Magnitude of COVID-19 deaths relative to other leading causes of death: a global analysis

Eunice Y S Chan,¹ Davy Cheng,¹ Janet Martin ²

ABSTRACT

Objectives To quantify the burden of death that COVID-19 contributes relative to the top three causes of death for all countries.

Design We performed uncertainty analyses and created contour plots for COVID-19 mortality to place the number of COVID-19 deaths in context relative to the top three causes of death in each country, across a plausible range of values for two key parameters: case fatality rate and magnitude of under-reporting.

Setting All countries that have reported COVID-19 cases to the WHO and are included in the Global Burden of Disease Study by the Institute of Health Metrics and Evaluation.

Main outcomes and measures Monthly number of deaths caused by COVID-19 and monthly number of deaths caused by the top three causes of death for every country.

Results For countries that were particularly hard hit during the outbreak in 2020, most combinations of model parameters resulted in COVID-19 ranking within the top three causes of death. For countries not as hard hit on a per-capita basis, such as China and India, COVID-19 did not rank higher than the third leading cause of death at any combination of the model parameters within the given ranges. Up-to-date ranking of COVID-19 deaths relative to the top three causes of death for all countries globally is provided in an interactive online application.

Conclusions Estimating the country-level burden of death that COVID-19 contributes relative to the top three causes of death is feasible through contour graphs, even when the actual number of deaths or cases is unknown. This method can help convey importance by placing the magnitude of COVID-related deaths in context relative to more familiar causes of death by communicating when COVID-related deaths rank among the top three causes of death.

INTRODUCTION

Since January 2020, SARS-CoV-2 has progressively moved across the world, leaving almost no country untouched. As resources are finite, reallocating resources to reduce COVID-19 deaths may divert resources from interventions effective for preventing deaths from other causes. Concerns have been raised about whether privileging resources for COVID-19 represents a worthy trade-off, or whether more common diseases are being neglected, resulting in overall net increased all-cause mortality.

In this manuscript, we aim to address the country-level impact of COVID-19 by estimating the magnitude of the disease relative to leading non-COVID-19 causes of death in each country across the world using an uncertainty analysis model. Since the number of cases and deaths due to COVID-19 is under-reported, especially in low-resource settings without testing and data reporting infrastructures, it is difficult to determine the true effects of COVID-19 relative to other non-COVID-19 causes of death, which makes it difficult for citizens to understand the relative importance of the pandemic compared with other more familiar causes of death, and for policymakers and healthcare providers to decide whether to privilege COVID-related care over routine care during a pandemic. Uncertainty analyses can be used to map the likely impact of COVID-19 relative to other causes of death.

Although we have performed the analysis for all countries that have reported COVID-19 cases (available in an up-to-date online app), we will focus the discussion in the main manuscript on eight countries: the USA, Brazil, South Africa, Iran, the UK, Russia, China and India. We chose these countries to provide a
broad representation across continents and across high-income and middle-income countries.

**METHODS**

Inspired by Helleringer and Noymer’s uncertainty analysis for Ebola,\(^1\) we performed uncertainty analyses for COVID-19 mortality to place the number of COVID-19 deaths in context relative to the top three causes of death in each country, across a plausible range of values for two key parameters: case fatality rate (CFR) and magnitude of under-reporting. Together, these two parameters provide plausible ranges of infection fatality rates (IFR). Online supplemental file 1 provides the extended explanation of the methodology. While we display the results for a subset of countries in this commentary, analyses for all countries are available in an online interactive application, which is updated: https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/.

To calculate the estimated number of deaths due to COVID-19 based on the number of reported COVID-19 cases and accounting for the correction factor (\(\beta\)) and the CFR (%), we use the following formula:

\[
\text{Estimated # of deaths} = \frac{\beta \times \text{CFR} \times \# \text{ of reported cases}}{100\%} \times \# \text{(1)}.
\]

While several CFR estimates are available, precise CFR estimation remains infeasible (incomplete ascertainment of cases and deaths), and is context dependent (dependent on age, comorbidities and access to healthcare). At the beginning of the pandemic, WHO\(^2\) reported that the crude CFR of COVID-19 ranged from less than 0.1% to over 25%. As the pandemic progressed, the uncertainty of CFR estimates has decreased. For example, Oke and Heneghan\(^3\) give 95% prediction intervals of the CFR for every country with 1000 cases or greater along with the CFR for each continent, with country-level estimates ranging from 0.06% to 19.64%. Since we focus on eight countries in this manuscript, we chose a range of 0.1%–5% as the plausible estimates of CFR as it encompasses the current CFR estimates for this subset of countries according to Oke and Heneghan.\(^3\) See table 1 for the CFR estimates as of 21 June 2021. Similarly, to estimate the extent of under-reporting due to the different testing strategies across countries and the lack of seroprevalence studies, it is also difficult to precisely determine the appropriate range for the under-reporting correction factor (defined as the ratio of true to confirmed cases). Therefore, we explored the proportion of people infected by the coronavirus who show few or no symptoms, which ranged from 17.9% to 80%,\(^4\)\(^-\)\(^7\) to provide an appropriate range for our model. Given this range in asymptomatic or paucisymptomatic infection, we assumed an under-reporting correction factor varying between 1 and 8. Using equation 1 for every combination of the ranges of values for CFR and \(\beta\) gives us a matrix of values that represents the range of estimates for COVID-19 deaths. In our analysis in this manuscript, we use the number of COVID-19 cases reported by WHO\(^8\) throughout the year 2020.

For our uncertainty analysis, our contour plots are coloured according to the rank of leading causes of deaths: the red regions indicate that COVID-19 would be the leading cause of death; the yellow regions indicate that COVID-19 would be the second leading cause of death; the cyan regions indicate that COVID-19 would be the third leading cause of death; and lastly, the dark blue regions indicate that COVID-19 would not rank in the top three causes of death. To determine where the contours are located on the plot, we compare the range of COVID-19 death estimates to the national estimates of the top three causes of death for each country. Using the data on level 3 causes of death from the Institute of Health Metrics and Evaluation’s (IHME) Global Burden of Disease (GBD) Study 2019,\(^9\) we estimated the latest available estimates for each cause of death using a cubic regression, and determined the top three causes of death and the estimation on the number of deaths due to each respective cause based on the extrapolated results. These estimates on the number of deaths due to the top three causes of death determine the location of the contours in our figures.

We also plotted the number of deaths due to COVID-19 reported by WHO on the contour plots, indicated in black. These contour lines allow users to see the ranking of COVID-19 (solely based on the reported value) relative to other more familiar causes of death by observing which coloured region the line is in, and allow users to determine possible combinations of the CFR and the correction factor that make up the reported number of deaths. This further allows us to determine the IFR of the disease.

