International Journal of System Dynamics Applications
Volume 10 • Issue 3 • July-September 2021

Research and Statistics:
Coronavirus Disease (COVID-19)

Nasir Mustafa, Istanbul Gelisim University, Turkey
https://orcid.org/0000-0002-5821-9297

ABSTRACT

A newly identified coronavirus, SARS-CoV-2, is spreading across the globe. Most of our work focuses on large problems that humanity has faced for a long time, such as child mortality, natural disasters, poverty, and almost 100 other problems. This article focuses on a new, emerging global problem: the ongoing outbreak of the coronavirus disease (COVID-19). The outbreak started in China in late 2019, and by March 2020 the disease has spread to countries around the world. The number of infections appearing each day has since plummeted in China, owing in large part to containment efforts, but the outbreak is now a global pandemic. Large outbreaks in South Korea, Iran, Italy, and elsewhere have propelled a spike in international cases across more than 180 countries.

KEYWORDS

Coronavirus Disease, COVID Outbreak, COVID-19, Research and Statistics COVID-19

THE NUMBER OF TOTAL CASES IS WHAT WE WANT TO KNOW, BUT THEIR NUMBER IS NOT KNOWN

To understand the scale of the COVID-19 outbreak, and respond appropriately, we would want to know how many people are infected by COVID-19. We want to know the total number of cases. However, the total number of COVID-19 cases is not known. It is unknown by us at Our World in Data, or any other research, governmental or reporting institution. There are several reasons why the total number is not known. Whilst for some the symptoms are very severe for a large share of the population the symptoms are mild. In such cases people may be unaware that they are infected with COVID-19, and therefore not get seen, and diagnosed, by a doctor. (Read et al., 2020; Koh 2020). The second reason that the confirmed cases are only a fraction of the total number is that many countries are struggling to test a large number of cases. Not every person that should be tested is able to. Since testing is crucial, we looked into this in more detail. Because of large problems with data availability on the scale of testing, we did a manual review of data on COVID-19 testing across national reports, and collated the most recent estimates that we could find as of 23 March 2020.

CONFIRMED CASES IS WHAT WE DO KNOW

What we do know is the number of confirmed cases. The World Health Organization (WHO) explains that a confirmed case is "a person with laboratory confirmation of COVID-19 infection, irrespective

DOI: 10.4018/IJSDA.20210701.oa1

of clinical signs and symptoms" (Maier & Brockmann 2020). The WHO also speaks of 'suspected cases' and 'probable cases', but the WHO Situation Reports do not provide figures on 'probable cases', and only report 'suspected cases' for Chinese provinces ('suspected cases' by country is not available) (Vaidya et al., 2016). The daily Situation Reports list the number of confirmed cases; we rely on these reported numbers for the regular updates of our own datasets presented below. As explained, the number of confirmed cases is lower than the number of total cases because not everyone is tested.

Growth of cases: How long did it take for the number of confirmed cases to double? In the section below we present the latest data on the number of confirmed cases by country, and how this has changed over time. But in an outbreak of an infectious disease it is important to not only focus on the number of cases, but also the growth rate at which the number of cases is increasing. The growth rate is an essential metric to understand and to monitor. This is because a fast growth rate can lead to very large numbers rapidly, even if the current numbers of cases and deaths are small when compared with other diseases. A helpful metric to measure the rate of change is to look at each country and ask: How long did it take for the number of confirmed cases to double? Let's take an example: if the number of confirmed cases as of today is 1000, and there were only 500 cases three days ago then we would say that it took three days for the number of confirmed cases to double (Danon et al., 2020). The doubling time of cases has changed and it will change in the future. It would be wrong to extrapolate current growth into the future. But it is important to keep focusing on the doubling time. As long as cases are doubling at a constant rate the growth is exponential. We humans tend to think in linear growth processes even when the growth is exponential, as psychological research has shown for decades. Below we give some intuition about exponential growth and provide the referenced psychological research on this.

UNDERSTANDING EXPONENTIAL GROWTH

The time it takes for the number of cases to double will change during the outbreak and it would be wrong to make projections based on the assumption that this stays constant. But it is important to remind ourselves of the nature of exponential growth. If during an outbreak the number of cases is in fact doubling and this doubling time stays constant, then the outbreak is spreading exponentially. Under exponential growth 500 cases grow to more than 1 million cases after 11 doubling times.[6] And after 10 more doubling times it would be 1 billion cases. This is in no way a prediction for the number of cases we should expect; it is simply a reminder that exponential growth leads to very large numbers very quickly, even when starting from a low base. And it is important to be reminded of the nature of exponential growth because most of us do not grasp exponential growth intuitively. Psychologists find that humans tend to think in linear growth processes (1, 2, 3, 4) even when this is not appropriately describing the reality in front of our eyes. This bias – to "linearize exponential functions when assessing them intuitively" – is referred to as 'exponential growth bias' (Stango & Zinman 2009). Psychological research shows that "neither special instructions about the nature of exponential growth nor daily experience with growth processes" improved the failure to grasp exponential growth processes (Wagenaar & Sangaria 1975).

THE GLOBAL AVERAGE HIDES MORE THAN IT REVEALS: WHY WE SHOW THIS DATA COUNTRY BY COUNTRY

It is important to not only look at the global number of cases and deaths. Some countries like China and Korea – have very substantial counter measures in place and new daily confirmed cases have declined. Many other countries do not have comparable measures in place and, as the table shows, numbers are rising fast. The global average does not allow us to understand this. The global average hides the differences between countries that are successfully reducing the number of confirmed cases and those that fail to achieve this – and this is what is most important in this early phase of the COVID-19 epidemic.

GROWTH: COUNTRY BY COUNTRY VIEW

As just explained, it is crucial to not just look at the number of cases, but also their growth over time. For this reason, the following table answers the following question for all countries: How long did it take for the number of confirmed cases to double?

CONFIRMED COVID-19 CASES BY COUNTRY

In our visualizations here you can explore the number of total confirmed cases and daily new confirmed cases for all countries with reported cases. These charts are interactive: the data is shown as the worldwide figures by default but can be explored by country.

Data: The data shown here is published by the World Health Organization (WHO). We from Our World in Data found several minor errors in the WHO data – we documented these errors, corrected them, reported them to the WHO.

EXPLANATION OF THE LARGE NUMBER OF CASES IN CHINA ON FEBRUARY 17TH DATA REPORTING

You will notice a dramatic increase in the number of confirmed cases on 17th February 2020. As the WHO notes in its Situation Report 27, this is the result of a change in reporting methodology to include clinically-diagnosed cases in addition to laboratory-confirmed cases as previously reported. This change in methodology only affected figures in the Hubei province in China, but due to the large number of cases in this region it had a significant impact on global figures too (Elfouly et al., 2017). The large apparent increase in confirmed cases in China, and globally, on 17th February therefore reflects this change in methodology rather than an actual sudden increase in cases on that particular day (Figures 1-2).

DEATHS FROM COVID-19

What We Want to Know Is the Total Number of Deaths and the Mortality Risk

To understand the mortality risk of the COVID-19 outbreak – the likelihood that someone who catches the disease will die from it – we would want to know the share of people who are infected with the disease that die from it. We want to know the final number of deaths that result for a given infected population. However, in an ongoing outbreak the final outcomes (death or recovery) for all cases is not yet known. The time from symptom onset to death ranges from 2 to 8 weeks for COVID-19 (Koh 2020). This means there are individuals who are in the early or mid-stages of infection who will die at a later date. We therefore cannot give a definitive figure for the mortality risk of the disease.

Confirmed Deaths to Date Is What We Do Know

What we do know is the total number of confirmed deaths to date. The World Health Organization (WHO) publishes updates on confirmed deaths to date in its 'Situation Reports'. This means we can track how the number of deaths is changing over time. It does not inform us of the probability that someone infected with the diseases dies from it – to know this we would need to know the final outcome of all cases of the disease. Some individuals who are currently infected with COVID-19 are likely to die at a later date.

