



RESEARCH ARTICLE

REVISED **Lessons learned and lessons missed: impact of the coronavirus disease 2019 (COVID-19) pandemic on all-cause mortality in 40 industrialised countries and US states prior to mass vaccination [version 2; peer review: 2 approved]**

Vasilis Kontis ¹, James E. Bennett¹, Robbie M. Parks ^{2,3}, Theo Rashid¹, Jonathan Pearson-Stuttard¹, Perviz Asaria¹, Bin Zhou¹, Michel Guillot^{4,5}, Colin D. Mathers⁶, Young-Ho Khang⁷, Martin McKee ⁸, Majid Ezzati ^{1,9}

¹MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK

²The Earth Institute, Columbia University, New York, NY, USA

³Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA

⁴Population Studies Center, Department of Sociology, University of Pennsylvania, Philadelphia, PA, USA

⁵French Institute for Demographic Studies (INED), Paris, France

⁶Independent Researcher, Geneva, Switzerland

⁷Institute of Health Policy and Management, Seoul National University, Seoul, South Korea

⁸Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK

⁹Regional Institute for Population Studies, University of Ghana, Legon, Ghana

V2 **First published:** 18 Oct 2021, 6:279
<https://doi.org/10.12688/wellcomeopenres.17253.1>
Latest published: 15 Feb 2022, 6:279
<https://doi.org/10.12688/wellcomeopenres.17253.2>

Abstract

Background: Industrialised countries had varied responses to the COVID-19 pandemic, which may lead to different death tolls from COVID-19 and other diseases.

Methods: We applied an ensemble of 16 Bayesian probabilistic models to vital statistics data to estimate the number of weekly deaths if the pandemic had not occurred for 40 industrialised countries and US states from mid-February 2020 through mid-February 2021. We subtracted these estimates from the actual number of deaths to calculate the impacts of the pandemic on all-cause mortality.

Results: Over this year, there were 1,410,300 (95% credible interval 1,267,600-1,579,200) excess deaths in these countries, equivalent to a 15% (14-17) increase, and 141 (127-158) additional deaths per 100,000 people. In Iceland, Australia and New Zealand, mortality was lower than would be expected in the absence of the pandemic, while South Korea and Norway experienced no detectable change. The USA, Czechia, Slovakia and Poland experienced >20% higher mortality. Within the USA, Hawaii experienced no detectable change in mortality and Maine a 5% increase, contrasting with New Jersey, Arizona,

Open Peer Review

Approval Status

	1	2
version 2 (revision) 15 Feb 2022	 view	 view
version 1 18 Oct 2021	 view	 view

- Rajeev Gupta** , Rajasthan University of Health Sciences, Jaipur, India
Eternal Heart Care Centre and Research Institute, Jaipur, India
- Virgilio Gómez-Rubio** , Universidad de Castilla-La Mancha, Ciudad Real, Spain

Any reports and responses or comments on the

Mississippi, Texas, California, Louisiana and New York which experienced >25% higher mortality. Mid-February to the end of May 2020 accounted for over half of excess deaths in Scotland, Spain, England and Wales, Canada, Sweden, Belgium, the Netherlands and Cyprus, whereas mid-September 2020 to mid-February 2021 accounted for >90% of excess deaths in Bulgaria, Croatia, Czechia, Hungary, Latvia, Montenegro, Poland, Slovakia and Slovenia. In USA, excess deaths in the northeast were driven mainly by the first wave, in southern and southwestern states by the summer wave, and in the northern plains by the post-September period.

Conclusions: Prior to widespread vaccine-acquired immunity, minimising the overall death toll of the pandemic requires policies and non-pharmaceutical interventions that delay and reduce infections, effective treatments for infected patients, and mechanisms to continue routine health care.

Keywords

Excess mortality, Covid-19, SARS-CoV-2, Bayesian ensemble modelling, autoregressive models, uncertainty.



This article is included in the [Coronavirus \(COVID-19\)](#) collection.

article can be found at the end of the article.

Corresponding author: Majid Ezzati (majid.ezzati@imperial.ac.uk)

Author roles: **Kontis V:** Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation; **Bennett JE:** Conceptualization, Data Curation, Methodology, Visualization; **Parks RM:** Data Curation, Methodology; **Rashid T:** Data Curation, Methodology; **Pearson-Stuttard J:** Writing – Review & Editing; **Asaria P:** Writing – Review & Editing; **Zhou B:** Data Curation, Writing – Review & Editing; **Guillot M:** Methodology, Writing – Review & Editing; **Mathers CD:** Data Curation, Writing – Review & Editing; **Khang YH:** Data Curation, Writing – Review & Editing; **McKee M:** Writing – Review & Editing; **Ezzati M:** Conceptualization, Funding Acquisition, Methodology, Supervision, Visualization, Writing – Original Draft Preparation

Competing interests: ME reports a charitable grant from the AstraZeneca Young Health Programme, outside the submitted work. JP-S is vice-chair of the Royal Society for Public Health and reports personal fees from Novo Nordisk A/S and Lane, Clark & Peacock LLP, outside of the submitted work.

Grant information: The development of methodology for estimating the impact of pandemic as an extreme event was supported by the Wellcome Trust [209376; Pathways to Equitable Healthy Cities]. Work on the US mortality data was partially supported by a grant from the US Environmental Protection Agency (EPA), as part of the Center for Clean Air Climate Solution [assistance agreement no. R835873]. This article has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the EPA. The EPA does not endorse any products or commercial services mentioned in this publication. Work on the UK mortality data was partially supported by the British Heart Foundation [RE/18/4/34215; Centre of Research Excellence grant]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Kontis V *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kontis V, Bennett JE, Parks RM *et al.* **Lessons learned and lessons missed: impact of the coronavirus disease 2019 (COVID-19) pandemic on all-cause mortality in 40 industrialised countries and US states prior to mass vaccination [version 2; peer review: 2 approved]** Wellcome Open Research 2022, 6:279 <https://doi.org/10.12688/wellcomeopenres.17253.2>

First published: 18 Oct 2021, 6:279 <https://doi.org/10.12688/wellcomeopenres.17253.1>

REVISED Amendments from Version 1

Based on the helpful comments from reviewers, we have added results on US states in the abstract, reorganised the text, included additional citations, and done a sensitivity analysis on how different models are weighted. We have also updated the section *Comparison with other estimates* to include the most recently published results from other sources.

Any further responses from the reviewers can be found at the end of the article

Introduction

Many industrialised countries experienced a rise in all-cause mortality in the first wave of the coronavirus disease 2019 (COVID-19) pandemic, while others avoided any excess deaths¹. These excess deaths were due to infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), delays and disruptions in the provision and use of healthcare for other diseases, loss of jobs and income, disruptions of social networks and support, and changes in nutrition, drug and alcohol use, transportation, crime, and violence^{2,3}.

Decline in infections following initial lockdowns and other restrictions, and advances in knowledge about the SARS-CoV-2 transmission and infection, presented a window of opportunity for countries to implement pandemic control measures and strengthen health and social care provision that would minimise the impacts of subsequent waves⁴⁻⁸. Comparative analysis of excess deaths prior to mass vaccination against COVID-19 helps understand how effectively these measures were implemented and how resilient the health and social care system was in each country. We quantified the weekly mortality impacts of the first year of the COVID-19 pandemic, from mid-February 2020 to mid-February 2021, in 40 industrialised countries, listed below. We used this period because mortality due to the pandemic was negligible before mid-February 2020¹, and vaccination rates against SARS-CoV-2 were still relatively low before mid-February 2021 in these countries (no more than 4% of the population had received both doses in any of these countries, as per [Our World in Data](#)⁹). After mid-February 2021, the effect of vaccines on mortality was expected to appear in some countries, which should be subject to a distinct analysis.

Methods**Data sources**

We included industrialised countries with complete or near-complete registration of deaths in our analysis if:

- Their total population in 2020 was more than 100,000. We excluded countries (e.g., Liechtenstein) with data but with smaller populations because, in many weeks, the number of deaths would be small or zero. This would, in turn, lead to either large uncertainty that would make it hard to differentiate between those places with and without an effect or unstable estimates because the model is fitted to many weeks with zero deaths.

- We could access up-to-date weekly data on all-cause mortality divided by age group and/or sex that extended through February 2021.
- The time series of data went back at least to the beginning of 2016 so that model parameters could be reliably estimated. For countries with longer time series, we used data starting in 2010.

The 40 countries in our analysis were divided into five geographical regions: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden). In addition to national estimates, we separately estimated excess deaths for all 50 US states and the District of Columbia. Some US states are larger than most countries included in our analysis (e.g., California's population of ~40 million is larger than those of 33 of the countries in the analysis), and the extent and temporal dynamics of the pandemic were heterogeneous across states due to their relative autonomy in policy formulation and implementation.

The sources of population and mortality data are provided in [Table 1](#). We calculated weekly population through interpolation of yearly population, consistent with the approach taken by national statistical offices for intra-annual population calculation¹⁰. Population for 2020 and 2021, where not available, was obtained through linear extrapolation from the last five years. We obtained data on temperature from ERA5¹¹, which uses data from global *in situ* and satellite measurements to generate a worldwide meteorological dataset, with full space and time coverage over our analysis period. We used gridded temperature estimates measured four times daily at a resolution of 30 km to generate weekly temperatures for each first-level administrative region, and [gridded population data](#) to generate population estimates by first-level administrative region in each country. We weighted weekly temperature by population of each first-level administrative region to create national level weekly temperature summaries.

Statistical methods

We used a probabilistic model averaging approach to estimate what deaths were expected to be between mid-February 2020 and mid-February 2021 had the pandemic not occurred, and compared these estimates with actual deaths from all causes in each country. The analytical method was designed to enhance comparison across countries and over time, and account for medium-long-term secular trends in mortality, the potential dependency of death rates in each week on those in preceding week(s) and in each year on those in preceding year(s), and factors that affect mortality including seasonality, temperature and public holidays.

Table 1. Sources of data on deaths and population.

	Data sources for deaths and population	Start of time series	Sex-specific analysis (see Methods for details)	Analysis age groups (see Methods for details)
Australia	ABS ¹ , UN ²	29/12/2014	Y	0-44, 45-64, 65+
Austria	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Belgium	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Bulgaria	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Canada	StatCan ^{5,6}	09/01/2010	Y	0-44, 45-64, 65+
Chile	MINSAL ⁷ , UN ²	01/01/2016	Y	0-44, 45-64, 65+
Croatia	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Cyprus	Eurostat ^{3,4}	29/12/2014	Y	0-64, 65+
Czechia	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Denmark	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
England and Wales ⁸	ONS ^{9,10}	02/01/2010	Y	0-44, 45-64, 65+
Estonia	Eurostat ^{3,4}	04/01/2010	Y	0-64, 65+
Finland	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
France	Eurostat ^{3,4}	31/12/2012	Y	0-44, 45-64, 65+
Germany	Destatis ¹¹ , Eurostat ⁴	04/01/2016	Y	0-44, 45-64, 65+
Greece	Eurostat ^{3,4}	29/12/2014	Y	0-44, 45-64, 65+
Hungary	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Iceland	Eurostat ^{3,4}	04/01/2010	N	All ages
Italy	Eurostat ^{3,4}	03/01/2011	Y	0-44, 45-64, 65+
Latvia	Eurostat ^{3,4}	04/01/2010	Y	0-64, 65+
Lithuania	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Luxembourg	Eurostat ^{3,4}	04/01/2010	Y	0-64, 65+
Malta	Eurostat ^{3,4}	03/01/2011	N	All ages
Montenegro	Eurostat ^{3,4}	04/01/2010	Y	0-64, 65+
Netherlands	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
New Zealand	Stats NZ ¹² , UN ²	02/01/2011	Y	0-64, 65+
Northern Ireland ⁸	NISRA ¹³ , Eurostat ⁴	01/01/2011	N	0-64, 65+
Norway	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Poland	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Portugal	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Romania	Eurostat ^{3,4}	29/12/2014	Y	0-44, 45-64, 65+
Scotland ⁸	NRS ¹⁴ , ONS ¹⁰	04/01/2010	Y	0-44, 45-64, 65+
Serbia	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Slovakia	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+

	Data sources for deaths and population	Start of time series	Sex-specific analysis (see Methods for details)	Analysis age groups (see Methods for details)
Slovenia	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
South Korea	KOSIS ¹⁵	03/01/2010	N	0-64, 65+
Spain	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Sweden	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
Switzerland	Eurostat ^{3,4}	04/01/2010	Y	0-44, 45-64, 65+
USA	CDC ^{16,17}	04/01/2015	N	0-44, 45-64, 65+ ¹⁸

¹ <https://www.abs.gov.au/statistics/health/causes-death/provisional-mortality-statistics/latest-release>

² <https://population.un.org/wpp>

³ <https://ec.europa.eu/eurostat/data/database> (table demo_r_mwk_05). Deaths with unknown age (0.03% of all deaths) were distributed across age groups proportional to the overall distribution of deaths for each year and month.

⁴ <https://ec.europa.eu/eurostat/data/database> (table demo_pjangroup)

⁵ <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310076801>. Death counts rounded to a neighbouring multiple of 5. There were no data for Yukon from 2017 to 2021 (before 2017, there were <10 deaths per week in Yukon).

⁶ <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501>

⁷ <https://deis.minsal.cl/#datosabiertos>. Deaths with unknown age and/or sex (0.02% of all deaths) were distributed across age groups and sexes proportional to the overall distribution of deaths for each year and month.

⁸ Data for the constituent nations in the UK are provided separately by NISRA for Northern Ireland, NRS for Scotland and ONS for England and Wales. These datasets use different reporting week definitions and could therefore not be combined into a single time series for the UK.

⁹ <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales>

¹⁰ <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>

¹¹ <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaeelle-Lebenserwartung/Tabellen/sonderauswertung-sterbefaeelle.html>

¹² <https://www.stats.govt.nz/experimental/covid-19-data-portal>

¹³ <https://www.nisra.gov.uk/publications/historical-weekly-deaths-data> and <https://www.nisra.gov.uk/publications/weekly-death-statistics-northern-ireland-2021>

¹⁴ <https://www.nrsotland.gov.uk/statistics-and-data/statistics-by-theme/vital-events/general-publications/weekly-and-monthly-data-on-births-and-deaths/deaths-involving-coronavirus-covid-19-in-scotland/related-statistics>

¹⁵ https://kosis.kr/covid_eng/statistics_excessdeath.do and <https://mdis.kostat.go.kr/index.do>

¹⁶ <https://data.cdc.gov/NCHS/Weekly-counts-of-deaths-by-jurisdiction-and-age-gr/y5bj-9g5w>. We used deaths adjusted for completeness by the CDC which account for potential underreporting in the most recent weeks. The adjustment methods are described at https://www.cdc.gov/nchs/nvss/vsrr/covid19/tech_notes.htm.

¹⁷ https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm

¹⁸ When analysing individual states, we merged 0–44 and 45–64 age groups into a single age group 0–64 years for Alaska, Delaware, DC, Hawaii, Idaho, Maine, Montana, Nebraska, New Hampshire, North Dakota, Rhode Island, South Dakota, Vermont West Virginia and Wyoming for reasons described in Methods.

The total mortality impact of the COVID-19 pandemic is the difference between the observed number of deaths from all causes of death and the number of deaths had the pandemic not occurred, which is not directly measurable. The most common approach to calculating the number of deaths had the pandemic not occurred has been to use the average number of deaths over previous years, e.g., the most recent five years, for the corresponding week or month when the comparison is made. This approach however does not take into account long- and short-term trends in mortality or time-varying factors like temperature that are largely external to the pandemic, but also affect death rates.

