

# UNDERSTANDING INTER-REGIONAL DIFFERENCES IN COVID-19 MORTALITY RATES

DIFFERENTIAL  
CAPITAL



PANDA  
PANDEMICS - DATA & ANALYTICS

9 June 2020

# UNDERSTANDING INTER-REGIONAL DIFFERENCES IN COVID-19 MORTALITY RATES

In this note, we present a model of deaths due to COVID-19 disease and attempt to explain differences between geographic regions using a few critical factors that are shaping policy decisions and scientific debates. Some of our key findings are:

- There are material differences in the COVID-19 mortality rates between countries and regions within the same country (the United Kingdom)
- We were unable to find any factors that explain a large proportion of the variation in peak mortality rates
- In particular, the average age, prevalence of hospital beds and changes in mobility – all of which are assumed to be critical for policy decisions – are not statistically significant determinants of the mortality trajectory
- We find strong evidence for spatial clustering of peak mortality rates

The results lead us to question the validity of a phased approach to relaxing lockdowns. According to most governments, an initial strict lockdown was necessary as a temporary means to curb the spread of the SARS-CoV-2 virus and prepare healthcare facilities. It remains debatable as to how these initial restrictions should be relaxed. Our finding that differences in mobility do not explain the large variability in mortality rates suggests that a gradual lifting of lockdown restrictions will not necessarily mitigate COVID-19-related deaths.

Given the spatial clustering of peak mortality and the fact that none of our variables are able to adequately explain the differences in observed mortality, it remains important to identify the relevant drivers of COVID-19 mortality in order to make effective policy decisions and save lives.

## DETAILED FINDINGS

### Country-Level Analysis

We began by modelling the daily COVID-19 death trajectories in 54 countries. Countries were included in the analysis if they exceeded 0.1 deaths per million for at least 50 days as at 21 May 2020. To align the curves, time was measured as the number of days since 0.1 deaths per million.

The variation in the maximum cumulative death rate between countries is extraordinary, ranging from 0.8 deaths per million in Thailand through to

789.5 deaths per million in Belgium (the mortality curves for both of these countries appear to be nearing their plateaus). The median deaths per million in our sample was 33.4 with an interquartile range of 11.2 to 94.7 (see Figure 1).

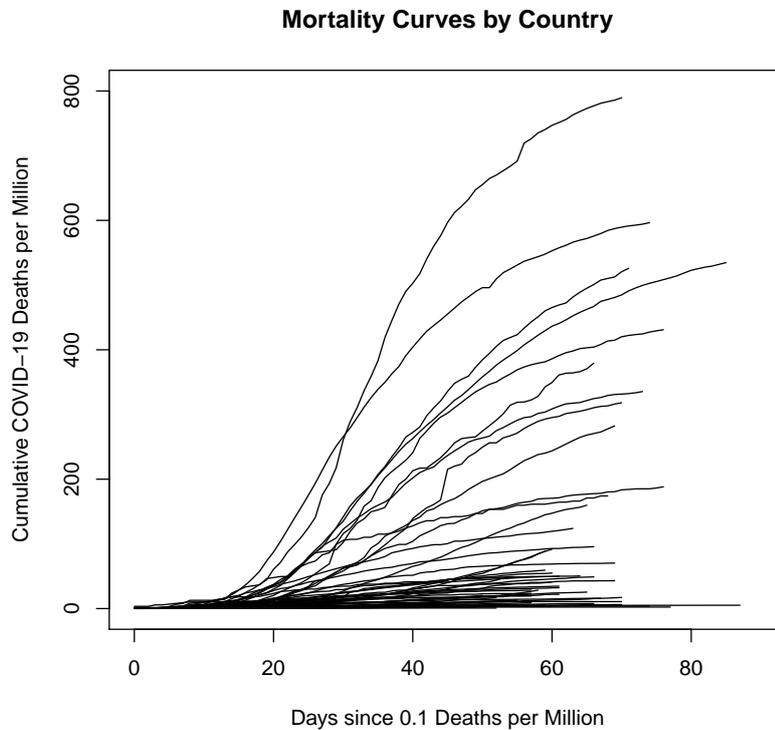


Figure 1: Observed mortality curves for each country.