Figure 1. Total confirmed deaths due to COVID-19

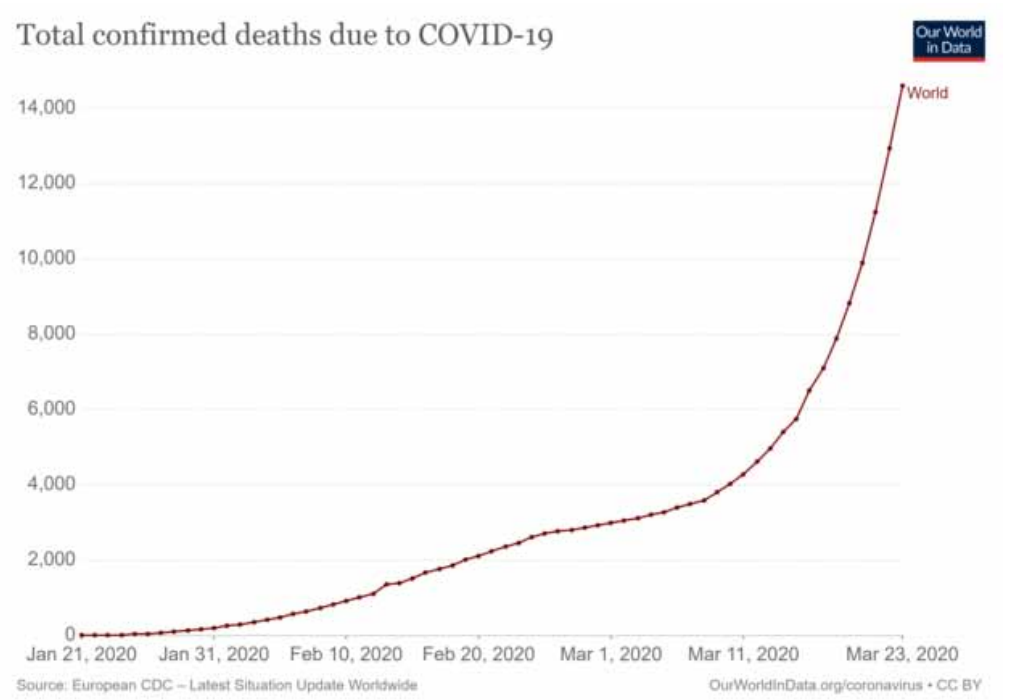
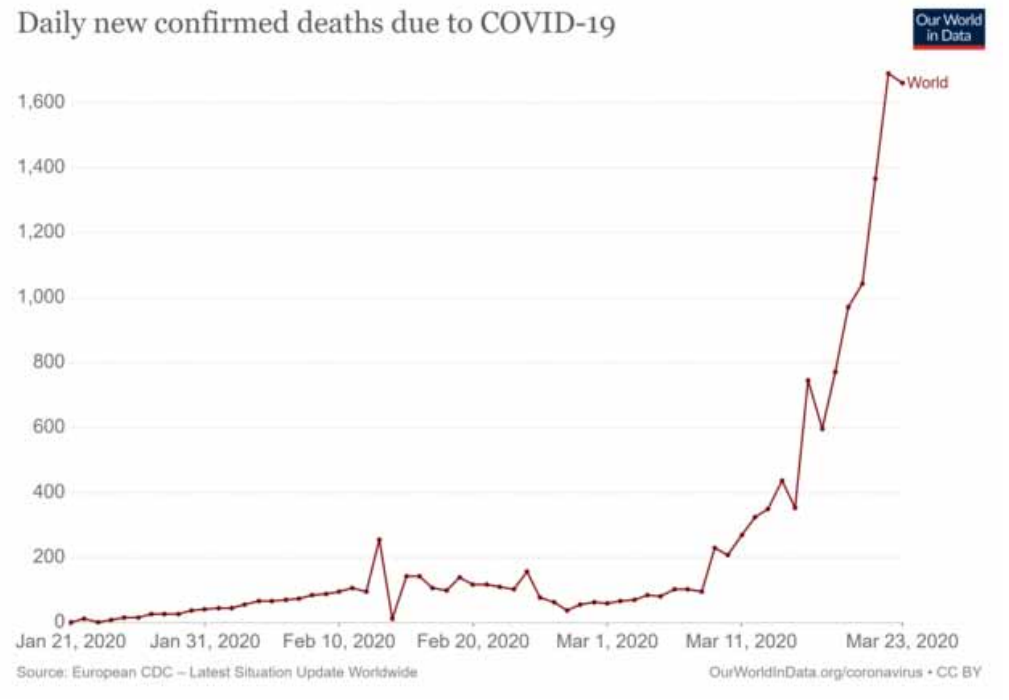


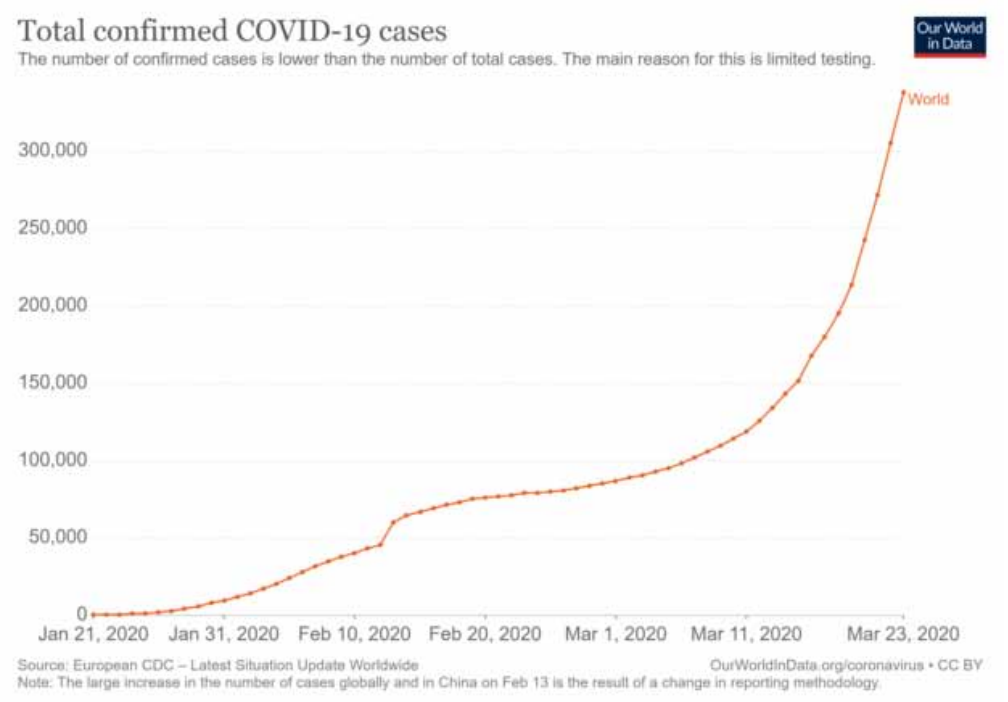
Figure 2. Daily new confirmed deaths of COVID-19



Confirmed COVID-19 Deaths by Country

In our visualizations, here you can explore the number of total deaths and daily new deaths for all countries with reported deaths. These charts are interactive: the data is shown as the worldwide figures by default. The data shown here is published by the World Health Organization (WHO). We from Our World in Data found several minor errors in the WHO data – we documented these errors, corrected them, reported them to the WHO, and are in close contact with colleagues at the WHO. The documentation of the data shown in Figure 3.

Figure 3. Total confirmed COVID-19 cases



Why Testing Is Important?

We want to know the total number of people infected with COVID-19. To know this, it is necessary to have widespread testing. When testing is too low we do not have a clear picture of what's going on. Testing is crucial as it allows the infected person to avoid infecting others and to quickly receive the care they need. And it is crucial for all of us to understand the prevalence of the disease, to understand how the disease evolves, and to allow us to take evidence-based decisions for counter measures that slow down the spread of the disease (Jin et al., 2020). This last point is very important: Testing is crucial to lower the rate of infection. When infected people do not know that they are infected, they might not stay at home – thereby running the risk of infecting others. Unfortunately, there are two important reasons why testing is still low in many countries affected by COVID-19. First, some people who are infected with COVID-19 have mild symptoms and therefore do not go to get checked (we explain what is known about the symptoms here and also discuss how many suffer from severe and critical symptoms). The second reason is that in many places the capacity for COVID-19 testing is low.

How Are COVID-19 Tests Done?

The most common diagnostic tests for COVID-19 are the so-called "PCR tests" that use swabbed samples from a patient's nose and throat. The first PCR tests were developed within two weeks of the disease being identified, and are currently part of the protocol recommended by the WHO (Pierce et al., 2019).

What Information About Test Coverage Do We Currently Have?

Ideally, we would want to know how many people in the world are being tested for COVID-19.

Every day, and how the available tests are being allocated. Unfortunately, there is no centralized database by the WHO on COVID-19 testing and many countries in the world currently do not publish official reports on tests performed. Several countries however do publish aggregate estimates on the total number of tests performed. These reports are published across individual websites, statistical reports and press releases – often in multiple languages and updated with different periodicity. Because a global overview was not available, we at Our World in Data brought together a large number of data sources from individual national reports (Cepaluni 2014).

Current COVID-19 Test Coverage Estimates

The two charts here show the most recent official estimates of tests we have been able to find as of 23 March 2020. Note that the estimates refer to different dates for each country, as indicated in the brackets. The first chart shows the total number of tests up to the specified date. The second chart shows the number of tests relative to the size of the population: it is the number of total COVID-19 tests per million people. The available data shows that South Korea has done many more tests than any other country. This suggests that the number of confirmed cases in Korea is closer to the total number of cases than in other countries. It is therefore particularly encouraging to see that the number of daily confirmed cases in South Korea has decreased – The US, on the other hand, has experienced big problems rolling out their testing strategy and according to the US Centers for Disease Control and Prevention, only a total number of 13,624 samples had been tested by 12 March, 2020. The total number of tests conducted in South Korea up to the same date was nearly 18-times larger. The low test coverage of the US is even starker if we look relative to the large population of the country. We see many smaller countries have been able to conduct more tests per million people.

