

RESEARCH

Open Access



# Country-specific determinants for COVID-19 case fatality rate and response strategies from a global perspective: an interpretable machine learning framework

Cui Zhou<sup>1,2</sup>, Åsa M. Wheelock<sup>3</sup>, Chutian Zhang<sup>1,2,7</sup>, Jian Ma<sup>1,2</sup>, Zhichao Li<sup>4</sup>, Wannian Liang<sup>1,2\*</sup>, Jing Gao<sup>1,3,5,6\*</sup> and Lei Xu<sup>1,2\*</sup>

## Abstract

**Background** There are significant geographic inequities in COVID-19 case fatality rates (CFRs), and comprehensive understanding its country-level determinants in a global perspective is necessary. This study aims to quantify the country-specific risk of COVID-19 CFR and propose tailored response strategies, including vaccination strategies, in 156 countries.

**Methods** Cross-temporal and cross-country variations in COVID-19 CFR was identified using extreme gradient boosting (XGBoost) including 35 factors from seven dimensions in 156 countries from 28 January, 2020 to 31 January, 2022. SHapley Additive exPlanations (SHAP) was used to further clarify the clustering of countries by the key factors driving CFR and the effect of concurrent risk factors for each country. Increases in vaccination rates was simulated to illustrate the reduction of CFR in different classes of countries.

**Findings** Overall COVID-19 CFRs varied across countries from 28 Jan 2020 to 31 Jan 31 2022, ranging from 68 to 6373 per 100,000 population. During the COVID-19 pandemic, the determinants of CFRs first changed from health conditions to universal health coverage, and then to a multifactorial mixed effect dominated by vaccination. In the Omicron period, countries were divided into five classes according to risk determinants. Low vaccination-driven class (70 countries) mainly distributed in sub-Saharan Africa and Latin America, and include the majority of low-income countries (95.7%) with many concurrent risk factors. Aging-driven class (26 countries) mainly distributed in high-income European countries. High disease burden-driven class (32 countries) mainly distributed in Asia and North America. Low GDP-driven class (14 countries) are scattered across continents. Simulating a 5% increase in vaccination rate resulted in CFR reductions of 31.2% and 15.0% for the low vaccination-driven class and the high disease burden-driven class, respectively, with greater CFR reductions for countries with high overall risk (SHAP value > 0.1), but only 3.1% for the ageing-driven class.

\*Correspondence:

Wannian Liang

liangwn@tsinghua.edu.cn

Jing Gao

jing.gao@ki.se

Lei Xu

xu\_lei@tsinghua.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusions** Evidence from this study suggests that geographic inequities in COVID-19 CFR is jointly determined by key and concurrent risks, and achieving a decreasing COVID-19 CFR requires more than increasing vaccination coverage, but rather targeted intervention strategies based on country-specific risks.

**Keywords** COVID-19, Global health, Strategy, Vaccination, Case fatality rate, Pandemics, XGBoost, SHAP

## Introduction

The severe disease burden caused by COVID-19 will continue to pose a challenge to global public health systems for the foreseeable future [1–3]. As of April 2023, the pandemic has caused more than 700 million confirmed infections and over six million deaths [4]. Vaccination programs have been widely implemented around the world, but while the surge in cases and deaths has been reduced to a certain extent, it is not yet fully controlled, and inequalities in vaccine distribution have emerged [5, 6]. Health outcomes for COVID-19, including case fatality rate (CFR), vary widely across countries and could be determined by country-specific risk factors. The determinants of cross-country variation in CFRs during a COVID-19 pandemic, in the context of multiple confounding factors, are unclear. Meanwhile, there is as yet a lack of evaluation of the benefits of vaccination across countries from a global perspective, and elucidating the extent to which countries will benefit from vaccination would provide the basis for global vaccine distribution. Therefore, understanding the risk features that affect COVID-19 CFRs is critical to guide global vaccine distribution to effectively reduce CFRs.

Notably, the cross-country variation in COVID-19 CFR differs from previous patterns of infectious disease, with even geographically contiguous countries exhibiting considerable difference in CFRs. Thus, COVID-19 CFRs are widely considered to be influenced by multidimensional factors. Previous studies have tried to explain cross-country variation in COVID-19 CFR using a variety of unidimensional factors such as population age structure [7, 8], comorbidities [9, 10], medical resources [11], environment [12], culture, and so on [13]. While these studies have found some associations, they have also ignored the important interaction effects of these factors on the risk of COVID-19 death within a single country. In addition, some studies have identified complex risk factors with relevance to a single region or time period, but their findings are difficult to generalise due to that same geographical or temporal specificity [14–16]. In addition, existing studies mostly used a linear approach to explain the effects of risk factors, thereby ignoring potential non-linear effects. Building on previous research, we recognise that COVID-19 CFRs are regulated by complex factors

and that identifying potential risk factors from mixed effects at the country level will provide complementary evidence for future pandemic responses.

Fast-evolving machine learning algorithms provide better analytical capabilities for real-world health emergencies. Extreme Gradient Boosting (XGBoost) is a highly optimised gradient boosting framework based on decision trees, where the algorithm iteratively combines the predictions of multiple weak learners to generate more powerful and robust models [17]. It has been widely used in medicine, chemistry, ecology, finance and other fields. Its diverse objective functions, ability to handle missing values, inclusion of regularisation terms, and easier identification of non-linear effects make it suitable for real-world health research [18]. SHapley Additive exPlanations (SHAP) is a well-established algorithm that provides a visual interpretation of the model results [19]. It can quantify the global contribution of each factor in a machine learning model, showing the direction and magnitude of each factor's effect, as well as breaking down a prediction to show how much each factor contributes to a predicted value. This enables both identification of universal risk factors in a global perspective and precise identification of each country-specific risk and its risk intensity.

Here, our study aims to identify national heterogeneity in risk factors for COVID-19 CFRs and quantify potential risks in 156 countries through the SHAP-interpreted XGboost algorithm, providing better exploratory insights into future joint interventions for the control of CFRs.

## Method

### Overview

The overall framework of this study is as follows. Firstly, we described the global distribution and epidemiological trends in CFRs, and further evaluated multidimensional features potentially affecting the heterogeneity of CFRs, including vaccination coverage, demographic factors, disease burden, behavioural risk factors, environmental risk factors, health services, and trust levels. Then, we constructed high-performance XGboost models and applied SHAP to explain those models and identify the important features affecting CFR across countries during different periods of the pandemic. After that, we clarified the country-specific risk factors for each country and

their protective and risk effects on the CFR, and grouped countries into five clusters according to key risk factors. Finally, to evaluate the benefit of increasing vaccination rate on future CFR, we further simulated the change in CFR following an increase of the vaccination rate in each country.

This study complies with the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) recommendations (Supplementary material 2.1).

### COVID-19 CFRs

Daily confirmed infections and deaths in 156 countries over the period of 28 Jan 2020 to 31 Jan 2022 were extracted from Our World in Data (OWID) [20]. Weekly CFRs were calculated from the number of new deaths and new cases per week. As there is a time-lag between deaths and cases, determined by cross-correlation analysis to be 12 days in length, we lagged the daily new deaths by 12 days to calculate the lag-adjusted weekly CFRs; we also removed countries for which less than 12 days of data were available (Supplementary material 3.1).