### RESULTS

Similar to Helleringer and Noymer,\(^1\) for all combinations of the model parameters, we mapped where COVID-19

<table>
<thead>
<tr>
<th>Country</th>
<th>CFR (%)</th>
<th>(95% prediction interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1.79</td>
<td>(1.79–1.80)</td>
</tr>
<tr>
<td>Brazil</td>
<td>2.80</td>
<td>(2.79–2.81)</td>
</tr>
<tr>
<td>South Africa</td>
<td>3.22</td>
<td>(3.19–3.25)</td>
</tr>
<tr>
<td>Iran</td>
<td>2.68</td>
<td>(2.66–2.69)</td>
</tr>
<tr>
<td>UK</td>
<td>2.76</td>
<td>(2.75–2.78)</td>
</tr>
<tr>
<td>Russia</td>
<td>2.43</td>
<td>(2.42–2.45)</td>
</tr>
<tr>
<td>China</td>
<td>5.06</td>
<td>(4.92–5.20)</td>
</tr>
<tr>
<td>India</td>
<td>1.30</td>
<td>(1.29–1.30)</td>
</tr>
</tbody>
</table>

CFR, case fatality rate.
deaths ranked relative to the top three causes of death in each respective country in 2020 (figure 1). To aid in figure interpretation, we use the USA as an example to demonstrate how one would determine the ranking of COVID-19 as one of the leading causes of death in the country. Despite knowing the number of reported deaths due to COVID-19 in the USA, these values are potentially under-reported for various reasons such as not testing the deceased for COVID-19 due to lack of resources (especially during the height of the pandemic), or differing
definitions as to what constitutes as death due to COVID-19. As shown in table 1, Oke and Heneghan3 predicted that the CFR for the USA is approximately 1.79%. If we also assumed that the number of reported cases is indeed the true number of COVID-19 infections that occurred in the USA in 2020 (ie, β=1), then we can see that the region corresponding to these parameters is yellow, indicating that COVID-19 would be at least the second leading cause of death in the country. However, as we recognise that the number of cases is under-reported, if the correction factor is greater than approximately 1.9 (and if we assume that the CFR is equal to the IFR), this will mean that COVID-19 would have been the leading cause of death in the USA in 2020.

For countries that struggled to manage the virus in 2020, such as the USA, Brazil, South Africa, Iran, the UK and Russia, most combinations of model parameters (CFR=0.1%–5% and β=1–8) resulted in COVID-19 ranking within the top three causes of death. Even though the contour plot may look similar for these countries, this does not necessarily mean similar severity of COVID-19, since different countries have different degrees of under-reporting—high-income countries are likely to have a lower correction factor as they have resources to test large number of people, whereas low-income countries are most likely to have a higher correction factor due to lack of testing kits. This suggests that despite what appears to be similar outcomes between countries, low-income countries are more likely to have more severe impact by COVID-19 (due to having high β) in comparison to high-income countries (which could have a similar CFR but low β). On the other hand, for countries not as hard hit on a per-capita basis in 2020, such as China and India, COVID-19 would only rank in the top three causes of death for combinations of large CFR and β. Lastly, for countries that were able to control the virus early in the pandemic, such as China, COVID-19 would not rank higher than the third leading cause of death at any combination of the model parameters within the given ranges.

Additionally, we can use this uncertainty analysis model to determine the possible combinations of CFR and correction factor based on the reported number of deaths. For example, in figure 1, we observe that COVID-19 was the second leading cause of death (since the black contour line is located in the yellow region of the contour plot) in the USA in 2020 (based on our projections on the number of non-COVID-19 deaths). Additionally, we observe that if we assumed that the reported number of cases is the true number of COVID-19 cases, then the CFR is very close to the estimate that Oke and Heneghan calculated. However, due to the number of COVID-19 cases being under-reported, by increasing the correction factor, we can observe the relationship between β and CFR. For example, according to Kalish et al.16 in the first 6 months of the COVID-19 pandemic, there were 4.8 undiagnosed SARS-CoV-2 infections for every diagnosed case. If we take β=4.8 and follow the contour line, we find that the corresponding CFR is approximately 0.26%. This could potentially give us an idea what the IFR is in the USA.

Our analysis provided in figure 1 gives a broad snapshot as to how COVID-19 has affected each country in 2020 but does not necessarily provide a true sense of the current circumstance, given continued pandemic growth. For updates, we have provided an interactive application (https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/) allowing users to observe how COVID-19 is currently affecting each country and to determine whether the virus is the leading cause of death in their country over adjustable periods of time. This provides more relatable benchmarks for users to understand when the burden of COVID-related deaths has rivalled the most common causes of death in their country, while allowing for users to vary the likely under-reporting and CFR, each of which can never be precisely known. When the COVID-related deaths have surpassed the contour for other most common causes of deaths for all combinations of plausible CFRs and cases (or ratio of unreported:reported cases), it can be assumed that regardless of the uncertainties in number reported, COVID-19 has overtaken all other leading causes of death.

**DISCUSSION**

Estimating the country-level burden of death that COVID-19 contributes relative to the top three causes of death is feasible through uncertainty analyses, even when the exact number of deaths or cases is unknown. This method can help place the importance and magnitude of COVID-related deaths in context relative to more familiar causes of death by communicating when COVID-related deaths rank among the top three causes of death.

This analysis should be interpreted with caveats since collateral increases in non-COVID-19 deaths due to decreased healthcare utilisation are unaccounted for. Additionally, at the time of writing, there was insufficient information about the number of deaths due to non-COVID-19 causes of deaths, and therefore, we estimated both the top three causes of death for each country and the value corresponding to each cause based on the data provided by IHME’s GBD Study. Due to preventative measures that were taken to slow the pandemic, the leading causes of non-COVID-19 deaths and the corresponding values that we estimated may not necessarily be accurate, especially over the longer term due to collateral pandemic-related effects on cardiovascular disease, stroke and cancer (changes in access to healthcare, and changes in lifestyle) which may not become apparent until significant time has passed. Lastly, once the pandemic has matured, plausible ranges for CFR and under-reporting may require adjustment.

Despite the limitations, uncertainty analyses provide opportunity to quantify the burden of disease that COVID-19 contributes to each country relative to other causes of death, across a broad range of possibilities. During the current pandemic, and in future pandemics,
timely application of this tool may better inform reflex assumptions that COVID-related burden of death is not worthy of the actions taken, or that under-reporting negates any credibility of the reported statistics. The online tool may also facilitate decision makers to determine the most appropriate course of action given limited healthcare resources at different stages of the outbreak.

Twitter Davy Cheng @DavyCheng1 and Janet Martin @janetmartinwhv

Contributors JM: conceptualisation, statistical analysis, manuscript revision, final approval of manuscript. EYSC: conceptualisation, statistical analysis, manuscript drafting, development of interactive application, final approval of manuscript and guarantor of overall content. DC: manuscript revision, final approval of manuscript.

Funding This work was supported by the Department of Anesthesia and Perioperative Medicine, Western University, through operational funding.

Disclaimer The funding source had no involvement in the design and conduct of the study, collection, management, analysis and interpretation of data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data and codes pertinent to this study are publicly available at the following link: https://github.com/MEDICI-UWO/COVID-19-uncertainty-analysis.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Janet Martin http://orcid.org/0000-0002-3599-8427

REFERENCES


7 Day M. Covid-19: four fifths of cases are asymptomatic, China figures indicate. BMJ 2020;368:m1375.


Supplementary Material: Magnitude of COVID-19 mortality relative to other causes of death in Italy, Spain, United Kingdom, Russia, China, Canada, USA, and Brazil

eAppendix: Extended explanation of methodology.