We developed an ensemble of 16 Bayesian mortality projection models that each make an estimate of weekly death rates that would have been expected if the COVID-19 pandemic had not occurred¹². We used multiple models because there is inherent uncertainty in the choice of model that best predicts death rates in the absence of pandemic. These models were formulated to incorporate features of weekly death rates, and how they

behave in the short-term (week to week) and medium-term (year to year), as follows:

- First, death rates may have a medium-to-long-term trend¹³ that would lead to a lower or higher mortality in 2020–2021 compared to earlier years. Therefore, all models included a linear trend term over weekly death rates.
- Second, death rates have a seasonal pattern^{14–17}. We included weekly random intercepts for each week of the year. To account for the fact that seasonal patterns “repeat” (i.e., late December and early January are seasonally similar) we used a seasonal structure^{18,19} for the random intercepts. The seasonal structure allows the magnitude of the random intercepts to vary over time, and implicitly incorporates time-varying factors such as annual fluctuations in flu season.
- Third, death rates in each week may be related to rates in preceding week(s), due to short-term phenomena such

as severity of the flu season. We formulated four sets of models to account for this relationship. The weekly random intercepts in these models had a first, second, fourth or eighth order autoregressive structure^{18,19}. The higher-order autoregressive models allow death rates in any week to be informed by those in a progressively larger number of preceding weeks. Further, trends not picked up by the linear or seasonal terms would be captured by these autoregressive terms.

- Fourth and additionally, mortality in one year may depend on mortality in the previous year, in a different way for each month, because phenomena such as seasonal flu may lead to longer-term dependencies in mortality. To allow for this possibility, we used two sets of models, with and without a first order autoregressive term over years for each month.
- Fifth, beyond having a seasonal pattern, death rates depend on temperature, and specifically on whether temperature is higher or lower than its long-term norm during a particular time of year²⁰⁻²⁵. The effect of temperature on mortality varies throughout the year, and may be in opposite directions for different times of year. We used two sets of models, one without temperature and one with a weekly term for temperature anomaly, defined as deviation of weekly temperature from the local average weekly temperature over the entire analysis period.
- Finally, death rates may be different around major holidays such as Christmas and New Year either because of changes in human activities and behaviour or, for the countries whose data are registration based, because of delays in registration. We included effects (as fixed intercepts) for the weeks containing Christmas and New Year in all countries. For England and Wales, Scotland and Northern Ireland, we also included effects for the week containing other public holidays, because reported death rates in weeks that contain a holiday were different from other weeks. This term was tested but not included for other countries because the effect was negligible.

These choices led to an ensemble of 16 Bayesian models (2 yearly autoregressive options \times 4 weekly autoregressive options \times 2 temperature anomaly options). The ensemble of models is shown in Table 2. In each model, the number of weekly deaths follows a Poisson distribution:

$$\text{deaths}_{\text{week}} \sim \text{Poisson}(\text{death rate}_{\text{week}} \cdot \text{population}_{\text{week}}).$$

Log-transformed death rates were modelled as a sum of components described above:

$$\log(\text{death rate}_{\text{week}}) = \alpha_0 + \alpha_{\text{holiday(week)}} + \beta \cdot \text{week} + \zeta_{\text{week}}^{(i)} + \eta_{\text{year}}^{\text{month}} + \theta_{\text{week}} + (\gamma + \nu_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}} + \varepsilon_{\text{week}}$$

The term α_0 denotes the overall intercept and $\alpha_{\text{holiday(week)}}$ is the holiday intercept, applied to weeks with a holiday. For example, if a week includes the 25th of December then $\alpha_{\text{holiday(week)}} = \alpha_{\text{Christmas}}$. For weeks that did not contain a holiday, this term did not appear in the above expression. All intercepts were assigned $\mathcal{N}(0,1000)$ priors. The term $\beta \cdot \text{week}$ represents the linear time trend. The coefficient β was also assigned a $\mathcal{N}(0,1000)$ prior.

The models used different orders (first, second, fourth or eighth) of the autoregressive term $\zeta_{\text{week}}^{(i)}$ with the superscript i denoting the order for weekly mortality patterns. The first-order autoregressive term is defined as $\zeta_{\text{week}}^{(1)} \sim \mathcal{N}(\phi \cdot \zeta_{\text{week}-1}^{(1)}, \sigma_{\zeta}^2)$ where the parameter ϕ lies between -1 and 1 and captures the degree of association between the number of deaths in each week and the preceding week. Hyperpriors are placed on the parameters $\kappa_1 = \log((1-\phi^2)/\sigma_{\zeta}^2)$ and $\kappa_2 = \log((1+\phi)/(1-\phi))$ which were assigned $\text{logGamma}(0.001, 0.001)$ and $\mathcal{N}(0, 1)$ distributions respectively. Similarly, an i^{th} order autoregressive term is given by $\zeta_{\text{week}}^{(i)} = \phi_1 \cdot \zeta_{\text{week}-1}^{(i)} + \dots + \phi_i \cdot \zeta_{\text{week}-i}^{(i)} + \varepsilon_{\text{week}}$ with $-1 < \phi_j < 1$. The parametrisation of these models was based on the partial autocorrelation function of the sequence ϕ_j ²⁶.

The term $\eta_{\text{year}}^{\text{month}}$ is an autoregressive term of order 1 over years and independent across months, indexed to the month and year to which each particular week belongs. For each month, the autoregressive prior for $\eta_{\text{year}}^{\text{month}}$ was the same as that for $\zeta_{\text{week}}^{(1)}$ described above. As described above, this term appeared in half of our models.

The term θ_{week} captures seasonality in mortality trends with a period of 52 weeks. The sums of every 52 consecutive terms $\theta_{\text{week}} + \theta_{\text{week}+1} + \dots + \theta_{\text{week}+51}$ were modelled as independent Gaussian with zero mean and variance σ_{θ}^2 . We used a $\text{logGamma}(0.001, 0.001)$ prior on the log precision $\log(1/\sigma_{\theta}^2)$. Each week is assigned an index between 1 and 52 depending on which week of the current year it is (the incomplete week 53 is mapped to either index 1 or 52 depending on whether it has greater overlap with week 52 of the current year or week 1 of the following year).

The effect of temperature anomaly on death rates is captured by the two terms γ and $\nu_{\text{week of year}}$. The term $\gamma \cdot \text{temperature anomaly}_{\text{week}}$ is the overall association between (log-transformed) death rates and temperature anomaly in a week. The term $\nu_{\text{week of year}} \cdot \text{temperature anomaly}_{\text{week}}$ captures deviations from the overall association for each week of the year. It consists of 52 terms with an independent and identically distributed prior defined via $\nu_{\text{week of year}} \sim \mathcal{N}(0, \sigma_{\nu}^2)$, and log-precision $\log(1/\sigma_{\nu}^2) \sim \text{logGamma}(0.001, 0.001)$.

Finally, the term $\varepsilon_{\text{week}}$ is a zero-mean term that accounts for additional variability. It is assigned an independent and identically distributed prior $\varepsilon_{\text{week}} \sim \mathcal{N}(0, \sigma_{\varepsilon}^2)$, and a $\text{logGamma}(0.001, 0.001)$ prior was placed on the log precision $\log(1/\sigma_{\varepsilon}^2)$. The

Table 2. Combination of terms used in each of the 16 models for estimating number of weekly deaths that would be expected had the pandemic not occurred. See Methods for an explanation of each term.

Model number	Global intercepts ¹	Time slope	Non-linear (autoregressive) term over weeks	Seasonal term	Non-linear (autoregressive) month-specific term over years	Temperature anomaly terms
1	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(1)}$	θ_{week}	-	-
2	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(1)}$	θ_{week}	-	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
3	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(1)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	-
4	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(1)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
5	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(2)}$	θ_{week}	-	-
6	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(2)}$	θ_{week}	-	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
7	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(2)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	-
8	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(2)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
9	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(4)}$	θ_{week}	-	-
10	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(4)}$	θ_{week}	-	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
11	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(4)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	-
12	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(4)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
13	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(8)}$	θ_{week}	-	-
14	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(8)}$	θ_{week}	-	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$
15	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(8)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	-
16	$\alpha_0 + \alpha_{\text{holiday(week)}}$	$\beta \cdot \text{week}$	$\zeta_{\text{week}}^{(8)}$	θ_{week}	$\eta_{\text{year}}^{\text{month}}$	$(\gamma + v_{\text{week of year}}) \cdot \text{temperature anomaly}_{\text{week}}$

¹ Due to the short duration of the time series, the holiday term was not identifiable for Chile (in the presence of the seasonal term) and was therefore not included.

components $\alpha_0, \alpha_{\text{holiday(week)}}$, $\beta \cdot \text{week}$, θ_{week} , $\varepsilon_{\text{week}}$ and $\zeta_{\text{week}}^{(i)}$ (for autoregressive order $i = 1, 2, 4$ or 8) appear in the expression for $\log(\text{death rate}_{\text{week}})$ in all models. The remaining components appear in some models only. Table 2 shows the terms included in each of the 16 models in the ensemble.

We used data on weekly deaths from the start of time series through mid-February 2020 to estimate the parameters of each model, which were then used to predict death rates for the subsequent 52 weeks as estimates of the counterfactual death rates if the pandemic had not occurred. These were then compared to reported deaths to calculate excess mortality due to the pandemic. For the projection period, we used recorded temperature so that our projections take into consideration actual temperature in 2020-2021. This choice of training and

prediction periods assumes that the number of deaths that are directly or indirectly related to the COVID-19 pandemic was negligible through mid-February 2020 in these countries¹, and separates the training data from subsequent weeks when impacts may have appeared.

All models were fitted using integrated nested Laplace approximation (INLA)²⁷, implemented in the R-INLA software (version 20.03). We used a model averaging approach to combine the predictions from the 16 models in the ensemble^{28,29}. Specifically, we took 2,000 draws from the posterior distribution of predicted deaths under each of the 16 models, and pooled the 32,000 draws to obtain the posterior distribution of deaths if the COVID-19 pandemic had not taken place. This approach generates a distribution of estimates that has equal samples from each model in the ensemble, and hence incorporates both the

uncertainty of estimates from each model and the uncertainty in the choice of model. The reported credible intervals represent the 2.5th and 97.5th percentiles of the resultant posterior distribution of the draws from the entire ensemble. We report the number of excess deaths, excess deaths per 100,000 people, and relative (percent) increase in deaths together with their corresponding 95% credible intervals. For the purpose of reporting, we rounded results on number of deaths that are 1,000 or more to the nearest hundred to avoid giving a false sense of precision in the presence of uncertainty; results less than 1,000 were rounded to the nearest ten. We also report the posterior probability that an estimated increase (or decrease) in deaths corresponds to a true increase (or decrease). Posterior probability represents the inherent uncertainty in how many deaths would have occurred in the absence of the pandemic. In a country and week in which the actual number of deaths is the same as the posterior median of the number expected in a no-pandemic counterfactual, an increase in deaths is statistically indistinguishable from a decrease; in such a situation, there is a 50% posterior probability of an increase and a 50% posterior probability of a decrease. Where the entire posterior distribution of the number of deaths expected without the pandemic is smaller than the actual number of deaths, there is a ~100% posterior probability of an increase and a ~0% posterior probability of a decrease and vice versa. For most countries and weeks, the posterior distribution of the number of deaths expected without the pandemic covers the observed number, but there is asymmetry in terms of whether much of the distribution is smaller or larger than the observed number. In such cases, there would be uneven posterior probabilities of an increase versus decrease in deaths, with the two summing to 100% (for example, 80% and 20%). Posterior probabilities more distant from 50%, toward either 0% or 100%, indicate more certainty. We also evaluated the sensitivity of our results to how the different models are weighted. Specifically, in the sensitivity analysis, the number of draws from each model was inversely proportional to the absolute error of prediction in the validation analyses described below. The results of the sensitivity analysis were virtually identical to those with equal draws, with weekly median excess deaths estimates differing by up to 2.4% for individual countries, and by 0.1% when averaged across all countries and weeks.

We did all analyses separately by sex and age group (0–44 years, 45–64 years, 65+ years) for countries with 2020 population of at least two million, where age- and sex-specific data were available (Table 1). For countries with 2020 population less than 2 million, we did our analyses for two age groups (0–64 years and 65+ years) because, in many weeks, the number of deaths in the age group 0–44 was small or zero, which would have led to either large uncertainty or unstable estimates. For the same reason, for countries with population under 500,000 (Iceland and Malta), we did our analyses for both sexes and all age groups combined. Models were also run for all ages and both sexes combined; the posterior medians of resultant estimates were nearly identical to the sum of the age-sex-specific ones, with a mean relative difference of 0.2%, ranging from -1.7% to 1.1%. For this reason, in figures and tables that are for all ages and both sexes, we report results from the combined

model so that the uncertainty of the estimates is correctly reported.

We report results for the entire year, as well as for three non-overlapping periods: the first wave of the pandemic (from mid-February 2020 through end of May), the (northern hemisphere) summer period (from beginning of June to mid-September 2020) and subsequent wave(s) (from mid-September 2020, when schools normally open in the northern hemisphere, to mid-February 2021).

Validation of no-pandemic counterfactual weekly deaths

We tested how well our model ensemble estimates the number of deaths expected had the pandemic not occurred by withholding data for 52 weeks starting from mid-February (i.e., the same projection period as done for 2020–2021) for an earlier year and using the preceding time series of data to train the models. In other words, we created a situation akin to 2020–2021 for an earlier year. We then projected death rates for the weeks with withheld data, and evaluated how well the model ensemble projections reproduced the known-but-withheld death rates. We repeated this for three different periods: 2017–2018 (i.e., train model using data from January 2010 to mid-February 2017 and test for the subsequent 52 weeks), 2018–2019 (i.e., train model using data from January 2010 to mid-February 2018 and test for the subsequent 52 weeks), and 2019–2020 (i.e., train model using data from January 2010 to mid-February 2019 and test for the subsequent 52 weeks). We performed these tests for each country using data for both sexes and all ages. We report the projection error (which measures systematic bias) and absolute projection error (which measures any deviation from the data). Additionally, we report coverage of the projection uncertainty; if projected death rates and their uncertainties are well estimated, the estimated 95% credible intervals should cover 95% of the withheld data.

The results of model validation (Table 3) show that the estimates of how many deaths would be expected had the pandemic not occurred from the Bayesian model ensemble were unbiased, with mean relative projection errors of 1.5% (between 0.5% and 2.2% in different years). The mean relative absolute error was between 8.0% and 8.7% in different years. 95% coverage, which measures how well the posterior distributions of projected deaths coincide with withheld data was 96% for all years, which shows that the posterior distribution is well estimated.

An earlier version of this manuscript can be found on medRxiv (doi: <https://doi.org/10.1101/2021.07.12.21260387>).

Results

Excess mortality between mid-February 2020 and mid-February 2021

Taken over the entire year, both sexes and all ages, an estimated 1,410,300 (95% credible interval 1,267,600–1,579,200) more people died in these 40 countries than would have been expected had the pandemic not taken place. This is equivalent to 141 (127–158) additional deaths per 100,000 people

Table 3. Results of the external predictive validity (out-of-sample validation) of the estimated no-pandemic counterfactual weekly deaths from the Bayesian model ensemble. Each number represents the total error over the validation period, averaged across countries.

Validation year	Projection error (relative projection error)	Absolute projection error (relative absolute projection error)	Percent covered by 95% credible interval
2017	1,893 (1.8%)	9,488 (8.5%)	97%
2018	1,107 (0.5%)	9,455 (8.7%)	94%
2019	3,306 (2.2%)	8,645 (8.0%)	98%
All three years	2,102 (1.5%)	9,196 (8.4%)	96%

* Australia, Chile, Cyprus, Germany, Greece, Romania and USA were not used for validation analysis because they had shorter time series. Hence leaving out the last three years of data would leave a time series that was too short for estimating model parameters.

and a 15% (14–17) increase in deaths over this period in all of these countries combined. The number of deaths assigned to COVID-19 in these countries over the same period was 1,256,861, which is 89% of the excess all-cause death toll (Table 4). The number of excess deaths were largest in the USA (623,100; 521,200–750,700), followed by Italy (118,800; 88,500–149,300) and England and Wales (102,100; 75,300–128,600) (Figure 1 and Table 4). Within the USA, California (71,800; 64,100–79,500) and Texas (57,400; 48,100–67,200) experienced the largest number of excess deaths, about the same as excess deaths in Spain and France, respectively (Figure 2).