We considered real GDP, hospital beds per 100,000 people, the average obesity rate, the Oxford lockdown stringency index, population density and a measure of density clustering as potential predictors. Of these, only GDP appeared to improve the fit of the model substantively, explaining 36% of the variation in growth rates between countries. In particular, we found that the mortality curves tend to reach their plateaus faster in wealthier countries than in poorer ones. Importantly though, none of the covariates we considered explained a significant proportion of the variation in the peak deaths per million. The fits of the best model are shown in Figure 2.

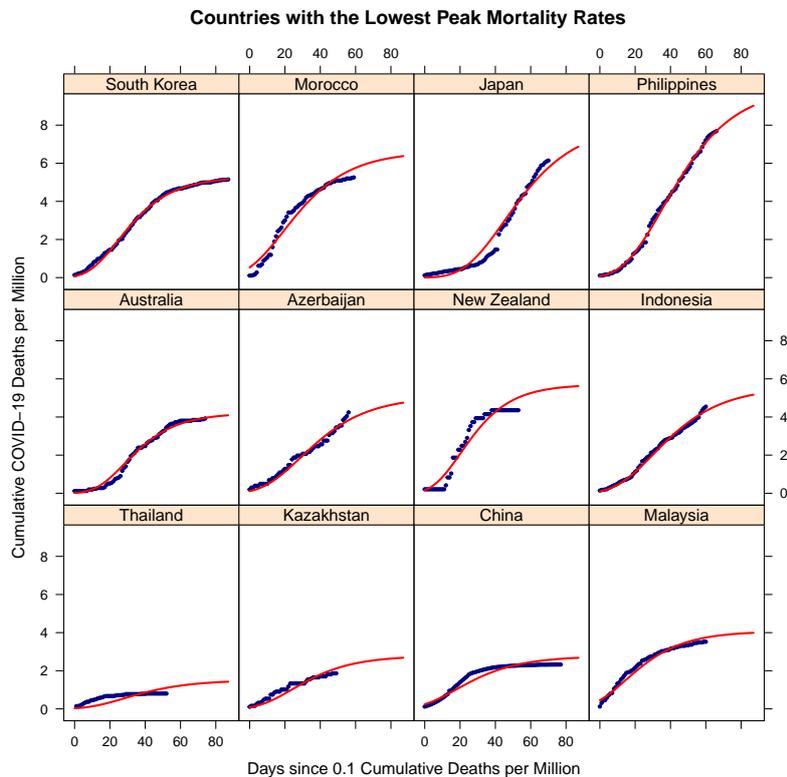
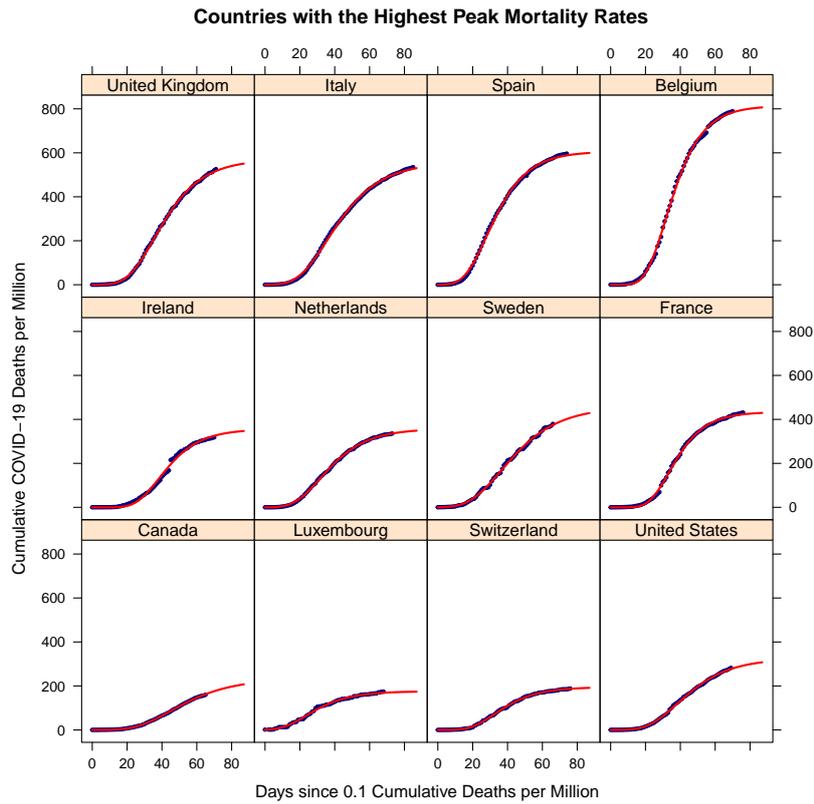