Model 1: Uncertainty analysis of COVID-19 mortality based on CFR and number of cases

Model 2: Uncertainty analysis of COVID-19 mortality based on the CFR and ratio of true and reported cases

eFigure 1: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Russia in 2020.

eFigure 2: Data of Canada’s top three causes of death from 1990 to 2017 (data from GBD 2017), fitted with degree 3 polynomials.

eFigure 3: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, on the basis of the case fatality rate and the ratio of true and reported COVID-19 cases as of December 16, 2020.

eFigure 4: Comparison of COVID-19 deaths and expected deaths from non-COVID-19 causes in the United States over different date ranges.

eTable 1: Top three causes of death in 2017 and estimated values for 2020 in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil. Data from 2017 were obtained from IHME’s Global Burden of Disease Study 2017.

eTable 2: Data on the cumulative number of cases as of December 16, 2020 for Canada, China, Italy, Spain, the UK, and USA, published in the WHO Coronavirus Disease (COVID-19) Dashboard.
In the following, we present two uncertainty analysis models of COVID-19 mortality. The first model derives the number of deaths due to the coronavirus on the basis of the case fatality rate (CFR) of COVID-19 and the number of cases. The second model is an extension of the first model in which it also derives the number of deaths due to COVID-19, but it also takes into account that the number of confirmed cases is likely to be underreported by using a correction factor, $\beta$. The second model is based on the model presented in Helleringer and Noymer\cite{1}, where they performed an uncertainty analysis on the number of deaths caused by Ebola in Liberia, Sierra Leone, and Guinea. Similar to the model in (1), we also do not account for increases in deaths unrelated to COVID-19 due to lower healthcare use during the outbreak.

Accompanying this manuscript, we present an interactive web-based application\footnote{https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/}, which contains both models for all countries that have reported COVID-19 cases to the World Health Organization (WHO) and are included in the Global Burden of Disease Study 2017 (GBD 2017) by the Institute for Health Metrics and Evaluation (IHME)\cite{2}. This application will be updated on a daily basis provided that the WHO continues to provide daily situation reports. In this manuscript, however, due to the vast number of countries, we will only present the results for eight countries: Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil.

**Model 1: Uncertainty analysis of COVID-19 mortality based on CFR and number of cases**

We are interested in knowing the number of monthly\footnote{Note that in the web application, we provide both the annual and monthly values. However, since we assume that the number of deaths are uniformly distributed throughout the year and how we have determined the range of the y-axis (which is explained below), the appearance of the contours will look the same regardless of whether we are looking at the annual or monthly numbers.} confirmed cases it would take for COVID-19 to become one of the leading causes of death in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, with respect to the case fatality rate (CFR) of the coronavirus. We can relate these components using the crude formula to calculate the CFR. By definition,

$$\text{CFR} = \frac{\text{Number of deaths due to a particular disease}}{\text{Total number of cases due to the same disease}} \times 100\%. \quad (1)$$

Since, in this manuscript, we are interested in the monthly number of deaths, we modify equation (1) by specifying that the number of deaths and cases are taken over a time interval of a month:

$$\text{CFR} = \frac{\text{Number of deaths due to a particular disease over a span of a month}}{\text{Number of cases due to the same disease over a span of a month}} \times 100\%. \quad (2)$$

This then allows us to calculate the theoretical number of deaths by simply rearranging the terms:

$$\text{Theoretical number of deaths due to a particular disease over a span of a month} = \text{CFR} \times \frac{\text{Total number of cases due to the same disease over a span of a month}}{100\%} \quad (3)$$

We use this formula and vary the CFR and the number of cases in our uncertainty analysis model, shown in eFigure 1.
eFigure 1: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Russia in 2020.

COPD = Chronic obstructive pulmonary disease. Ranking of COVID-19 deaths relative to non-COVID-19 cause of death: Red = COVID-19 is the leading cause of death, Yellow = COVID-19 is the second leading cause of death, Cyan = COVID-19 is the third leading cause of death, Blue = COVID-19 has a lower ranking than third cause of death. The white contours represent the estimated number of deaths due to COVID-19 in 2020 for specific combinations of model parameters (where 1k = 1000). Calculations of the number of deaths due to COVID-19 and comparisons with other causes of death are described in the section describing uncertainty analysis model 1. To aid figure interpretation, we illustrate the case of Canada. The red region indicates the combination of model parameters where there are more theoretical deaths due to COVID-19 over a span of a month in Canada than the leading cause of mortality, which is ischemic heart disease, and thus, COVID-19 would be the leading cause of death based on the number of confirmed cases and the case fatality rate. Similarly, the yellow region indicates that COVID-19 would be the second leading cause of death in Canada, surpassing Alzheimer’s disease, and lastly, the blue region indicates that COVID-19 would not be one of the top three leading causes of death in Canada. The boundary between the red and yellow regions represents the combination of model parameters where the number of number of deaths due to COVID-19 is equal to the number of deaths due to ischemic heart disease. Similarly, the boundary between the yellow and cyan regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to Alzheimer’s disease, and the boundary between the cyan and blue regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to lung cancer.

Since the COVID-19 outbreak is still ongoing, we decided to take a particularly broad range for the CFR because the crude estimate of the case fatality rate (by simply dividing the number of known deaths by the number of confirmed cases) could potentially be off by one or two orders of magnitude and it is currently not possible to...
provide a precise estimate for the CFR of the disease. (3) Many articles have provided CFR estimates for COVID-19. The World Health Organization (WHO) has estimated that the global CFR for COVID-19 is about 3.4% as of March 3, 2020. (4) Wilson et al. (5) calculated the CFR by using a 13-day time lag and estimated that the CFR for COVID-19 is 3.5% in China and 4.2% in 82 countries, territories and areas. They also note that lower estimates are likely to be closer to the true CFR. Oke and Heneghan (6), on the other hand, computed the CFR for COVID-19 using fixed-effect inverse-variance weighting and predicted that the CFR (as of December 16, 2020) ranges from 0.61% to 4.8% for countries in Africa, 0.33% to 4.4% for countries in Asia, 0.007% to 7.5% for countries in Australia/Oceania, 0.89% to 3.4% for countries in Europe, 0.21% to 14.7% for countries in North America, and 1.16% to 4.7% for countries in South America. Conversely, the authors of Worldometer (7) calculated the CFR using the method described in Ghani et al. (8) and suggested a worldwide CFR estimate of 20% as of April 21, 2020. With such varied CFR estimates, we decided to choose a range that contains the most common CFR estimates at the time of writing, and therefore, we vary the CFR between 0.1% and 10%. However, for eFigure 1, due to how we calculate the vertical range for the contour plots (explained below), we restrict the CFR range to 2% to 10% to provide a better visualization for each contour.

To assess the magnitude of COVID-19 mortality relative to other diseases and conditions, we used estimates of mortality from other causes of death obtained from the Global Burden of Disease Study 2017 by the Institute for Health Metrics and Evaluation (IHME). (2) The top three causes of death in Italy, Spain, UK, Russia, China, Canada, USA, and Brazil and their respective values are given in eTable 1. Assuming that the top three causes of death in the countries are the same in 2020 as for 2017, we extrapolated the data (between the years 1990 and 2017, inclusive) from the GBD 2017 to estimate the number of deaths in the year 2020 for the top three causes in each respective country. We extrapolated the data by first using R’s lm() function (in combination with R’s poly() function) to fit a degree 3 polynomial regression model through the data points and then using R’s predict() function to estimate the values based on the regression model. eFigure 2 shows the plots for the number of deaths due to ischemic heart disease (eFigure 2a), Alzheimer’s disease (eFigure 2b), and lung cancer (eFigure 2c) in Canada with their respective regression lines and corresponding $R^2$. eTable 1 gives the estimated values for the top three causes of death in the aforementioned countries in the year 2020. Since we are interested in the monthly mortality rate rather than the annual mortality rate (given in eTable 1) and assuming that the number of deaths is uniformly distributed throughout the year, for the calculations for our contour plots (eFigure 1), we divided the estimated values by 12.

eFigure 2a: Ischemic heart disease, $R^2 = 0.9672836$

eFigure 2b: Alzheimer’s disease and other dementias, $R^2 = 0.9993398$

eFigure 2c: Tracheal, bronchus, and lung cancer, $R^2 = 0.9858434$

*eFigure 2: Data of Canada’s top three causes of death from 1990 to 2017 (data from GBD 2017), fitted with degree 3 polynomials.*