In Iceland, Australia and New Zealand, mortality was 3–6% lower over this period than what would be expected if the pandemic had not occurred, with posterior probabilities of the estimated decrease being a true decrease ranging 82–94% (Figure 3). South Korea and Norway experienced no detectable change in mortality (54% and 74% probability of an increase respectively, with posterior median estimated increases <2%), and Finland, Greece, Cyprus and Denmark experienced increases of 2–5% (Figure 3A), with posterior probabilities that these changes represent an increase in death ranging from 84% to 97%. At the other extreme, the populations of the USA, Czechia, Slovakia and Poland experienced at least 20% higher mortality over these 52 weeks than they would have had the pandemic not occurred; the increase was between 15% and 20% in England and Wales, Italy, Portugal, Spain, Romania, Slovenia, Lithuania, Bulgaria, Chile, Belgium and Switzerland; the posterior probabilities that these countries experienced an increase in deaths were >99%. Because baseline mortality (i.e., death rates expected without the pandemic) varied across countries, the ordering of countries in terms of excess deaths per 100,000 people (Figure 3B) differed from the ranking of percent increase. Bulgaria, Romania, Lithuania, Czechia, Poland, Slovakia and Portugal experienced more than 200 excess deaths per 100,000 people and Italy, USA, England and Wales, Slovenia, Spain, Croatia, Belgium and Montenegro between 150 and 200, all with posterior probabilities of an

increase in deaths >99%. There was as much variation in excess mortality across US states as across the 40 countries together, with Hawaii having experienced the same level of mortality as would have been expected without the pandemic, Maine a 5% increase, and, at the other extreme, New Jersey, Arizona, Mississippi, Texas, California, Louisiana and New York at least 25% higher mortality over this year (Figure 4).

Dynamics of excess mortality

There was substantial heterogeneity across countries in terms of the patterns and dynamics of excess mortality over time (Figure 5 and Figure 6). Some countries in Central and Eastern Europe – Bulgaria, Lithuania, Poland, Romania, Serbia and Montenegro – had no or little excess mortality in the first wave of the pandemic (mid-February 2020 to end of May 2020), but experienced between 5% and 13% increase in mortality during the (northern hemisphere) summer (June 2020 to mid-September 2020; Figure 7A). In contrast, some countries with medium or high levels of excess mortality in the first wave returned to death rates in the summer that were about the same as the no-pandemic baseline (England and Wales, Belgium, Scotland, Northern Ireland, Sweden, Netherlands, France, Switzerland, Luxembourg and Cyprus) or only slightly higher than this baseline (Canada, Italy and Spain). Portugal and the USA experienced a similar increase in mortality over the summer – 10% (1–21) and 17% (12–24), respectively – to what they had in the first wave. During the same period, Australia, New Zealand and Iceland had a mortality deficit compared to levels that would have been expected without a pandemic. In Australia and New Zealand, which were in winter season in this period, this reduction has been attributed to fewer deaths from seasonal flu due to reduced contact among people^{30–33}. Chile, the other southern hemisphere country in our analysis, had 12% (8–17) higher mortality in the first wave, followed by an even larger increase of 21% (15–26) during the (southern hemisphere) winter period.

The subsequent wave(s) of the pandemic (mid-September 2020 to mid-February 2021) saw yet more changes in excess deaths

Table 4. Number of excess deaths from any cause and deaths assigned to coronavirus disease 2019 (COVID-19) due from mid-February 2020 to mid-February 2021, by country.

Excess deaths $\geq 1,000$ are rounded to the nearest hundred and excess deaths $< 1,000$ to the nearest ten. Deaths assigned to COVID-19 were taken directly from the cited sources and not rounded.

Country	Number of excess deaths (95% credible interval)	Number of deaths assigned to COVID-19 as underlying cause ¹
Australia	-4,090 (-11,590 to 2,300)	909
Austria	9,100 (4,900 to 13,200)	8,385
Belgium	17,900 (12,300 to 23,600)	22,077
Bulgaria	18,200 (12,800 to 23,500)	9,854
Canada	19,800 (8,800 to 31,300)	21,723
Chile	18,400 (15,200 to 21,600)	20,126
Croatia	6,400 (3,300 to 9,400)	5,449
Cyprus	140 (-520 to 840)	232
Czechia	24,000 (18,500 to 29,300)	19,777
Denmark	2,400 (-170 to 5,100)	2,343
England and Wales	102,100 (75,300 to 128,600)	128,077
Estonia	790 (140 to 1,500)	535
Finland	1,200 (-450 to 2,800)	756
France	62,700 (33,200 to 95,100)	84,306
Germany	64,100 (-1,870 to 135,400)	67,903
Greece	3,700 (-4,180 to 11,000)	6,297
Hungary	14,300 (7,500 to 21,300)	14,347
Iceland	-140 (-340 to 40)	29
Italy	118,800 (88,500 to 149,300)	95,718
Latvia	2,200 (1,100 to 3,400)	1,542
Lithuania	6,600 (4,500 to 8,700)	3,178
Luxembourg	390 (150 to 640)	625
Malta	320 (20 to 640)	304
Montenegro	950 (530 to 1,400)	950
Netherlands	17,300 (9,400 to 25,300)	15,231
New Zealand	-1,050 (-3,390 to 1,300)	26
Northern Ireland	2,300 (1,300 to 3,300)	2,751
Norway	490 (-1,090 to 2,100)	608
Poland	82,300 (62,500 to 101,400)	42,171

Country	Number of excess deaths (95% credible interval)	Number of deaths assigned to COVID-19 as underlying cause ¹
Portugal	20,700 (14,100 to 27,200)	15,962
Romania	45,600 (30,800 to 61,500)	19,894
Scotland	7,100 (3,600 to 10,600)	9,355
Serbia	10,300 (5,600 to 14,800)	4,337
Slovakia	11,300 (8,200 to 14,200)	6,671
Slovenia	3,600 (2,700 to 4,400)	4,057
South Korea	560 (-10,870 to 11,000)	1,562
Spain	76,100 (52,400 to 100,100)	67,636
Sweden	9,900 (6,400 to 13,700)	12,914
Switzerland	10,200 (7,100 to 13,300)	9,174
USA	623,100 (521,200 to 750,700)	529,070

¹ Data are from Office for National Statistics for England and Wales (<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales>), NRS for Scotland (<https://www.nrscotland.gov.uk/covid19stats>), NISRA for Northern Ireland (https://www.nisra.gov.uk/system/files/statistics/Weekly_Deaths%20-%20w%20e%203rd%20September%202021.XLSX) and the European Centre for Disease Prevention and Control (ECDC) for other countries (<https://opendata.ecdc.europa.eu/covid19/nationalcasedeath/csv>; accessed on 21 September 2021).

patterns across countries. While New Zealand, Australia, Iceland, Finland, Norway, Cyprus and South Korea remained resilient to the rise in mortality (i.e., no or $< 2\%$ increase in mortality compared to the no-pandemic baseline), many countries in Europe, especially in Central Europe, experienced a rise in mortality compared to the no-pandemic baseline: by $> 40\%$ in Slovakia, Czechia and Poland, and by 20–40% in England and Wales, Italy, Austria, Hungary, Montenegro, Croatia, Portugal, Switzerland, Romania, Lithuania, Bulgaria and Slovenia, all with posterior probabilities of positive excess mortality greater than 99%. Excess deaths also reappeared in other countries that had experienced a medium to large toll in the first wave including Belgium, Spain, Scotland, Northern Ireland, Sweden, Canada, France and the Netherlands – some at the same level (France and Northern Ireland) and others at lower levels (Canada, Scotland, Spain, Belgium, Sweden) than the first wave but all lasting for many weeks during this period. The USA had an even larger increase in mortality compared to the no-pandemic baseline after mid-September than it had in the first wave and summer months, making it the only country to maintain a steady burden of excess mortality. There were nonetheless variations in excess deaths over time across different states in the USA (Figure 8), reflecting the autonomy

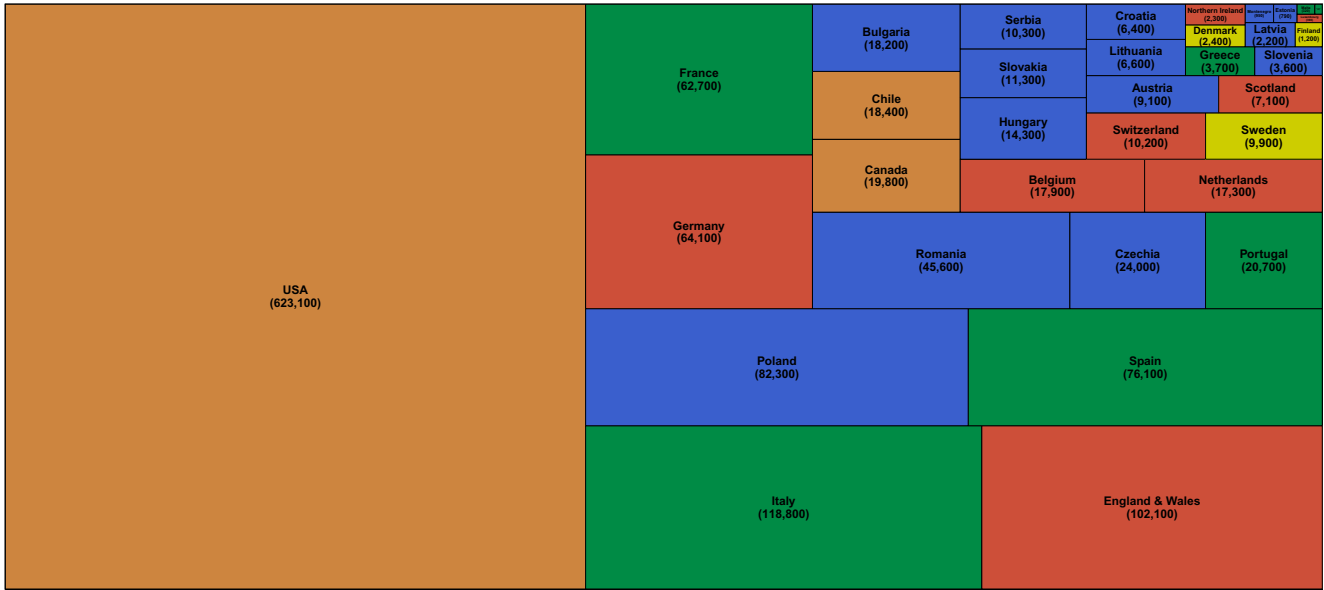


Figure 1. Number of excess deaths due to the first year of the coronavirus disease 2019 (COVID-19) pandemic by country. The size of each rectangle shows the number of deaths from all causes in excess of what would be expected if there had been no COVID-19 pandemic from mid-February 2020 through mid-February 2021 for each country. There are no segments for Australia, New Zealand, Norway, Iceland and South Korea because we estimated no detectable excess deaths or a potential reduction in mortality compared to the no-pandemic baseline. Colour for each country indicates its geographical region: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden).

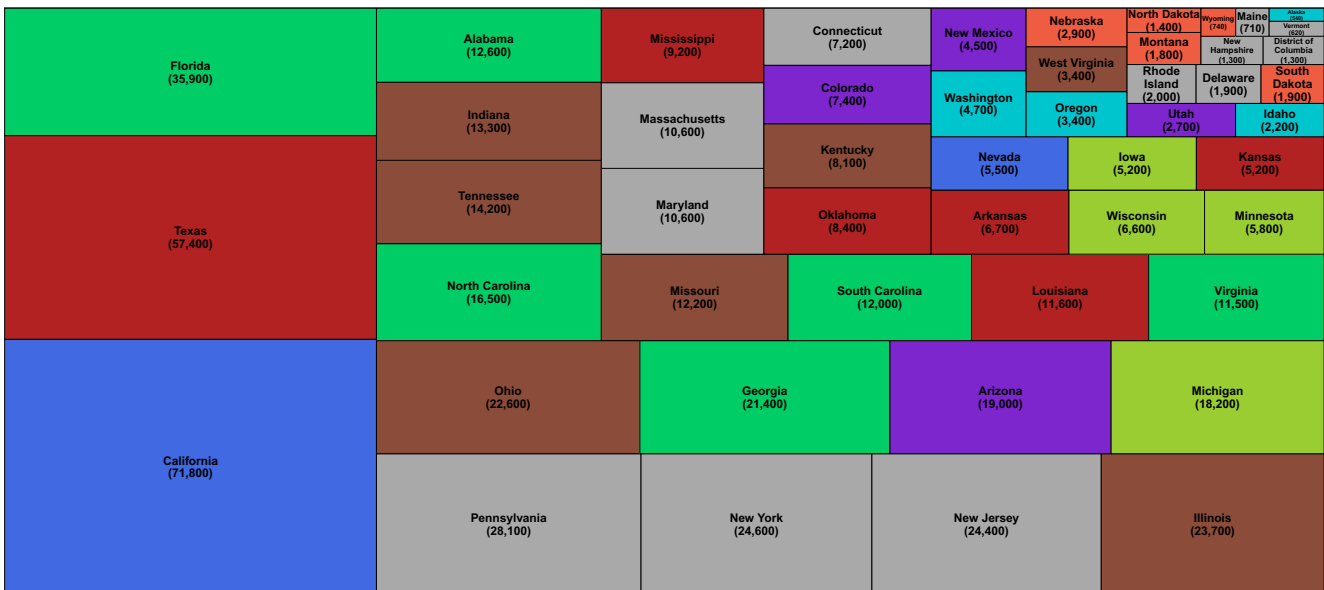


Figure 2. Number of excess deaths due to the first year of the coronavirus disease 2019 (COVID-19) pandemic by US state. The size of each rectangle shows the number of deaths from all causes in excess of what would be expected if there had been no COVID-19 pandemic from mid-February 2020 through mid-February 2021 for each state and the District of Columbia. There is no segment for Hawaii because we estimated no detectable excess deaths. The colour of each state indicates its geographical region: Southeast (Alabama, Florida, Georgia, North Carolina, South Carolina, Virginia); Northwest (Alaska, Idaho, Oregon, Washington); Southwest (Arizona, Colorado, New Mexico, Utah); South (Arkansas, Kansas, Louisiana, Mississippi, Oklahoma, Texas); West (California, Nevada); Northeast (Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont); Central (Illinois, Indiana, Kentucky, Missouri, Ohio, Tennessee, West Virginia); East North Central (Iowa, Michigan, Minnesota, Wisconsin); and West North Central (Montana, Nebraska, North Dakota, South Dakota, Wyoming).