Figure 2: Observed (navy dots) and predicted (red lines) mortality curves for countries with the highest (top figure) and lowest (bottom figure) peak mortality rates in our sample.

## UK Analysis

To control for differences between countries (for example, the quality of health care and the stringency of lockdown measures), we also analysed the weekly COVID-19 mortality data for 292 local authorities in England and Wales. Local authorities, which we will henceforth refer to as counties, were included in the analysis if they had exceeded 0.1 cumulative deaths per million for at least 8 weeks as at 20 May 2020.

Again, we observed a huge amount of variation in peak deaths, ranging from 75.5 deaths per million in Hastings to 1601.2 in Hertsmere (see Figure 3). While we did expect more variation due to the smaller population sizes within each UK county, we thought that eliminating country-level differences would at least offset this.

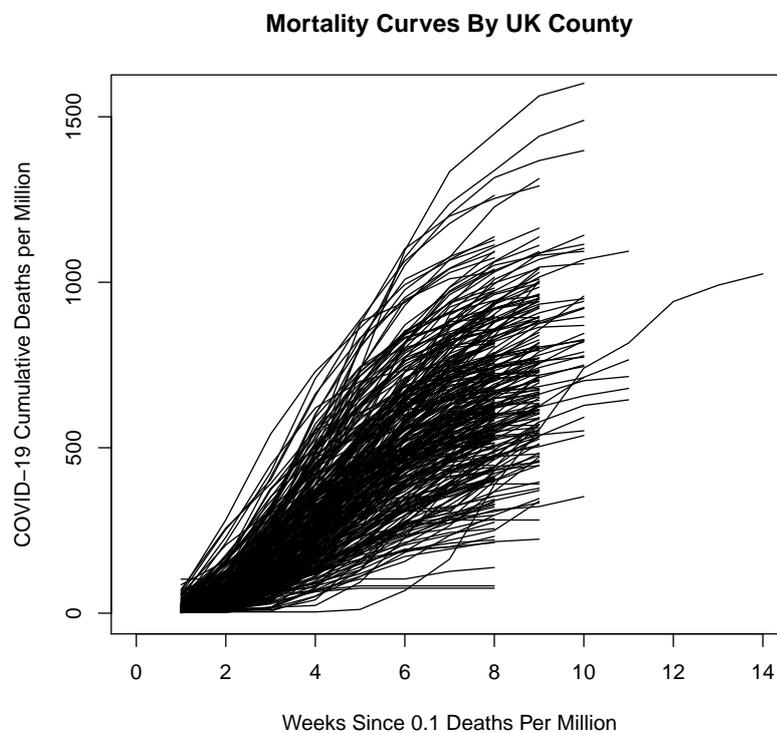


Figure 3: Observed mortality curves for UK counties.