3 In the online interactive application, for countries with HIV/AIDS as one of their top 3 causes of death in 2017, since the number of deaths due to HIV/AIDS is decreasing, we excluded this cause from their top 3 causes of death and replaced it with their fourth leading cause of death.
eTable 1: Top three causes of death in 2017 and estimated values for 2020 in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil. Data from 2017 were obtained from IHME's Global Burden of Disease Study 2017.(2)

<table>
<thead>
<tr>
<th>Country</th>
<th>Rank</th>
<th>Cause of death</th>
<th>2017 Value</th>
<th>Est. 2020 Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>175792</td>
<td>187504</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Stroke</td>
<td>122783</td>
<td>129451</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lower respiratory infections</td>
<td>84073</td>
<td>99135</td>
</tr>
<tr>
<td>Canada</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>48381</td>
<td>55270</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>25219</td>
<td>28368</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lung cancer</td>
<td>21440</td>
<td>21878</td>
</tr>
<tr>
<td>China</td>
<td>1st</td>
<td>Stroke</td>
<td>2109772</td>
<td>2285914</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Ischemic heart disease</td>
<td>1750038</td>
<td>1906529</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Chronic obstructive pulmonary disease</td>
<td>965873</td>
<td>1082575</td>
</tr>
<tr>
<td>Italy</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>96461</td>
<td>102243</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>73339</td>
<td>80328</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>59355</td>
<td>63998</td>
</tr>
<tr>
<td>Russia</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>562921</td>
<td>504328</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Stroke</td>
<td>317548</td>
<td>309653</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Alzheimer's disease</td>
<td>84369</td>
<td>102266</td>
</tr>
<tr>
<td>Spain</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>55234</td>
<td>56321</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>51759</td>
<td>54956</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>32382</td>
<td>32292</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>87019</td>
<td>100632</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>63894</td>
<td>72675</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>47022</td>
<td>49209</td>
</tr>
<tr>
<td>United States</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>533166</td>
<td>591513</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>258587</td>
<td>288770</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lung cancer</td>
<td>190698</td>
<td>199251</td>
</tr>
</tbody>
</table>
Lastly, we are particularly interested in determining the minimum number of confirmed COVID-19 cases (monthly) it takes to surpass other causes of death in Italy, Spain, UK, Russia, China, Canada, USA, and Brazil at any given case fatality rate. Visually (see eFigure 1), these minima occur at the boundaries between two different colors. Once again, we rearrange equation (2), but this time, we isolate the variable representing the number of cases:

\[
\text{Number of cases due to a particular disease over a span of a month} = \frac{\text{Number of deaths due to the same disease over a span of a month}}{\text{CFR}} \times 100\% \quad (4)
\]

In order for COVID-19 to surpass other diseases in their ranking in terms of mortality, the number of deaths due to COVID-19 must be at least equal to the number of deaths due to the disease of interest. Theoretically, this means that the minimum number of COVID-19 cases (over a span of a month) can be determined if we set the number of deaths due to COVID-19 to the number of deaths due to the other disease and/or condition of interest. Therefore, by substituting the appropriate terms into equation (4), we derive the formula

\[
\text{Minimum number of confirmed cases of COVID-19 over a span of a month} = \frac{\text{Number of deaths due to disease of interest over a span of a month}}{\text{CFR}} \times 100\% \quad (5)
\]

As an example, if we wanted to determine the minimum number of monthly cases of COVID-19 in order to surpass ischemic heart disease as the leading cause of mortality in Canada in 2020 (based on the value given in eTable 1) and assuming that the case fatality rate is 3%, using equation (5), the minimum number of cases is equal to \(\frac{55270}{3\%} \times 100\% \approx 1842333.3\) (1842934 if you want a whole number).

We also use equation (5) to determine the appropriate vertical display range (the number of cases) for the contour plots shown in eFigure 1. We want to obtain a vertical range that displays all of the contours/boundaries within the smallest vertical range possible. With this in mind, the upper bound of the vertical range should be the minimum number of COVID-19 cases for it to be the leading cause of death when the CFR is equal to its respective lower bound. This is because the region above this value (regardless of what the CFR is) will obviously be red on the contour plot, which does not provide us with any more interesting insights. We can apply a similar logic to find the lower bound of the vertical range: the lower bound should be the minimum number of COVID-19 cases for it to be the third leading cause of death because the region below this will be blue (since we are only interested in the top three leading causes of death). Therefore, for our particular contour plots, we can use equation (5) to calculate both the upper and lower vertical bound. To calculate the upper bound, we substitute the number of deaths due to the leading cause of mortality in each country into the numerator of the fraction on the right hand side of the equation and set CFR = 2% (which is the lower bound of the range that we chose for CFR). As for the lower bound, we substitute the number of deaths due to the third leading cause of mortality into the numerator of the fraction and set the case fatality rate to our upper bound, 10%. By doing this, we can observe the number of cases it takes for COVID-19 to be one of the top three causes of mortality for all values of CFR within our range of interest.
Model 2: Uncertainty analysis of COVID-19 mortality based on the CFR and ratio of true and reported cases

We extend the first model to take into account that the number of cases of COVID-19 is most likely underreported by introducing a correction factor, $\beta$. The correction factor, $\beta$, is a ratio between the true number of COVID-19 cases and the number of reported cases, which gives the relation

$$\text{True number of cases} = \beta \times \text{Number of reported cases}. \quad (6)$$

Therefore, when $\beta > 1$, the true number of cases is greater than the number of reported cases (meaning that the number of COVID-19 cases are underreported). Conversely, when $\beta < 1$, the true number of cases is less than the number of reported cases (meaning that the number of COVID-19 cases are overreported). Lastly, when $\beta = 1$, the number of reported cases is the same as the true number of cases (meaning that the number of COVID-19 cases are neither underreported nor overreported). Since it is widely agreed that the number of reported cases are underreported, the values of $\beta$ that are of interest to us would be greater than one.

To estimate the potential number of deaths based on the true number of cases of COVID-19, we substitute equation (6) into equation (3) to acquire

$$\text{Theoretical number of COVID-19 deaths} = \frac{\beta \times \text{CFR} \times \text{Number of reported COVID-19 cases}}{100\%} \quad (7),$$

which is the formula we use for this uncertainty analysis model, shown in eFigure 3.

We turn to the literature to determine the range we should choose for $\beta$ in our model. Li et al. (9) used a networked dynamic metapopulation model and Bayesian inference and estimated that 86% of all infections in China were undocumented before the travel restriction was put in place on January 23, 2020. Since there is a limited number of articles that explicitly estimate the number of unreported cases at the time of writing, we also consulted studies that estimate the proportion of COVID-19 cases that show mild or no symptoms, since these cases would most likely go undetected. An article from Nature (10) states that there are estimates that suggest around 60% of people with COVID-19 show mild or no symptoms. On the other hand, Nishiura et al. (11) looked at data of the Japanese citizens who were evacuated from Wuhan, China in early February and were repeatedly tested and closely monitored; the authors estimated based on this data that the asymptomatic proportion is 30.8%. Lastly, Mizumoto et al. (12) looked at data collected from the COVID-19 outbreak on the Diamond Princess cruise ship and estimated that the asymptomatic proportion among all infected cases is 17.9%. However, this estimate is not representative for a general population because many, if not most, of the passengers on the cruise ship were elderly. With such varied estimates, we, therefore, decided to take a large range and vary $\beta$ from 1 to 8.