The fact that South Korea has managed to expand testing so quickly shows that it is possible. Because testing is crucial it is important that in the coming days other countries follow (Figure 4).

The low test coverage of the US is even starker if we look relative to the large population of the country. We see many smaller countries have been able to conduct more tests per million people. The fact that South Korea has managed to expand testing so quickly shows that it is possible. Because testing is crucial it is important that in the coming days other countries follow.

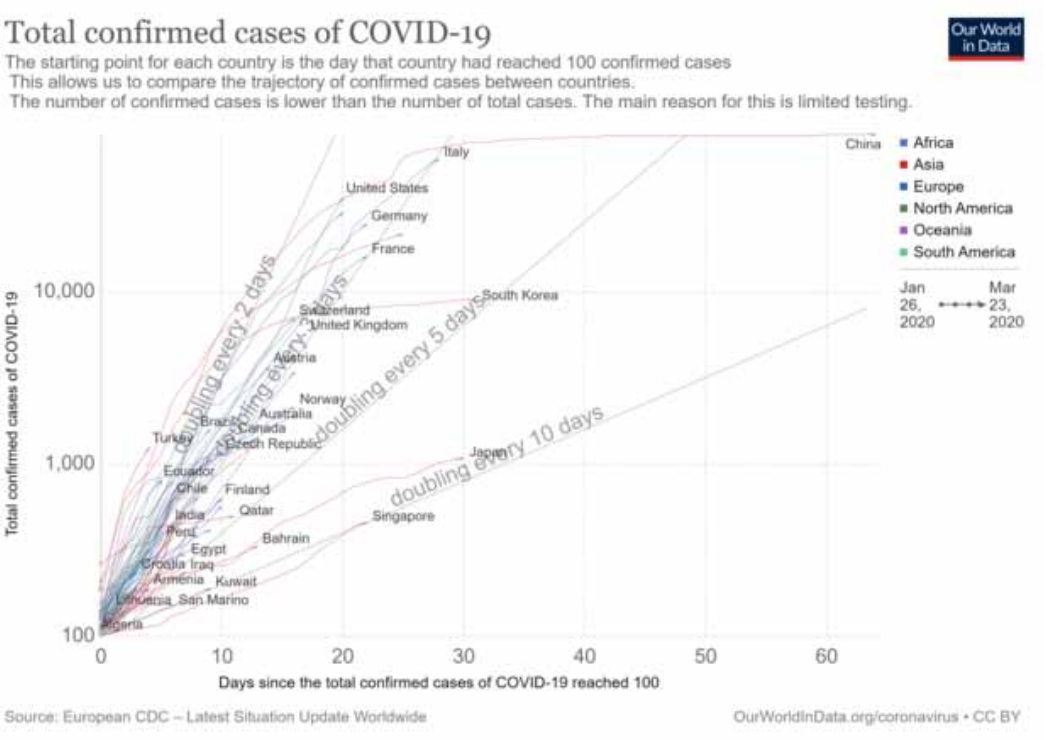
THE NAME OF THE DISEASE AND VIRUS?

The names for the virus and the disease it causes have been announced by the World Health Organization and the International Committee on Taxonomy of Viruses (Topcuoglu, 2020).

The Disease Is Called Coronavirus Disease - It Is Abbreviated as COVID-19

The virus is called severe acute respiratory syndrome coronavirus 2 and it is abbreviated as SARS-CoV-2. In the same statement, the WHO also explains that they themselves also refer to the virus as "the virus responsible for COVID-19" or "the COVID-19 virus" when communicating with the public. We follow the same conventions here.

Figure 4. Total confirmed cases of COVID-19 with starting points and countries



How Did the Outbreak Start?

On 29 December 2019, Chinese authorities identified a cluster of similar cases of pneumonia in the city of Wuhan in China. Wuhan is a city with 11 million inhabitants and is the capital of the Hubei Province. These cases were soon determined to be caused by a novel coronavirus that was later named SARS-CoV-2. (Niederberger 2020). Coronaviruses are a group of viruses that are common in humans and are responsible for up to 30% of common colds (Mesel-Lemoine et al., 2012). Corona is Latin for "crown" – this group of viruses is given its name due to the fact that its surface looks like a crown under an electron microscope. Two outbreaks of new diseases in recent history were also caused by coronaviruses – SARS in 2003 that resulted in around 1,000 deaths (Smith 2006) and MERS in 2012 that resulted in 862 deaths (Erasmus 2020). The first cases of COVID-19 outside of China were identified on January 13 in Thailand and on January 16 in Japan. On January 23rd, the city of Wuhan and other cities in the region were placed on lockdown by the Chinese Government. Since then COVID-19 has spread to many more countries – cases have been reported in all world regions. You can see the latest available data in the dashboards of cases and deaths which are kept up-to-date by Johns Hopkins University.

Related Work by Our World in Data

Pneumonia: Severe cases of COVID-19 can progress to pneumonia (Wang et al., 2020). Our entry on pneumonia provides an overview of the data and research on this disease that kills 2.6 million annually.

The Spanish Flu (1918-20): The global impact of the largest influenza pandemic in history – We look at the global death toll and mortality rate of the Spanish flu and compare it with three other large influenza pandemics in the last century (Wang et al., 2020).

Strategies to Respond to COVID-19

Early containment allows the healthcare system to provide care for all who need it.

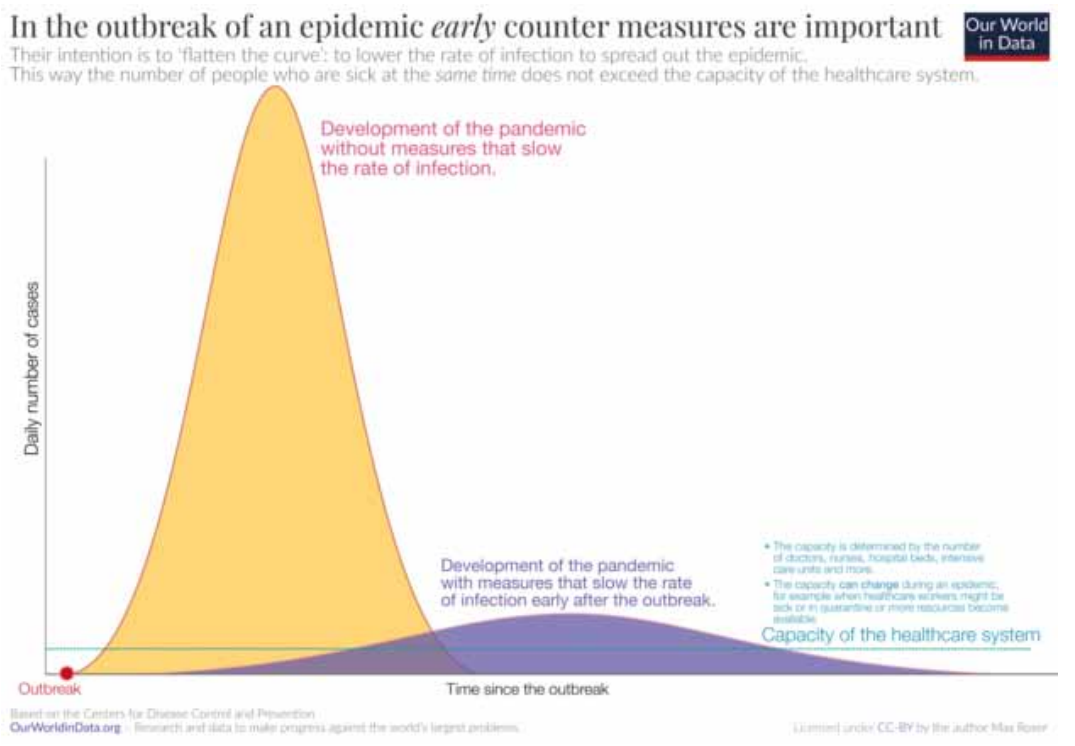
The total mortality of an epidemic can be high even if the symptoms for the vast majority are mild. While it might not seem intuitive, it is possible for the following two things to be true at the same time:

- For the majority of people, symptoms are mild and in some cases similar to the common flu. An epidemic of the same disease can cause a very high number of deaths;
- As we discuss here, the symptoms of COVID-19 can be very severe in many cases. Many of these patients require treatment in intensive care units (ICUs). The WHO reports that "about a quarter of severe and critical cases require mechanical ventilation" (Koh 2020).