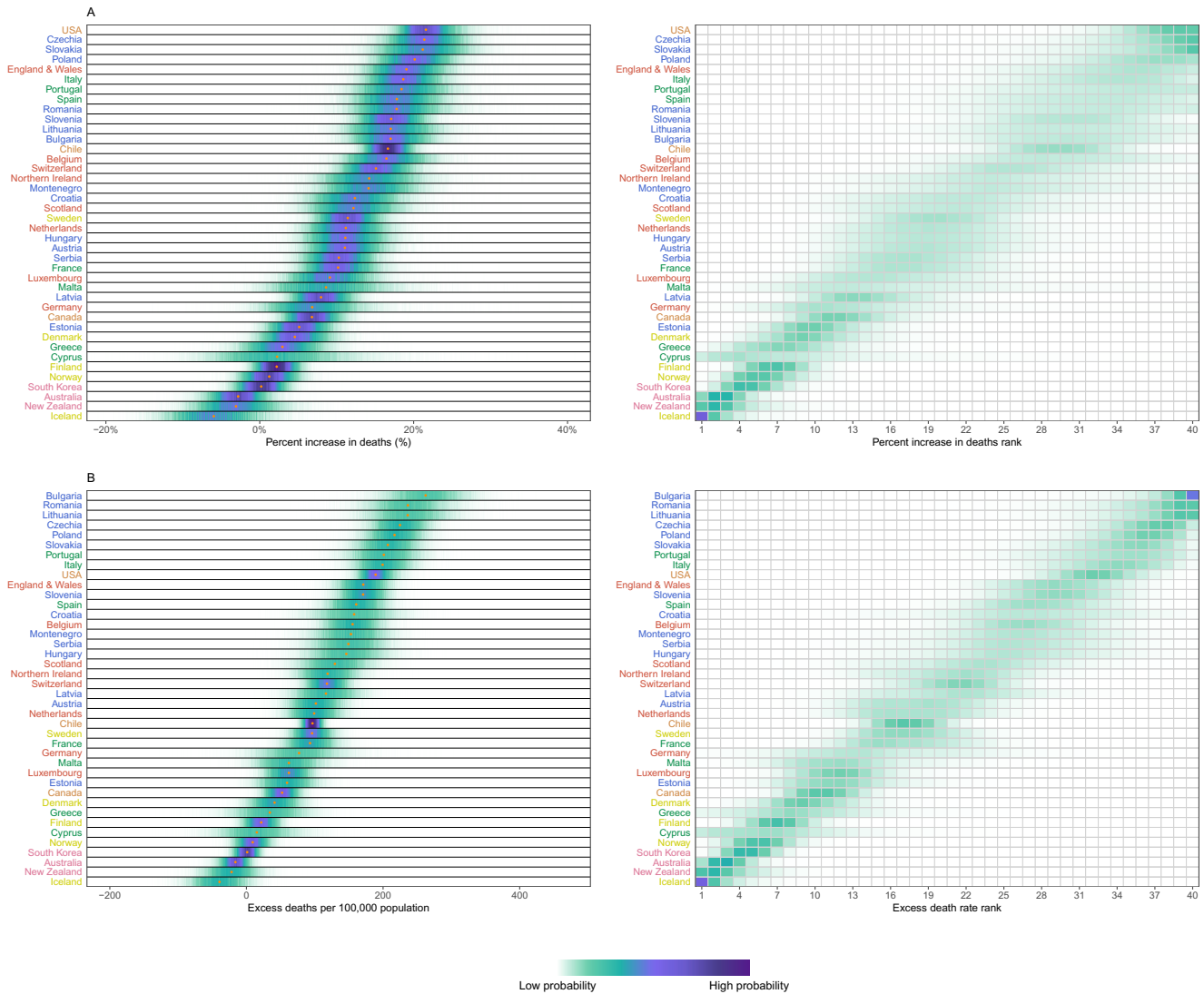


Figure 3. Excess mortality due to the first year of the coronavirus 2019 (COVID-19) pandemic, by country. (A) Posterior distribution of percent increase in deaths from any cause from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. **(B)** Posterior distribution of excess deaths from any cause per 100,000 people from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. In both panels, the right-hand side shows the probability distribution for the country's rank. Countries are ordered vertically by median increase from smallest (at the bottom) to the largest (at the top). Colour for each country's name indicates its geographical region: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden). See Figure 10 for results by sex.

that states, and their governors and legislatures, had with regard to key responses^{6,34}.

As a result of these heterogeneous dynamics, there was virtually no correlation between excess mortality in the first wave and the summer period among countries (correlation coefficient of percent increase in the two periods = 0.03), and weakly negative correlation between excess mortality in the first wave and mid-September and later (correlation coefficient = -0.15). This was translated to a variable distribution of excess mortality

burden across the three periods (Figure 7B). For example, the first wave accounted for over half of excess deaths in Scotland, Spain, England and Wales, Canada, Sweden, Belgium, the Netherlands and Cyprus. At the other extreme, the period between mid-September 2020 and mid-February 2021 accounted for over 90% of excess deaths in Bulgaria, Croatia, Czechia, Hungary, Latvia, Montenegro, Poland, Slovakia and Slovenia. A similar variation was seen across the US states, with excess deaths along the north-eastern coast (Massachusetts, New Jersey, Connecticut, New York and District of Columbia)

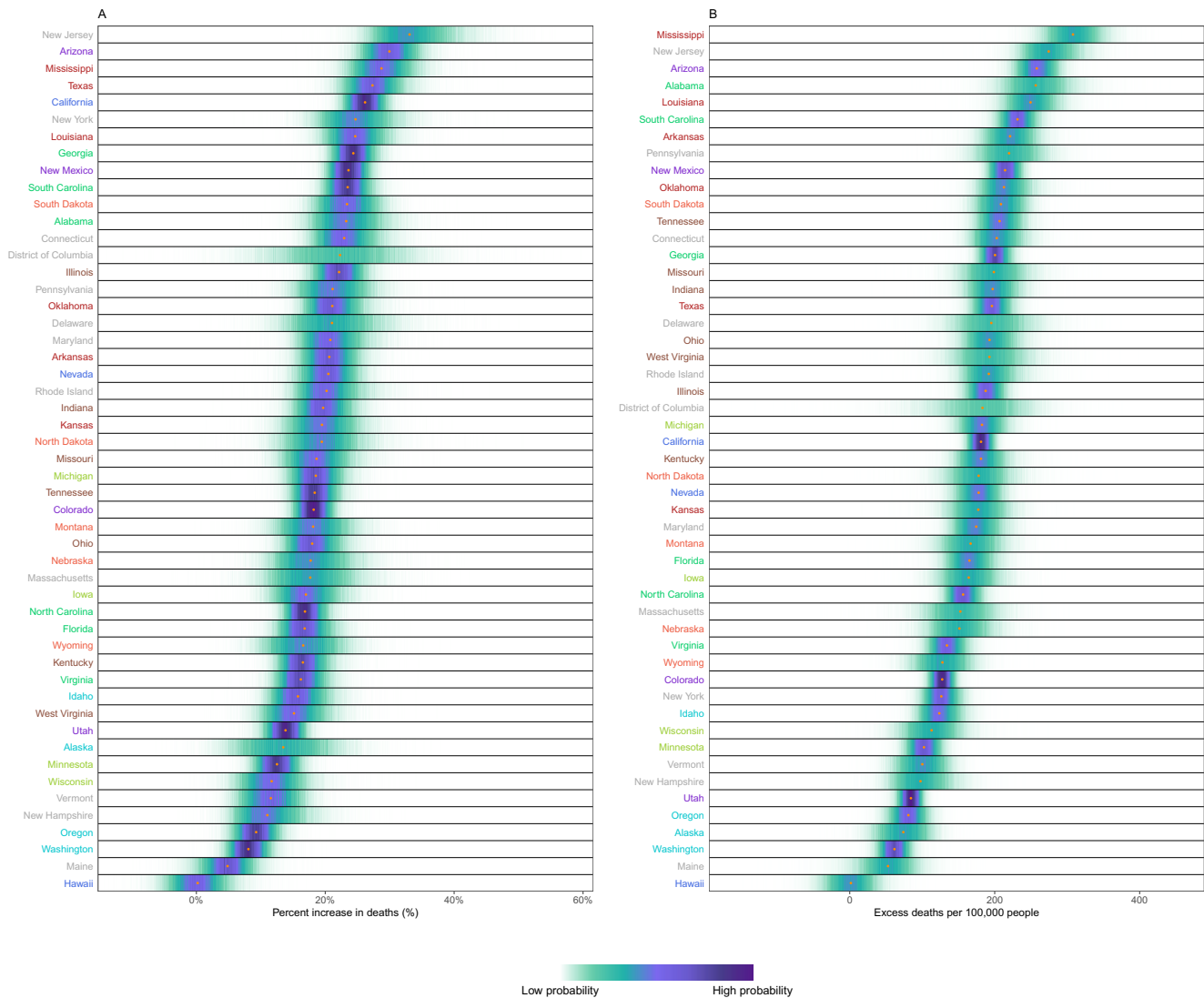


Figure 4. Excess mortality due to the first year of the coronavirus disease 2019 (COVID-19) pandemic, by US state. (A) Posterior distribution of percent increase in deaths from any cause from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. **(B)** Posterior distribution of excess deaths from any cause per 100,000 people from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. States are ordered vertically by median increase from smallest (at the bottom) to the largest (at the top). Colour for each state indicates its geographical region: Southeast (Alabama, Florida, Georgia, North Carolina, South Carolina, Virginia); Northwest (Alaska, Idaho, Oregon, Washington); Southwest (Arizona, Colorado, New Mexico, Utah); South (Arkansas, Kansas, Louisiana, Mississippi, Oklahoma, Texas); West (California, Nevada); Northeast (Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont); Central (Illinois, Indiana, Kentucky, Missouri, Ohio, Tennessee, West Virginia); East North Central (Iowa, Michigan, Minnesota, Wisconsin); and West North Central (Montana, Nebraska, North Dakota, South Dakota, Wyoming).

being dominated by the first wave, in some southern states (Florida, Arizona, Texas and South Carolina) by the summer, and in the northern plains (Wisconsin, North and South Dakota and Montana) by the post-September period.

Age and sex-distribution of excess mortality

Countries differed in how excess deaths were distributed across age groups (Figure 9). In Denmark, Sweden, France, Switzerland, Belgium and Slovenia >95% of all excess deaths were in those

aged 65 years and older. On the other hand, Estonia, Finland (which had the smallest detectable excess mortality of any country), USA, Canada, Lithuania and Chile had the largest share of excess deaths in people aged younger than 65 years. Of the 35 countries with a detectable increase in mortality (defined as median estimated increase of >2%) and sufficient data to analyse by age group, Canada experienced the largest share of excess deaths in those aged younger than 45 years (16% of all excess deaths), followed by the USA (5%) and Finland (5%;

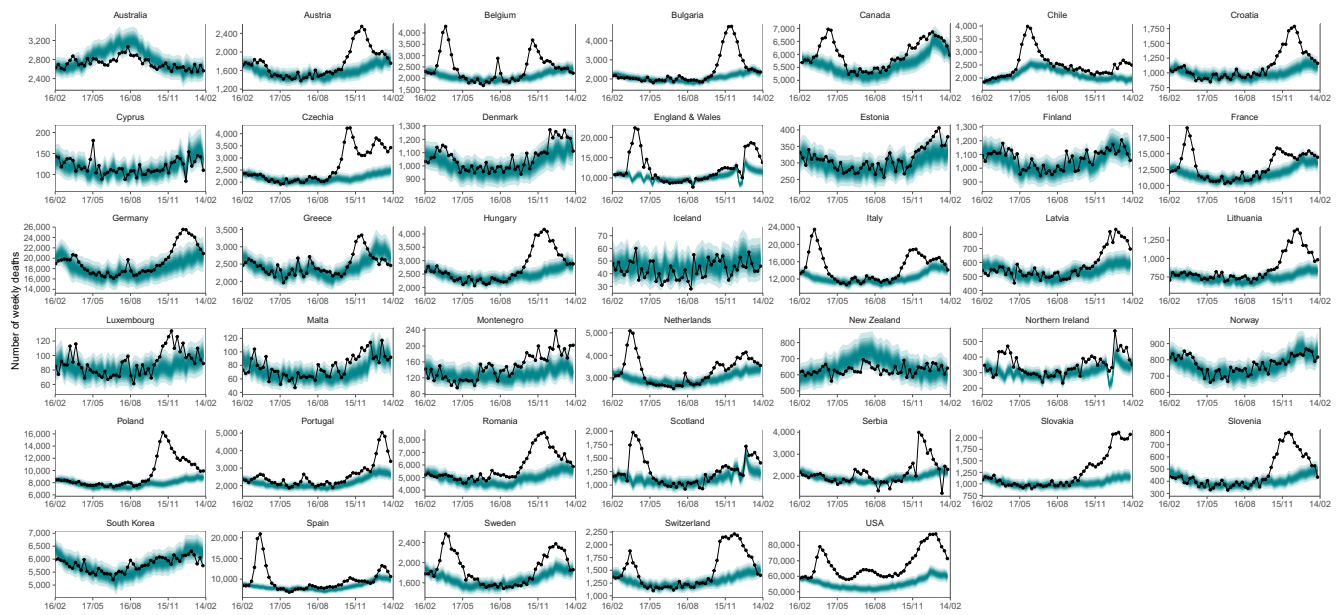


Figure 5. Weekly number of deaths from mid-February 2020 through mid-February 2021. The points show reported deaths. The turquoise shading shows the credible intervals around the median prediction, from 5% (dark) to 95% (light) in 10% increments.

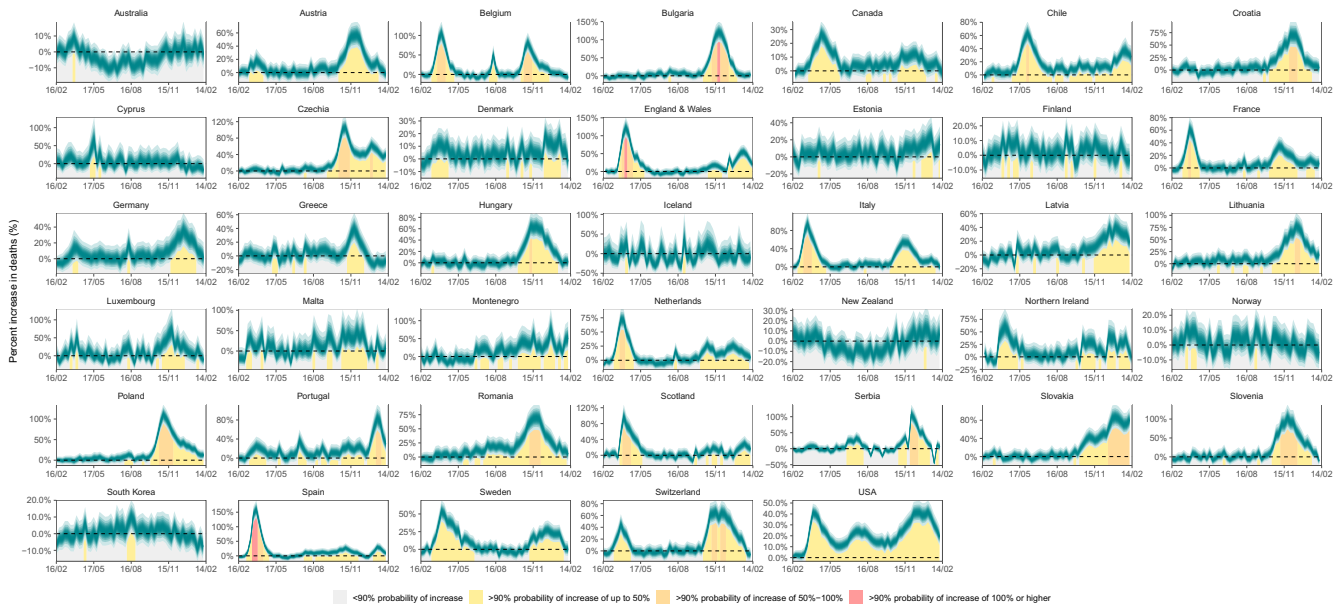


Figure 6. Weekly percent increase in mortality due to the coronavirus disease 2019 (COVID-19) pandemic by country. The turquoise shading shows the credible intervals around the median prediction, from 5% (dark) to 95% (light) in 10% increments. The background shading (grey/yellow/orange/red shading under the turquoise ribbons) indicates the magnitude of the weekly increase that was detectable with a posterior probability of at least 90%.