### SARS-CoV-2 lineage data

SARS-CoV-2 lineage data were obtained from an integrated global SARS-CoV-2 database, the China National Center for Bioinformatics (CNCB), which includes data from the Global Initiative on Sharing All Influenza Data (GISAID), NCBI GenBank, National Genomics Data Center (NGDC), National Microbiology Data Center (NMDC), and China National GeneBank (CNGB). This database also provides variants identified from these sequences [21]. For each day over the study period, we determined which variant types accounted for more than 70% of all detected sequences globally, and we classified variants that met that standard as having a worldwide dominance. We defined the period of a variant's dominance as spanning from the time when the WHO defined it as a variant of concern (VOC) to the time when the next VOC appeared in no more than 10% of countries. The COVID-19 pandemic was thus divided into four periods, including the ancestral variant dominance period (original period) from 28 January to 17 December 2020, the Alpha variant dominance period (Alpha period) from 18 December 2020 to 6 April 2021, the Delta variant dominance period (Delta period) from 11 May to 21 November 2021, and the Omicron variant dominance period (Omicron period) from 26 November 2021 to 31 January 2022.

### Vaccination data

Daily vaccination data from January 28, 2020 to January 3, 2022 were extracted from OWID and pre-processed by linear interpolation in 156 countries [22]. Vaccination

status was defined according to whether the last dose had been received within six months, since the protection offered by the COVID-19 vaccine drops sharply after six months [23, 24]. Vaccination rates were further organised into two categories: the proportion of the population having completed the initial vaccination protocol within six months (fully vaccinated) and that having received a booster within six months (booster given).

### Multi-dimensional explanatory variables

To comprehensively assess the risk factors influencing COVID-19 CFR, we included 35 features in six dimensions that are known or thought to affect CFRs (Table 1): demographic characteristics, national disease burden, behavioural risk factors, environmental risk factors, level of national health services, and level of trust.

### XGboost

#### Model building

To develop explanatory and predictive models, we employed XGBoost algorithm to capture the non-linear associations between COVID-19 CFRs and multiple dimensional features. XGBoost is an ensemble machine learning method based on decision trees that applies a gradient boosting framework [18]. It creates a robust, more accurate prediction model from an ensemble of weak prediction models and incorporates a penalty term for model complexity to improve performance. The objective function of the XGBoost algorithm is as follows:

$$Obj(\theta) = L(\theta) + \Omega(\theta) = \sum_i L(\hat{y}_i, y_i) + \sum_k \Omega(f_k), f_k \in F$$

where  $L$  is the training loss function.  $L(\hat{y}_i, y_i)$  corresponds to the training loss function for each sample, where  $y_i$  indicates the true value of the  $i$  sample and  $\hat{y}_i$  indicates the estimated value of the  $i$  sample.  $\Omega$  is regularization function that measures the model's complexity, where  $k$  is the number of trees,  $F$  is the set of all possible regression trees.

#### Feature selection

We filter the main features using the Recursive Feature Elimination (RFE) algorithm, which aims to capture CFR variations while retaining as few features as possible. The RFE strategy uses all the features to train the supervised model and then evaluates the features according to their importance in the model [25]. The detailed steps include: (1) Initialisation: all features are used to train the supervised model. (2) Feature importance evaluation: based on the importance of the features in the model, the least important features are selected for elimination. (3) Model update: retrain the model using the dataset with one feature removed. (4) Determine stopping condition: check

**Table 1** The list of covariates used in analyses

Short name	Dimension	Definition	Temporal coverage	Data source	Missing	Median (IQR)
Fully vaccinated	Vaccination coverage	Proportion of the population completing the initial vaccination protocol within six months	2020–2022	Our World in Data [22]	0/156	27.3 (11.4–40.0)
Booster given		Proportion of the population that received a booster dose within six months	2020–2022	Our World in Data [22]	0/156	3.9 (0–18.6)
Aged 65 +	Demographic characteristics	Proportion of population aged 65 and over	2017	UN Population Division [45]	3/156	6.6 (3.5–14.7)
Gender ratio		Number of males born per 100 females	2020	World Development Indicators—World Bank [46]	0/156	105.1 (103.9–105.9)
Average years of schooling		Average number of years the population older than 25 participated in formal education	2017	Our World in Data [47]	2/156	9.2 (6.3–11.3)
GDP per capita		Per capita gross domestic product, a measure of a country's economic output per person	2018	World Bank [48]	4/156	13,567.3 (4796.5–28,942.6)
Lower respiratory infections	Disease burden	Age-standardised prevalence of lower respiratory infections per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	130.7 (84.2–184.6)
Upper respiratory infections		Age-standardised prevalence of upper respiratory infections per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	3228 (2774–3712)
Chronic obstructive pulmonary disease		Age-standardised prevalence of chronic obstructive pulmonary disease per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	2054.3 (1713.3–2665.0)
Cardiovascular diseases		Age-standardised prevalence of cardiovascular diseases per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	6477 (5616–7132)
Stroke		Age-standardised prevalence of stroke per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	1233.6 (958.2–1411.3)
Cancers		Age-standardised prevalence of cancers per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	6157 (4421–9517)
Diabetes		Age-standardised prevalence of diabetes per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	5490 (4243–7292)
Chronic kidney disease		Age-standardised prevalence of chronic kidney disease per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	7871 (6384–9957)
Hypertension		Prevalence of hypertension among adults aged 30–79 years	2019	Global Burden of Disease Study 2019 [49]	1/156	37.3 (31.6–42.1)

**Table 1** (continued)

Short name	Dimension	Definition	Temporal coverage	Data source	Missing	Median (IQR)
Mental disorders		Age-standardised prevalence of mental disorders per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	12,651 (11,365–13,971)
Noncommunicable diseases		Age-standardised total mortality from noncommunicable diseases per 100,000 population	2019	WHO [50]	0/156	90,447 (89,596–91,749)
HIV		Age-standardised prevalence of HIV infection per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	176.8 (46.4–589.7)
Tuberculosis		Age-standardised prevalence of tuberculosis per 100,000 population	2019	Global Burden of Disease Study 2019 [49]	0/156	19,036 (13,148–24,781)
Overweight	Behavioural risk factor	Prevalence of overweight among adults, BMI ≥ 25 (age-standardized estimate) (%)	2016	WHO [51]	2/156	55.6 (30.9–60.6)
Low physical activity		Disability-adjusted life years attributed to low physical activity	2019	Global Burden of Disease Study 2019 [49]	0/156	188.2 (134.2–330.3)
Smoking		Disability-adjusted life years attributed to smoking	2019	Global Burden of Disease Study 2019 [49]	0/156	2024.9 (1517.8–2721.2)
Dietary risks		Disability-adjusted life years attributed to dietary risks	2019	Global Burden of Disease Study 2019 [49]	0/156	2210.2 (1607.3–2950.3)
Trees per capita	Environmental risk factor	Number of trees per capita	2014	Mapping tree density at a global scale [52]	0/156	191.0 (52.2–695.8)
PM <sub>2.5</sub>		Population-weighted exposure to ambient PM <sub>2.5</sub> pollution (µg/m <sup>3</sup> ) is defined as the average level of exposure of a nation's population to concentrations of suspended particles measuring less than 2.5 microns in aerodynamic diameter, which are capable of penetrating deep into the respiratory tract and causing severe health damage	2017	Global Burden of Disease Study 2019 [49]	1/156	22.2 (14.5–36.7)
Mean temperature		Mean temperature (Celsius)	1991–2020	World Bank Climate Change Knowledge Portal [53]	0/156	22.1 (11.3–25.5)
Population density		Population density (people per sq. km of land area)	2020	World Bank [54]	1/156	81.7 (31.8–160.8)
HAQ index	Health service	Healthcare Access and Quality Index	2018	GBD 2016 Healthcare Access and Quality Collaborators [32]	8/156	66.8 (41.5–81.0)