'Flattening the Curve'

This is why early counter measures are important in an epidemic. Their intention is to lower the rate of infection so that the epidemic is spread out over time such that the peak demand on the healthcare system is lower. Containment measures are intended to avoid an outbreak trajectory in which a large number of people get sick at the same time. This is what the visualization shows. This is the reason that limiting the magnitude of peak incidence of an outbreak is important. Health systems can care for more patients across an outbreak when the number of cases is spread out over a long period rather than condensed in a very short period. What such counter measures to the pandemic attempt to avoid is that the number of patients at one point in time is so large that health systems fail to provide the required care for some patients (Galli 2018) (Figure 5).

Figure 5. Early counter measures

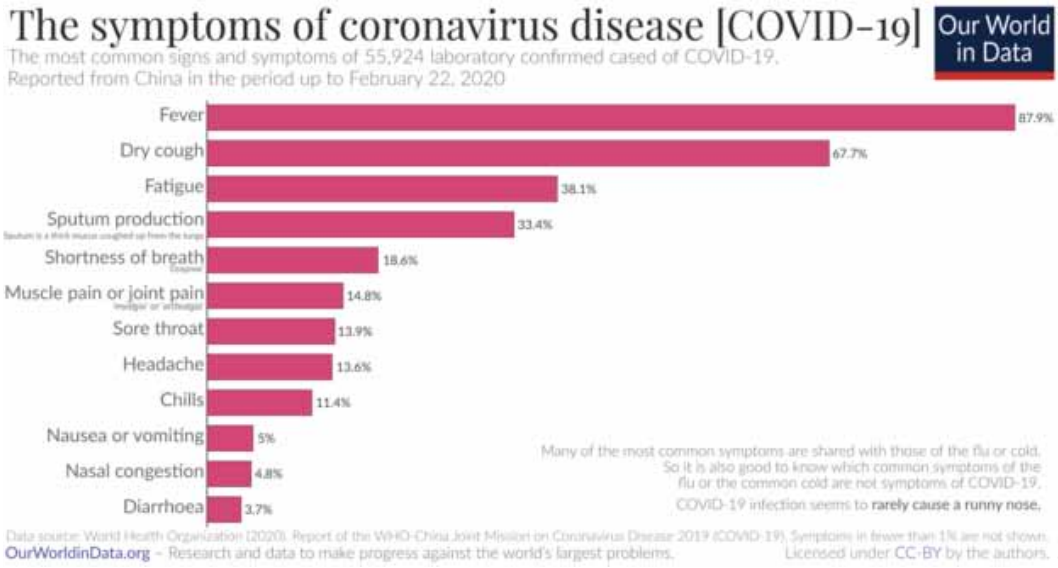


COVID-19: WHAT ARE THE SYMPTOMS? HOW DOES THE DISEASE PROGRESS?

Why We Need to Study Data to Know the Symptoms of COVID-19

COVID-19 leads to a number of symptoms, but from what is known currently some symptoms are much more common than others and for this reason we need to look at the available data. The danger of relying solely on text and not on numbers is that crucial nuance can get lost. This is the case for the media coverage of the symptoms of COVID-19. Coverage of the disease, even in reputable sources, includes long lists of symptoms without conveying to the reader how common or rare the listed symptoms are here is a poor example from the BBC. It is crucial to know how common the various symptoms of COVID-19 are, as it allows a better assessment of whether one suffers from the disease or not. This is lost in reporting that relies on text– especially if the list of potential symptoms is long, and overlaps strongly with many other types of illness. In a simple list of COVID-19 symptoms the reader might see that muscle pain is listed as a symptom and then wrongly conclude that they do not have the disease if they are not suffering from muscle pain. Knowing the frequency of symptoms means knowing that the vast majority of known cases (85% in the sample below) did not suffer from this symptom. The symptoms of COVID-19. The WHO described the symptoms of 55,924 laboratory confirmed cases of COVID-19 in China in the period up to February 20 (Kizito & Semwanga 2020). The visualization here shows this data. It is most crucial to know the common symptoms: fever and a dry cough. As the visualization shows, close to 90% of cases had a fever and two-thirds had a dry cough. The third most common symptom was fatigue. Almost 40% of cases suffered from it. 'Sputum production' was experienced by every third person. Sputum is not saliva. It is a thick mucus which is coughed up from the lungs. Of the 55,924 cases, less than 1-in-5 (18.6%) experienced shortness of breath ('dyspnoea'). An earlier study, reported that a much higher share (55%) of cases suffered from dyspnoea, but this was based on a much smaller number of cases (835 patients) (Wang et al., 2020). Many of the most common symptoms are shared with those of the common flu or cold. So it is also good to know which common symptoms of the common flu or the common cold are not symptoms of COVID-19. COVID-19 infection seems to rarely cause a runny nose (Figure 6).

Figure 6. Symptoms



How Long Is the Incubation Period of COVID-19?

The WHO writes "people with COVID-19 generally develop signs and symptoms, including mild respiratory symptoms and fever, on an average of 5-6 days after infection" (Koh 2020). While the mean incubation period is 5 to 6 days, the WHO adds that the incubation period can vary in a wide range of between 1 to 14 days (Salzberger et al., 2020). This is based on the 55,924 confirmed cases in China.

There are reports of cases with longer incubation periods in the media.

How Long Does COVID-19 Last?

On average, the disease lasts two weeks. The WHO reports that "the median time from onset to clinical recovery for mild cases is approximately 2 weeks" (Dawson 2020). Again, this is based on the 55,924 confirmed cases in China. For severe and critical cases, it is 3 to 6 weeks according to the same study. And for those who eventually died, the time from symptom onset to death ranged from 2 to 8 weeks. This is important when interpreting the case fatality rate (see below). Measures of the CFR of an ongoing outbreak do (obviously) not include deaths of patients who will eventually die, but have not died yet at the time of measurement. This means that the current CFR would be lower than the eventual CFR.

How Does COVID-19 Progress?

The symptoms of the disease develop and change over time. It seems to be common that symptoms start with a fever, followed by a dry cough (Caines 1936). After several days, some patients experience shortness of breath. Symptoms can increase in severity as emphasized in the following section. In severe and critical cases, it can lead to severe pneumonia, respiratory failure, septic shock, and multiple organ dysfunction or failure. As we discuss in detail below, for some cases COVID-19 leads to death.

The Severity of the Symptoms of COVID-19

This visualization shows the severity of symptoms suffered by 44,415 Chinese patients confirmed to have coronavirus in the early period up to February 11 (Gao et al., 2020). It is likely that many more cases were so mild that they were not identified as COVID-19. Estimates published by Read et al. (2020) suggest that only around 5% of cases in China have been diagnosed and recorded (Read et al., 2020). Symptoms were categorized as mild, severe, or critical and the research article describes these as follows:

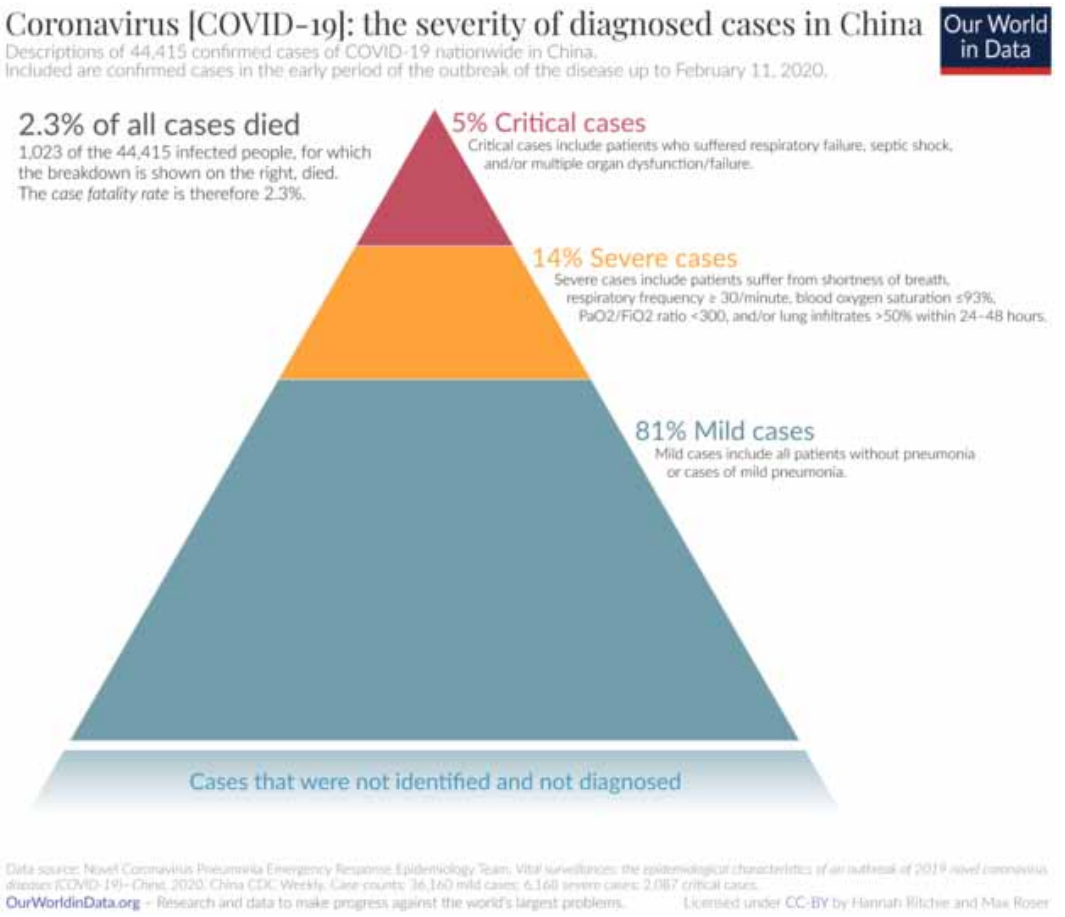
- **Critical Cases:** Critical cases include patients who suffered respiratory failure, septic shock, and/or multiple organ dysfunction or failure;
- **Severe Cases:** This includes patients who suffered from shortness of breath, respiratory frequency $\geq$ 30/minute, blood oxygen saturation $\leq$93%, PaO2/FiO2 ratio <300 (The Permanente Journal 2020), and/or lung infiltrates >50% within 24–48 hours;
- **Mild Cases:** The majority (81%) of these coronavirus disease cases were mild cases. Mild cases include all patients without pneumonia or cases of mild pneumonia (Figure 7).

