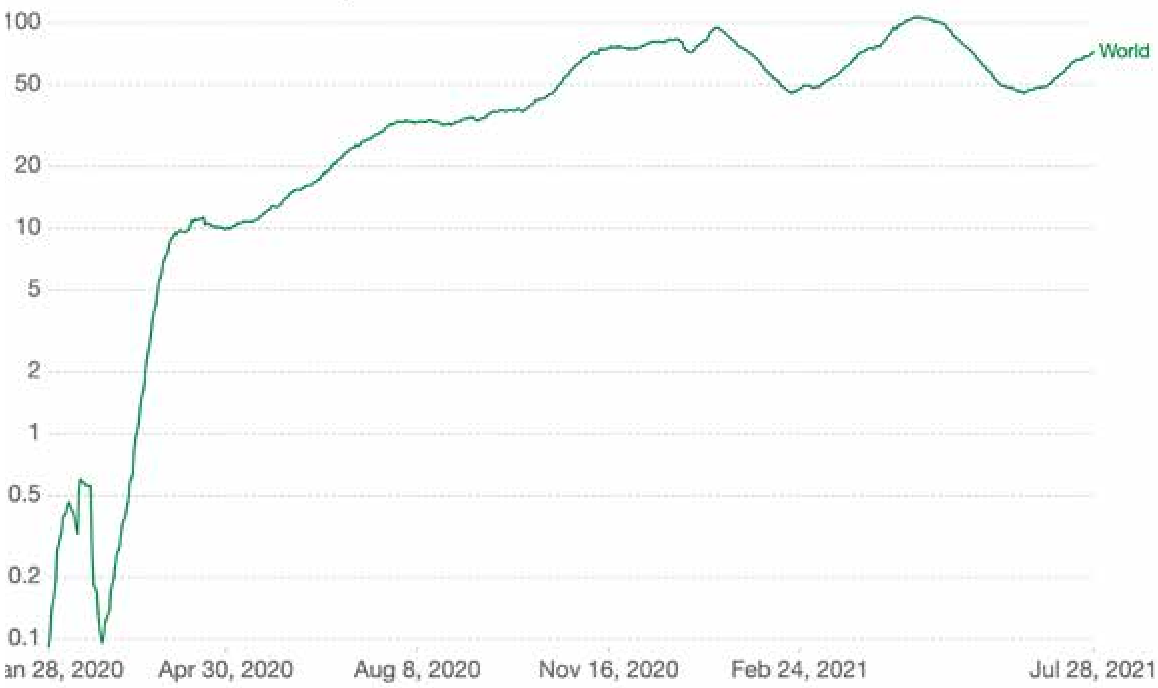


Figure 5.3: Daily new cases worldwide (since July 2021)

Source: https://ourworldindata.org/covid-cases
Creative commons license

and, consequently, makes it into the official statistics. But it would be reasonable to assume that the under-reporting takes place almost everywhere, so some predictions can still be made from reported cases. Covid-19, at least shortly after someone has initially been infected, does not have very specific symptoms. Consequently, the only reliable diagnostic test is detection of the viral RNA in the nasal mucus through an RT-PCR test, which, at least initially, was not widely available. Consequently, we have every reason to believe that the actual number of infections is being undercounted by some factor not so easy to determine. Because of limited testing capacity, many people who believe that they may have the disease are not able to be tested. In any case, because of costs, some governments only test people who show symptoms. Moreover, as we have recently learned, some people who become infected do not have any symptoms, yet are able to infect others and, thus, spread the disease. Despite these challenges, the imperfect data available to us is better than no data at all. Otherwise, we would have a situation where everyone is left to speculate according to their whims and personal preferences.

We shall first look at the data for a few representative countries for the number of new cases against time, later looking at some other types of complementary data. There are many ways to plot the data and each method highlights a different