**Table 1** (continued)

Short name	Dimension	Definition	Temporal coverage	Data source	Missing	Median (IQR)
IHR score		Average of 13 International Health Regulations core capacity scores, 1st version of the questionnaire	2019	WHO [55]	0/156	66.5 (49.8–83.0)
Hospital beds		The number of hospital beds available per 10,000 inhabitants	2017	WHO [56]	3/156	20.6 (10.0–36.0)
Expenditure		Current health expenditure per capita in US\$	2019	WHO [57]	4/156	392.5 (75.2–1407.3)
Hospitals		Total density of hospitals per 100,000 population	2013	WHO [58]	49/156	1.0 (0.5–1.8)
Trust the national government	Trust	Trust in the national government of the country	2020	Wellcome Global Monitor Survey [59]	55/156	52.5 (43.0–69.5)
Trust journalists		Trust in journalists in this country	2020	Wellcome Global Monitor Survey [59]	51/156	55.2 (47.4–64.8)
Trust science		Trust in science	2020	Wellcome Global Monitor Survey [59]	50/156	80.5 (70.0–88.8)

GDP, gross domestic product; PM<sub>2.5</sub>, particles measuring less than 2.5 microns in aerodynamic diameter; HAQ, Healthcare Access and Quality Index; IHR, International Health Regulations core capacity; WHO, World Health Organization

whether the stopping condition is satisfied; if not, return to step 2; if it is satisfied, go to the next step. (5) Feature selection: select features from the model with better fit. In each iteration, root mean square error (RMSE) is used to evaluate the fit of the model. The model that performs best in the feature elimination process is selected as the final model. Overall, RFE finds the best subset of features for a model by progressively eliminating unimportant features, thereby reducing the number of the features while maintaining the predictive power of the model.

#### **Hyperparameter tuning**

The optimal set of hyperparameter values was selected using a ten-fold cross-validation grid search. The tuned parameters consisted of learning rate (from 0.05 to 0.2 with an interval of 0.05) and the maximum depth of the tree (from 4 to 10 with an interval of 1). Since our dependent variable of interest was zero-inflated right-skewed data, the objective function was set as 'reg:tweedie'. The training process was stopped when more training cycles failed to enhance the validation dataset's performance. The dataset was split into three parts: 60% for training, 20% for validation, and 20% for testing.  $R^2$  and RMSE were used to assess the model's accuracy.

#### **Simulation**

We predicted the change in CFR under scenarios where booster vaccination rate was increased by 5% in each country. We used the best model parameters derived from the training and validation dataset, and then held all other variables constant, and changed the booster vaccination rate for each country to predict the CFRs. The principle of increasing booster vaccination is based on each country's actual full and booster vaccination rates, so we predicted CFRs for increasing booster vaccination rates within the range of a country's booster vaccination rate not exceeding the cumulative proportion of the population fully vaccinated. This approach ensured that our predictions remained within realistic limits, which reflected the actual limitations of booster vaccination coverage.

#### **Model interpretation**

We used the SHAP framework to rank features according to their importance and explain how features affect the CFR. SHAP is a game theoretic approach that can explain the output of the XGBoost model. It connects the optimal credit allocation with a local explanation using the classical Shapley values from game theory and their associated extensions [19]. The variability of the predictions is assigned to the available features, allowing evaluation of the contribution of each feature to each prediction point. SHAP provides valuable insights into a model's

behaviour by overcoming the main drawback of inconsistency in classical global feature importance measures, minimizes the possibility of underestimating the importance of a feature with a certain attribution value, shows consistency and accuracy in its importance ordering, and interpreting the model's global behaviour while retaining local faithfulness. The overall importance of a feature was scored as the mean absolute value of all SHAP values for that feature, and we considered features scoring 0.1 or higher as important [26–28]. The association between CFR and each key feature was examined via partial dependence plots, which were adjusted for all other confounding variables.

#### **Statistical analysis**

Continuous data are presented as a mean with standard deviation (SD) where normally distributed and as a median with the 25th and 75th percentiles where non-normally distributed. We used Spearman's rank correlation to measure the correlation of CFR with each continuous features, such as booster vaccination rate. Differences in CFRs among four groups of countries with different income levels were tested using analysis of variance (ANOVA), and then differences between pairs of country groups were tested by post-hoc tests using the Bonferroni method.

Analyses were performed in the R 4.1.1 and Python 3.8 environments.