We attempted to explain the observed variation with the following set of county-level covariates: average age, income, population density, number of people per hospital, number of surrounding counties within a fixed radius and the week number when each county first observed  $\geq 0.1$  deaths per million (did this happen earlier or later in the pandemic?). We also included Google's community mobility data, which tracks changes in visits to places

such as grocery stores and parks. These data were also available at county level.

We found that only the number of surrounding counties and the week number when deaths per million first exceeded 0.1 were significant. In particular, the growth rate of cumulative deaths was significantly higher for counties that had a larger number of nearby counties, explaining 27.7% of this variation. For counties that passed the 0.1 deaths per million threshold later in the pandemic, the mortality curves rise up sooner and to significantly lower asymptotes. Including the week number of  $\geq 0.1$  deaths per million explains 62.6% of the variation in the placement of the curve along the time axis, but only 8% of the variation in peak deaths between counties. The observed and predicted mortality curves are shown in Figure 4.

It is interesting that differences in mobility between UK counties did not explain the observed variation in mortality rates. The average change in workplace mobility from mid-March to the end of May was  $-56\%$ , ranging from  $-71\%$  to  $-45\%$ . For transit stations, the range was even wider from  $-83\%$  to  $-30\%$ , with an average change of  $-55\%$ . Therefore, we can only say that for these observed ranges, there appears to be no impact on the mortality experience.

An example is the Hertfordshire area, where the Hertsmere district has the worst COVID-19 mortality rate in the UK, yet the mobility of residents was lower than average at  $-61\%$  and  $-67\%$  for workplaces and transit stations, respectively. In contrast, North East Lincolnshire had one of the lowest COVID-19 mortality rates in the UK, but above average workplace and transit station mobility changes of  $-49\%$  and  $-50\%$ , respectively.

Due to software limitations, the analyses above did not account for spatial correlation in mortality between UK counties or countries. Indeed, we found significant evidence of spatial clustering in the predicted peak deaths at both the country and UK county levels (Moran's  $I$ ,  $p$ -value  $\approx 0$  for both analyses). Looking at the UK heatmaps in Figure 5, there are clear clusters of neighbouring counties with high COVID-19 death rates, even within London. Interestingly, neighbouring counties can also have dramatically different peak death rates.

The London counties present quite an intriguing case study. This is because the population sizes are broadly similar, and the number of hospitals relative to the population in each county is also broadly similar. However, we note the extreme differences between the Harrow and Brent counties compared to the adjacent areas and the fact that the central counties have much lower mortality rates. A variable that would be of interest is the ethnic profiles of the various counties and we found that differences in the ethnic compositions of the counties also did not explain the differences in mortality experience in the case of London. We did not obtain the data for ethnicities for the rest of