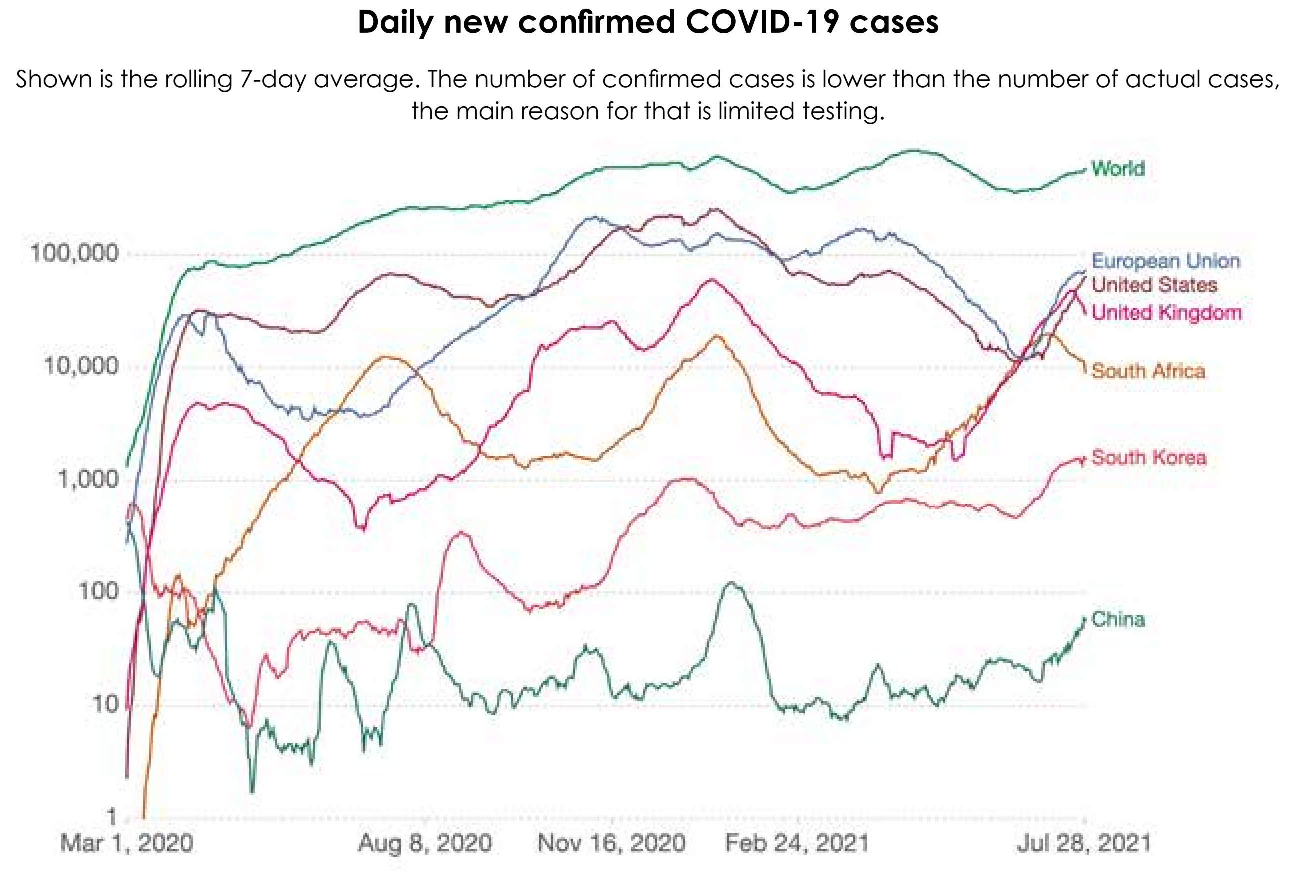


Figure 5.4: Daily new cases–A few representative countries and regions

Source: https://ourworldindata.org/covid-cases
Creative commons license.

aspect of the data. Nevertheless, one common tool is using a logarithmic scale for the number of infections. Logarithms (regardless of their base) have the property log(ab)=log(a)+log(b). This means that multiplying a curve by a constant factor slides it either up or down, depending on whether the factor is greater than or less than one, but the shape of the curve is not changed. A logarithmic vertical scale allows us to focus on the evolution of the growth (or decay) rate of the infection. Exponential growth on a log scale is represented by a straight line of positive slope, the magnitude of the slope being equal to the growth rate. Similarly, exponential decay is represented by a straight line of negative slope, the magnitude of the slope representing the decay rate. In general, the curve will not be straight, indicating a changing growth (or decay) rate.

With some understanding of the mathematics used in modelling, let us look at some data, like the daily infection rate for the entire world, in order to identify a few general trends. We begin at 23 Jan 2020, when there were 100 cases per day (in order to avoid noise and other sampling artifacts). Noise refers to changes in the data caused

Daily new confirmed COVID-19 cases

Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases, the main reason for that is limited testing.

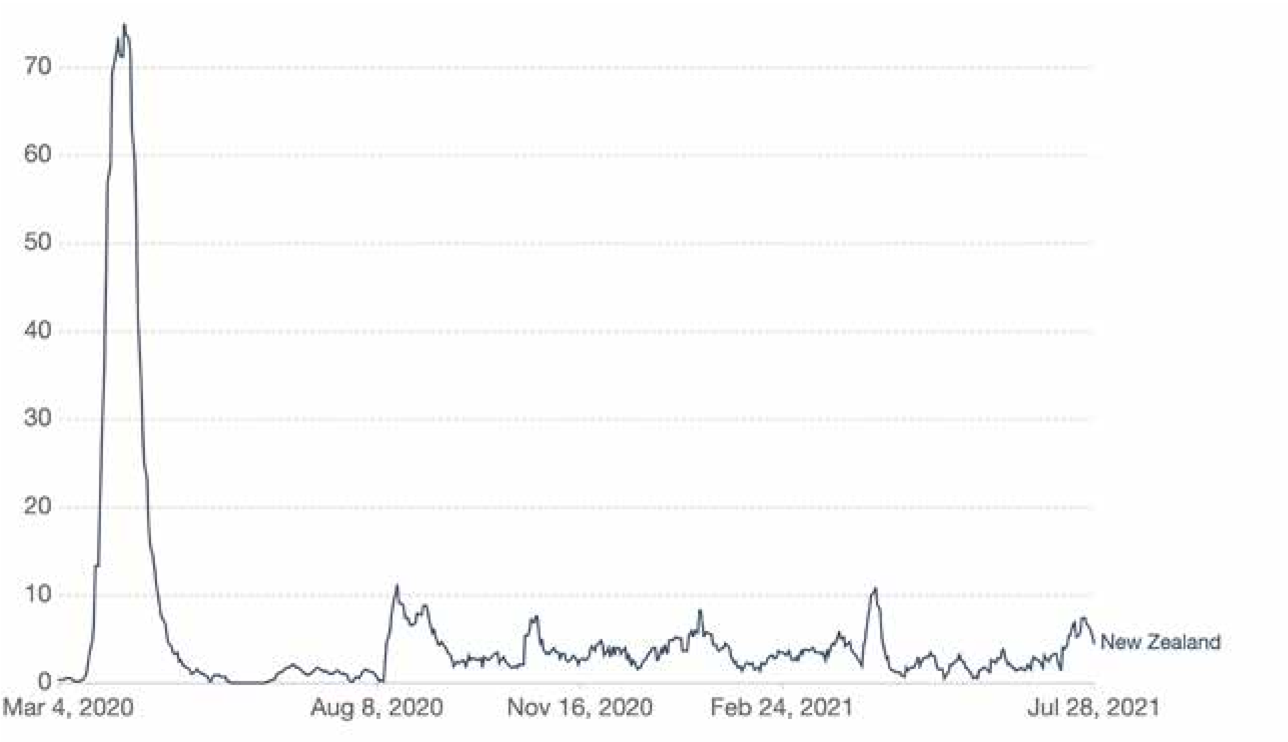


Figure 5.5: New Zealand daily new confirmed COVID-19 cases

Source: https://ourworldindata.org/covid-cases by other influences not related to the disease. From 24 Feb 2020 through to 24 March 2020, we observe an 83-fold increase, corresponding to a doubling of the infection rate every 4.4 days. During this time, China was able to bring down the number of infections through drastic measures. However, the pandemic had already spread to Europe, which was slow to react, leading to a massive increase in infections, until Europe put in place a lockdown, leading to a massive drop in the number of daily infections. The US, however, was slow to react and now leads the world in new infections and deaths.