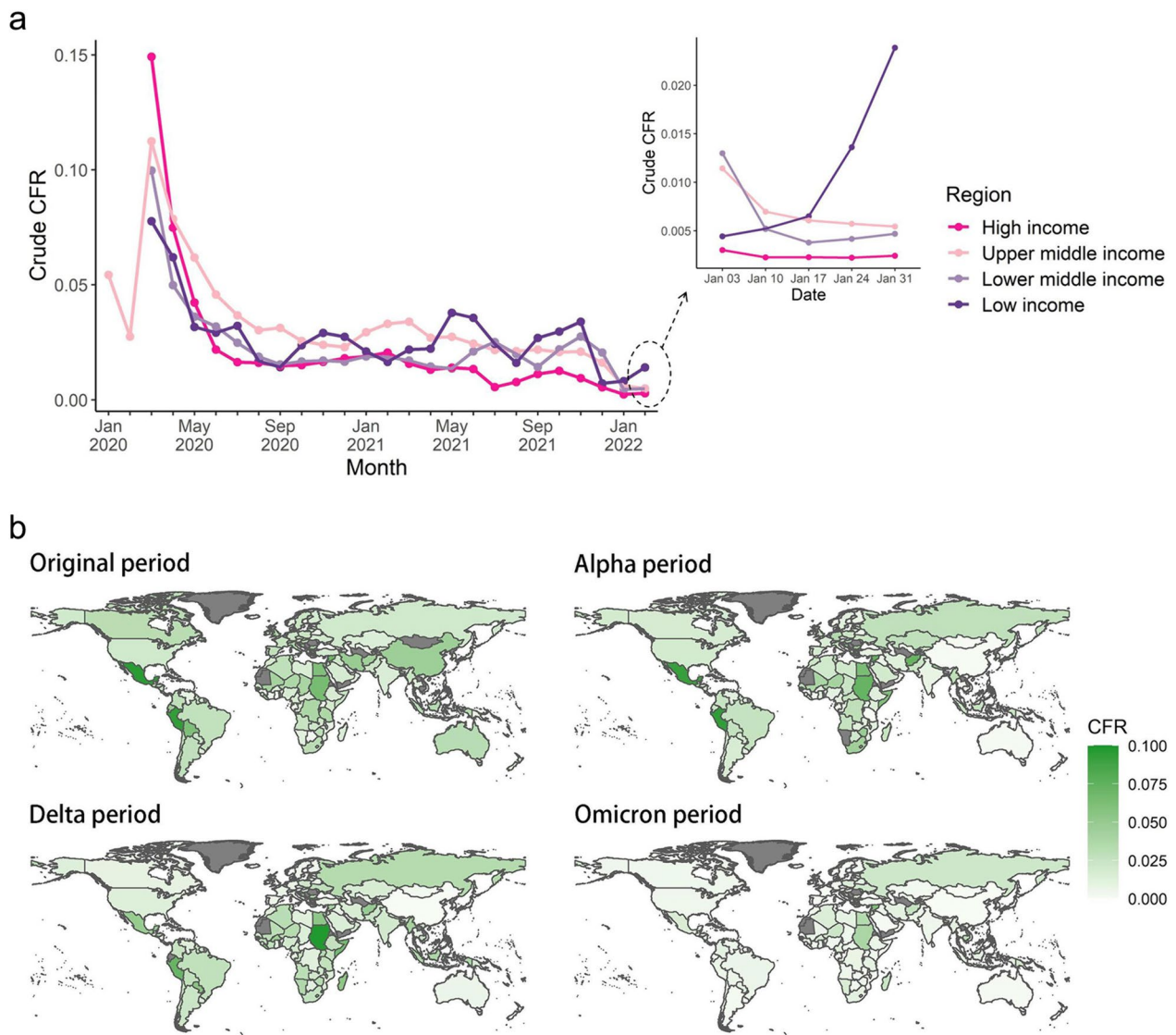
## **Results**

### **Temporal and regional heterogeneity of COVID-19 CFRs**

Overall COVID-19 CFRs varied significantly across countries, ranging from 68 per 100,000 population to 6,373 per 100,000 population. The global CFR exhibited a decreasing trend from January 2020 to January 2022, with respective values of 2.26%, 1.95%, 1.92%, and 0.74% for the original, Alpha, Delta, and Omicron periods (Fig. 1a, b). During the pandemic, CFRs gradually dropped in the high income countries after the first outbreak, while low income countries had relatively high CFRs through the end of the study period. Univariate analyses revealed significant associations with CFR for some factors such as cumulative vaccination rate, but did not satisfactorily explain the differences in CFRs across countries, for example the observation that countries with low vaccination rates always exhibit higher CFRs, but so do some countries with high vaccination rates such as Peru, Ecuador, and Mexico (Supplementary material 3.3).

### **Changes in the determinants of COVID-19 CFRs over the four periods of the pandemic**

Most cross-country variation in CFRs in the Alpha, Delta, and Omicron periods could be well explained by



**Fig. 1** Trends in and distributions of CFR. **a** Epidemiological curves of COVID-19 CFR by WHO region from 28 January 2020 to 31 January 2022. **b** Global distribution of CFR in the original, Alpha, Delta, and Omicron periods

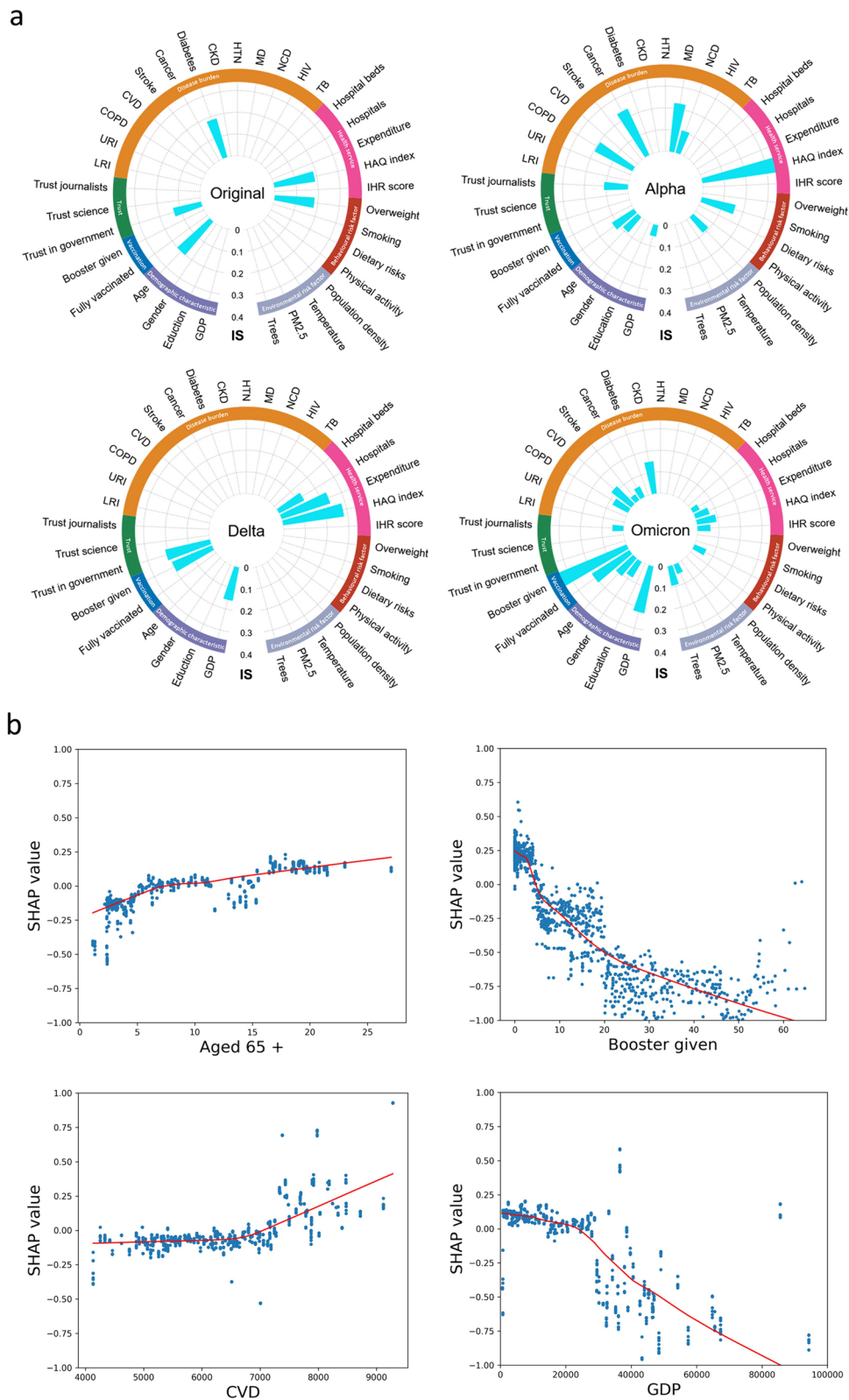
the SHAP-interpreted XGboost model ( $R^2$ : 0.76, 0.62, 0.58, respectively), but only limited interpretation was achieved for the original period ( $R^2$ : 0.33). Important determinants of CFR and their number were found to

vary across periods. From the Alpha period to the Omicron period, the important determinants first changed from health conditions to universal health coverage, and

(See figure on next page.)