For this uncertainty analysis model, we also decided to choose a smaller range for the case fatality rate (0.1%–5%) compared to what was chosen for our first model. We reduced the range to provide better visualization of the contours as they are mostly situated where the CFR is small.
### Table 2: Data on the cumulative number of cases as of December 16, 2020 for Canada, China, Italy, Spain, the UK, and USA, published in the WHO Coronavirus Disease (COVID-19) Dashboard.

<table>
<thead>
<tr>
<th>Country</th>
<th>Date of first COVID-19 case reported to WHO</th>
<th>Number of reported cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>February 27, 2020</td>
<td>6927145</td>
</tr>
<tr>
<td>Canada</td>
<td>January 27, 2020</td>
<td>468862</td>
</tr>
<tr>
<td>China</td>
<td>January 4, 2020</td>
<td>95279</td>
</tr>
<tr>
<td>Italy</td>
<td>January 31, 2020</td>
<td>1870576</td>
</tr>
<tr>
<td>Russia</td>
<td>February 1, 2020</td>
<td>2734454</td>
</tr>
<tr>
<td>Spain</td>
<td>February 1, 2020</td>
<td>1762212</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>February 1, 2020</td>
<td>1888120</td>
</tr>
<tr>
<td>United States</td>
<td>January 23, 2020</td>
<td>16245376</td>
</tr>
</tbody>
</table>
eFigure 3: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, on the basis of the case fatality rate and the ratio of true and reported COVID-19 cases as of December 16, 2020.

COPD = Chronic obstructive pulmonary disease. Ranking of COVID-19 deaths relative to non-COVID-19 cause of death: Red = COVID-19 is the leading cause of death, Yellow = COVID-19 is the second leading cause of death, Cyan = COVID-19 is the third leading cause of death, Blue = COVID-19 has a lower ranking than third cause of death. The white contours represent the estimated number of deaths due to COVID-19 in 2020 for specific combinations of model parameters (where 1k = 1000). Calculations of the number of deaths due to COVID-19 and comparisons with other causes of death are described in the section describing uncertainty analysis model 2. To aid figure interpretation, we illustrate the case of Italy. The red region indicates the combination of model parameters where there are more theoretical deaths due to COVID-19 over the duration of the outbreak in Italy than the leading cause of mortality, which is ischemic heart disease, and thus, COVID-19 would be the leading cause of death based on the number of confirmed cases and the case fatality rate. Similarly, the yellow region indicates that COVID-19 would be the second leading cause of death in Italy, surpassing Alzheimer's disease, the cyan region indicates that COVID-19 would be the third leading cause of death in Italy, surpassing stroke, and lastly, the blue region indicates that COVID-19 would not be one of the top three leading causes of death in Italy. The boundary between the red and yellow regions represents the combination of model parameters where the number of number of deaths due to COVID-19 is equal to the number of deaths due to ischemic heart disease. Similarly, the boundary between the yellow and cyan regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to Alzheimer's disease, and the boundary between the cyan and blue regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to lung cancer. In China, the entire plot is blue because the number of deaths due to COVID-19 is less than the expected number of deaths from the third leading cause of death in the country for all model parameters.
Now that we have determined the range that we want to vary the CFR and $\beta$, we need to choose which data set we want to use for each country in order to be able to calculate the theoretical number of deaths—we decided to use the cumulative number of cases provided by the WHO published on the date of writing (December 16, 2020), given in eTable 2. eFigure 3 gives the uncertainty analyses on the basis of the case fatality rate and the ratio of the true to reported number of cases of the eight countries of interest.

In order to be able to compare the magnitude of mortality due to COVID-19 to the top three causes of death in each country fairly, we need to ensure that the number of deaths are taken over the same duration. Similar to the first model, we use the values that we estimated for the year 2020, which are presented in eTable 1 (details on how these values were estimated are provided in the previous section). We adjust these annual values to the appropriate time frame, which, in this case, is the duration, in days, that the coronavirus has been present in the country $^4$, by using the following formula:

$$\text{Number of deaths due to certain cause during certain time frame} = \frac{\text{Annual number of deaths due to same cause}}{365} \times \text{Number of days that COVID-19 has been present in country},$$

again assuming that the number of deaths are uniformly distributed throughout the year. Despite the fact that there is a time lag between the onset of symptoms and death, using this duration should be acceptable since the duration of people dying from COVID-19 is similar in length to the duration of the same group of people contracting the coronavirus; the time frame is just delayed. We use these corrected values to determine how each region is colored in the contour plots presented in eFigure 3. Although the first reported case for China occurred in 2019, we still used our estimated value for the year 2020 for the adjustment rather than taking a linear combination of the estimated values for the years 2019 and 2020 based on the proportion of days that the coronavirus has been in the country each year. Not only is this highly unnecessary (as we are dealing with estimated values to begin with), but it actually makes no difference to the contour plot, since the maximum value of the theoretical number of deaths calculated using equation (7) is much smaller than the values given for the leading causes of death in China. Based on the reported number of cases provided by the World Health Organization (WHO) on December 16, 2020, our model (eFigure 3) predicts that the estimated number of deaths due to COVID-19 ranges from 1856 to 742295 for Italy, 1752 to 700754 for Spain, 1870 to 747868 for the United Kingdom, 2708 to 1083178 for Russia, 96 to 38111 for China, 469 to 187545 for Canada, 16246 to 6498151 for the United States, and 6928 to 2770858 for Brazil.

$^4$ Note that we are counting the day that the WHO reports the first occurrence of the virus in the country as the first day rather than the day that the virus physically arrived in the country. (For e.g., although there are reports that COVID-19 first arrived in China on December 1, 2019, since the coronavirus was reported to the WHO on December 31, 2019, we will use the latter date as the first day.)
**eFigure 4: Comparison of COVID-19 deaths and expected deaths from non-COVID-19 causes in the United States over different date ranges.**

Since the spread of COVID-19 is exponential, this means that the number of daily cases grows over a certain period of time after the virus arrives to the country. Therefore, eFigure 3 does not necessarily give us a true sense of where COVID-19 is ranked in leading causes of mortality, especially at the present. For example, in eFigure 4, we compare the uncertainty analysis model for the United States over different date ranges: the plot on the left takes the entire duration that COVID-19 has been present in the United States, the plot second to the left takes the number of cases confirmed in the first 30 days of the arrival of the virus, and the other two plots take the duration of 60 days approximately when the first and second wave of the pandemic occurred, respectively. This suggests that limiting the date range over a certain period of time can provide a better depiction of the severity of COVID-19 at a particular time.

- Code to reproduce these results are publicly available at the following link: https://github.com/MEDICI-UWO/COVID-19-uncertainty-analysis
- R Shiny Application is available at the following link (updated daily): https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/
**References**


Supplementary Material: Magnitude of COVID-19 mortality relative to other causes of death in Italy, Spain, United Kingdom, Russia, China, Canada, USA, and Brazil

eAppendix: Extended explanation of methodology.