What Do We Know About the Risk of Dying From COVID-19?

What we want to know is the risk of dying once you are infected with COVID-19.

As of early March 2020, this is not known to researchers, and we therefore have to rely on a number of metrics that allow us to get a perspective on the risk of mortality. However, to make sense of these metrics we need to understand their definitions and the challenges in measuring them. This is the focus of the section below. We provide an overview of what we know and what these measures mean, to allow an interpretation of the current knowledge.

Figure 7. The security of diagnosed cases in China



THE DEFINITION OF THE CASE FATALITY RATE (CFR)

Most current discussions of the mortality risk of COVID-19 refer to the case fatality rate (CFR). This is the metric we will focus on, but it is crucial to understand the caveats to this data, and how it differs from alternative measures (Baker et al., 2015). The case fatality rate is the share who died from the disease among individuals diagnosed with the disease. The CFR is calculated by dividing the total number of deaths from a disease by the number of confirmed cases. It is expressed as a percentage and used as a measure of disease severity. In the following section, we look at the challenges of estimating the CFR. One of the key challenges is that the number of confirmed cases is often smaller than the number of total cases. The trouble is that often many cases of a disease are never diagnosed. This could be because cases with mild symptoms are often not tested or because not everyone who is sick goes to a hospital where such cases could be diagnosed, or because testing facilities are limited. The case fatality rate is sometimes called case fatality risk or case fatality ratio. The case fatality rate should not be confused with the crude mortality rate from the disease. The crude mortality rate measures the probability that any individual in the population will die from the disease – not just those who are confirmed cases.

It is a very different measure. It's calculated by dividing the number of deaths from the disease by the total population. This crude mortality rate is sometimes also referred to as the crude death

rate. This is important to differentiate, because unfortunately people sometimes confuse case fatality rates with crude death rates. A common example is the Spanish flu pandemic in 1918. The often cited estimate by Johnson and Mueller (2002) is that 50 million people died globally from this pandemic and this implies that 2.7% of the world population at the time died. This means the crude mortality rate was 2.7%. But 2.7% is often misreported as the case fatality rate (Livingston & Bucher 2020), If it was in fact the case that the crude mortality rate was 2.7% then the case fatality rate was much higher, since not everyone in the world was infected with the Spanish flu (Pasqualoni 2020).

Measuring and Interpreting the Case Fatality Rate

It is important to understand the measurement challenges to understand what the case fatality rate can and cannot tell us about a disease outbreak.

There is no single case fatality rate for a disease – it is context-specific, changing with time and location (Youbin & Bao Hong 2020).

Unfortunately, it is common to report the CFR as a single value. But the CFR is not a biological constant. The CFR is not a value which is tied to the given disease, but is instead reflective of the severity of the disease in a particular context, at a particular time, in a particular population. The probability that someone dies from a disease is not only dependent on the disease itself, but also the social and individual response to it: the level and timing of treatment they receive, and the ability of the given individual to recover from it. This means that the CFR can decrease or increase over time, and that it can vary by location and by the characteristics of the infected population (age, sex, pre-existing conditions). The CFR of COVID-19 differs by location and has changed during the early period of the outbreak. In the chart, here we see that the case fatality rate of COVID-19 is not constant. This chart was published in the Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), in February 2020 (Nesteruk 2020). The plotted values of the COVID-19 CFR here refer to several locations in China during the early stages of the outbreak, from the beginning of January 2020 to 20th February 2020 (Chen et al., n.d.). We see that in the earliest stages of the outbreak the CFR was much higher: 17.3% across China as a whole (in yellow) and greater than 20% in the centre of the outbreak, in Wuhan (in blue). In the weeks that followed, the CFR declined. The WHO reports that "the standard of care has evolved over the course of the outbreak". The CFR fell to 0.7% for patients with the onset of symptoms after February 1st. We also see that the CFR was different in different locations (Nesteruk 2020). By 1st February the CFR in Wuhan was still 5.8% while it was 0.7% across the rest of China. This makes clear that what we said about the CFR more generally is true for the CFR of COVID-19 specifically. The CFR is not only reflective of the disease itself, but specific to where and when people are diagnosed with the disease. It's therefore incorrect to report it as a single point value for the disease, and instead requires us to also report the time and place (Wilder-Smith & Freedman 2006) (Figure 8).

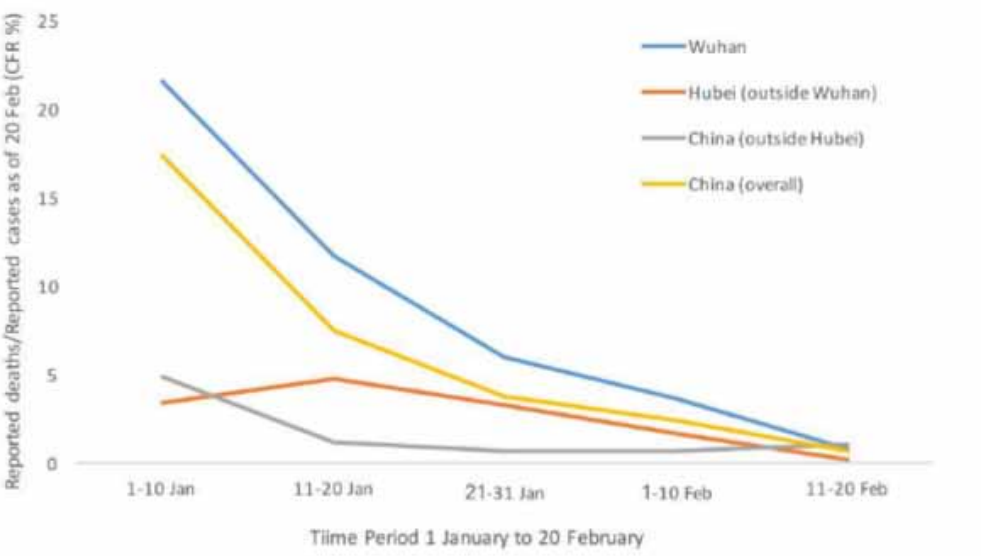
Global Case Fatality Rate of COVID-19

Based on the discussion of the definition of the case fatality rate (CFR), we should stress again that there is no single figure of CFR for any particular disease. The CFR varies by location, and is typically changing over time. However, with a good understanding of the measure and its limitations, CFR is helpful for understanding what we currently know about the severity of the disease and for responding accordingly (Xu et al., 2020).

In the period up to and including 9th March 2020, the global Case Fatality Rate for COVID19 are as follows:

Case fatality rate globally = 3.48%
[based on 109,578 confirmed cases and 3809 deaths]

Figure 8. Case fatality ratio



Case fatality ratio (reported deaths among total cases) for COVID-19 in China over time and by location, as of 20 February 2020

Case fatality rate in China: 3.86%
[based on 80,904 confirmed cases and 3123 deaths]

Case fatality rate for the rest of the world: 2.39%
[based on 28,674 confirmed and 686 deaths]

As explained above, this number has changed and it will continue to change. It's currently higher than the estimates of a CFR of around 2% that were published until early February. As we've discussed above, this does not necessarily represent a worsening of the situation: as we saw during the SARS outbreak, the CFR can rise during an outbreak because the outcome of more cases becomes known. As we also explained above, it would be wrong to assume that this CFR would be true everywhere, because it is a global average of confirmed deaths and cases. The early CFR in Wuhan was very high as we see here; the large number of deaths there in the early period impacts the average. Other studies for the Zhejiang province suggest that the CFR in China outside of Wuhan was likely lower (Galli 2020).