The data exhibits weekly ripples, primarily explained by delays in reporting cases over the weekend. In some countries there were delays in getting test results as these were not released in a regular manner. By contrast, in the 76 days from 11 May 2020 through to 26 July 2020, the worldwide daily infection rate increased by a factor of 3.33, corresponding to a doubling every 44 days. Since then, the worldwide daily number of new cases has continued to climb. [There are many compilations of worldwide Covid-19 data (see references at end). The plots here are taken from the website: https://ourworldindata.org/covid-cases]

China, where the initial outbreak occurred, seems to have brought down and controlled its numbers. The European Union was the next region hit by the pandemic and, for a while, led the world numbers, before it was overtaken by the United States. South Africa (as well as the rest of Africa) was a latecomer and, after peaking in mid-June, saw a substantial decline followed by a new peak, or 'second wave'.

New Zealand: An Apparent Success Story.

Some would argue that China's success in reducing and reversing the spread of the virus may be due to its unique social system and, thus, cannot be repeated elsewhere. Very strict government control steps were followed by a compliant population. However, New Zealand, which has a social system quite unlike China, has, at least to date, also been able to almost extinguish the virus. It is worth noting that the population of China was about 1400 million in 2020, while the population of New Zealand is merely about 5 million, so the scales are vastly different. A small population is obviously easier to manage than a huge one. Other examples where the numbers have been brought down close to zero include Australia and Vietnam, although Australia did suffer a second wave. Israel was successful in bringing down the numbers at the end of May 2020, but has since been less successful.

We could show more plots and different ways of plotting the data, and there is a lot more to be learned from the data. But, instead, we invite readers to explore the data by themselves using the internet and to draw their own conclusions. In any case, by the time this booklet goes to print, the above data will be out of date, so the various cited websites are the best source of up-to-date information.

Good websites are: https://www.worldometers.info/coronavirus/, https://ourworldindata.org/covid-cases for worldwide data, and https://sacoronavirus.co.za/, https://www.covid19sa.org/ for data from South Africa, nationally as well as by province.

Above, we focused on daily case numbers based on test results, but it is good practice to look at as many different types of data as possible to look for corroboration (where there is agreement) as well as possible inconsistencies (where there are differences). This is necessary when we know the data is not perfect but we want to put together the most reliable picture of what is going on.

Other types of data are number of hospitalisations due to Covid-19, deaths as a result of Covid-19, number of people admitted to intensive care for Covid-19, fraction of positive test results, and excess mortality (number of deaths). Each tells us something slightly different and has its own pros and cons concerning accuracy. While there might be great variations on whether people with less serious Covid-19 are tested, those requiring hospitalisation are more likely to be counted in the same way, and counting the more serious cases, requiring more scarce medical resources is of great interest in its own right. The fraction of positive test results is, to some extent, less dependent on how much testing takes place. Assuming that a varying fraction of people with Covid-19-like symptoms gets tested, the fraction of positives is more likely to reflect the true increase in the prevalence of new infections than the raw number of positive test results. Mortality data (death rate), which may be expected to lag the number of new infections by several weeks, is perhaps more reliable for measuring the shape of the curve, assuming that the cause of death is usually determined correctly. These data, however, are not completely unambiguous, meaning they can be interpreted in different ways. Many, but not all, people dying from Covid-19 have co-morbidities (meaning other diseases at the same time). Consequently, there may be several causes of death, and no right answer in how to identify the most important cause of death. Whatever the actual cause of death, hospitals would have the ability to tell if the patient was Covid positive or not.

Looking for excess mortality (higher than normal death rates) sidesteps this problem and also gives a way to ensure our deductions are sensible. In situations where hospitals are overwhelmed, as was the case in Italy in late March 2020, comparing the rate of deaths with previous years allows for unrecorded Covid-19 deaths to be detected, as can be seen by comparing the 2020 curve to previous years.

Excess mortality during COVID-19:
Deaths from all causes compared to previous years, all ages

Shown is how the number of weekly deaths in 2020 - 2021 differs as a percentage from the average number of deaths in the same week over the years 2015-2019. The metric is called the P-score. We do not show data from the most recent weeks because it is incomplete due to delays in death reporting.