**Fig. 2** The importance of each factor affecting CFR and its effects in the original, Alpha, Delta, and Omicron periods. **a** ISs for each feature affecting CFR in each period model, obtained by taking the absolute mean of the SHAP values. The 35 features represent seven distinct dimensions: vaccination coverage, demographic factors, disease burden, behavioural risk factors, environmental risk factors, health services, and trust levels. **b** SHAP dependence plots for proportion of population aged over 65, booster vaccination rate, CVD, and GDP per capita in the XGBoost models. SHAP values above zero represent an increased risk of higher COVID-19 CFR. Abbreviations: IS, important score; LRI, lower respiratory infections; URI, upper respiratory infections; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular diseases; CKD, chronic kidney disease; HTN, hypertension; MD, mental disorders; NCD, noncommunicable diseases; HIV, HIV infection; TB, tuberculosis





then to a multifactorial mixed effect dominated by vaccination (Fig. 2a).

The explanatory plots for each factor affecting CFR (Fig. 2b) indicate vaccination to have been an evident determinant of cross-country variation in CFRs since the Alpha period, and especially important in the Omicron period, with fully vaccinated (importance score (IS): 0.21) and booster given (IS: 0.37) status both showing a strong protective effect. From the Alpha period to the Omicron period, the protective effect of GDP on CFR gradually increased, while the importance of the HAQ index gradually decreased. In addition, ageing (IS: 0.09 and 0.11, respectively) and disease burden (IS: 0.12–0.24) were identified as important factors for increased CFR in the Alpha and Omicron periods, but not in the Delta period. A variety of disease burdens also exhibited important impacts on CFR: chronic obstructive pulmonary disease (COPD), cancers, and mental illness in the Alpha period, and cardiovascular diseases (CVD) and chronic kidney disease (CKD) in the Omicron period. Trust in government and journalists evidenced relative importance to the CFR over all four time periods (IS: 0.05–0.21). In addition, tree cover first appeared as a relatively important factor in the Omicron model.

#### Country-specific determinants and concurrent risks of COVID-19 CFR

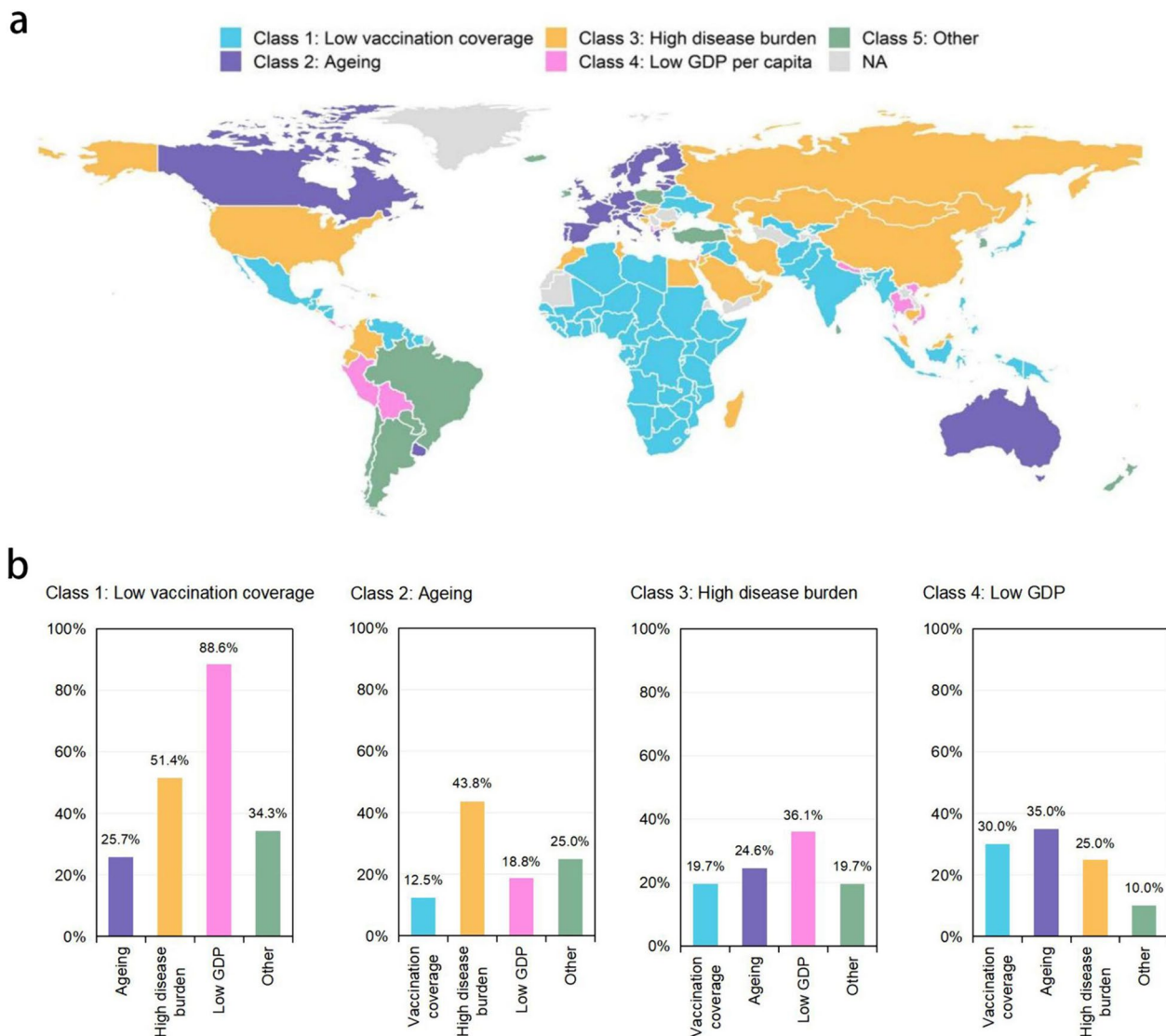
The Omicron period model revealed that of the various determinants of CFR, the main contributors (IS > 0.1) were the population receiving booster doses and full vaccination, GDP per capita, prevalence of chronic kidney disease and cardiovascular disease, and the proportion of the population aged 65 and over. We subsequently grouped the countries into five classes based on these risks: low vaccine coverage, ageing, high disease burden, low GDP, and other (Fig. 3a). For most of the high-income countries the main risk factor is ageing (n = 26, 48.1%), in addition to 10 countries where the main risk factor is high burden of disease (18.5%), while for most of the low-income countries the main risk factor was low vaccination coverage (n = 22, 95.7%).