noting that excess death rates in Finland, although detectable, were lower than in other countries). The high mortality toll in younger Canadians may have been due to COVID-19 death at home³⁵ and an increase in deaths from drug overdose³⁶. This division arises largely from how much specific segments of the

society, such as workers or care home residents, were exposed to infection. Percent increase in mortality was similar between men and women in most countries (Figure 10). There were nonetheless some exceptions, e.g. in Chile, Montenegro, Serbia and the Netherlands deaths increased by a larger percent

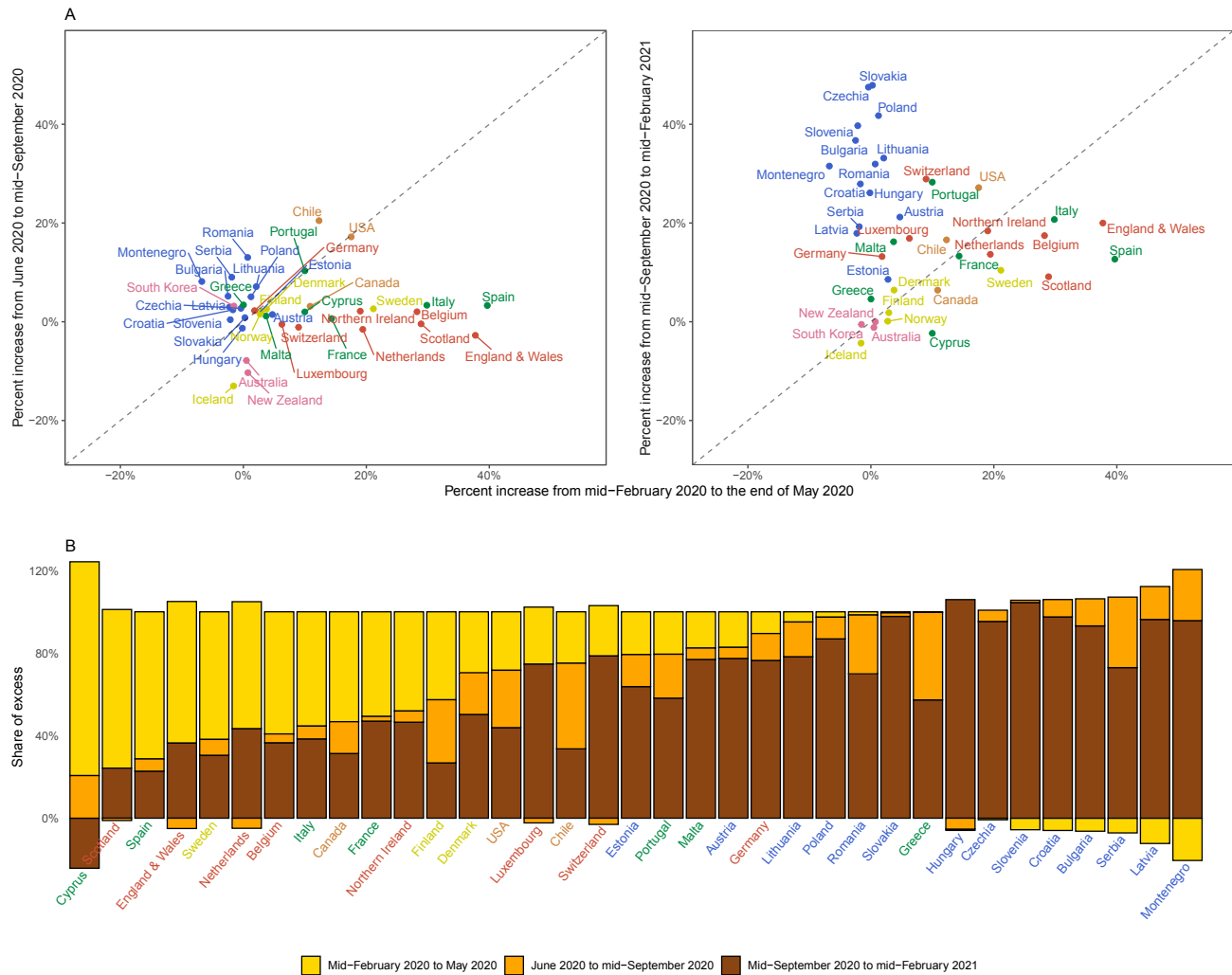


Figure 7. Excess mortality due to the coronavirus disease 2019 (COVID-19) pandemic in different time periods. (A) Comparison of percent increase in mortality from any cause in excess of what would be expected if there had been no COVID-19 pandemic in summer (beginning of June 2020 to mid-September 2020) and subsequent waves (mid-September 2020 to mid-February 2021) with the first wave (mid-February 2020 to end of May 2020) in each country. (B) Proportion of excess deaths in each of the above three periods in each of the countries. There are no bars for Australia, New Zealand, Norway, Iceland and South Korea in panel B because we estimated no detectable excess deaths or a potential reduction in mortality compared to the no-pandemic baseline. Colour for each country indicates its geographical region: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden). In some countries, there was a reduction in mortality relative to a no-pandemic baseline in some weeks, shown as negative numbers. The country's total excess death toll is the net effect of these reductions and increases in other periods, with all bars adding to 100%. See Figure 6 for weekly percent increase in mortality.

in men (12%–16%) than women (6%–9%); in contrast, in Slovenia, women (15%) experienced a slightly larger percent increase than men (14%).

Strengths and limitations

The main strength of our work is the development and application of a method to systematically and consistently use time

series data from previous years to estimate how many deaths would be expected in the absence of pandemic through early 2021. The models incorporated important features of mortality, including seasonality of death rates, how mortality in one week or year may depend on previous week(s) and year(s), and the seasonally-variable role of temperature. To our knowledge, our models are the only ones that formally incorporated

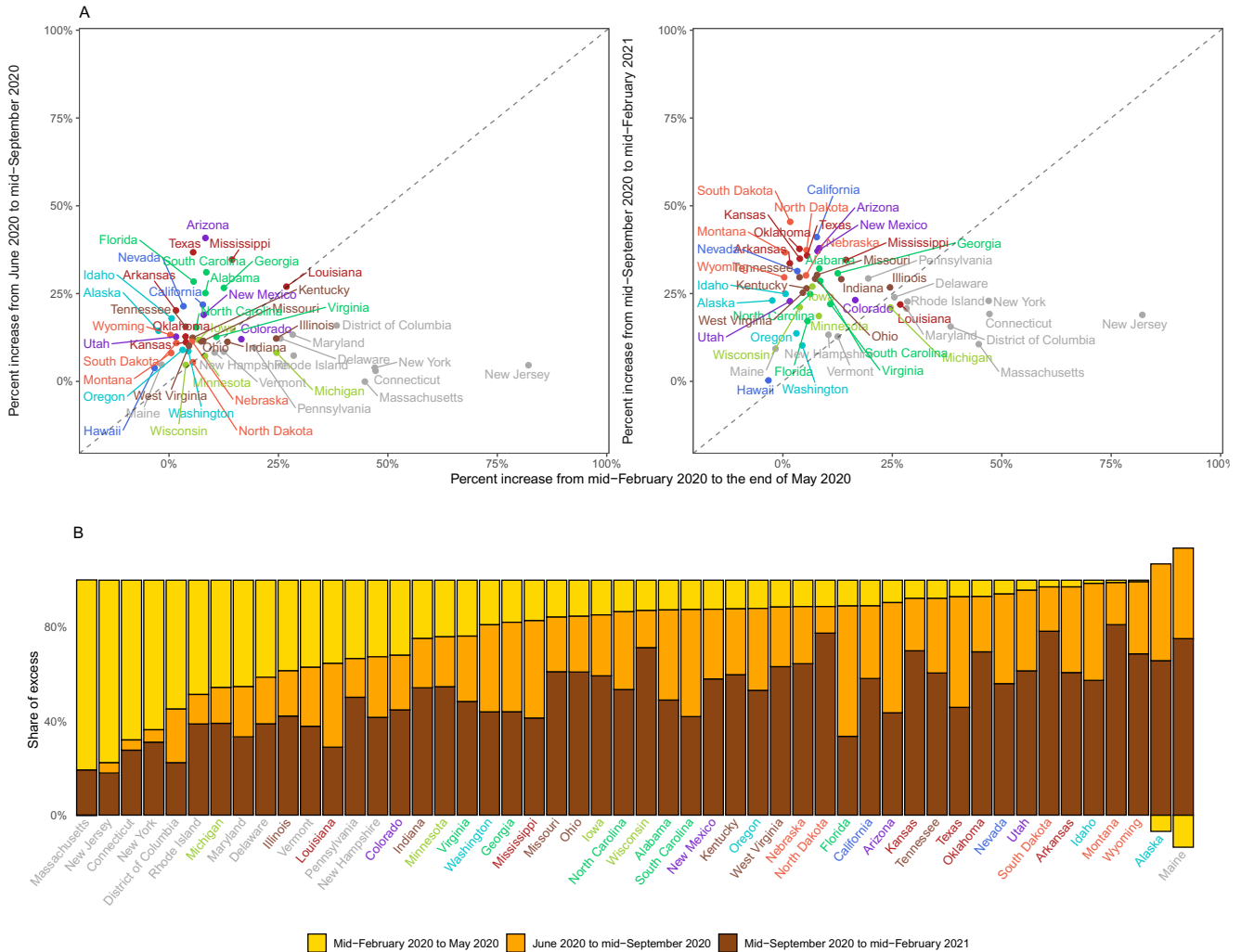


Figure 8. Excess mortality due to the coronavirus disease 2019 (COVID-19) pandemic in different time periods for US States. (A) Comparison of percent increase in mortality from any cause in excess of what would be expected if there had been no COVID-19 pandemic in summer (beginning of June 2020 to mid-September 2020) and subsequent waves (mid-September 2020 to mid-February 2021) with the first wave (mid-February 2020 to end of May 2020) in each state. (B) Proportion of excess deaths in each of the above three periods in each state. There is no bar for Hawaii because we estimated no detectable excess deaths. In some states, there was a reduction in mortality relative to a no-pandemic baseline in some weeks, shown as negative numbers. The state's total excess death toll is the net effect of these reductions and increases in other periods, with all bars adding to 100%. Colour for each state indicates its geographical region: Southeast (Alabama, Florida, Georgia, North Carolina, South Carolina, Virginia); Northwest (Alaska, Idaho, Oregon, Washington); Southwest (Arizona, Colorado, New Mexico, Utah); South (Arkansas, Kansas, Louisiana, Mississippi, Oklahoma, Texas); West (California, Nevada); Northeast (Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont); Central (Illinois, Indiana, Kentucky, Missouri, Ohio, Tennessee, West Virginia); East North Central (Iowa, Michigan, Minnesota, Wisconsin); and West North Central (Montana, Nebraska, North Dakota, South Dakota, Wyoming).

the role of temperature on weekly mortality, and accounted for dependency of mortality in one week on preceding week(s) and in one year on preceding year(s). This methodology allows more robust estimation of the total impacts of the pandemic, especially as more time elapses since the beginning of the pandemic. It also enables comparisons of excess deaths across countries on a real-time basis. By modelling death rates, rather than simply the number of deaths as is done in most other analyses, we account for changes in population size and age structure. We used an ensemble of models which typically leads to more robust projections and better accounts for both the uncertainty associated with each individual model and that of

model choice. As a result, our approach gives a more complete picture of the inherent uncertainty in how many excess deaths the pandemic has caused than approaches that are not probabilistic or use a single model.

A limitation of our work is that we did not have data on underlying cause of death. Having a breakdown of deaths by underlying cause will help develop cause-specific models and understand which causes have exceeded or fallen below the levels expected. Nor did we have data on total mortality by individual or community sociodemographic status to understand inequalities in the impacts of the pandemic beyond deaths

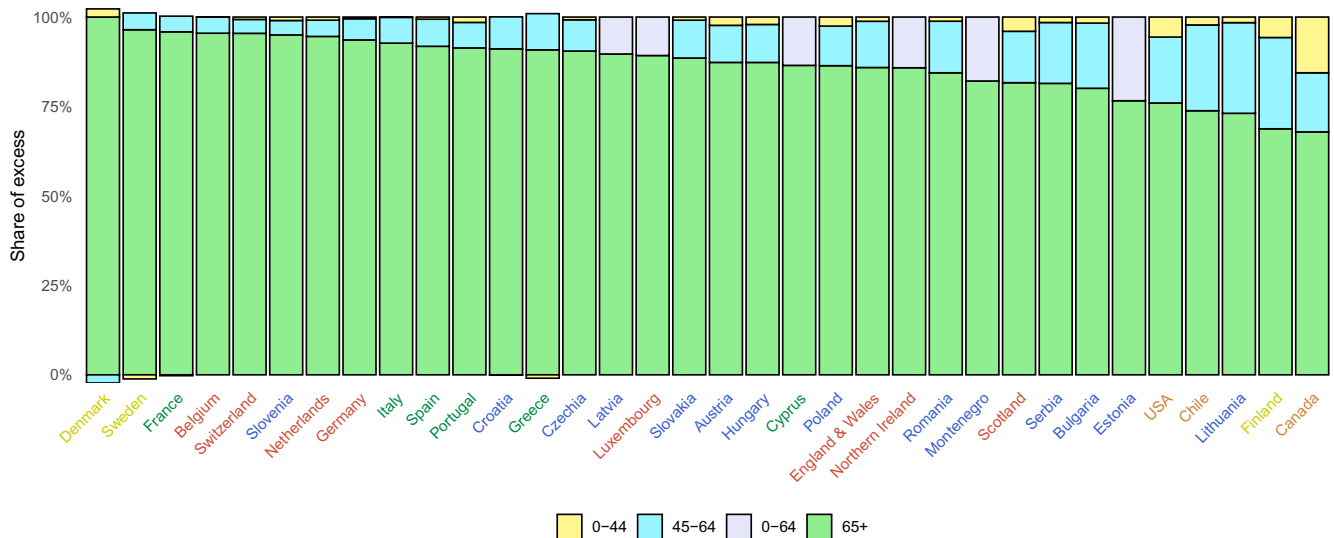


Figure 9. Distribution of excess deaths due to the first year of the coronavirus disease 2019 (COVID-19) pandemic by age group. The figure shows the share of excess deaths in each age group by country. There are no bars for Australia, New Zealand, Norway, Iceland and South Korea because we estimated no detectable excess deaths or a potential reduction in mortality compared to the no-pandemic baseline. There is no bar for Malta because we only made all-age estimates for reasons described in Methods. For Luxembourg, Cyprus, Latvia, Northern Ireland, Montenegro and Estonia, analysis was done for 0–64 years without a further split into 0–44 years and 45–64 years for reasons described in Methods. Colour for each country indicates its geographical region: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden).

assigned to COVID-19 as the underlying cause of death. Where data have been analysed for population subgroups, excess mortality tends to be higher in marginalised individuals and communities^{37–39}. More detailed data will allow more granular analysis of the impacts of the pandemic, which can in turn inform resource allocation and a more targeted approach to mitigating both the direct and indirect effects of the COVID-19 pandemic.