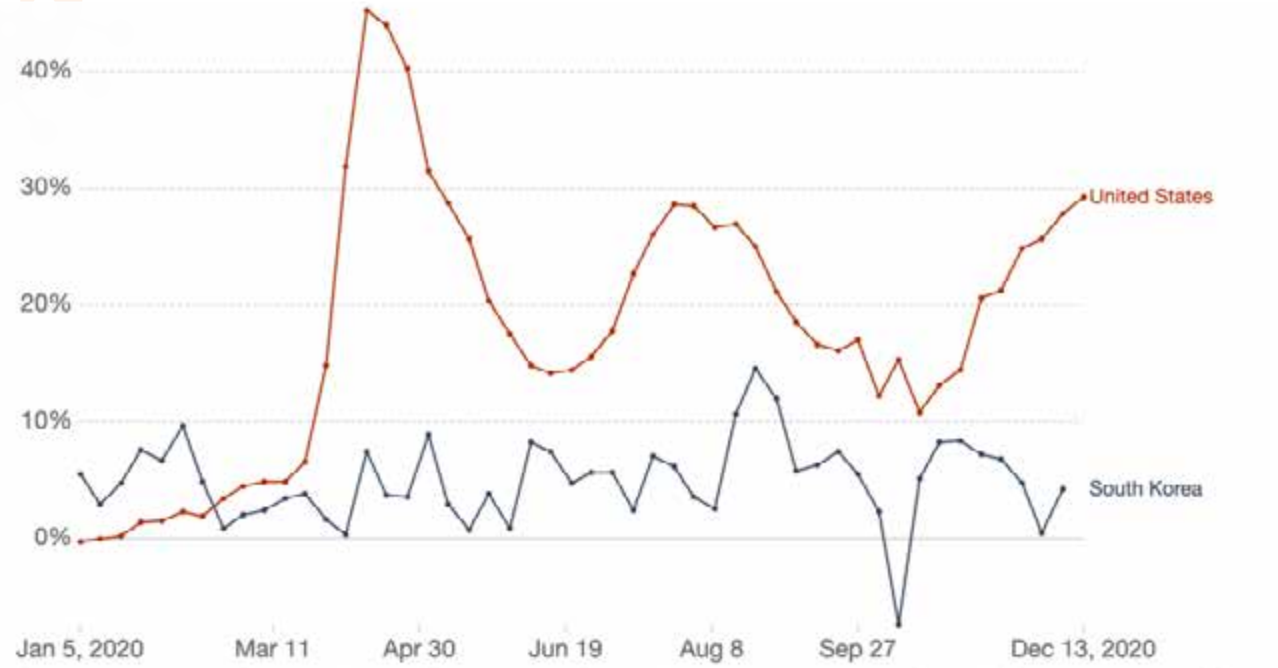


Figure 5.6: Comparison of excess deaths in the United States and South Korea

Source: https://ourworldindata.org/covid-cases
Creative commons license. Human mortality database (2021), UK Office forNational Statistics (2020)

Weekly mortality data can be compared with previous years to estimate the number of deaths due to Covid-19 by another method. We observe three peaks of excess mortality in the United States, the highest reaching 45%, whereas in South Korea, even though 2020 had some excess mortality, the rate averages to around 5%. Examining excess death data, rather than the statistics of deaths formally attributed to Covid-19, includes people who died of Covid-19 without being diagnosed as such. However, part of this excess mortality may be due to people who died of other causes, perhaps because they did not access healthcare, either because of fear of getting infected or because of medical resources being diverted to dealing with Covid-19 patients.

South Africa weekly Deaths from all causes 1+ years: 29 Dec 2019 - 16 Jan 2021

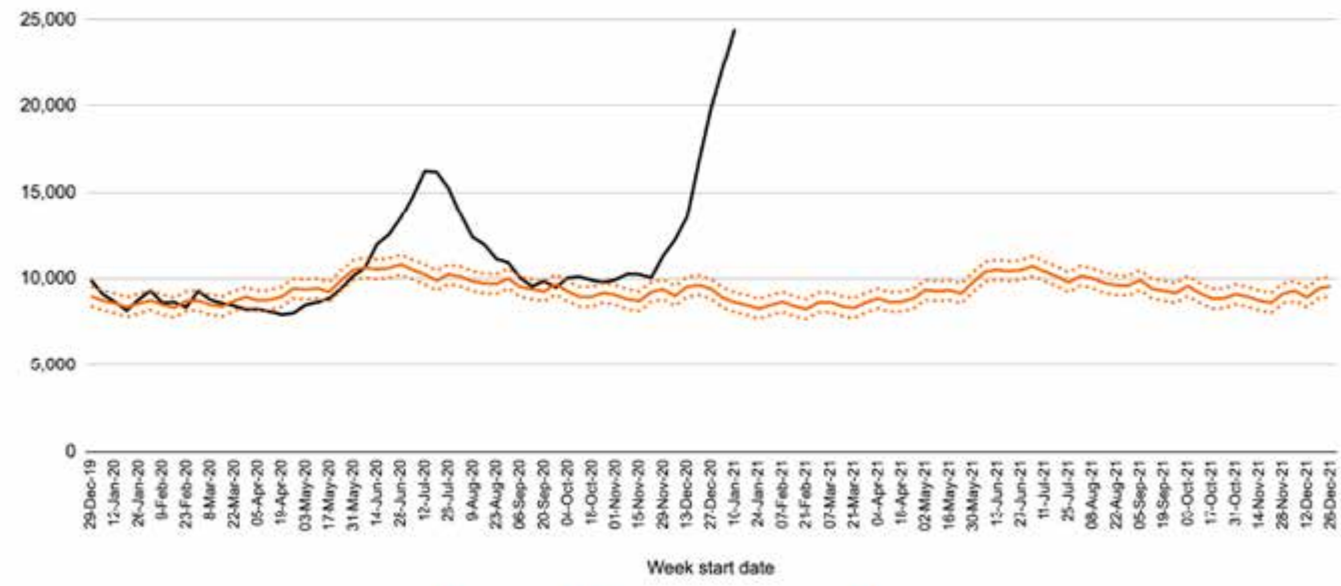


Figure 5.7: South African excess mortality data

Source: South African Medical Research Council https://www.samrc.ac.za/reports/report-weeklydeaths-south-africa

The solid red line indicates the weekly mortality prediction based on previous years, with the range expected from statistical fluctuations indicated by the dotted red lines. The solid black line indicates the 2020 data. R_0 is defined to be the average number of people newly infected by each infected person. As explained above, when a sizable fraction of the population is presently infected or has become immune, the number of new infections is smaller and denoted by R.
In the simplest model:

$$R=(1-f)R_0$$

where f is the fraction of people already infected or immune.

Because R_0 depends on people's behaviour, it can be reduced by the various measures outlined in the previous chapter: masks, quarantines, avoiding congregating in large groups, social distancing, testing and isolating.

What is 'herd immunity'?

Above, we explained how the virus will eventually go away if R drops below one. Happily, making the virus go away does not require every transmission of the virus from one person to another to be stopped. It is enough to lower the rate of infection so that each infected person infects, on average, fewer than one other person. Mathematically, this means that the reproduction number drops below one. In this case, with time, fewer and fewer people are infected, and eventually the number of people infected will drop to zero. If R is only slightly less than one, this will take a very long time. Many generations of infection will be needed to make the virus go away. But if R is much less than one, a small number of generations will be needed.