Figure 4 showed the total risk and the risk of each contributor for each country respectively, with SHAP values less than zero as the protective effect and greater than zero as the risk effect. For countries in Class 1 (n = 70), the main determinant of CFR was low vaccination coverage. This class was mainly comprised of countries in Africa, South East Asia and Latin America. Across all Class 1 countries only 17.1% and 0.4% of people were fully vaccinated and booster given, respectively. The highest risk due to low booster vaccination was in Sudan (SHAP value: 0.40) and due to low full vaccination was in

Niger (SHAP value: 0.48) (Fig. 4). In addition, most countries in Class 1 featured multiple concurrent risk factors: 88.6% were also at risk of low GDP, and some countries (51.4%) such as Syria, Sudan, Afghanistan, and Iraq were at risk of high disease burden (Fig. 3b). For countries in Class 2, the main determinant of CFR was ageing. There are 26 countries in this class, including 23 European high-income countries such as Portugal, Germany, and Finland, as well as Canada, Australia, and Uruguay. On average, the proportion of people aged over 65 was around 19%. Countries in Class 2 had fewer concurrent risks; only seven countries, including Czechia, Estonia, and Lithuania, evidenced risk of high disease burden as a secondary determinant (Fig. 4). For countries in Class 3 (n = 32), the main determinant of CFR was high disease burden, including a high burden of CVD and CKD. Within the class, the average cardiovascular disease prevalence was 7915 per 100,000 and the average chronic kidney disease prevalence was 9,548 per 100,000. The highest risk due to CVD was in Egypt (SHAP value: 0.92), and due to CKD was in Syria (SHAP value: 0.18) (Fig. 4). Countries in Class 3 also faced more concurrent risks, with 68.8% and 46.9% being at risk of low GDP and ageing, respectively (Fig. 3b). For countries in Class 4 (n = 14), the main determinant of CFR was low GDP. This class of countries were scattered globally and characterized by fewer concurrent risks. Finally, for countries in Class 5, the main determinants of CFR comprised other factors of lesser global importance such as health expenditure, trust in journalists, and dietary risks.

#### Future benefits of a 5% increase in vaccination vary by country

When simulating a 5% increase in vaccination, countries showed differing degrees of reduction in CFR (Fig. 5a). For countries in Class 1 and Class 3, where low vaccination rates and high disease burden constitute the main risk factors (Fig. 5b), increasing vaccination produced a greater change in CFR, with median values of 31.2% and 15.0%, respectively. Although most Class 1 countries had a significant reduction in CFRs after modelling increased vaccination rates, there were still some countries where the reduction in CFR was not significant (change rate < 0.1), e.g. Burundi, due to their lower overall risk (median SHAP value for overall risk: -0.79) compared to other countries (median SHAP value for overall risk: 0.19). Conversely, continued increases in vaccination were of limited benefit in ageing countries (Class 2) where vaccination rates were already high, achieving a median change of 3.1%, and also in the low GDP-driven Class 4, for which the median change was 4.8%.

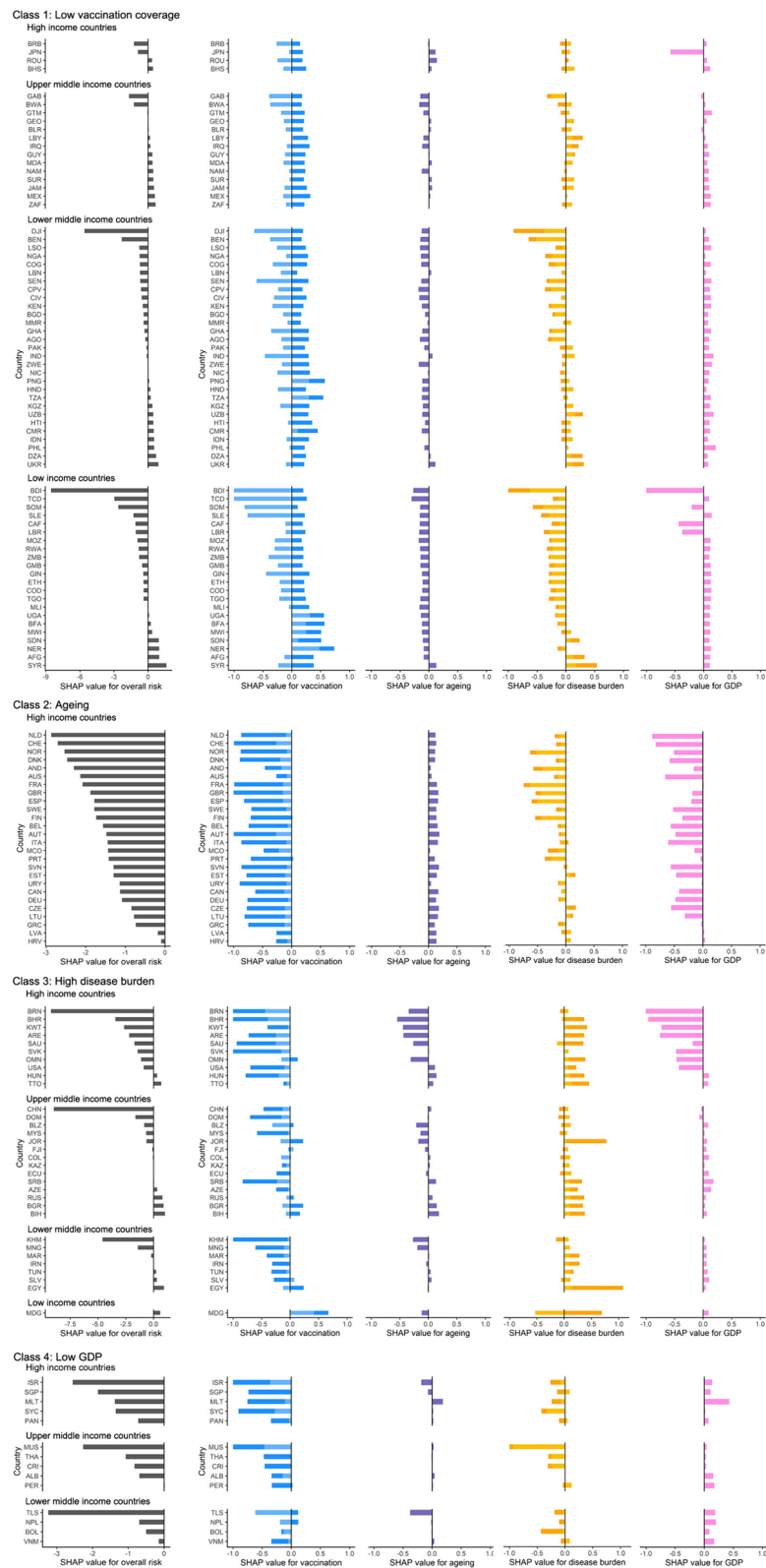


**Fig. 3** Country classification according to the most important risk factors and concurrent risks influencing COVID-19 CFR. **a** Grouping of countries into five classes based on the most important risk factors in the Omicron model. Class 1: low vaccine coverage; Class 2: ageing; Class 3: high disease burden; Class 4: low GDP; Class 5: other. **b** Percentage of countries with certain concurrent risks in each class of countries