Comparison with other estimates

The *Financial Times* and *The Economist's* [excess deaths tracker](#) report the number of excess deaths for various countries based on comparisons of deaths in 2020 and 2021 with 2015–2019 averages. This approach does not account for general trends in mortality nor for factors like temperature that affect mortality and vary from year to year. *The Economist* has also recently published a set of [excess deaths estimates](#) using data from the [Human Mortality Database](#) and the World Mortality Dataset⁴⁰, and an ensemble of gradient boosted decision trees. Countries with small, medium and large number of excess deaths are largely consistent between our analysis and these sources. There are nonetheless some differences. For example, we estimated ~76,100 excess deaths for Spain, compared to ~81,300 by *Financial Times* and ~86,300 and ~89,600 respectively by the two models from *The Economist*. Our median excess death estimate for Denmark was about twice as large as that of *Financial Times*, and those for Greece and Serbia about one third smaller. Similarly, *The Economist's* two models predicted a mortality deficit of ~2,900 and ~3,200 deaths, respectively, for South Korea,

while we estimated no detectable change in mortality. Nonetheless, the 95% credible intervals of our estimates contained those of *Financial Times* and *The Economist*.

The Institute for Health Metrics and Evaluation has released [numbers of excess deaths](#) by fitting a model for seasonality (the details of the seasonal model are not currently available) and projecting the residuals for pre-2020 using a spline model. The models do not account for temperature, as ours do, but hot summer weeks with particularly large deaths were excluded. Several sources have commented that the estimates are likely an overestimate, especially in their earlier version^{41–43}. For example, the Institute estimated ~156,800 excess deaths for the UK for the same period as our analysis, compared to ~111,500 by us and ~118,500 by the national statistical offices for England, Wales, Scotland and Northern Ireland. They estimated ~573,500 excess deaths for the USA (revised downwards from an estimate of ~760,000 in September 2021) compared to ~623,100 by us after accounting for temperature. They estimated ~33,900 deaths for Canada, compared to ~19,800 by us and ~20,900 by [Statistics Canada](#), and ~31,100 excess deaths for Portugal, compared to ~20,700 by us. [EuroMoMo](#) fits a sinusoidal seasonal model to death counts but does not report country-specific excess deaths and hence could not be compared with our results.

The UK Office for National Statistics (ONS) calculated a number of age-standardised measures of excess mortality for 15 European countries based on comparisons of deaths in 2020 with

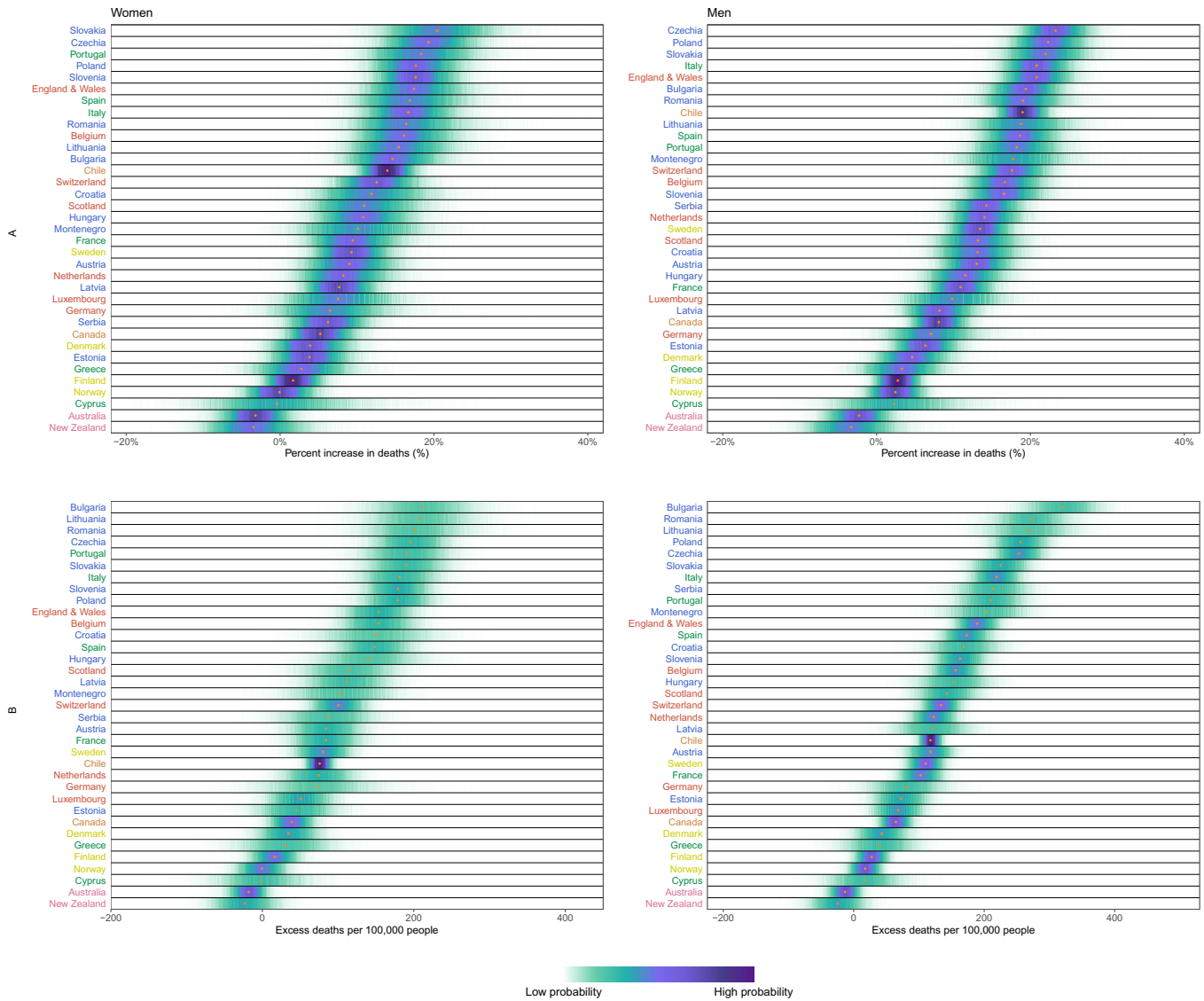


Figure 10. Excess mortality due to the first year of the coronavirus disease 2019 (COVID-19) pandemic, by country and sex. (A) Posterior distribution of percent increase in deaths from any cause from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. (B) Posterior distribution of excess deaths from any cause per 100,000 people from mid-February 2020 to mid-February 2021. Gold dots show the posterior medians. Countries are ordered vertically by median increase from smallest (at the bottom) to the largest (at the top). Data for Northern Ireland, South Korea and USA were only available for both sexes combined and did not allow sex-specific results. There are no segments for Malta and Iceland because estimates for these countries were only made for both sexes combined, for reasons described in Methods. Colour for each country indicates its geographical region: the Pacific (Australia, New Zealand, South Korea), the Americas (Canada, Chile, the USA), Central and Eastern Europe (Austria, Bulgaria, Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia), Southwestern Europe (Cyprus, France, Greece, Italy, Malta, Portugal, Spain), Northwestern Europe (Belgium, England and Wales, Germany, Luxembourg, the Netherlands, Northern Ireland, Scotland, Switzerland) and Nordic (Denmark, Finland, Iceland, Norway, Sweden).

2015-2019 averages⁴⁴, as did Eurostat for the monthly number of deaths. These analyses did not account for temperature and holidays, and the Eurostat analysis did not account for changes in population. The ONS concluded that Norway, Finland, Denmark and Latvia, Cyprus and Estonia had a mortality deficit whereas our estimates indicated no detectable excess mortality for Norway, and increases from 2 to 8% for the other countries. Differences between our results and those of the ONS may be partly related to the fact that ONS analysis also

included the pre-pandemic months and did not account for inter-annual variations in temperature. For example, in the northern hemisphere, the first and last three months of 2020 were on average warmer than the average of the past five years but weeks 13-40 were on average slightly cooler.

Discussion

The magnitude of excess mortality in the first wave of the COVID-19 pandemic was related to two factors, as seen in

quantitative and qualitative studies on the response to, and impacts of, the pandemic's early waves^{1,6,8,45-61}. First, how well countries, and subnational entities such as US states, managed the early months of the pandemic – specifically the agility of imposing timely lockdowns and other social distancing measures and border controls (e.g., complete or partial travel restrictions and/or quarantine for travellers) and adequate and effective testing, contact tracing and isolation of infected individuals and their contacts. Second, how prepared and resilient the health and social care system was to control the spread of infection, in the community as well as in health facilities and care homes, while continuing routine care.

Countries eased or maintained travel restrictions and distancing measures of the first wave to different extents and at different paces^{5,62}. They also differed in terms of testing for surveillance and identifying infected individuals, how well and how fast they traced contacts, and how they supported the isolation of infected individuals and their contacts^{51,59}. Australia and New Zealand took advantage of their island geographies and pursued an approach of disease elimination⁶³ – following strict lockdowns they imposed tight border control which kept cases to sporadic small numbers and allowed careful contact tracing and isolation. Iceland, Norway and South Korea did not close their borders but put in place various forms and durations of quarantine/isolation and testing for travellers. They also effectively integrated their well-coordinated public health capabilities⁶⁴ with modern biomedical (e.g., genomics) and digital technologies (e.g., data from credit card transactions, mobile phones and CCTV [closed-circuit television] footage), and did widespread symptomatic and asymptomatic testing to identify, track and isolate infected individuals and their contacts, and to successfully suppress the epidemic^{47,65-69}, with additional restrictions only when there was a surge in infections. All three countries also have a strong healthcare system that continued to provide routine care alongside care for COVID-19 patients.

At the other extreme, many countries in Central and Eastern Europe, which had put strict measures in place and had experienced no detectable excess mortality during the first half of 2020, removed restrictions on travel and social contact in summer of 2020, at times to a greater extent or at a faster pace than their Western European counterparts^{58,62,70,71}. With virtually the entire population still susceptible to infection, this set into motion community transmission, which coincided with the introduction of more transmissible variants of SARS-CoV-2 which were not controlled as fast and as strictly as earlier in 2020, leading to their true 'first wave' in Autumn 2020 which was equivalent to or worse than those in their Western European counterparts in magnitude and duration (Figure 6 and Figure 7). Some Mediterranean countries, such as Malta and Greece, and Northwestern European countries, such as Austria and Germany, were also largely spared during the first half of 2020, only to see an increase in deaths in autumn and winter, due to a combination of (tourism-related) travel and increased local mobility and social interactions⁷².

Between these extremes, other countries in Europe and Canada mandated or encouraged masks and face coverings, continued some forms of distancing measures (including occasional

lockdowns), increased their testing capacity to various extents, and restarted routine healthcare. There were also improvements in treatments and protocols following large-scale trials and analyses of routine care data⁷³⁻⁷⁵. These changes meant that, despite the repeated rise in infections, the mortality toll from COVID-19 and other diseases was lower than the first wave but nonetheless considerable in these countries⁷³. The continued death toll in these countries may have been because distancing measures were not as stringent as those in the first wave, and because testing, contact tracing and isolation support did not reach the coverage or depth needed to contain transmission, as did those in Iceland and South Korea^{51,76}. This was compounded by more transmissible variants and that the second wave occurred in winter when more time is spent indoors with less ventilation. The experience of the USA did not resemble that of any of the other countries. Rather, different states saw a rise in infections and deaths at different times⁷⁷, because there was little coordinated national response and because periods of extensive travel, such as Thanksgiving and Christmas holidays, led to spread of infection across states.

The observed patterns of excess mortality in the first year of the pandemic indicate that the pandemic's death toll in the next year is likely to depend on three factors: The first, and most important factor in the countries analysed here will be the breadth and pace of vaccination, including whether vaccination is extended to school-aged children and the use of boosters to enhance immunity especially against new variants of SARS-CoV-2, because vaccines have been shown to be highly effective in preventing (severe) COVID-19 and deaths in trials and in real-world settings⁷⁸⁻⁸¹. Even with high vaccine coverage, some adherence to other measures may be needed when the number of infections rises, because vaccine efficacy is less than 100%, especially against new variants of SARS-CoV-2⁸¹, and because the morbidity and longer-term health morbidity impacts of infection may be significant. Second, as the direct impacts of the COVID-19 pandemic are reduced through vaccination, the indirect impacts will become more visible. These include how much the backlog of routine care and persistently high health system pressure impacts deaths from other conditions, and the impacts on jobs and income. Mitigating these requires economic and social policies that generate secure employment and income support, and strengthening health and social care. A third, and perhaps more uncertain factor, is the magnitude of direct COVID-19 deaths that might be expected in (northern hemisphere) winter 2021-2022, because retraction of non-pharmaceutical interventions that mandate or facilitate social distancing and use of masks before the entire population is vaccinated may lead to circulating SARS-CoV-2 infections in countries as a whole as well as in specific geographical and sociodemographic subgroups of the population. In mid-February 2021, vaccination rates were still low in the countries included in our analysis, with the highest rates in the UK (22% of adults with one dose and 1% with two doses), Serbia (12% and 3%, respectively), the USA (11% and 4%, respectively) and Chile (11% and 0.3%, respectively). Since then, vaccination accelerated in industrialised countries and emerging economies and in many countries 70% or more of the population have been vaccinated. Even in those, specific geographical or social subgroups of the

population may have lower vaccination rates. Further, for countries where supply and access limit the pace of vaccination, the coming year could look as it did for the countries in this paper: a choice between lockdowns and a large death toll. To avoid these scenarios, which both have adverse health and well-being impacts, vaccine access and roll out must be accelerated and accompanied with actions to both delay and contain infections, especially new variants of concern, through effective and timely testing, contact tracing and isolation support and measures such as mask wearing in crowded indoor settings.

Data availability

Underlying data

Input data on deaths, population and temperature are available at <https://doi.org/10.5281/zenodo.5535829>¹².

This repository contains the following underlying data:

- data/data.csv (data on deaths, temperature and population by age group, sex, country and week)
- output/result_summaries.csv (weekly estimates of predicted deaths, excess deaths, excess death rates per 100,000 and relative increase in deaths)

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

The original data sets used in the study are publicly available from the following locations:

Data on deaths and population

UN World Population prospects: [https://population.un.org/wpp/Download/Files/1_Indicators%20\(Standard\)/CSV_FILES/WPP2019_PopulationByAgeSex_Medium.csv](https://population.un.org/wpp/Download/Files/1_Indicators%20(Standard)/CSV_FILES/WPP2019_PopulationByAgeSex_Medium.csv)

Australia: <https://www.abs.gov.au/statistics/health/causes-death/provisional-mortality-statistics/latest-release> (*Provisional Mortality Statistics, Weekly Dashboard, Jan 2020-May 2021.xlsx* and *Doctor certified deaths by week of occurrence, 2015-19.xlsx*)

Eurostat: <https://ec.europa.eu/eurostat/data/database> (tables demo_r_mwk_05 and demo_pjangroup)

Canada: <https://www150.statcan.gc.ca/t1/tb11/en/tv.action?pid=1310076801> and <https://www150.statcan.gc.ca/t1/tb11/en/tv.action?pid=1710000501>

Chile: <https://deis.minsal.cl/#datosabiertos> (*DEFUNCIONES_FUENTE_DEIS_2016_2021_23092021.zip*)

England and Wales: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales> and

<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>

Germany: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Sterbefaelle-Lebenserwartung/Tabellen/sonderauswertung-sterbefaelle.html>

New Zealand: <https://www.stats.govt.nz/experimental/covid-19-data-portal> (time series for category “Total deaths (all causes)” and indicator “Weekly deaths by age and sex”)

Northern Ireland: <https://www.nisra.gov.uk/publications/historical-weekly-deaths-data> (*Historical Weekly Deaths, 2011–2020*) and <https://www.nisra.gov.uk/publications/weekly-death-statistics-northern-ireland-2021> (*Weekly Deaths Tables – Week ending 3 September 2021*)

Scotland: <https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/weekly-and-monthly-data-on-births-and-deaths/deaths-involving-coronavirus-covid-19-in-scotland/related-statistics> (*Weekly deaths by location of death, age group, sex and cause, 2020 and 2021* and *Weekly deaths by sex and age group, 2000 to 2019*)

South Korea: https://kosis.kr/covid_eng/statistics_excessdeath.do and <https://mdis.kostat.go.kr/index.do>

United States: <https://data.cdc.gov/NCHS/Weekly-counts-of-deaths-by-jurisdiction-and-age-gr/y5bj-9g5w> and https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm

Data on temperature and gridded population

<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5> and

<https://sedac.ciesin.columbia.edu/data/collection/gpw-v4>

Extended data

The computer code for the Bayesian model ensemble used in this study is available at: https://github.com/vkontis/excess_mortality/tree/pub2

Archived analysis code at time of publication: <https://doi.org/10.5281/zenodo.5535829>¹².