Globalisation complicates the eradication of the virus by such means. A country may get their numbers down to zero or close to zero by a strict lockdown over several months, as occurred in China. If people from other countries where the infection persists were not travelling into the country and creating new infections, the lockdown could be lifted and people could go back to their old ways without the virus. R_0 can be much greater than one as long as the number of people infected is zero. But introducing one new infection can restart the exponential growth described above. Consequently, there is great interest for all countries to work together on a worldwide strategy for eliminating the virus everywhere.

As we mentioned above, it is most likely that once someone has been infected with Covid-19 and recovers, they are immune from being re-infected by the virus for a very long time, possibly for life. This is the case for many (but not all) viruses, so this is our present best guess. Moreover, immunity can be conferred by a vaccine, as described above, rather than by actual infection. The more people are immune, the lower R becomes compared to R_0 so R can be made smaller than one once a certain critical fraction of the population has acquired an immunity, either from actual infection or from having taken a vaccine. If everyone is the same, the fraction of people who need to be immune is not 100%. It is sufficient for (1-f) R_0 to be smaller than one. When at least this fraction of immunity has been achieved, the rest of the population acquires what is sometimes called 'herd immunity'.

In the case of vaccines, the fraction vaccinated never approaches exactly one (or 100%), and all people do not respond to the vaccine in the same way. For many vaccines, some people acquire an immunity while others do not, especially older people whose immune

Covid-19 on the Diamond Princess Cruise Ship

The fateful voyage of the Diamond Princess, a British cruise ship that carried approximately 3700 passengers and crew, starting in Yokohama on 20 January 2020, demonstrated how easily the disease could spread in a closed environment. This incident also provides an estimate of the fraction of asymptomatic cases because everyone aboard was eventually tested for Covid-19. When the ship departed, one passenger who had been to China had developed a cough but boarded anyway. After disembarking in Hong Kong, he tested positive for Covid-19. Despite being notified immediately by the Hong Kong public health authorities, the cruise ship line delayed taking action and notifying the passengers. Partying in close quarters on the ship continued. Upon learning of the outbreak, the Japanese government decided to quarantine the passengers and crew on the ship for 14 days. Initially 30 ill passengers and crew were tested, giving 10 positives.

(continued on page 57)

systems are typically less responsive. But thanks to 'herd immunity', a vaccine can protect an entire population without being 100% effective. It is enough for a certain critical fraction of the population to be vaccinated, and then everyone is protected. It is also important that there not be pockets of sub-populations with a sub-critical fraction of people vaccinated, as has sometimes occurred due to rumours of the dangers of vaccination causing parents to fail to vaccinate their children. There have recently been a number of outbreaks of measles because of this. Someone travels to a country were measles is still prevalent and returns, infecting a sub-community, which then suffers an outbreak.

How is Covid-19 likely to end?

This is the big question and nobody knows the answer for sure. However, a number of scenarios can be sketched of how this on-going story may play out. With time we will find out which scenario unfolds.

Scenario I–A good vaccine comes along soon

Let us start with the happy ending–or best case scenario. Presently many different vaccines are in the process of being developed and tested, some of which are now available.

Nevertheless, a number of challenges remain. Even in the wealthiest countries, owing to various challenges, it will be many months until enough people have been vaccinated to substantially stem the continuation of the pandemic.

The vaccine must be produced in massive quantities, with billions of doses, and distributed throughout the world. Having just a few countries vaccinate its population will not suffice to eradicate the disease. The immunity may also wear off, meaning that with time infections will return to the countries initially benefiting from the vaccination.

Covid-19 on the Diamond Princess Cruise Ship

(continued)

In the ensuing days, the isolation measures on the ship proved inadequate and, as more and more people were tested, the number of positives climbed. In the end everyone was tested, which showed a 19.2% positive rate. Among the positives, 42% were asymptomatic. Over 700 people were infected and 14 people ultimately died.

Another danger in this case has to do with the possibility of the virus mutating. All viruses mutate to some extent, and some mutations are harmless. But the virus may mutate into a new form against which the vaccine no longer works.

For decades, scientists have sought to develop a vaccine against HIV-1, and there have been some partial successes. But each time the virus has mutated into a new form against which the vaccine proved to be powerless. This is because HIV-1 mutates particularly quickly. This property has made it challenging to develop effective