Case Fatality Rate of COVID-19 By Age

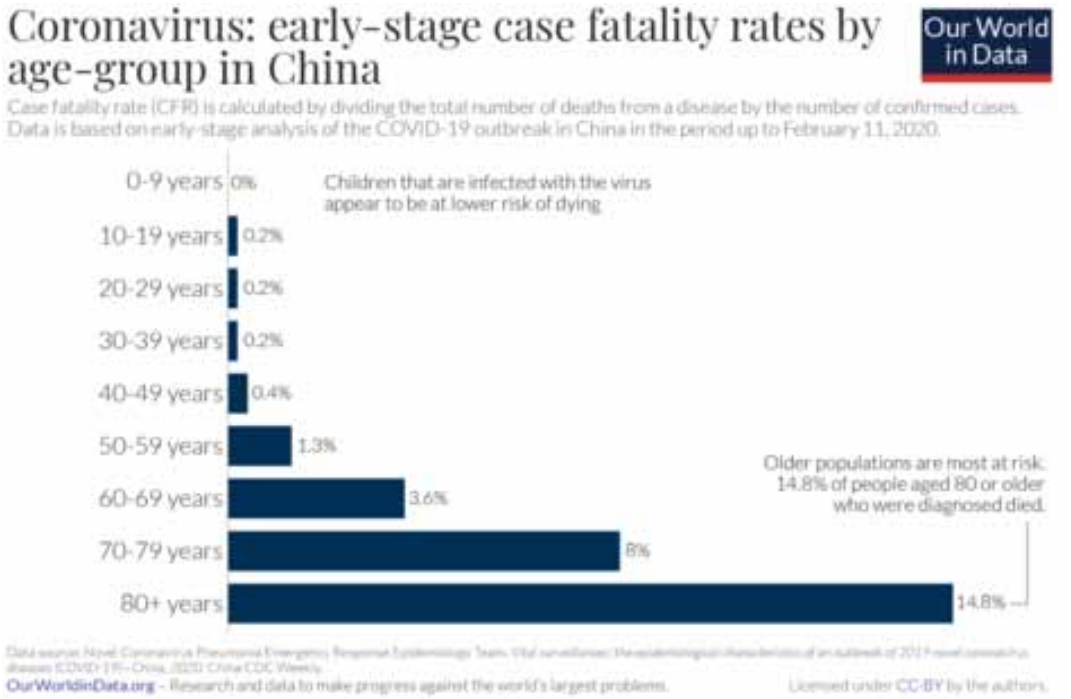
The total population-level estimate of the case fatality rate (CFR) above is useful for understanding the average severity of an outbreak, but does not tell us who within a population is most at risk. But this understanding is crucial in an outbreak. Understanding the relative risk to different sections of a population allows us to focus on the most vulnerable, and improve the allocation of health resources to those who need them most.

The Chinese Center for Disease Control and Prevention has published system analysis recorded cases and deaths in China for the period until February 11th, 2020 which provides a breakdown of all known cases, deaths and the CFR by specific demographics (age, sex, preexisting condition etc.) (Jin et al., 2020). A breakdown of the CFR by age group is shown in the visualization. It shows very

large differences of the CFR by age. For many infectious diseases young children are most at risk. We see this for malaria: the majority of deaths (57% globally) are in children under five years of age. The same was true for the largest pandemic in recorded history: During the 'Spanish flu' in 1918 it was primarily children and young adults who died from the pandemic.

For the COVID-19 cases in China the opposite seems to be true, at least based on the information available at the time of writing. The elderly are at the greatest risk of dying if infest (Figure 9).

Figure 9. Early-stage case fatality

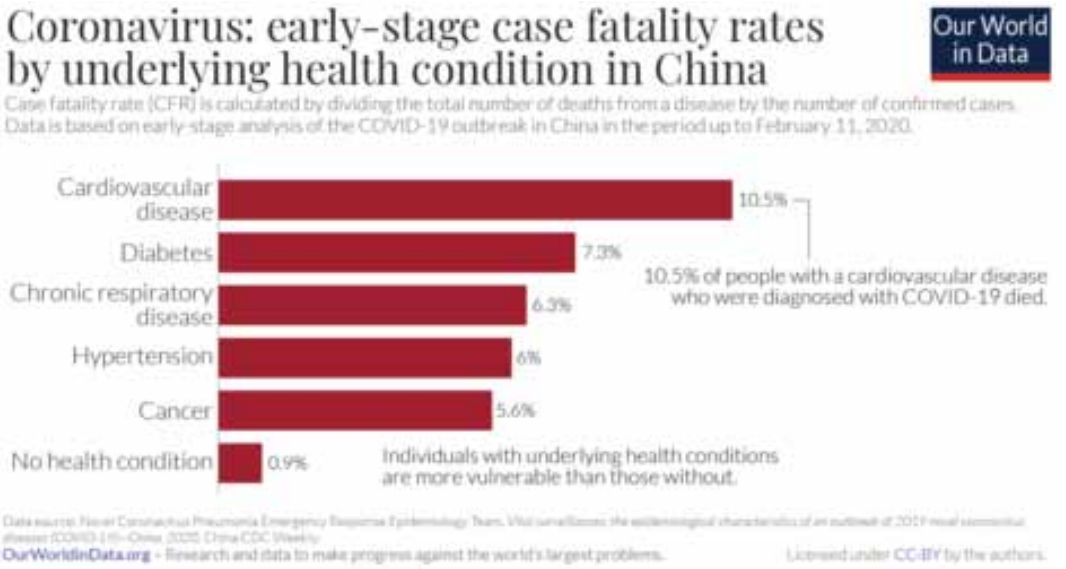


Case Fatality Rate of COVID-19 by Preexisting Health Conditions

Early data from China suggests that those with underlying health conditions are at a higher risk.

The visualization here shows the case fatality rate for populations within China based on their health status or underlying health condition. This is based on the same data from the Center for Disease Control and Prevention's initial breakdown of cases, deaths and CFR among specific demographics in the population (Munster et al., 2020). This analysis was based on recorded deaths and cases in China in the period up to February 11th 2020. The researchers found that the CFR for those with an underlying health condition is much higher than for those without. More than 10% of those diagnosed with COVID-19 who already had a cardiovascular disease, died as a result of the virus. Diabetes, chronic respiratory diseases, hypertension, and cancer were all risk factors as well, as we see in the chart. The CFR was 0.9% for those without a preexisting health condition. Above we saw that the elderly are most at risk of dying from COVID-19. This might be partly explained by the fact that they are also most likely to have underlying health conditions such as cardiovascular disease, respiratory disease and diabetes; these health conditions make it more difficult to recover from the COVID-19 infection (Dutta 2017) (Figure 10).

Figure 10. Early-stage fatality rates

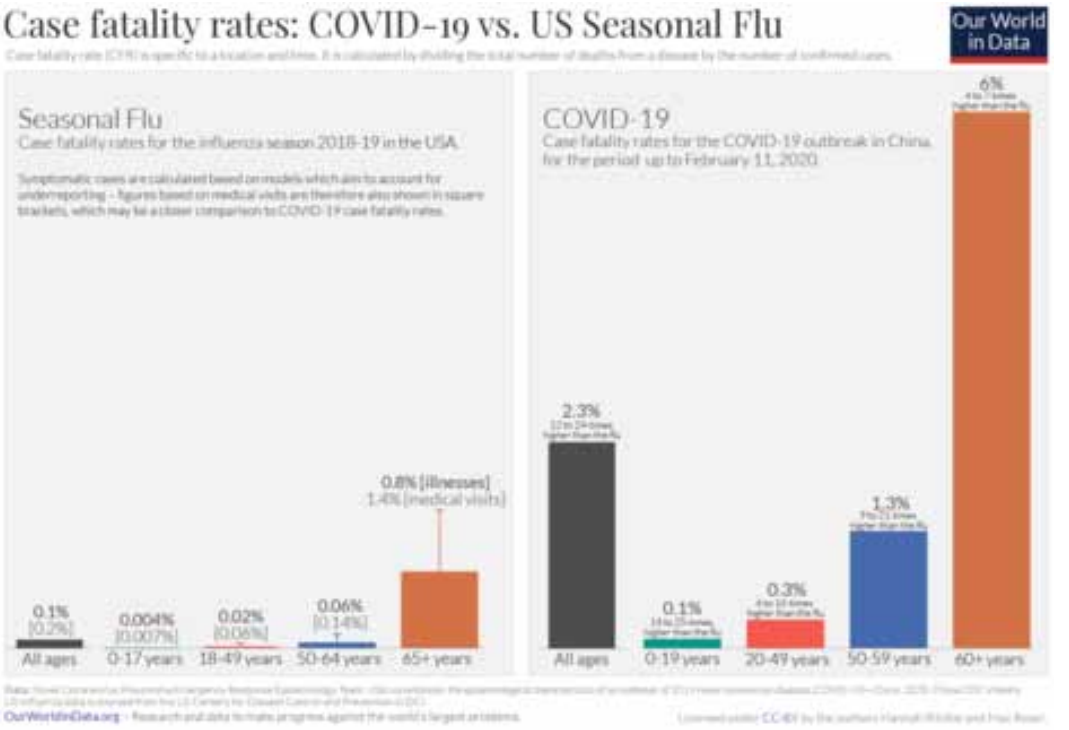


How Do Case Fatality Rates From COVID-19 Compare to Those of the Seasonal Flu?