License: GNU Affero General Public License v3.0

References

- Kontis V, Bennett JE, Rashid T, *et al.*: **Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries.** *Nat Med.* 2020; **26**(12): 1919–28. [PubMed Abstract](#) | [Publisher Full Text](#)
- Beaney T, Clarke JM, Jain V, *et al.*: **Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide?** *J R Soc Med.* 2020; **113**(9): 329–34. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Douglas M, Katikireddi SV, Taulbut M, *et al.*: **Mitigating the wider health effects of covid-19 pandemic response.** *BMJ.* 2020; **369**: m1557. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- The Academy of Medical Sciences: **Coronavirus: preparing for challenges this winter.** 2020. [Reference Source](#)
- Han E, Tan MMJ, Turk E, *et al.*: **Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe.** *Lancet.* 2020; **396**(10261): 1525–34. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jung AS, Haldane V, Neill R, *et al.*: **National responses to covid-19: drivers, complexities, and uncertainties in the first year of the pandemic.** *BMJ.* 2021; **375**: e068954. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tan MMJ, Neill R, Haldane V, *et al.*: **Assessing the role of qualitative factors in pandemic responses.** *BMJ.* 2021; **375**: e067512. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wu S, Neill R, De Foo C, *et al.*: **Aggressive containment, suppression, and mitigation of covid-19: lessons learnt from eight countries.** *BMJ.* 2021; **375**: e067508. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mathieu E, Ritchie H, Ortiz-Ospina E, *et al.*: **A global database of COVID-19 vaccinations.** *Nat Hum Behav.* 2021; **5**(7): 947–953. [PubMed Abstract](#) | [Publisher Full Text](#)
- Office for National Statistics: **Quarterly mortality report, England: October to December 2019 and year-end review.** 2019. [Reference Source](#)
- European Centre for Medium-Range Weather Forecasts: **ERA5 Reanalysis.** Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. 2017. [Publisher Full Text](#)
- Vkontis: **vkontis/excess_mortality: covid_impacts_2020-21 (v1.0).** *Zenodo.* 2021. <http://www.doi.org/10.5281/zenodo.5535829>
- Kontis V, Bennett JE, Mathers CD, *et al.*: **Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble.** *Lancet.* 2017; **389**(10076): 1323–35. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Parks RM, Bennett JE, Foreman KJ, *et al.*: **National and regional seasonal dynamics of all-cause and cause-specific mortality in the USA from 1980 to 2016.** *eLife.* 2018; **7**: e35500. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Feinstein CA: **Seasonality of Deaths in the U.S. by Age and Cause.** *Demogr Res.* 2002; **6**(17): 471–88. [Publisher Full Text](#)
- Fowler T, Southgate RJ, Waite T, *et al.*: **Excess Winter Deaths in Europe: a multi-country descriptive analysis.** *Eur J Public Health.* 2015; **25**(2): 339–45. [PubMed Abstract](#) | [Publisher Full Text](#)
- McKee CM: **Deaths in winter: Can Britain learn from Europe?** *Eur J Epidemiol.* 1989; **5**(2): 178–82. [PubMed Abstract](#) | [Publisher Full Text](#)
- Hyndman RJ, Athanasopoulos G: **Forecasting: principles and practice.** 2nd edition. Melbourne, Australia: OTexts; 2018. [Reference Source](#)
- Congdon P: **Applied Bayesian Modelling.** John Wiley & Sons; 2003. [Reference Source](#)
- Basu R: **High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008.** *Environ Health.* 2009; **8**: 40. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Basu R, Samet JM: **Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence.** *Epidemiol Rev.* 2002; **24**(2): 190–202. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gasparrini A, Guo Y, Hashizume M, *et al.*: **Mortality risk attributable to high and low ambient temperature: a multicountry observational study.** *Lancet.* 2015; **386**(9991): 369–75. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Song X, Wang S, Hu Y, *et al.*: **Impact of ambient temperature on morbidity and mortality: An overview of reviews.** *Sci Total Environ.* 2017; **586**: 241–54. [PubMed Abstract](#) | [Publisher Full Text](#)
- Bennett JE, Blangiardo M, Fecht D, *et al.*: **Vulnerability to the mortality effects of warm temperature in the districts of England and Wales.** *Nat Clim Change.* 2014; **4**(4): 269–73. [Publisher Full Text](#)
- Parks RM, Bennett JE, Tamura-Wicks H, *et al.*: **Anomalously warm temperatures are associated with increased injury deaths.** *Nat Med.* 2020; **26**(1): 65–70. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sørbye SH, Rue H: **Penalised Complexity Priors for Stationary Autoregressive Processes.** *J Time Ser Anal.* 2017; **38**(6): 923–35. [Publisher Full Text](#)
- Rue H, Martino S, Chopin N: **Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.** (with discussion) *J R Stat Soc Series B.* 2009; **71**(2): 319–92. [Publisher Full Text](#)
- Timmermann A: **Chapter 4 Forecast combinations.** In: Elliott G, Granger C, Timmermann A, eds. *Handbook of Economic Forecasting.* 2006; **1**: 135–96. [Publisher Full Text](#)
- Graefe A, Küchenhoff H, Stierle V, *et al.*: **Limitations of Ensemble Bayesian Model Averaging for forecasting social science problems.** *Int J Forecast.* 2015; **31**(3): 943–51. [Publisher Full Text](#)
- Jones N: **How coronavirus lockdowns stopped flu in its tracks.** *Nature.* 2020. [PubMed Abstract](#) | [Publisher Full Text](#)
- Klein A: **Australia sees huge decrease in flu cases due to coronavirus measures.** *NewScientist.* 2020. [Reference Source](#)
- Olsen SJ, Azziz-Baumgartner E, Budd AP, *et al.*: **Decreased influenza activity during the COVID-19 pandemic - United States, Australia, Chile, and South Africa, 2020.** *MMWR Morb Mortal Wkly Rep.* 2020; **69**(37): 1305–9. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Peek K: **Flu Has Disappeared Worldwide during the COVID Pandemic.** Scientific American. 2021. [Reference Source](#)
- Lewis M: **The Premonition: A Pandemic Story.** Penguin Books, Limited; 2021. [Reference Source](#)
- Young C: **'Significant number' of young people are dying at home with COVID-19.** *National Post.* 2021. [Reference Source](#)
- Statistics Canada: **Provisional death counts and excess mortality, January 2020 to June 2021.** 2021. [Reference Source](#)
- Davies B, Parkes BL, Bennett J, *et al.*: **Community factors and excess mortality in first wave of the COVID-19 pandemic in England.** *Nat Commun.* 2021; **12**(1): 3755. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Suleman M, Sonthalia S, Webb C, *et al.*: **Unequal pandemic, fairer recovery.** The Health Foundation, 2021. [Reference Source](#)
- Stokes AC, Lundberg DJ, Elo IT, *et al.*: **COVID-19 and excess mortality in the United States: A county-level analysis.** *PLoS Med.* 2021; **18**(5): e1003571. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Karlinsky A, Kobak D: **Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset.** *eLife.* 2021; **10**: e69336. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gelman A: **Doubting the IHME claims about excess deaths by country.** Statistical Modeling, Causal Inference, and Social Science; 2021. [Reference Source](#)
- Dyer O: **Covid-19: Study claims real global deaths are twice official figures.** *BMJ.* 2021; **373**: n1188. [PubMed Abstract](#) | [Publisher Full Text](#)
- There have been 7m-13m excess deaths worldwide during the pandemic.** The Economist. 2021. [Reference Source](#)
- Office of National Statistics (ONS): **Comparisons of all-cause mortality between European countries and regions 2020.** 2021. [Reference Source](#)
- Okeill LC, Verity R, Watson OJ, *et al.*: **Have deaths from COVID-19 in Europe plateaued due to herd immunity?** *Lancet.* 2020; **395**(10241): e110–e111. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Haldane V, De Foo C, Abdalla SM, *et al.*: **Health systems resilience in managing the COVID-19 pandemic: lessons from 28 countries.** *Nat Med.* 2021; **27**(6): 964–80. [PubMed Abstract](#) | [Publisher Full Text](#)
- Lewis D: **Why many countries failed at COVID contact-tracing — but some got it right.** *Nature.* 2020; **588**(7838): 384–387. [PubMed Abstract](#) | [Publisher Full Text](#)
- Legido-Quigley H, Asgari N, Teo YY, *et al.*: **Are high-performing health systems resilient against the COVID-19 epidemic?** *Lancet.* 2020; **395**(10227): 848–50. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

49. Lee JK, Bullen C, Ben Amor Y, *et al.*: **Institutional and behaviour-change interventions to support COVID-19 public health measures: a review by the Lancet Commission Task Force on public health measures to suppress the pandemic.** *Int Health.* 2021; **13**(5): 399–409.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Oh J, Lee HY, Khuong QL, *et al.*: **Mobility restrictions were associated with reductions in COVID-19 incidence early in the pandemic: evidence from a real-time evaluation in 34 countries.** *Sci Rep.* 2021; **11**(1): 13717.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
51. Moon J, Chekar CK, Barberá D, *et al.*: **Optimising 'Test and Trace' Systems: Early lessons from a comparative analysis of six countries.** *SSRN Electronic Journal.* 2020.
[Publisher Full Text](#)
52. Violato C, Violato EM, Violato EM: **Impact of the stringency of lockdown measures on covid-19: A theoretical model of a pandemic.** *PLoS One.* 2021; **16**(10): e0258205.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. Hirsch C: **Europe's coronavirus lockdown measures compared.** *Politico.* 2020.
[Reference Source](#)
54. Horton R: **The COVID-19 Catastrophe: What's Gone Wrong and How To Stop It Happening Again.** Wiley; 2021.
[Reference Source](#)
55. Woolhouse M: **The Year the World Went Mad: A Scientific Memoir from the Pandemic.** Sandstone Press Limited; 2022.
[Reference Source](#)
56. House of Commons Health and Social Care, Science and Technology Committees: **Coronavirus: lessons learned to date.** House of Commons; 2021.
[Reference Source](#)
57. Hale T, Angrist N, Hale AJ, *et al.*: **Government responses and COVID-19 deaths: Global evidence across multiple pandemic waves.** *PLoS One.* 2021; **16**(7): e0253116.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
58. Webb E, Winkelmann J, Scarpetti G, *et al.*: **Lessons learned from the Baltic countries' response to the first wave of COVID-19.** *Health Policy.* 2021; **S0168-8510(21)00293-1**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
59. Rajan S, Cylus JD, Mckee M: **What do countries need to do to implement effective 'find, test, trace, isolate and support' systems?** *J R Soc Med.* 2020; **113**(7): 245–250.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
60. Lewis T: **How the U.S. Pandemic Response Went Wrong—and What Went Right—during a Year of COVID.** Scientific American. 2021.
[Reference Source](#)
61. Henley J: **'A once-in-a-generation event': lessons from a year of lockdown in Europe.** The Guardian. 2021.
[Reference Source](#)
62. Hale T, Webster S, Petherick A, *et al.*: **Oxford COVID-19 Government Response Tracker.** Blavatnik School of Government, University of Oxford. 2020.
[Reference Source](#)
63. Baker M, Kvalsvig A, Verrall AJ, *et al.*: **New Zealand's elimination strategy for the COVID-19 pandemic and what is required to make it work.** *N Z Med J.* 2020; **133**(1512): 10–14.
[PubMed Abstract](#)
64. Yoo KJ, Kwon S, Choi Y, *et al.*: **Systematic assessment of South Korea's capabilities to control COVID-19.** *Health Policy.* 2021; **125**(5): 568–576.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. Scudellari M: **How Iceland hammered COVID with science.** *Nature.* 2020; **587**(7835): 536–539.
[PubMed Abstract](#) | [Publisher Full Text](#)
66. Zastrow M: **South Korea is reporting intimate details of COVID-19 cases: has it helped?** *Nature.* 2020.
[PubMed Abstract](#) | [Publisher Full Text](#)
67. Kim JH, An JAR, Oh SJ, *et al.*: **Emerging COVID-19 success story: South Korea learned the lessons of MERS.** Our World in Data - Exemplars in Global Health platform. Our World in Data; 2021.
[Reference Source](#)
68. You J: **Lessons From South Korea's Covid-19 Policy Response.** *Am Rev Public Adm.* 2020; **50**(6–7): 801–8.
[Publisher Full Text](#)
69. Lim B, Kyoungseo Hong E, Mou J, *et al.*: **COVID-19 in Korea: Success Based on Past Failure.** *Asian Economic Papers.* 2021; **20**(2): 41–62.
[Publisher Full Text](#)
70. Klimovsky D, Nemeč J: **The Impacts of the COVID-19 Pandemic and Responses from Various Policy Actors in the Czech Republic and Slovakia in 2020: An Introduction to a Special Issue.** *SciPap.* 2021; **29**(1): 1255.
[Publisher Full Text](#)
71. **Coronavirus: Czechs hold 'farewell party' for pandemic.** BBC News, 2021.
[Reference Source](#)
72. Lemey P, Ruktanonchai N, Hong SL, *et al.*: **Untangling introductions and persistence in COVID-19 resurgence in Europe.** *Nature.* 2021; **595**(7869): 713–717.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
73. Ledford H: **Why do COVID death rates seem to be falling?** *Nature.* 2020; **587**(7833): 190–2.
[PubMed Abstract](#) | [Publisher Full Text](#)
74. Dennis JM, McGovern AP, Vollmer SJ, *et al.*: **Improving Survival of Critical Care Patients With Coronavirus Disease 2019 (COVID-19): a National Cohort Study, March to June 2020.** *Crit Care Med.* 2021; **49**(2): 209–14.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
75. Chalmers JD, Crichton ML, Goeminne PC, *et al.*: **Management of hospitalised adults with coronavirus disease 2019 (COVID-19): a European Respiratory Society living guideline.** *Eur Respir J.* 2021; **57**(4): 2100048.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
76. Chung SC, Marlow S, Tobias N, *et al.*: **Lessons from countries implementing find, test, trace, isolation and support policies in the rapid response of the COVID-19 pandemic: a systematic review.** *BMJ Open.* 2021; **11**(7): e047832.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. The National Academy for State Health Policy: **2020 COVID-19 State Restrictions, Re-openings, and Mask Requirements.** 2021.
[Reference Source](#)
78. Public Health England, Cambridge University MRC Biostatistics Unit: **Direct and indirect impact of the vaccination programme on COVID-19 infections and mortality.** Public Health England; 2021.
[Reference Source](#)
79. Thompson MG, Burgess JL, Naleway AL, *et al.*: **Prevention and Attenuation of Covid-19 with the BNT162b2 and mRNA-1273 Vaccines.** *N Engl J Med.* 2021; **385**(4): 320–329.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
80. Jara A, Undurraga EA, González C, *et al.*: **Effectiveness of an Inactivated SARS-CoV-2 Vaccine in Chile.** *N Engl J Med.* 2021; **385**(10): 875–884.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
81. Eyre DW, Taylor D, Purver M, *et al.*: **Effect of Covid-19 Vaccination on Transmission of Alpha and Delta Variants.** *N Engl J Med.* 2022; **NEJMoa2116597**.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 2

Reviewer Report 21 February 2022

<https://doi.org/10.21956/wellcomeopenres.19577.r48703>

© 2022 Gupta R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Rajeev Gupta 

¹ Academic Research Development Unit, Rajasthan University of Health Sciences, Jaipur, Rajasthan, India

² Eternal Heart Care Centre and Research Institute, Jaipur, India

The article is now acceptable for indexing.