Model 1: Uncertainty analysis of COVID-19 mortality based on CFR and number of cases

Model 2: Uncertainty analysis of COVID-19 mortality based on the CFR and ratio of true and reported cases

eFigure 1: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Russia in 2020.

eFigure 2: Data of Canada’s top three causes of death from 1990 to 2017 (data from GBD 2017), fitted with degree 3 polynomials.

eFigure 3: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, on the basis of the case fatality rate and the ratio of true and reported COVID-19 cases as of December 16, 2020.

eFigure 4: Comparison of COVID-19 deaths and expected deaths from non-COVID-19 causes in the United States over different date ranges.

eTable 1: Top three causes of death in 2017 and estimated values for 2020 in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil. Data from 2017 were obtained from IHME’s Global Burden of Disease Study 2017.

eTable 2: Data on the cumulative number of cases as of December 16, 2020 for Canada, China, Italy, Spain, the UK, and USA, published in the WHO Coronavirus Disease (COVID-19) Dashboard.
eAppendix: Extended Explanation of Methodology

In the following, we present two uncertainty analysis models of COVID-19 mortality. The first model derives the number of deaths due to the coronavirus on the basis of the case fatality rate (CFR) of COVID-19 and the number of cases. The second model is an extension of the first model in which it also derives the number of deaths due to COVID-19, but it also takes into account that the number of confirmed cases is likely to be underreported by using a correction factor, $\beta$. The second model is based on the model presented in Helleringer and Noymer(1), where they performed an uncertainty analysis on the number of deaths caused by Ebola in Liberia, Sierra Leone, and Guinea. Similar to the model in (1), we also do not account for increases in deaths unrelated to COVID-19 due to lower healthcare use during the outbreak.

Accompanying this manuscript, we present an interactive web-based application\(^1\), which contains both models for all countries that have reported COVID-19 cases to the World Health Organization (WHO) and are included in the Global Burden of Disease Study 2017 (GBD 2017) by the Institute for Health Metrics and Evaluation (IHME).\(^2\) This application will be updated on a daily basis provided that the WHO continues to provide daily situation reports. In this manuscript, however, due to the vast number of countries, we will only present the results for eight countries: Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil.

Model 1: Uncertainty analysis of COVID-19 mortality based on CFR and number of cases

We are interested in knowing the number of monthly\(^2\) confirmed cases it would take for COVID-19 to become one of the leading causes of death in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, with respect to the case fatality rate (CFR) of the coronavirus. We can relate these components using the crude formula to calculate the CFR. By definition,

\[
\text{CFR} = \frac{\text{Number of deaths due to a particular disease}}{\text{Total number of cases due to the same disease}} \times 100\%. \tag{1}
\]

Since, in this manuscript, we are interested in the monthly number of deaths, we modify equation (1) by specifying that the number of deaths and cases are taken over a time interval of a month:

\[
\text{CFR} = \frac{\text{Number of deaths due to a particular disease over a span of a month}}{\text{Number of cases due to the same disease over a span of a month}} \times 100\%. \tag{2}
\]

This then allows us to calculate the theoretical number of deaths by simply rearranging the terms:

\[
\text{Theoretical number of deaths due to a particular disease over a span of a month} = \text{CFR} \times \frac{\text{Total number of cases due to the same disease over a span of a month}}{100\%} \tag{3}
\]

We use this formula and vary the CFR and the number of cases in our uncertainty analysis model, shown in eFigure 1.

\(^{1}\)https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/

\(^{2}\)Note that in the web application, we provide both the annual and monthly values. However, since we assume that the number of deaths are uniformly distributed throughout the year and how we have determined the range of the y-axis (which is explained below), the appearance of the contours will look the same regardless of whether we are looking at the annual or monthly numbers.
**Figure 1: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Russia in 2020.**

COPD = Chronic obstructive pulmonary disease. Ranking of COVID-19 deaths relative to non-COVID-19 cause of death: Red = COVID-19 is the leading cause of death, Yellow = COVID-19 is the second leading cause of death, Cyan = COVID-19 is the third leading cause of death, Blue = COVID-19 has a lower ranking than third cause of death. The white contours represent the estimated number of deaths due to COVID-19 in 2020 for specific combinations of model parameters (where 1k = 1000). Calculations of the number of deaths due to COVID-19 and comparisons with other causes of death are described in the section describing uncertainty analysis model 1. To aid figure interpretation, we illustrate the case of Canada. The red region indicates the combination of model parameters where there are more theoretical deaths due to COVID-19 over a span of a month in Canada than the leading cause of mortality, which is ischemic heart disease, and thus, COVID-19 would be the leading cause of death based on the number of confirmed cases and the case fatality rate. Similarly, the yellow region indicates that COVID-19 would be the second leading cause of death in Canada, surpassing lung cancer, and lastly, the blue region indicates that COVID-19 would not be one of the top three leading causes of death in Canada. The boundary between the red and yellow regions represents the combination of model parameters where the number of number of deaths due to COVID-19 is equal to the number of deaths due to ischemic heart disease. Similarly, the boundary between the yellow and cyan regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to Alzheimer’s disease, and the boundary between the cyan and blue regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to lung cancer.

Since the COVID-19 outbreak is still ongoing, we decided to take a particularly broad range for the CFR because the crude estimate of the case fatality rate (by simply dividing the number of known deaths by the number of confirmed cases) could potentially be off by one or two orders of magnitude and it is currently not possible to
provide a precise estimate for the CFR of the disease.(3) Many articles have provided CFR estimates for COVID-19. The World Health Organization (WHO) has estimated that the global CFR for COVID-19 is about 3.4% as of March 3, 2020.(4) Wilson et al.(5) calculated the CFR by using a 13-day time lag and estimated that the CFR for COVID-19 is 3.5% in China and 4.2% in 82 countries, territories and areas. They also note that lower estimates are likely to be closer to the true CFR. Oke and Heneghan(6), on the other hand, computed the CFR for COVID-19 using fixed-effect inverse-variance weighting and predicted that the CFR (as of December 16, 2020) ranges from 0.61% to 4.8% for countries in Africa, 0.33% to 4.4% for countries in Asia, 0.007% to 70.5% for countries in Australia/Oceania, 0.89% to 3.4% for countries in Europe, 0.21% to 14.7% for countries in North America, and 1.16% to 4.7% for countries in South America. Conversely, the authors of Worldometer(7) calculated the CFR using the method described in Ghani et al.(8) and suggested a worldwide CFR estimate of 20% as of April 21, 2020. With such varied CFR estimates, we decided to choose a range that contains the most common CFR estimates at the time of writing, and therefore, we vary the CFR between 0.1% and 10%. However, for eFigure 1, due to how we calculate the vertical range for the contour plots (explained below), we restrict the CFR range to 2% to 10% to provide a better visualization for each contour.

To assess the magnitude of COVID-19 mortality relative to other diseases and conditions, we used estimates of mortality from other causes of death obtained from the Global Burden of Disease Study 2017 by the Institute for Health Metrics and Evaluation (IHME).(2) The top three causes of death in Italy, Spain, UK, Russia, China, Canada, USA, and Brazil and their respective values are given in eTable 1. Assuming that the top three causes of death in the countries are the same in 2020 as for 2017, we extrapolated the data (between the years 1990 and 2017, inclusive) from the GBD 2017 to estimate the number of deaths in the year 2020 for the top three causes in each respective country. We extrapolated the data by first using R’s lm() function (in combination with R’s poly() function) to fit a degree 3 polynomial regression model through the data points and then using R’s predict() function to estimate the values based on the regression model. eFigure 2 shows the plots for the number of deaths due to ischemic heart disease (eFigure 2a), Alzheimer’s disease (eFigure 2b), and lung cancer (eFigure 2c) in Canada with their respective regression lines and corresponding $R^2$. eTable 1 gives the estimated values for the top three causes of death in the aforementioned countries in the year 2020. Since we are interested in the monthly mortality rate rather than the annual mortality rate (given in eTable 1) and assuming that the number of deaths is uniformly distributed throughout the year, for the calculations for our contour plots (eFigure 1), we divided the estimated values by 12.

\[ R^2 = 0.9672836 \]
\[ R^2 = 0.9993398 \]
\[ R^2 = 0.9858434 \]

**eFigure 2:** Data of Canada’s top three causes of death from 1990 to 2017 (data from GBD 2017), fitted with degree 3 polynomials.