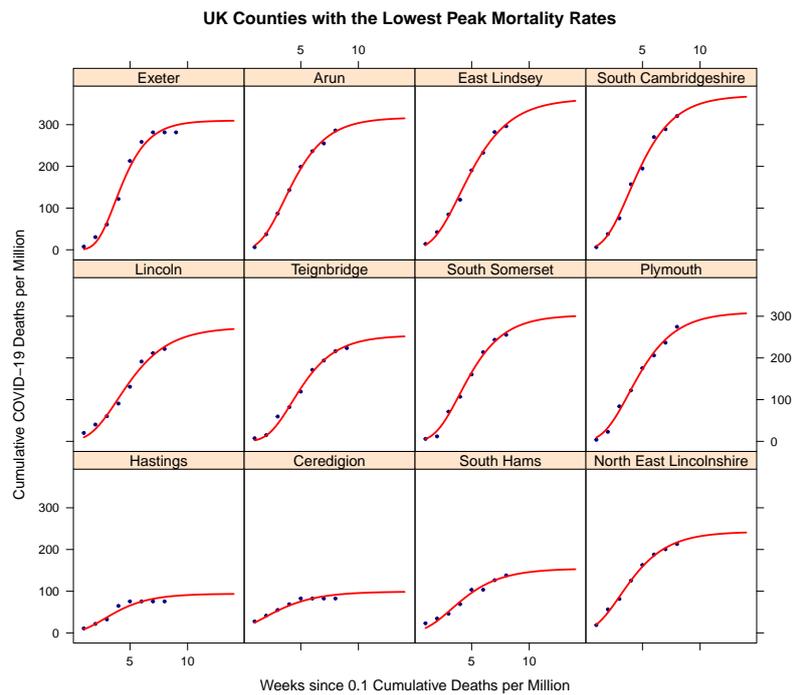
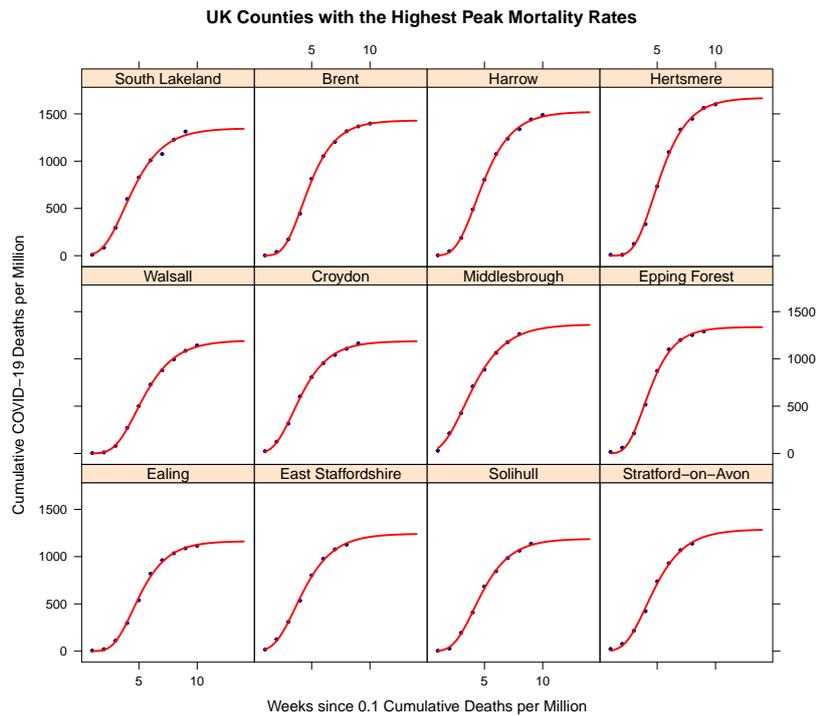


Figure 4: Observed (navy dots) and predicted (red lines) mortality curves for countries with the highest (top figure) and lowest (bottom figure) peak mortality rates in our sample.

England and Wales at a county level.

We also explored whether the spatial clustering of COVID-19 mortality that we observed at a country level might be explained by differences in the average air temperature. However, temperature too turned out to be an insignificant predictor of mortality.