Measles cases in the United States 1944 - 2007

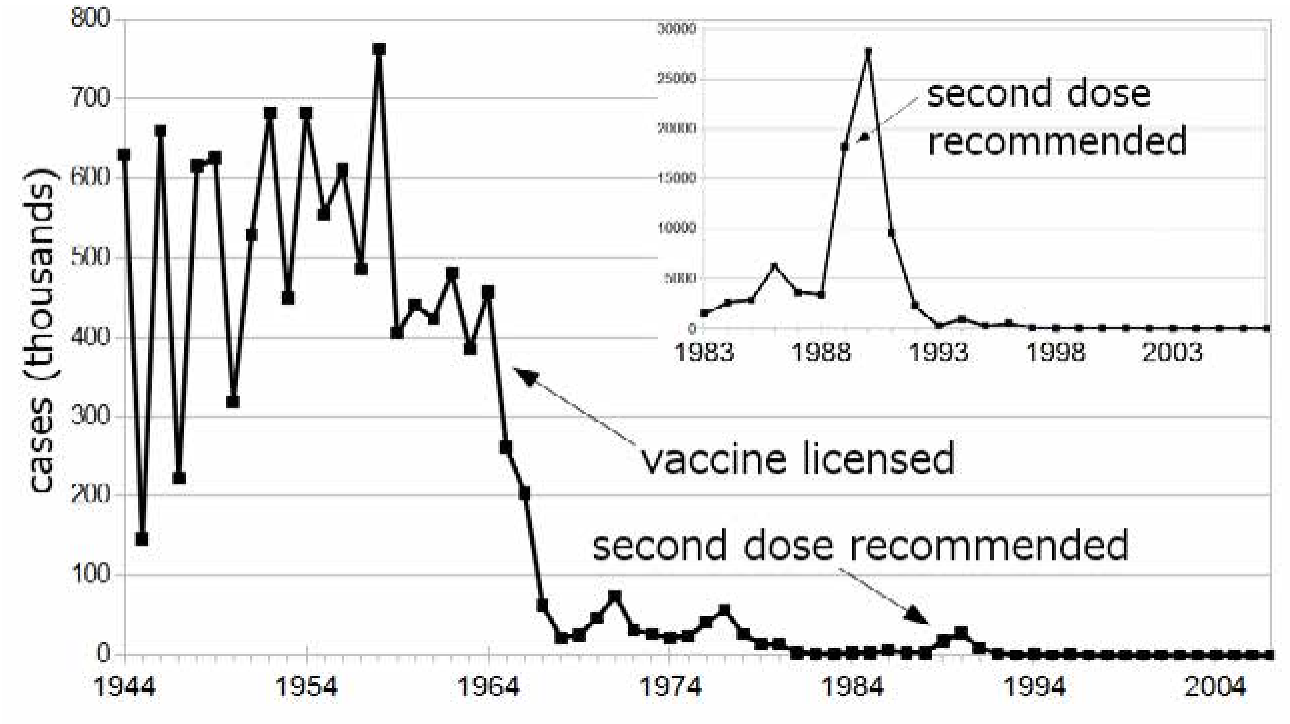


Figure 5.8: Incidence of measles in the United States

Source: https://commons.wikimedia.org/wiki/File:Measles_US_1944-2007_inset.png

drugs against HIV. The early drugs proved effective at first. But after some time a resistant strain would develop within the patient against which the monotherapy proved powerless. It is worth noting that South African scientists have been at the forefront of research in HIV research largely because of the local population having a high rate of infection. Happily, coronaviruses mutate more slowly. Whereas the reverse transcription of HIV is extremely error prone, producing many mutations, the RNA replication process of the coronavirus includes an error correction mechanism leading to a comparatively low mutation rate.

Thus, we see that there is hope for this optimistic outcome to be realised eventually but it is by no means guaranteed.

Scenario II–Covid-19 becomes endemic

If the population remained fixed and infected people became immune for life, the disease would eventually die out as a result of herd immunity, as described above, although the final stage of extinction would occur very slowly. It is possible that seasonal effects, of little relevance when R is much greater than one, could speed up a possible extinction of the virus. But a population over a long timescale is not a fixed population. Older people die and are replaced by new births, even if the total population number does not grow. Consequently, the new births will drive the susceptible fraction above the critical value, and new outbreaks will occur.

This is not just theoretical speculation but describes the situation for measles before a vaccine came along in the mid-1960s. Abundant, high quality data is available for measles, as shown in Figure 5.8.

The number of cases saw a seasonal periodicity until a vaccine was discovered in the mid-1960s followed by a program of mass vaccination, which virtually eliminated the disease. However, recently, due to false information claiming that the vaccine is not safe and leading to pockets of unvaccinated children, new outbreaks have resulted.

The somewhat irregular spikes showing a periodicity of one or two years are most likely due to seasonal effects, which cause R to oscillate with a yearly pattern. We have not yet observed any seasonal variation in Covid-19 despite speculation that it would "go away like the flu", which away from the tropics is a wintertime phenomenon. In the summer of each hemisphere there are hardly any cases of the flu, but in winter the new strains propagate away from the tropical regions with a remarkable regularity. In the case of Covid-19, if R hovers around one, seasonal effects could become important.

This scenario would not be a happy ending. But neither would it be the worst-case outcome. People would

Covid-19 in France

After China, Europe (particularly France, Italy, Spain, and later the United Kingdom) was hit hard by the Covid-19 pandemic. The number of new cases saw a peak around the beginning of April 2020 but, by means of lockdowns and other measures, European countries were successful in greatly reducing the number of daily new cases toward Summer 2020. However, starting in Autumn, the number of cases has steadily risen, resulting in a new series of lockdowns.

After the number of new cases was doubling every few days, the French government announced a lockdown on 11 March 2020 that lasted till mid-May. All people except essential workers were confined to their homes. All stores, except grocery stores and pharmacies, were closed. These measures succeeded in bringing the daily new case numbers down to about a tenth of their peak value, at which point a gradual reopening was put in place.

(continued on page 60)

be infected at an earlier age, when Covid-19 is least serious. But to reach this steady-state situation, a lot of people would have to become infected, and many people would die. And as we are learning, those who recover from the more serious cases of Covid-19 often suffer serious health problems that do not go away.

This scenario assumes that the virus retains its present character over the years to come.

Scenario III—The virus mutates into a new strain that is much more infectious and deadly

Covid-19 is the worst disease outbreak in the last 100 years. Competition includes the so-called Spanish Influenza of 1918. Although many of the infectious diseases of the past, against which effective vaccines are now commonplace, have been around for at least a few hundred years, viruses evolve and new strains come along. For influenza, for example, each year one or more new strains appear due to mutations and different combinations of viral genes. This is why each year a new vaccine against influenza must be formulated, tailored to the strains thought to become prevalent during that year. Vaccinating against influenza, because of its rapidly mutating character, is akin to striking a moving target. Other viruses, prevalent among other species such as birds, pigs and bats, through mutation, change species and become infectious for humans. This is what happened for Covid-19, which has been shown to be genetically extremely similar to coronaviruses prevalent among bats, and which possibly jumped to humans via some other intermediate species.

We thus see that viruses evolve, and they evolve quickly compared to other kinds of living organisms. SARS- and MERS-CoVs are other coronaviruses similar to SARS-CoV-2, but much more deadly. The case fatality rates for MERS, SARS, and SARS-CoV-2 are given as 37%, 9.2%, and 3.8%, respectively, although the last figure is the least certain due to difficulties in identifying all Covid-19 cases. Although all SARS and MERS always

Covid-19 in France

(continued)

One of the drivers for the lockdown was the fact that hospitals were being overrun, in terms of number of beds, trained medical personnel, and capacity in intensive care wards. Patients in the East of France had to be sent to hospitals elsewhere in the country for lack of local hospital capacity. In the Paris metropolitan area, hospitals were very close to exceeding capacity. At first, there was also a shortage of masks and other personal protective equipment for medical personnel.