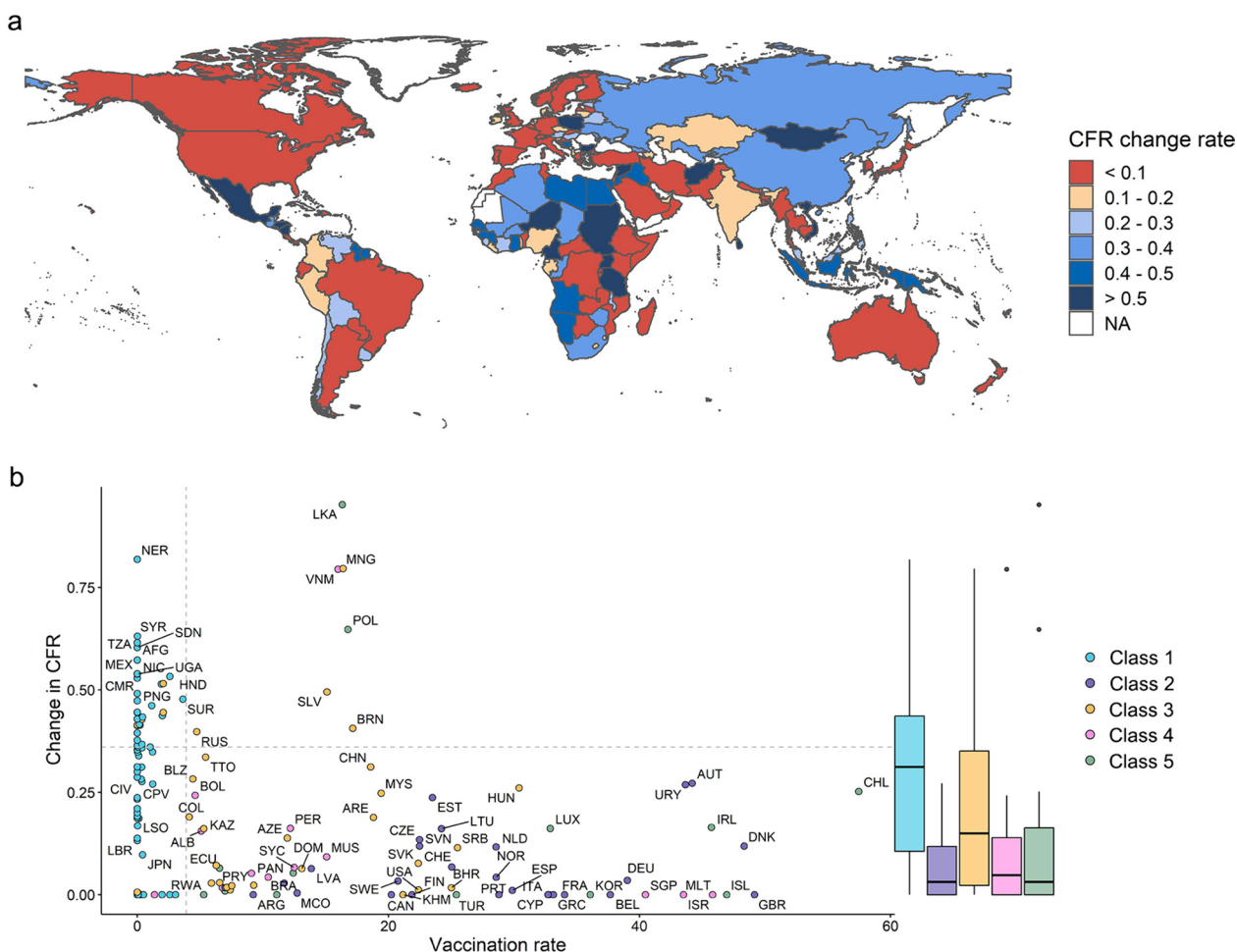
### Discussion

We draw three conclusions from this study. First, across the different variant dominance periods of the pandemic, the important determinants of COVID-19 CFRs changed from health conditions to universal health coverage, and then to a multifactorial mixed effect dominated by vaccination. This different weighting of factors may be due to the distinct characteristics of the respectively dominant SARS-CoV-2 strains. The higher transmissibility of the Delta variant compared to the Alpha variant may lead to its easy transmission even in healthy populations rather than a greater susceptibility in individuals with underlying disease [29]. Thus, changes in the infected population

during the Delta variant period may reduce the impact of disease burden on CFR. Moreover, Delta variants result in a significant increase in the risk of hospitalisation and death in infected individuals, placing a greater burden on the healthcare system [30, 31]. Our analyses suggest that the level of the national health service is a key predictor of CFR during this period, replacing the effects of the disease burden. Adjusting investments to improve access and quality across healthcare needs will not only benefit routine care, but also improve overall health coverage in preparation for the next pandemic [32]. Furthermore, social determinants and public health interventions also affect the association between the disease burden and



**Fig. 4** Overall risk and contributions of main risk factors to the CFR for each country in Classes 1-4. Country abbreviations use the ISO 3166 ALPHA-3 codes [44]



**Fig. 5** Distribution of and cross-class differences in the change in CFR after a simulated 5% increase in vaccination. **a** Global distribution of the predicted change in CFR after a 5% increase in vaccination coverage. **b** Scatter plot showing the change in CFR following increased vaccination versus current booster vaccination rate for each country. The box plot shows the distribution of change in CFR for each cluster, with boxes indicating the median and 25th and 75th percentiles

the CFR. Vulnerable population, such as the elderly and those with underlying diseases, are prioritised for vaccination, they have reduced CFR, which may also result in a reduction in the impact of the disease burden on the country’s CFR [33]. Meanwhile, the COVID-19 pandemic took on a new pattern as a result of the emergence of the Omicron variant [34]. The immune escape characteristics of Omicron make it more contagious than earlier strains, but it also seems to be gentler, typically resulting in less severe disease [35]. In addition to the characteristics of the virus itself, patients during the Omicron period also benefited from the strong protection against severe disease and death still afforded by the COVID-19 vaccine [36]. Our study thus confirms the importance of vaccination, especially booster doses, in reducing the risk of death in Omicron pandemics. Especially in this present stage dominated by the ‘Stealth’ Omicron, BA.2,

during which strict prevention policies are challenged by insidious transmission and the number of infections has become difficult to control, improving vaccination coverage is a cost-effective approach for reducing severe health outcomes and relieving pressure on the healthcare system.

The second major conclusion of this study is that differences in CFRs between countries are driven by effects of country-specific risk factors. Our findings highlight the noteworthy risk factors of COVID-19 death for each country at the current stage, with the most important risks being low vaccination, ageing, high disease burden, and low GDP. Based on the leading risks, we further categorized countries into four classes. Grouping countries in this way will provide joint intervention strategies for real-world policymakers and also help further a coordinated response to the pandemic that balances global and

national benefits. Notably, ageing as a major risk factor was mainly found in high-income developed countries, where vaccination rates are already high and CFRs relatively low; accordingly, in addition to sustaining vaccination rates, policies in the post-COVID era may need to prioritise vulnerable populations such as older people. Similarly, countries with a high disease burden as the main risk, including some like Egypt, Madagascar, and Jordan where vaccine supply is relatively limited, would be better served by adjusting vaccine priority distribution programmes to protect the large number of vulnerable people with underlying diseases. It is also important to provide health education to these populations to enable them to accept vaccines. In another consideration, although the protective effect of vaccines has been widely demonstrated, our results suggest that in countries where low vaccination is a major risk factor, CFRs are also affected by a broad range of concurrent risks; consequently, we believe that a joint intervention would be an effective measure for reducing CFRs in this class of countries. In the short term, in addition to vaccination, a promising area for interventionists to work on is raising the level of national trust. Our findings support previous research that trust in government and science can increase risk perceptions of COVID-19 among the population, promote cooperation with outbreak prevention and control efforts, and more quickly control the number of cases and deaths [37]. Pandemics have always posed a challenge to trust between the public and the government, and maintaining and rebuilding trust during a crisis is crucial to maintaining political participation and social cohesion [38]. In the long term, behavioural factors such as smoking, obesity, diet, and nutrition, along with environmental factors such as tree cover and PM<sub>2.5</sub>, are all risk factors that can be changed through health education and policy development, and are areas in which advance preparation is needed in order to mitigate the effects of future epidemics. Regulating taxes on tobacco, tightening restrictions on smoking places, and setting a legal age for smoking would contribute to reducing the potential harm from smoking at a national level. Obesity and malnutrition are long-standing health challenges and risk factors for a range of chronic diseases, the dangers of which are already well known. However, governments also need to guide people towards healthy eating habits through policies such as requiring calorie labelling on foods and restricting the promotion of high-sugar and high-fat foods. In addition, environmental factors are of increasing concern to epidemiologists, and our research suggests that tree cover and PM<sub>2.5</sub> have some impact on severe health outcomes in COVID-19. It has also been suggested that PM<sub>2.5</sub> may potentially serve as a carrier

for the virus [39]. Therefore, an improved environment with less air pollution would benefit both patients with COVID-19 and healthy populations.