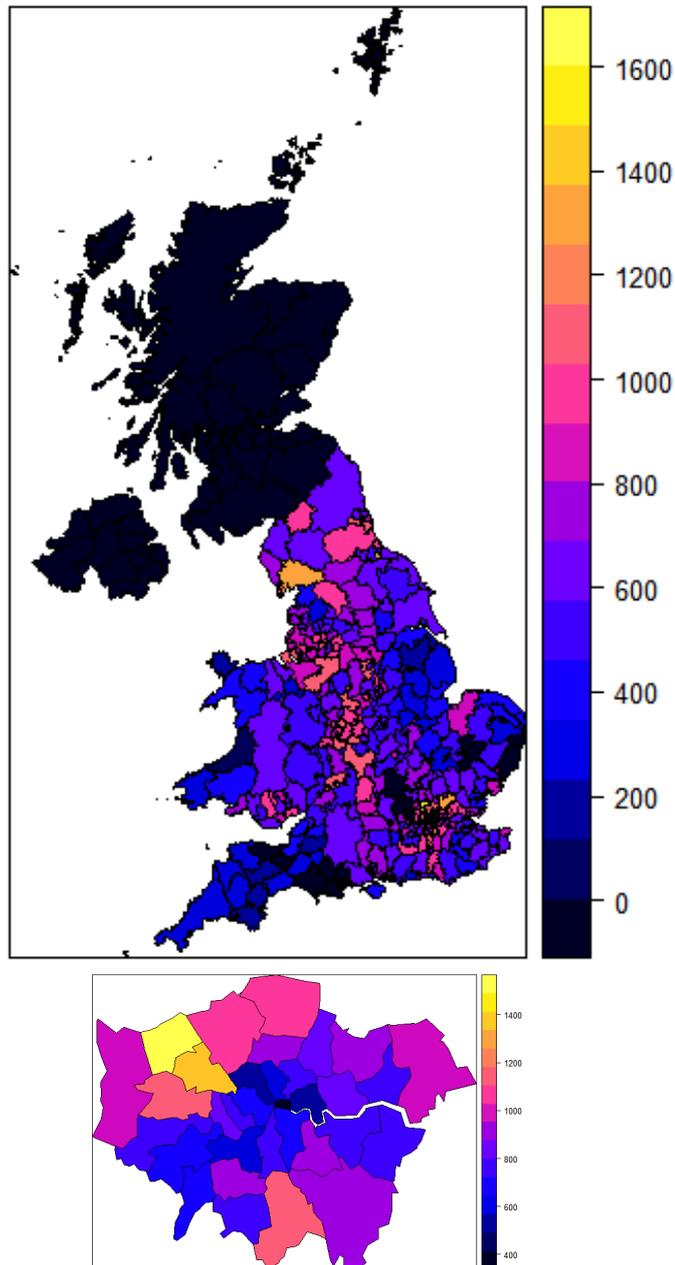


Figure 5: Peak deaths per million in England and Wales (top) and London (bottom). All counties were included, regardless of whether peak mortality was  $\geq 0.1$  deaths per million.

## DISCUSSION

### Are Phased Lockdowns Warranted?

We have shown that differences in mobility levels do not appear to explain the variability in the COVID-19 mortality experience of different regions. Moreover, in the UK analysis, we observed that counties with very low mobility can have extremely high peak mortality rates and vice versa. Changing the restrictions on mobility, at least within the ranges we have observed, is therefore unlikely to have the desired effect on mortality. This raises doubts as to the efficacy of a phased lockdown approach. Assuming that reducing mobility from normal levels does reduce mortality risk, it is also debatable as to whether such reductions in mobility should be achieved by coercion.

To be clear, we are not saying that lockdowns and social distancing do not work. However, we cannot argue that the phased adoption of these measures has any impact on risk mitigation. This is an important consideration for policy makers who must carefully balance the benefits of a phased lockdown strategy with the economic harm caused by such an intervention.

It is also worth noting that our findings cannot be used on their own to argue the case for no lockdowns. In the absence of a strict lockdown, the mortality rates of all regions may very well have been higher. Indeed, we have not studied regions where mobility remained normal during the COVID-19 pandemic. Without conducting such an analysis, we cannot comment on the merit of lockdowns compared to no lockdowns.

### Other Important Considerations

We acknowledge that there are many other important explanatory variables that we did not consider in our analyses. Of note is the prevalence of comorbidities, which are widely believed to increase mortality risk. It is interesting that hospital availability and the average age of the population – both of which are commonly assumed to influence mortality rates – are not able to explain mortality differences between regions in both our country-level and UK analyses. Indeed, it is clear from the case fatality ratios that age is an important factor, with older people within a region being disproportionately affected. We note, however, that the case fatality ratio in countries like Italy (high peak mortality) are materially higher at each age compared to countries like South Korea (low peak mortality) (see Figure 6).