This question is answered in the visualization here. We compare the CFR during the outbreak of COVID-19 in China with the CFR of the US seasonal flu in 2018-19. The case fatality rate of the seasonal flu in the US is around 0.1% to 0.2%, while the case fatality rate for COVID-19, measured in the cited study, was 2.3%. The US data is sourced from the US CDC. Here we present an upper and lower estimate for the 2018-19 flu season. These two figures reflect whether we look at the percentage of deaths out of the number of symptomatic illnesses (giving us 0.1%), or the number of medical visits (giving us 0.2%). In the traditional calculation of CFR, we would tend to focus on the number of symptomatic illnesses. This is analogous to the number of confirmed cases, on which the COVID-19 figures are based. However, the US CDC derives these figures based on disease outbreak modelling which attempts to account for underreporting – you can read more about how it derives its annual flu figures. This means that some of the biases which tend to underestimate the actual number of cases have been corrected for. This is not the case for the COVID-19 figures, so it may be an unfair comparison. Looking at estimates based on the number of medical visits may discount from the US seasonal flu data many of the kind of mild cases that may have been missed in the COVID-19 confirmed cases. However, this is likely to skew the comparison slightly in the other direction: we know that not all of the confirmed cases included in COVID-19 figures were of a severity such that they would have received a medical visit in the absence of the heightened surveillance of the outbreak. So, here we present both figures of the US seasonal flu figures: the CFR based on symptomatic illnesses, and those based on medical visits (shown in square brackets). It's likely that the fairest comparison to COVID-19 lies somewhere between these two values. You can find the data for the reported cases, medical visits and deaths from the US Centers for Disease Control and Prevention (CDC) here. The CDC reports 35,520,883 symptomatic cases of influenza in the US and 34,157 deaths from the flu. To calculate the CFR based on symptomatic illnesses, we divide the number of deaths by the number of confirmed cases and find a case fatality rate of 0.1% (Dyer 2020). The CFRs for COVID-19 are again based on the numbers reported by the Chinese Center for Disease Control and Prevention (Hon & Granville 1884). As before, the Chinese data refers to recorded deaths and confirmed cases in China as of February 11th, 2020. As calculated above, the global CFR for COVID-19 continues to change over time,

and the global average CFR based on the WHO data is 3.4% (as of 9th March 2020). While the CFR for COVID-19 is much higher than the CFR of the seasonal flu the two diseases are similar in the profile of the fatality rate by age: elderly populations have higher case fatality rates. However, the CFR of COVID-19 is much higher for all age groups, including young people. On top of each bar we have indicated how much higher the CFR for COVID-19 is for each age group (Figure 11).

Figure 11. COVID-19 vs. Seasonal flu



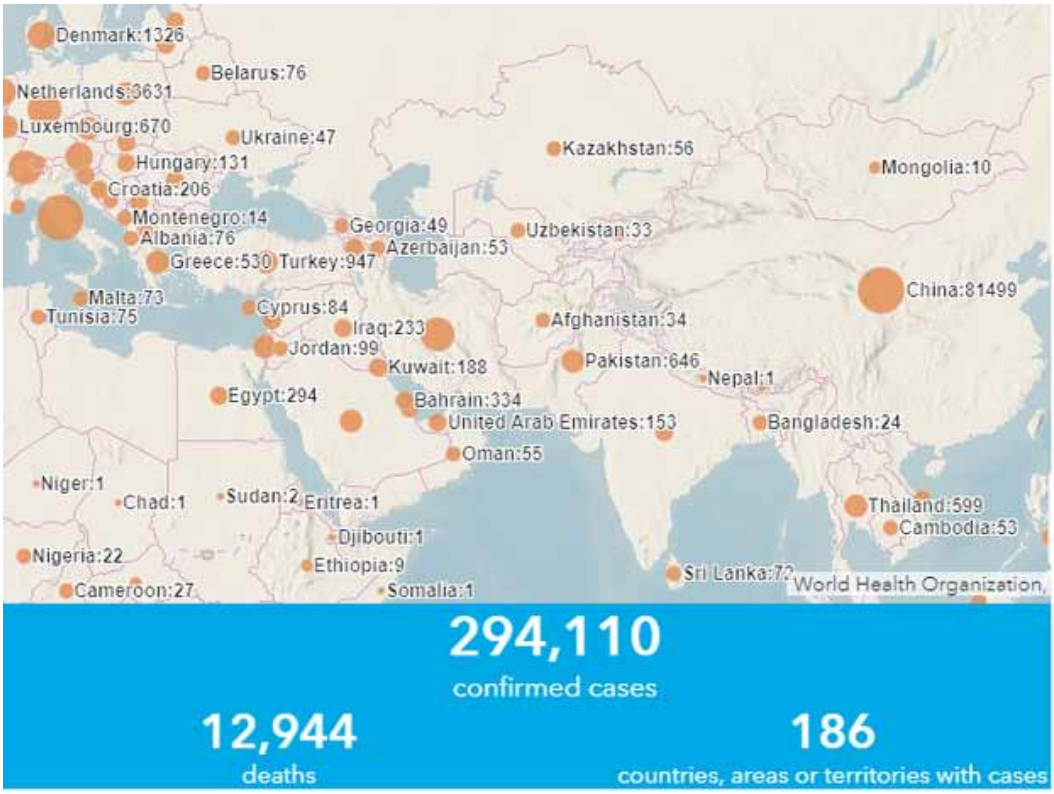
WHO Data on COVID-19

The World Health Organization (WHO) publishes a dashboard similar to that of Johns Hopkins above. The WHO dashboard on global cases and deaths is embedded here. In this dashboard, it is possible to see up-to-date country specific data by selecting the country in the top right. In addition to this dashboard, the WHO publishes daily Situation Reports which can be found here. It is the daily Situation reports that we rely on in our own published datasets on case and death numbers. Unlike the daily Situation Reports, the WHO dashboard is updated three times per day: any inconsistencies between the WHO dashboard and the data we present will be explained by this fact. As we explained above, the Our World in Data team found several minor errors in the WHO data. We documented these errors, corrected them, reported them to the WHO, and are in close contact with colleagues at the WHO (Dong et al., 2020) (Figure 12).

nCoV-2019 Data Working Group Data

The nCoV-2019 Data Working Group, which includes colleagues from the University of Oxford, publishes epidemiological data from the outbreak via this global dashboard. From

Figure 12. Confirmed cases



this dashboard, it is possible to obtain the underlying data which includes demographic and epidemiological descriptions of a long list of individual cases. Their data on the list of cases includes individual travel history and key dates for each patient – date of onset of symptoms, date of hospitalization and date of laboratory confirmation of whether the person was infected with the COVID-19 virus or not. This data is intended to be helpful in the estimation of key statistics for the disease: Incubation period, basic reproduction number (R0), age-stratified risk, risk of importation. In previous disease outbreaks such global individual data was not openly available.

REFERENCES

Anderson, R. M., Fraser, C., Ghani, A. C., Donnelly, C. A., Riley, S., Ferguson, N. M., Leung, G. M., Lam, T. H., & Hedley, A. J. (2005). Epidemiology, transmission dynamics, and control of SARS: The 2002–2003 epidemic. *SARS*, 61-80. 10.1093/acprof:oso/9780198568193.003.0010

Association between peripheral blood oxygen saturation (SpO2)/Fraction of inspired oxygen (FiO2) ratio time at risk and hospital mortality in mechanically ventilated patients. (2020). *The Permanente Journal*. Advance online publication. doi:10.7812/tpp/19.113

Baker, M., Yokoe, D. S., Stelling, J., Kaganov, R. E., Letourneau, A. R., O'Brien, T., Kulldorff, M., Babalola, D., Barrett, C., Drees, M., Platt, R., & Huang, S. S. (2015). Automated outbreak detection of hospital-associated infections. *Open Forum Infectious Diseases, 2*(suppl_1). 10.1093/ofid/ofv131.60

Caines, D. E. (2007). *(born 27 Feb. 1936), independent researcher*. Who's Who., doi:10.1093/ww/9780199540884.013.9853

Cepaluni, G. (2014). *'United we stand, divided we fall.' which countries join coalitions more often in GATT/WTO negotiations?* SSRN Electronic Journal. doi:10.2139/ssrn.2599867