Competing Interests: I have contributed to data on non-communicable disease risk factors to NCDRISC (collaboration). Some of the authors in the article are members of that group. I have not been a member of the writing group of any of the articles emanating from NCDRISC.

Reviewer Expertise: Cardiovascular epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 17 February 2022

<https://doi.org/10.21956/wellcomeopenres.19577.r48702>

© 2022 Gómez-Rubio V. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Virgilio Gómez-Rubio 

Department of Mathematics, School of Industrial Engineering-Albacete, Universidad de Castilla-La Mancha, Ciudad Real, Spain

Thanks for your reply and the changes in the manuscript that address all my concerns raised in my previous review.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 January 2022

<https://doi.org/10.21956/wellcomeopenres.19069.r47314>

© 2022 Gómez-Rubio V. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Virgilio Gómez-Rubio 

Department of Mathematics, School of Industrial Engineering-Albacete, Universidad de Castilla-La Mancha, Ciudad Real, Spain

The authors have presented an interesting analysis of the excess mortality in 40 countries during the first year of the COVID-19 pandemic. They have used publicly available data and Bayesian hierarchical models for the analysis. They have also provided links to the data sources as well as a link to a GitHub repository with the R code and datasets to make sure that the results can be replicated. As part of the analysis, they conduct a validation of the model which makes me think that the results obtained (and their conclusions) are reliable. The conclusions are supported by the data and the analysis carried out.

My only concern is about the "model averaging approach" followed by the authors. First of all, why not select a single model using the different model selection criteria available in INLA? Some of the models are nested so it makes me wonder whether the 'extra' components are needed. Most likely the answer is yes, and then the authors should choose the model with the extra components.

Secondly, I agree that following a Bayesian model averaging is useful but I am not sure that this can be achieved by simply "pooling" the samples obtained from INLA (see, for example, Gómez-Rubio et al., 2020¹ where we illustrate how to do BMA with INLA by weighting the different models according to the marginal likelihood). The authors should have weighted the samples according to the values of the marginal likelihood for each model (which are reported by INLA). These values could have been used to compute the posterior probability of each model for model selection as well. Having said this, it looks like the results produced are similar to those provided by other analyses, which may indicate that weighting the samples have no major impact (i.e., all models produce very similar death estimates), but I think this is something that the authors should check.

References

1. Gómez-Rubio V, Bivand R, Rue H: Bayesian Model Averaging with the Integrated Nested Laplace Approximation. *Econometrics*. 2020; **8** (2). [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Feb 2022

Majid Ezzati, Imperial College London, London, UK

My only concern is about the "model averaging approach" followed by the authors. First of all, why not select a single model using the different model selection criteria available in INLA? Some of the models are nested so it makes me wonder whether the 'extra' components are needed. Most likely the answer is yes, and then the authors should choose the model with the extra components.

This is a good point. We used a model ensemble and model averaging approach for three reasons: First, we have shown in prior work¹ that for forecasts and projections, which by definition involve out-of-sample estimation, a model ensemble/average outperforms the best available model when the projection timeline is longer than a few time units (which would be a few weeks here) – see Appendix Figure 3 in Kontis *et al.*¹. This is consistent with findings from other works that involve forecasting and projections²⁻⁴. Second, which single

model performs best in out-of-sample prediction depends on the country and year, even among those that have related terms, as seen in Appendix Figure 1 of Kontis *et al.*¹. This year-to-year and country-to-country variability in what the best model is, is itself due to the complex behaviour of mortality data over space and time, and means that the best model for the pandemic period is uncertain. In our validation analyses, 13 out of 16 models ranked best (i.e., had the smallest absolute projection error) for at least one country and one validation year. Even when looking at individual countries, in only 2 out of 33 countries did the same model emerge as the best model in all three validation years. What model is best is not only about choosing the one with the most components, but also the choice of prior (e.g., the order of the autoregressive term). Finally, and as a consequence of the aforementioned issues, the use of multiple models and model averaging provides a more complete picture of the inherent uncertainty in how many excess deaths the pandemic has caused than approaches that are not probabilistic or use a single model.

Secondly, I agree that following a Bayesian model averaging is useful but I am not sure that this can be achieved by simply "pooling" the samples obtained from INLA (see, for example, Gómez-Rubio et al., 2020 where we illustrate how to do BMA with INLA by weighting the different models according to the marginal likelihood). The authors should have weighted the samples according to the values of the marginal likelihood for each model (which are reported by INLA). These values could have been used to compute the posterior probability of each model for model selection as well. Having said this, it looks like the results produced are similar to those provided by other analyses, which may indicate that weighting the samples have no major impact (i.e., all models produce very similar death estimates), but I think this is something that the authors should check.

Although in-sample measures of fit are well-suited for applications where there is an interest in the parameters of the model, for projection tasks like ours, models must be evaluated and weighted based on their out-of-sample performance, which is what they have been designed for. Projections of individual models can be combined in a fully Bayesian approach, by assigning weights based on out-of-sample performance, or simply by assigning uniform weights (i.e., using an unweighted average). In this work, we used the latter approach for the following reasons: First, methods for combining projections in a fully Bayesian manner typically use simpler base models (e.g. simple linear regressions)^{4,5} and in our case, such an approach would not have been feasible due to its computational complexity. Second, calculating weights based on out-of-sample performance requires further validation to decide how weights should be defined (e.g., as the reciprocal of absolute projection error, or in some other way that penalises models with larger errors), as we have done in prior work¹. Validating the choice of model weights in turn requires an additional hold-out period which in our case was not available for many countries, due to the short duration of their time series. Choosing uniform model weights avoids this problem and has been shown to produce equal or superior results to using non-equal weights in many applications^{2,6}.

Nonetheless, based on this comment, we now include a sensitivity analysis on how the different models are weighted. Specifically, in the sensitivity analysis, the number of draws from each model was inversely proportional to the absolute error of prediction in the

validation analyses. As seen, the results are robust to this choice, with the current approach having the advantage of being usable for countries with shorter time series.

References

1. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *Lancet* 2017; **389** (10076): 1323-35.
2. Clemen RT. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* 1989; **5**(4): 559-83.
3. Makridakis S, Winkler RL. Averages of Forecasts: Some Empirical Results. *Management Science* 1983; **29**(9): 987-96.
4. Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review* 2005; **133**(5): 1155-74.
5. Eklund J, Karlsson S. Forecast Combination and Model Averaging Using Predictive Measures. *Econometric Reviews* 2007; **26**(2-4): 329-63.
6. Timmermann A. Chapter 4 Forecast Combinations. In: Elliott G, Granger CWJ, Timmermann A, eds. *Handbook of Economic Forecasting*: Elsevier; 2006: 135-96.

Competing Interests: No competing interests were disclosed.

Reviewer Report 18 November 2021

<https://doi.org/10.21956/wellcomeopenres.19069.r46937>

© 2021 Gupta R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Rajeev Gupta** 

¹ Academic Research Development Unit, Rajasthan University of Health Sciences, Jaipur, Rajasthan, India

² Eternal Heart Care Centre and Research Institute, Jaipur, India

General comments:

1. This is an interesting article on the spread of COVID-19 in 40 industrialized countries. However, the article has limited information on the lessons learned from these studies. This is the major weakness of the study. Qualitative studies should be mentioned.
2. Many books have been written on the topic and it would be useful for the authors to look at qualitative information from some of them. I would specifically mention the following (which I have read (sic) and many others that I could not):
 1. Michael Lewis. *The premonition: a pandemic story*.

2. Niall Ferguson: Doom: the politics of catastrophe.
3. Jeremy Farrar and Anjana Ahuja. Spike: The virus vs the people, the inside story.
4. Adam Kucharski. The rules of contagion.
5. Many journal as well as newspaper articles also provide insights into the failures and successes.

3. Would it be worthwhile to add the delta wave epidemic that swept through India, UK and US and is currently creating havoc in Europe?

Specific comments:**Abstract:**

1. Data from US and US states have been prominently displayed in the Results section but are missing.
2. The conclusion statement does not mention the lessons missed (NPI, delays, etc).

Methods:

1. Reference to the "Our World in Data" website is not provided.
2. The authors have extensively adjusted for the temperature variation data. This is important as this is an important determinant. Two other variables could be important: (a) Age distribution of populations in these countries; and (b) Ambient pollution. Data of both are widely available and it would be interesting to evaluate whether these variables influence the outcomes.
3. Page 7, para 2, line 2: is it mid-February 2021? Please correct.

Results:

1. Do we have data on excess mortality for US states? May be interesting to add (unless it is part of a forthcoming study).
2. No data are presented regarding "lessons learned". May be useful to study the social determinants such as GDPs, income, and health systems.

Discussion:

1. Please discuss the data presented and their limitations, and then focus on the implications of the study.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: I have contributed to data on non-communicable disease risk factors to NCDRISC (collaboration). Some of the authors in the article are members of that group. I have not been a member of the writing group of any of the articles emanating from NCDRISC.

Reviewer Expertise: Cardiovascular epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Feb 2022

Majid Ezzati, Imperial College London, London, UK

1. This is an interesting article on the spread of COVID-19 in 40 industrialized countries. However, the article has limited information on the lessons learned from these studies. This is the major weakness of the study. Qualitative studies should be mentioned.

The revised paper refers to over 25 qualitative and evidence synthesis studies, either across countries or for specific countries.

2. Many books have been written on the topic and it would be useful for the authors to look at qualitative information from some of them. I would specifically mention the following (which I have read (sic) and many others that I could not):

1. **Michael Lewis. The premonition: a pandemic story.**
2. **Niall Ferguson: Doom: the politics of catastrophe.**
3. **Jeremy Farrar and Anjana Ahuja. Spike: The virus vs the people, the inside story.**
4. **Adam Kucharski. The rules of contagion.**
5. **Many journal as well as newspaper articles also provide insights into the failures and successes.**

We are aware of the large number of books, and the even larger number of newspaper articles and opinion pieces, on the COVID-19 pandemic. We have cited some key ones that are directly related to our paper including books as well as newspaper articles.

3. Would it be worthwhile to add the delta wave epidemic that swept through India, UK and US and is currently creating havoc in Europe?

We believe that the Delta and Omicron waves should be the subject of distinct analyses for two reasons: First, as stated in the Introduction of our paper, the effectiveness of vaccination in reducing deaths motivates distinct analyses and interpretation prior to v versus after mass vaccination, with the focus of our paper on the former. Second, because the number of deaths that would be expected had the pandemic not occurred is not directly measurable, its estimation requires projection modelling based on past time-series of weekly deaths. The longer the period of projection since the beginning of the pandemic, the more difficult and uncertain the projections, possibly requiring methods that are different from those used for the first year.

Specific comments:**Abstract:**

Data from US and US states have been prominently displayed in the Results section but are missing.

We have included in the abstract as correctly suggested.

The conclusion statement does not mention the lessons missed (NPI, delays, etc).

We have made the use of NPIs explicit in the abstract's conclusion, from its original implicit use.

Methods:

Reference to the "Our World in Data" website is not provided.

We had initially done this as a hyperlink. We have now also cited the corresponding publication.

The authors have extensively adjusted for the temperature variation data. This is important as this is an important determinant. Two other variables could be important: (a) Age distribution of populations in these countries; and (b) Ambient pollution. Data of both are widely available and it would be interesting to evaluate whether these variables influence the outcomes.

Our statistical models were designed to estimate weekly death rates, had the pandemic not occurred. With this objective, the models do not adjust for temperature or any other variable. Rather the models include, as covariates, variables that are predictors of mortality, vary from week to week (because those that do not vary from week to week are captured by other model terms), and are themselves not affected by the pandemic (so that we make unbiased estimates of counterfactual, no-pandemic mortality).

Temperature was included in our model because it is an established predictor of short-term mortality, and was itself not affected by the pandemic. We did not include air pollution in the model, because the pandemic and the associated policy and behavioural responses, in

particular the reduction in vehicular traffic and industrial activity, led to changes in air pollution^{1,2}. This in turn means that the observed pollution is different from what pollution levels would have been without the pandemic. In contrast, researchers have estimated a short-term pandemic impact of only $\sim 0.01^\circ\text{C}$ on ambient temperature³, which is orders of magnitude smaller than seasonal and interannual variation.

The age distribution of each country is already embedded in past data on deaths and population, and is therefore implicitly used to produce our projections. More details provided below.

As the Reviewer correctly points out, our main results are based on analysis that used data from all ages together, for three reasons: First, at least three terms in our model implicitly take into account any change in age distribution that might occur from week to week, which is the relevant time interval for our work: (1) The model has a random intercept for each week of the year (θ_{week}) which allows deaths to change from week to week, some of which may be partly due to differential changes in age groups; the weekly terms repeat across years although their magnitude can change due to variations in temperature. (2) The autoregressive structure of the model allows these terms to be related to deaths in the previous week which means that phenomena that gradually change the age structure of deaths on a weekly scale (e.g., weekly dynamics of flu) are captured. (3) Similarly, we have a term that captures monthly mortality, also with autoregressive structure at annual scale, which implicitly captures phenomena that change deaths at annual scale in a correlated way. Second, when the time unit of analysis is the week, rather than the year as is the case in most mortality projections, the number of deaths in many country-week-age units is zero or small. Using all ages together can make model fits more stable, hence trading off age-related precision with how well model parameters are estimated. Third, running a single model for all ages means that the uncertainty of the estimates is correctly reported.

That being said, we have also repeated the analysis by age group and compare the estimates with those of the combined model in the section titled *statistical methods*. This comparison shows that the two approaches generate similar estimates of counterfactual no-pandemic mortality, for the aforementioned reasons.

Page 7, para 2, line 2: is it mid-February 2021? Please correct.

Assuming the comment refers to the sentence starting with “We used data on weekly deaths from the start of time series through mid-February 2020 ...”, “mid-February 2020” is correct as estimation of model parameters was based on data before the pandemic.

Results:

Do we have data on excess mortality for US states? May be interesting to add (unless it is part of a forthcoming study).

Figures 2, 4 and 8, and the associated text, present results on excess mortality for US states.

No data are presented regarding “lessons learned”. May be useful to study the social

determinants such as GDPs, income, and health systems.

The countries studied here are all high-income although they differ in health care spending and health service infrastructure⁴. It would be a separate body of work, and a distinct paper, to quantify how various features of health system – from local public health and primary care to specialist care – may have affected the outcome of the pandemic. We now discuss this issue at the beginning of the Discussion section, with reference to a number of studies that have qualitatively and quantitatively analysed the role of both health system resilience and early social distancing measures on the impacts of pandemic.

Discussion:

Please discuss the data presented and their limitations, and then focus on the implications of the study.

We have reorganised the text as suggested.

References

1. Venter ZS, Aunan K, Chowdhury S, Lelieveld J. COVID-19 lockdowns cause global air pollution declines. *Proc Natl Acad Sci U S A* 2020; **117**(32): 18984-90.
2. Liu F, Page A, Strode SA, et al. Abrupt decline in tropospheric nitrogen dioxide over China after the outbreak of COVID-19. *Sci Adv* 2020; **6**(28): eabc2992.
3. Forster PM, Forster HI, Evans MJ, et al. Current and future global climate impacts resulting from COVID-19. *Nature Climate Change* 2020; **10**(10): 913-9.
4. Haldane V, De Foo C, Abdalla SM, et al. Health systems resilience in managing the COVID-19 pandemic: lessons from 28 countries. *Nat Med* 2021; **27**(6): 964-80.

Competing Interests: No competing interests were disclosed.