---

3 In the online interactive application, for countries with HIV/AIDS as one of their top 3 causes of death in 2017, since the number of deaths due to HIV/AIDS is decreasing, we excluded this cause from their top 3 causes of death and replaced it with their fourth leading cause of death.
eTable 1: Top three causes of death in 2017 and estimated values for 2020 in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil. Data from 2017 were obtained from IHME’s Global Burden of Disease Study 2017.(2)

<table>
<thead>
<tr>
<th>Country</th>
<th>Rank</th>
<th>Cause of death</th>
<th>2017 Value</th>
<th>Est. 2020 Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>175792</td>
<td>187504</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Stroke</td>
<td>122783</td>
<td>129451</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lower respiratory infections</td>
<td>84073</td>
<td>99135</td>
</tr>
<tr>
<td>Canada</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>48381</td>
<td>55270</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>25219</td>
<td>28368</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lung cancer</td>
<td>21440</td>
<td>21878</td>
</tr>
<tr>
<td>China</td>
<td>1st</td>
<td>Stroke</td>
<td>2109772</td>
<td>2285914</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Ischemic heart disease</td>
<td>1750038</td>
<td>1906529</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Chronic obstructive pulmonary disease</td>
<td>965873</td>
<td>1082575</td>
</tr>
<tr>
<td>Italy</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>96461</td>
<td>102243</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>73339</td>
<td>80328</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>59355</td>
<td>63998</td>
</tr>
<tr>
<td>Russia</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>562921</td>
<td>504328</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Stroke</td>
<td>317548</td>
<td>309653</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Alzheimer's disease</td>
<td>84369</td>
<td>102266</td>
</tr>
<tr>
<td>Spain</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>55234</td>
<td>56321</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>51759</td>
<td>54956</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>32382</td>
<td>32292</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>87019</td>
<td>100632</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>63894</td>
<td>72675</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Stroke</td>
<td>47022</td>
<td>49209</td>
</tr>
<tr>
<td>United States</td>
<td>1st</td>
<td>Ischemic heart disease</td>
<td>533166</td>
<td>591513</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>Alzheimer's disease</td>
<td>258587</td>
<td>288770</td>
</tr>
<tr>
<td></td>
<td>3rd</td>
<td>Lung cancer</td>
<td>190698</td>
<td>199251</td>
</tr>
</tbody>
</table>
Lastly, we are particularly interested in determining the minimum number of confirmed COVID-19 cases (monthly) it takes to surpass other causes of death in Italy, Spain, UK, Russia, China, Canada, USA, and Brazil at any given case fatality rate. Visually (see eFigure 1), these minima occur at the boundaries between two different colors. Once again, we rearrange equation (2), but this time, we isolate the variable representing the number of cases:

\[
\text{Number of cases due to a particular disease over a span of a month} = \frac{\text{Number of deaths due to the same disease over a span of a month}}{\text{CFR}} \times 100\% \quad (4)
\]

In order for COVID-19 to surpass other diseases in their ranking in terms of mortality, the number of deaths due to COVID-19 must be at least equal to the number of deaths due to the disease of interest. Theoretically, this means that the minimum number of confirmed COVID-19 cases (over a span of a month) can be determined if we set the number of deaths due to COVID-19 to the number of deaths due to the other disease and/or condition of interest. Therefore, by substituting the appropriate terms into equation (4), we derive the formula

\[
\text{Minimum number of confirmed cases of COVID-19 over a span of a month} = \frac{\text{Number of deaths due to disease of interest over a span of a month}}{\text{CFR}} \times 100\% \quad (5)
\]

As an example, if we wanted to determine the minimum number of monthly cases of COVID-19 in order to surpass ischemic heart disease as the leading cause of mortality in Canada in 2020 (based on the value given in eTable 1) and assuming that the case fatality rate is 3%, using equation (5), the minimum number of cases is equal to \(\frac{55270}{3\%} \times 100\% \approx 184233.3\) (1842934 if you want a whole number).

We also use equation (5) to determine the appropriate vertical display range (the number of cases) for the contour plots shown in eFigure 1. We want to obtain a vertical range that displays all of the contours/boundaries within the smallest vertical range possible. With this in mind, the upper bound of the vertical range should be the minimum number of COVID-19 cases for it to be the leading cause of death when the CFR is equal to its respective lower bound. This is because the region above this value (regardless of what the CFR is) will obviously be red on the contour plot, which does not provide us with any more interesting insights. We can apply a similar logic to find the lower bound of the vertical range: the lower bound should be the minimum number of COVID-19 cases for it to be the third leading cause of death because the region below this will be blue (since we are only interested in the top three leading causes of death). Therefore, for our particular contour plots, we can use equation (5) to calculate both the upper and lower vertical bound. To calculate the upper bound, we substitute the number of deaths due to the leading cause of mortality in each country into the numerator of the fraction on the right hand side of the equation and set CFR = 2% (which is the lower bound of the range that we chose for CFR). As for the lower bound, we substitute the number of deaths due to the third leading cause of mortality into the numerator of the fraction and set the case fatality rate to our upper bound, 10%. By doing this, we can observe the number of cases it takes for COVID-19 to be one of the top three causes of mortality for all values of CFR within our range of interest.
Model 2: Uncertainty analysis of COVID-19 mortality based on the CFR and ratio of true and reported cases

We extend the first model to take into account that the number of cases of COVID-19 is most likely underreported by introducing a correction factor, $\beta$. The correction factor, $\beta$, is a ratio between the true number of COVID-19 cases and the number of reported cases, which gives the relation

$$\text{True number of cases} = \beta \times \text{Number of reported cases}. \quad (6)$$

Therefore, when $\beta > 1$, the true number of cases is greater than the number of reported cases (meaning that the number of COVID-19 cases are underreported). Conversely, when $\beta < 1$, the true number of cases is less than the number of reported cases (meaning that the number of COVID-19 cases are overreported). Lastly, when $\beta = 1$, the number of reported cases is the same as the true number of cases (meaning that the number of COVID-19 cases are neither underreported nor overreported). Since it is widely agreed that the number of reported cases are underreported, the values of $\beta$ that are of interest to us would be greater than one. To estimate the potential number of deaths based on the true number of cases of COVID-19, we substitute equation (6) into equation (3) to acquire

$$\text{Theoretical number of COVID-19 deaths} = \frac{\beta \times \text{CFR} \times \text{Number of reported COVID-19 cases}}{100\%} \quad (7),$$

which is the formula we use for this uncertainty analysis model, shown in eFigure 3.

We turn to the literature to determine the range we should choose for $\beta$ in our model. Li et al.(9) used a networked dynamic metapopulation model and Bayesian inference and estimated that 86% of all infections in China were undocumented before the travel restriction was put in place on January 23, 2020. Since there is a limited number of articles that explicitly estimate the number of unreported cases at the time of writing, we also consulted studies that estimate the proportion of COVID-19 cases that show mild or no symptoms, since these cases would most likely go undetected. An article from Nature(10) states that there are estimates that suggest around 60% of people with COVID-19 show mild or no symptoms. On the other hand, Nishiura et al.(11) looked at data of the Japanese citizens who were evacuated from Wuhan, China in early February and were repeatedly tested and closely monitored; the authors estimated based on this data that the asymptomatic proportion is 30.8%. Lastly, Mizumoto et al.(12) looked at data collected from the COVID-19 outbreak on the Diamond Princess cruise ship and estimated that the asymptomatic proportion among all infected cases is 17.9%. However, this estimate is not representative for a general population because many, if not most, of the passengers on the cruise ship were elderly. With such varied estimates, we, therefore, decided to take a large range and vary $\beta$ from 1 to 8.