Chen, Wei, Geng, & Jing. (n.d.). *Chest CT of the 2019 novel coronavirus (COVID-19): A case report.* 10.1148/cases.20201486

Danon, L., Brooks-Pollock, E., Bailey, M., & Keeling, M. J. (2020). *A spatial model of COVID-19 transmission in England and Wales: Early spread and peak timing*. 10.1101/2020.02.12.20022566

Dawson, S. (2020). *Novel coronavirus disease COVID-19*. 10.14293/s2199-1006.1.sor-med.clkmnzy.v1

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet. Infectious Diseases*, *20*(5), 533–534. Advance online publication. doi:10.1016/S1473-3099(20)30120-1 PMID:32087114

Dutta, P. (2017). Decision making in medical diagnosis via distance measures on interval valued fuzzy sets. *International Journal of System Dynamics Applications*, *6*(4), 63–83. doi:10.4018/IJSDA.2017100104

Dyer, O. (2020). COVID-19: Hospitals brace for disaster as US surpasses China in number of cases. *BMJ (Clinical Research Ed.)*, *m1278*, m1278. doi:10.1136/bmj.m1278 PMID:32220826

Elfouly, F. H., Ramadan, R. A., Mahmoud, M. I., & Dessouky, M. I. (2017). Efficient data reporting in a multi-object tracking using WSNs. *International Journal of System Dynamics Applications*, *6*(1), 38–57. doi:10.4018/IJSDA.2017010103

Erasmus, Z. (2020). "Who was here first?", or "Who lives here now?": Indigeneity, a difference like no other. *Persistence of Race*, 155-165. 10.18820/9781928480457/09

Galli, B. J. (2018). Systems thinking and systems analysis in Six Sigma. *International Journal of System Dynamics Applications*, *7*(4), 98–112. doi:10.4018/IJSDA.2018100106

Galli, B. J. (2020). Measurement system analysis and system thinking in Six Sigma. *International Journal of System Dynamics Applications*, *9*(1), 44–62. doi:10.4018/IJSDA.2020010103

Gao, Q., Hu, Y., Dai, Z., Wu, J., Xiao, F., & Wang, J. (2020). *The epidemiological characteristics of 2019 novel coronavirus diseases (COVID-19) in Jingmen, Hubei, China*. 10.1101/2020.03.07.20031393

Jin, H., Liu, J., Cui, M., & Lu, L. (2020). Novel coronavirus pneumonia emergency in Zhuhai: Impact and challenges. *The Journal of Hospital Infection*, *104*(4), 452–453. Advance online publication. doi:10.1016/j.jhin.2020.02.005 PMID:32068012

Jin, H., Liu, J., Cui, M., & Lu, L. (2020). Novel coronavirus pneumonia emergency in Zhuhai: Impact and challenges. *The Journal of Hospital Infection*, *104*(4), 452–453. Advance online publication. doi:10.1016/j.jhin.2020.02.005 PMID:32068012

Kizito, A., & Semwanga, A. R. (2020). Modeling the complexity of road accidents prevention. *International Journal of System Dynamics Applications, 9*(2), 24-41. 10.4018/ijsda.2020040102

Koh, G. (2020). F1000Prime recommendation of report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). *F1000 - Post-publication Peer Review of the Biomedical Literature*. 10.3410/f.737509210.793572110

Koh, G. (2020). F1000Prime recommendation of report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). *F1000 - Post-Publication Peer Review of the Biomedical Literature*. 10.3410/f.737509210.793572110

Koh, G. (2020). F1000Prime recommendation of report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). *F1000 - Post-Publication Peer Review of the Biomedical Literature*. 10.3410/f.737509210.793572110

Koh, G. (2020). F1000Prime recommendation of report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). *F1000 - Post-Publication Peer Review of the Biomedical Literature*. 10.3410/f.737509210.793572110

Livingston, E., & Bucher, K. (2020). undefined. *Journal of the American Medical Association*. Advance online publication. doi:10.1001/jama.2020.4344 PMID:32181795

Maier, B. F., & Brockmann, D. (2020). *Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in Mainland China*. 10.1101/2020.02.18.20024414

Mesel-Lemoine, M., Millet, J., Vidalain, P., Law, H., Vabret, A., Lorin, V., Escriou, N., Albert, M. L., Nal, B., & Tangy, F. (2012). A human coronavirus responsible for the common cold massively kills dendritic cells but not monocytes. *Journal of Virology*, *86*(14), 7577–7587. doi:10.1128/JVI.00269-12 PMID:22553325

Munster, V. J., Koopmans, M., Van Doremalen, N., Van Riel, D., & De Wit, E. (2020). A novel coronavirus emerging in China— Key questions for impact assessment. *The New England Journal of Medicine*, *382*(8), 692–694. doi:10.1056/NEJMp2000929 PMID:31978293

Nesteruk, I. (2020). *Statistics based predictions of coronavirus 2019-nCoV spreading in Mainland China.* 10.1101/2020.02.12.20021931

Niederberger, C. (2020). Re: Rapid sperm capture: high-throughput Flagellar waveform analysis. *The Journal of Urology*, *203*(4), 648. Advance online publication. doi:10.1097/JU.0000000000000722 PMID:31909665

Pasqualoni, S. E. (2020). *2019 novel coronavirus (COVID-19).* 10.2172/1602724

Pierce, D., Shepherd, S., & Johnson, D. (2019). Modelling the impacts of inter-city connectivity on city specialization. *International Journal of System Dynamics Applications*, *8*(4), 47–70. doi:10.4018/IJSDA.2019100104

Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., & Jewell, C. P. (2020). *Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions*. 10.1101/2020.01.23.20018549

Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., & Jewell, C. P. (2020). *Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions*. 10.1101/2020.01.23.20018549

Salzberger, B., Glück, T., & Ehrenstein, B. (2020). Successful containment of COVID-19: The WHO-report on the COVID-19 outbreak in China. *Infection*, *48*(2), 151–153. doi:10.1007/s15010-020-01409-4 PMID:32185635

Smith, R. D. (2006). Responding to global infectious disease outbreaks: Lessons from SARS on the role of risk perception, communication and management. *Social Science & Medicine*, *63*(12), 3113–3123. doi:10.1016/j.socscimed.2006.08.004 PMID:16978751

Stango, V., & Zinman, J. (2009). Exponential growth bias and household finance. *The Journal of Finance*, *64*(6), 2807–2849. doi:10.1111/j.1540-6261.2009.01518.x

Taubenberger, J. K., & Morens, D. M. (2006). 1918 Influenza: The mother of all pandemics. *Revista Bioética*, *17*(1), 69–79.

Theodore, hon. Edward Granville, (1884–9 Feb. 1950), chairman of directors of consolidated press limited, Sydney, which publishes the Daily telegraph, Sunday telegraph, and the Australian women's weekly. (2007). *Who Was Who*. 10.1093/ww/9780199540884.013.u232560

Topcuoglu, N. (2020). Public health emergency of international concern: Coronavirus disease 2019 (COVID-19). *The Open Dentistry Journal*, *14*(1), 71–72. doi:10.2174/1874210602014010071

Vaidya, S. R., Raut, C. G., & Jadhav, S. M. (2016). Laboratory confirmation of rubella infection in suspected measles cases. *Journal of Medical Virology*, *88*(10), 1685–1689. doi:10.1002/jmv.24535 PMID:27018071

Wagenaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics*, *18*(6), 416–422. doi:10.3758/BF03204114

Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *Lancet*, *395*(10223), 470–473. doi:10.1016/S0140-6736(20)30185-9 PMID:31986257

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., & Peng, Z. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *Journal of the American Medical Association*, *323*(11), 1061. doi:10.1001/jama.2020.1585 PMID:32031570

Wilder-Smith, A., & Freedman, D. O. (2006). Confronting the new challenge in travel medicine: SARS. *Journal of Travel Medicine*, *10*(5), 257–258. doi:10.2310/7060.2003.2669 PMID:14531976

Wilkinson, D. (1982). Physics: The next 1 000 000 000 000 000 000 000 000 000 000 000 years. *Physics Bulletin*, *33*(4), 135–138. doi:10.1088/0031-9112/33/4/033

Xu, X., Wu, X., Jiang, X., Xu, K., Ying, L., Ma, C., Li, S., Wang, H., Zhang, S., Gao, H., Sheng, J., Cai, H., Qiu, Y., & Li, L. (2020). Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: Retrospective case series. *BMJ (Clinical Research Ed.)*, *m606*, m606. doi:10.1136/bmj.m606 PMID:32075786

Youbin, L., & Bao hong, L. (2020). *Epidemic situation of novel coronavirus pneumonia in China Mainland*. 10.1101/2020.02.17.20024034

Nasir Mustafa is working as an Assistant Professor in the Department of Health Science at Istanbul Gelisium University, Turkey. https://orcid.org/0000-0002-5821-9297