France, however, has experienced a second wave, as has most of Europe and the United States, leading to a new lockdown, followed by curfews and other restrictions.

France's second wave peaked around the beginning of November 2020, followed by a third wave that peaked around mid-April 2021. As of writing we seem to be seeing the start of a fourth wave driven by the delta variant. Currently 42% of the French population is fully vaccinated, and the current government strategy focuses on increasing the fraction of the population vaccinated.

lead to severe symptoms, as discussed above, this is not the case for SARS-CoV-2, which has proved much more contagious. A mutation or recombination event by which Covid-19 becomes more deadly cannot be ruled out. Something of this sort could be described as the worst case scenario.

New strains

Recently new strains of the virus that spread more rapidly than the old strains have appeared in the United Kingdom, Brazil, and South Africa. The replication of viruses implies the copying of its genetic information encoded in the viral RNA. During this process, errors may occur by chance. Amino acids may be substituted, deleted, or inserted by mistake. Such deviations from the original in the copied RNA are known as mutations. Some mutations may be advantageous to the virus's survival, by improving the 'fitness' of the virus. Others will hardly affect the virus's survival, and yet others will be harmful to its survival and consequently will disappear. Of course, scientists are worried about the mutations that make the virus 'stronger' and more dangerous to the host–that is, us.

Since the beginning of the pandemic, scientists have already observed a few thousand mutations of the virus. Three mutations, however, have alarmed the medical community, as they have increased the virus's ability to infect the hosts. These are the mutations that help the virus spread more easily and have induced a surge in Covid-19 cases worldwide.

> ***The higher transmissibility of the new variant is undisputable. Early evidence suggests that the new variant does not cause more severe forms of the disease.***

One of these variants, originally called 501Y.V2, then B.1.351 and now Beta, was discovered in South Africa and described by Prof Tulio de Oliveira's team at KRISP in Durban. Doctors had noticed a surge in Covid-19 cases in the Eastern Cape and the genomic investigation of its cause revealed the new variant of the virus. Beta has now become the dominant version of the virus in South Africa. The higher transmissibility of the new variant is indisputable. Early evidence suggests that the new variant does not cause more severe forms of the disease. Scientists are also investigating if the available vaccines will offer protection from the new variant and if antibodies from previous infection with the original variant will suppress the new variant.

In the meantime, two more variants have been observed that share similar characteristics to Beta. The variant detected first in the UK is called Alpha (formerly 501Y.V1 or B.1.1.7) and has been detected in many other countries since then. Another variant, P.1 (501Y.V3), was observed first in Brazil. It is imperative to monitor the evolution of these variants globally and understand their characteristics.

The Delta variant (B.1.617.2), first discovered in India in late 2020 because of its enhanced ability to spread, has become dominant in many countries already and may become the dominant variant everywhere.

Modelling the spread of infectious diseases is difficult. Professional modellers often disagree with each other due to the use of different simplifying assumptions and interpretations of the data. Moreover, individual models often predict a range of outcomes rather than making a single precise prediction. The range of predictions is often more reliable than the predictions of any single model.

Above we described exponential growth, which provides good short-term predictions, but applied mathematicians have developed more sophisticated models called SIR models, in which individuals are pooled into susceptible, infected, and recovered classes. Mathematical models predict how individuals pass from one class to another. Although these models are helpful, many of the details that enter into the models rely on imperfect data, making it difficult to establish reliable long-term predictions. Nevertheless, such modelling helps us better understand the range of possibilities, and the impact of policy on outcomes.

Not being able to predict the future accurately is not unique to epidemiology.

Not being able to predict the future accurately is not unique to epidemiology. In economics and business, predicting inflation, exchange rates, prices and demand for particular goods, etc. is indispensable to making sound business decisions. Yet, precise predictions cannot be made, and the projections made are often way off.

If one catches Covid-19, what is the probability of death?
Which factors affect mortality?

These are important questions but not so easy to answer precisely. This is because the data available to us is not exactly the data that we would like to have. We would like to count all the cases of persons who are infected. However, the reporting of more serious cases is more complete, and cases of those who are infected but show no symptoms are almost never reported. This causes the total number of people infected to be under-reported. On the other hand, deaths lag new cases by several weeks. When the daily number of cases is growing, this causes the mortality (death rate) fraction to be under counted. Various attempts can be made to model the mortality (death rate) for those infected. Moreover, beyond these problems arising from the data, the mortality fraction per infection depends on a number of factors: age, sex, co-morbidities, quality of medical care, to name just a few examples. Mortality rates from different countries will differ, in part because of different mixes of these factors. It also appears that the chance of dying has been going down, likely because doctors are learning how to better treat Covid-19 patients.

Some drugs used for other diseases have been useful in reducing deaths in Covid-19 patients. These include Remdesevir and dexamethasone, where good outcomes for patients were found.