For this uncertainty analysis model, we also decided to choose a smaller range for the case fatality rate (0.1%–5%) compared to what was chosen for our first model. We reduced the range to provide better visualization of the contours as they are mostly situated where the CFR is small.
eTable 2: Data on the cumulative number of cases as of December 16, 2020 for Canada, China, Italy, Spain, the UK, and USA, published in the WHO Coronavirus Disease (COVID-19) Dashboard.

<table>
<thead>
<tr>
<th>Country</th>
<th>Date of first COVID-19 case reported to WHO</th>
<th>Number of reported cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>February 27, 2020</td>
<td>6927145</td>
</tr>
<tr>
<td>Canada</td>
<td>January 27, 2020</td>
<td>468862</td>
</tr>
<tr>
<td>China</td>
<td>January 4, 2020</td>
<td>95279</td>
</tr>
<tr>
<td>Italy</td>
<td>January 31, 2020</td>
<td>1870576</td>
</tr>
<tr>
<td>Russia</td>
<td>February 1, 2020</td>
<td>2734454</td>
</tr>
<tr>
<td>Spain</td>
<td>February 1, 2020</td>
<td>1762212</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>February 1, 2020</td>
<td>1888120</td>
</tr>
<tr>
<td>United States</td>
<td>January 23, 2020</td>
<td>16245376</td>
</tr>
</tbody>
</table>
eFigure 3: Comparison of COVID-19 deaths and the top three categories of expected deaths from non-COVID-19 causes in Italy, Spain, United Kingdom, Russia, China, Canada, United States, and Brazil, on the basis of the case fatality rate and the ratio of true and reported COVID-19 cases as of December 16, 2020.

COPD = Chronic obstructive pulmonary disease. Ranking of COVID-19 deaths relative to non-COVID-19 cause of death: Red = COVID-19 is the leading cause of death, Yellow = COVID-19 is the second leading cause of death, Cyan = COVID-19 is the third leading cause of death, Blue = COVID-19 has a lower ranking than third cause of death. The white contours represent the estimated number of deaths due to COVID-19 in 2020 for specific combinations of model parameters (where 1k = 1000). Calculations of the number of deaths due to COVID-19 and comparisons with other causes of death are described in the section describing uncertainty analysis model 2. To aid figure interpretation, we illustrate the case of Italy. The red region indicates the combination of model parameters where there are more theoretical deaths due to COVID-19 over the duration of the outbreak in Italy than the leading cause of mortality, which is ischemic heart disease, and thus, COVID-19 would be the leading cause of death based on the number of confirmed cases and the case fatality rate. Similarly, the yellow region indicates that COVID-19 would be the second leading cause of death in Italy, surpassing Alzheimer's disease, the cyan region indicates that COVID-19 would be the third leading cause of death in Italy, surpassing stroke, and lastly, the blue region indicates that COVID-19 would not be one of the top three leading causes of death in Italy. The boundary between the red and yellow regions represents the combination of model parameters where the number of number of deaths due to COVID-19 is equal to the number of deaths due to ischemic heart disease. Similarly, the boundary between the yellow and cyan regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to Alzheimer's disease, and the boundary between the cyan and blue regions is where the number of deaths due to COVID-19 is equal to the number of deaths due to lung cancer. In China, the entire plot is blue because the number of deaths due to COVID-19 is less than the expected number of deaths from the third leading cause of death in the country for all model parameters.
Now that we have determined the range that we want to vary the CFR and \( \beta \), we need to choose which data set we want to use for each country in order to be able to calculate the theoretical number of deaths—we decided to use the cumulative number of cases provided by the WHO published on the date of writing (December 16, 2020), given in eTable 2. eFigure 3 gives the uncertainty analyses on the basis of the case fatality rate and the ratio of the true to reported number of cases of the eight countries of interest.

In order to be able to compare the magnitude of mortality due to COVID-19 to the top three causes of death in each country fairly, we need to ensure that the number of deaths are taken over the same duration. Similar to the first model, we use the values that we estimated for the year 2020, which are presented in eTable 1 (details on how these values were estimated are provided in the previous section). We adjust these annual values to the appropriate time frame, which, in this case, is the duration, in days, that the coronavirus has been present in the country\(^4\), by using the following formula:

\[
\text{Number of deaths due to certain cause during certain time frame} = \frac{\text{Annual number of deaths due to same cause}}{365} \times \text{Number of days that COVID-19 has been present in country.}
\]

again assuming that the number of deaths are uniformly distributed throughout the year. Despite the fact that there is a time lag between the onset of symptoms and death, using this duration should be acceptable since the duration of people dying from COVID-19 is similar in length to the duration of the same group of people contracting the coronavirus; the time frame is just delayed. We use these corrected values to determine how each region is colored in the contour plots presented in eFigure 3. Although the first reported case for China occurred in 2019, we still used our estimated value for the year 2020 for the adjustment rather than taking a linear combination of the estimated values for the years 2019 and 2020 based on the proportion of days that the coronavirus has been in the country each year. Not only is this highly unnecessary (as we are dealing with estimated values to begin with), but it actually makes no difference to the contour plot, since the maximum value of the theoretical number of deaths calculated using equation (7) is much smaller than the values given for the leading causes of death in China. Based on the reported number of cases provided by the World Health Organization (WHO) on December 16, 2020, our model (eFigure 3) predicts that the estimated number of deaths due to COVID-19 ranges from 1856 to 742295 for Italy, 1752 to 700754 for Spain, 1870 to 747868 for the United Kingdom, 2708 to 1083178 for Russia, 96 to 38111 for China, 469 to 187545 for Canada, 16246 to 6498151 for the United States, and 6928 to 2770858 for Brazil.

\(^4\) Note that we are counting the day that the WHO reports the first occurrence of the virus in the country as the first day rather than the day that the virus physically arrived in the country. (For e.g., although there are reports that COVID-19 first arrived in China on December 1, 2019, since the coronavirus was reported to the WHO on December 31, 2019, we will use the latter date as the first day.)
**eFigure 4: Comparison of COVID-19 deaths and expected deaths from non-COVID-19 causes in the United States over different date ranges.**

Since the spread of COVID-19 is exponential, this means that the number of daily cases grows over a certain period of time after the virus arrives to the country. Therefore, eFigure 3 does not necessarily give us a true sense of where COVID-19 is ranked in leading causes of mortality, especially at the present. For example, in eFigure 4, we compare the uncertainty analysis model for the United States over different date ranges: the plot on the left takes the entire duration that COVID-19 has been present in the United States, the plot second to the left takes the number of cases confirmed in the first 30 days of the arrival of the virus, and the other two plots take the duration of 60 days approximately when the first and second wave of the pandemic occurred, respectively. This suggests that limiting the date range over a certain period of time can provide a better depiction of the severity of COVID-19 at a particular time.

- Code to reproduce these results are publicly available at the following link: [https://github.com/MEDICI-UWO/COVID-19-uncertainty-analysis](https://github.com/MEDICI-UWO/COVID-19-uncertainty-analysis)
- R Shiny Application is available at the following link (updated daily): [https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/](https://medicimagic.shinyapps.io/COVID-19_uncertainty_analysis/)
References