We must also point out that the population of COVID-19 cases in South Korea was much younger than that of Italy. Therefore, an adjustment to the age distribution by referencing the infected population may yield more significant results when it comes to explaining the differing mortalities across regions. This would be a useful area for future research. Such an analysis

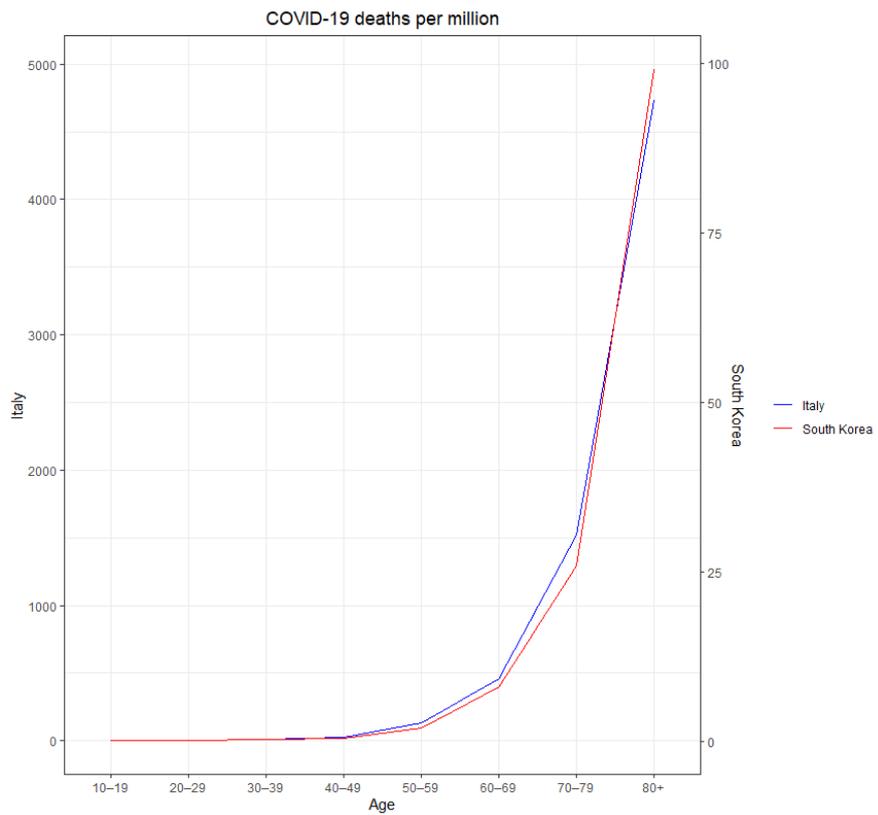
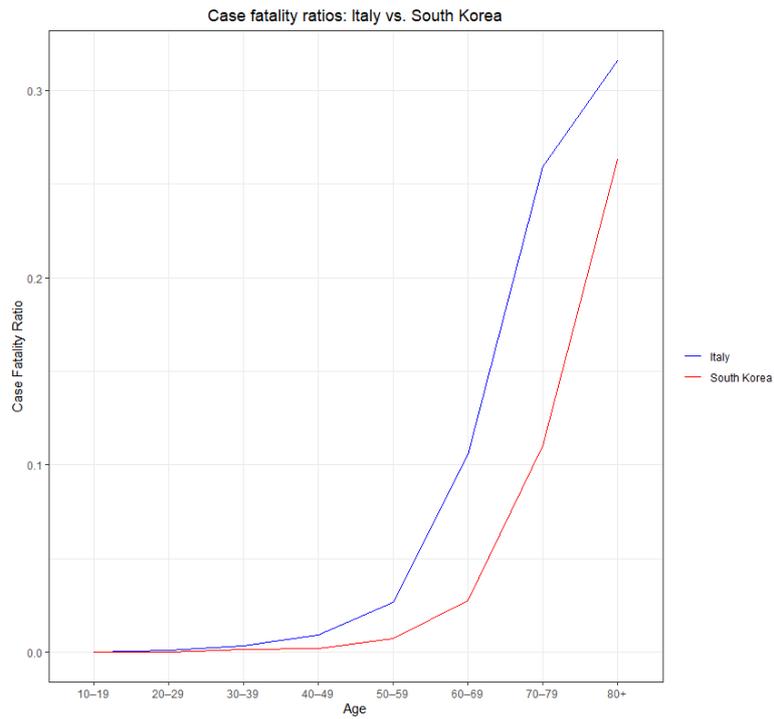