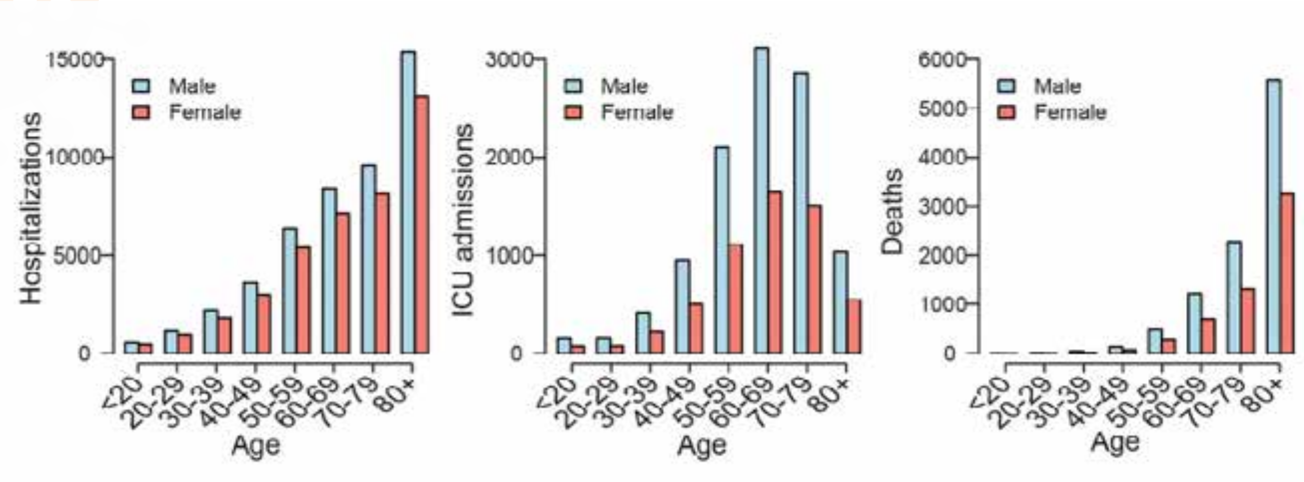


Figure 5.9: French Covid-19 data and age
The left, middle, and right panels show the distribution of hospitalisations, intensive care unit (ICU) admissions, and deaths in France, respectively, classified according to age and gender.

Reprinted with permission: H. Salje et al., "Estimating the Burden of SARS-CoV-2 in France," Science 369 (2020) 208 (10July 2020)

Age and co-morbidities are important factors in determining how sick someone who gets infected will become and whether they are likely to die. As a rule, older people are more likely to become very sick, and also to die. Moreover, people who suffer from other health problems (also known as 'co-morbidities') are more likely to become severely sick and to die. Such health problems include: diabetes, obesity, respiratory disease, reduced immune system function, to name a few examples. These conditions can increase the chance of a bad outcome by a factor of up to 3 or more. A study published in the prestigious scientific journal *Nature* using data collected in the United Kingdom, explored how different factors increased or decreased the chance of dying from Covid-19. This study found that men infected are about 1.6 times as likely to die as women infected, and found the following increase in the chance of dying as a function of age, using the 50-59 age group as a reference.

Age group	Increase in probability of death
18-39	0.05
40-49	0.28
50-59	1.0 (reference)
60-69	2.79
70-79	8.62
80+	38.29

In summary, young people and people without other medical conditions have less to worry about regarding their outcome if they get infected. This however does not mean that they should not be careful, because they can infect the older people around them and, in rare situations, very young people have mysteriously died.

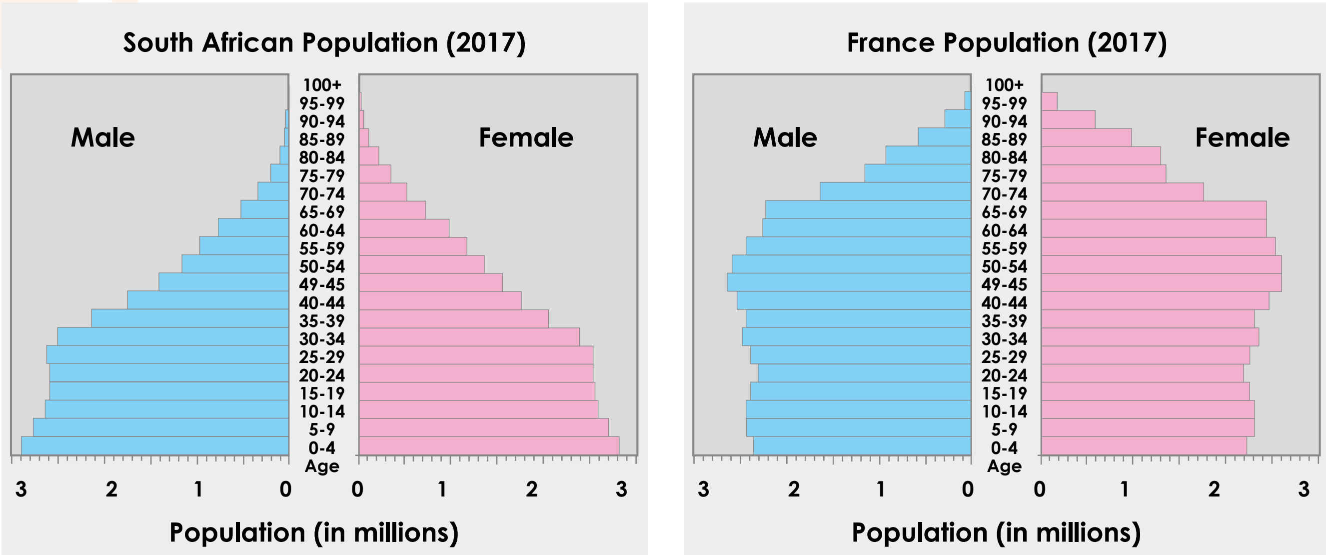


Figure 5.10: Age pyramid: South Africa vs France

Source: https://commons.wikimedia.org/wiki/File:South_Africapop.svg: https://commons.wikimedia.org/wiki/File:Francepop.svg

ISBN 978-1-928496-37-3
DOI http://dx.doi.org/10.17159/assaf.2021/0072
July 2021
Published by: Academy of Science of South Africa (ASSAf)
PO Box 72135, Lynnwood Ridge, Pretoria, South Africa, 0040
Tel: +27 12 349 6600 • Fax: +27 86 576 9520
E-mail: admin@assaf.org.za

The Academy of Science of South Africa (ASSAf) was inaugurated in May 1996. It was formed in response to the need for an Academy of Science consonant with the dawn of democracy in South Africa: activist in its mission of using science and scholarship for the benefit of society, with a mandate encompassing all scholarly disciplines that use an open-minded and evidence-based approach to build knowledge. ASSAf thus adopted in its name the term 'science' in the singular as reflecting a common way of enquiring rather than an aggregation of different disciplines. Its Members are elected on the basis of a combination of two principal criteria - academic excellence and significant contributions to society.

The Parliament of South Africa passed the Academy of Science of South Africa Act (No 67 of 2001), which came into force on 15 May 2002. This made ASSAf the only academy of science in South Africa officially recognised by government and representing the country in the international community of science academies and elsewhere.

https://www.assaf.org.za/