Figure 6: Age-specific case fatality ratios (top figure) and deaths per million (bottom figure) for Italy (blue) and South Korea (red).

would, however, need to take into account biases in the criteria for testing and differing test methods across populations. This is also why part of this investigation focused on counties within the UK, where we assume that the testing methodology and quality and access to healthcare is fairly consistent. Note, however, that adjusting the age distribution by referencing the age of the infected population would not explain why different populations have differing distributions of infection by age. This is also important to understand if the aim of a policy is to mitigate mortality risk.

One issue with disease-specific mortality data is that it is not always clear whether a person died *from* the disease or *with* the disease, and the guidelines for this assessment differ between regions and over time. For example, deaths that were previously assigned to strokes and heart attacks in individuals who later tested positive for COVID-19 are now being assigned to COVID-19 after it was discovered that the disease causes small blood clots that lead to these conditions. One potential solution to the inconsistent reporting of COVID-19-related deaths is to study the mortality in excess of what would be expected in a normal year. We hope to pursue this in our future research.

## APPENDIX: METHODOLOGY

We modelled the cumulative deaths per million  $y_i(t)$  in region  $i$  at time  $t$  with a Gompertz growth curve of the form

$$y_i(t) = A_i \exp\{-K_i e^{-R_i t}\} + \varepsilon_i(t),$$

where, for region  $i$ ,  $A_i > 0$  is the asymptote or peak deaths,  $K_i$  controls the displacement along the time axis, and  $R_i > 0$  controls the growth rate of deaths per million. The random departures from the Gompertz curve,  $\varepsilon_i(t)$ , are assumed to be independent and Gaussian with a constant variance  $\sigma^2$ . Rather than specify a separate set of parameters for each region, we adopt a mixed-effects approach and model each parameter as the sum of a fixed effect and a random effect drawn from a centred Gaussian distribution with covariance matrix  $\Sigma$ ; that is,

$$\begin{aligned} A_i &= A + a_i \\ K_i &= K + k_i \\ R_i &= R + r_i \end{aligned}$$

and

$$\begin{pmatrix} a_i \\ k_i \\ r_i \end{pmatrix} \sim \mathcal{N}(\vec{0}, \Sigma).$$

The advantage of this formulation is that it only has 10 parameters (the fixed effects  $A$ ,  $K$  and  $R$ , and the variance and covariance parameters  $\sigma^2$  and  $\Sigma$ ) and these do not increase with the number of regions.

In this model, the regional variation in peak deaths, for example, is captured by the  $\Sigma_{11}$  entry in the random effect covariance matrix. Our objective is to determine if any covariates can explain this variation. We test this by adding covariates  $X_1, \dots, X_p$  to the parameter equation

$$A_i = A + \beta_1 X_1 + \dots + \beta_p X_p + a_i$$

and determining whether there is a significant reduction in the variance of  $a_i$  relative to that of the model without covariates. This can be done for any or all of the three fixed effect parameters. We used the Bayesian information criterion (BIC) to determine which covariates offer a significant improvement in model fit.