The third major conclusion of this study is that the health benefits of continued vaccination vary between countries having different driving factors for death. On the issue of vaccine allocation, as advocated by Jeremy Bentham's Utilitarianism, a rule for society should be established that has the best outcome for the greatest amount of people in society, in the sense that a cost-effective vaccine allocation scheme should be developed in a global perspective that reduces the risk of death for the greatest proportion of people worldwide. The WHO has worked to this end by convening COVAX [40], a ground-breaking global collaboration aimed at accelerating the development and production of and equitable access to the COVID-19 vaccine, ensuring that every country has access to the vaccine and is able to promote vaccination to protect their whole population, starting with the most vulnerable. Progress on this project has not been smooth, with most early supplies of vaccine having been promptly purchased by wealthy countries and the supply shortages further exacerbated by vaccine nationalism, hoarding, and export bans. Even though COVAX has delivered more than 1.4 billion doses of vaccine to 142 countries, and 65.2% of the world's population has received at least one dose, only a cumulative 15.3% of people in low-income countries are included in that fraction [41]. This is insufficient to reach vulnerable populations such as health workers, the elderly, and people with chronic diseases. In times of inadequate vaccine supply, our model allows for real-time assessment of the risk of COVID-19 death in countries in need and of the health benefits of vaccination so as to guide vaccine allocation more rationally.

Our ecological studies based on country-level data provide a global perspective on the risk assessment of COVID-19 CFR. Country-level studies provide a more comprehensive understanding of the consistent impacts of risk factors across countries worldwide than more granular studies. We draw more generalisable conclusions at larger geographical scales, and identify key risk factors that are specific to each country, complementing the more granular studies within countries that together support policy decisions. Meanwhile, our studies provide insights into the allocation of health resources, such as vaccines, in a global perspective. Population-based and individual-based studies focus on different dimensions and issues that complement each other and contribute to a comprehensive understanding of disease development and control. For example, while there are a large number of individual-level studies across time periods

that show that underlying disease is always a good predictor of death in patients with COVID-19 [42], consistent with the findings of other country-level studies that risk factors differ in importance across time periods for the national CFR, with the burden of disease from the underlying disease becoming less important during the Delta period [43]. This variation in risk factors between time periods supports policymakers in considering different intervention strategies at different times. While individual-level studies provide insights into direct health impacts, country-level studies better explain differences in disease outcomes between countries, providing a broader view of how macro-factors, such as healthcare policies and economic conditions, impact public health outcomes.

There are several limitations in our analysis. First, the study design is a country-level ecological analysis based on retrospective data, and care should be taken regarding ecological fallacies in the interpretation and generalisation of the results. Our findings do not explain CFR differences within countries, and targeted COVID-19 intervention strategies within countries may need to be supported by more fine-grained data. Second, our data were sourced from multiple publicly available data sources, and after comparing them we selected the more credible sources and also applied outlier treatment, but the credibility of our analysis relies greatly on the quality of the data. Third, COVID-19 cases and deaths are from national self-reported data and do not consider excess deaths from COVID-19. Fourth, we considered as many country-level COVID-19-related factors as possible, but due to data limitations, we were unable to adjust for differences in vaccine type and ethnicity. Fifth, the original period model has a low  $R^2$  value and does not capture the variation in CFR well. As the model can only explain the features we included, there may be some unknown features that we have not been able to identify.

The cross-temporal and cross-country variation in COVID-19 CFRs illustrates the importance of conducting further research on risk assessment. Our exploratory study reminds policy makers to consider risk factors holistically and assess whether their countries can rebuild policy trust, face the challenges of vaccine hesitancy, revitalize primary healthcare, and strengthen behavioural and environmental risk management and investment in the post-COVID era. At present, consideration of COVID-19 as an endemic disease has also entered the plans of some countries; that is, SARS-CoV-2 will not be eradicated and is instead expected to persist in a less lethal pattern, placing greater demands on healthcare systems and cyclical vaccination.

## Conclusions

Evidence from this study suggests that cross-temporal and cross-country variation in COVID-19 CFR is jointly determined by key and concurrent risks. Across the different variant dominance periods of the pandemic, the important determinants of COVID-19 CFRs changed from health conditions to universal health coverage, and then to a multifactorial mixed effect dominated by vaccination. We quantified the country-specific risk of COVID-19 CFR for 156 countries along seven dimensions: vaccination coverage, demographic factors, disease burden, behavioural risk factors, environmental risk factors, health services, and trust levels, and clarify the extent to which countries will benefit from increased vaccination. The findings suggested that achieving a decreasing COVID-19 case fatality rate requires more than increasing vaccination coverage, but rather targeted intervention strategies based on country-specific risks. In countries where low vaccination coverage is a major risk factor for COVID-19 deaths, increased vaccination is more effective in reducing CFR, especially in countries with high overall risk. In countries where high disease burden and ageing are major risk factors for COVID-19 deaths, it is important to focus on protection of vulnerable populations in the short term, and on interventions targeting age structure and population health status in the long term. Some risk factors that influence CFRs, such as GDP, cannot be controlled by policymakers or changed in the short term, underlining the importance of global public health efforts to strengthen cross-border cooperation to mitigate inequities.

## Abbreviations

COVID-19	Coronavirus disease 2019
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
CFRs	Case fatality rates
VOC	Variant of concern
XGBoost	Extreme Gradient Boosting
LASSO	Least absolute shrinkage and selection operator
SHAP	SHapley Additive exPlanations
WHO	World Health Organization
RFE	Recursive feature elimination
RMSE	Root-mean-square error
SD	Standard deviation
IQR	Interquartile range
ANOVA	Analysis of variance
HAQ Index	Healthcare access and quality index
IHR	International Health Regulations core capacity
GDP	Gross domestic product
LRI	Lower respiratory infections
URI	Upper respiratory infections
COPD	Chronic obstructive pulmonary disease
CVD	Cardiovascular diseases
CKD	Chronic kidney disease
HTN	Hypertension
MD	Mental disorders
NCD	Noncommunicable diseases
HIV	Human immunodeficiency virus infection
TB	Tuberculosis

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12963-024-00330-4>.

Supplementary material 1.

### Acknowledgements

The authors would like to thank the participants in this study.

### Author contributions

CuiZ: data collection, conceptualization, investigation, data analysis, and writing original draft and revision. ÅW: conceptualization supervision, and investigation. ChuZ, JM and LZ: data analysis. KD: revision. WL: conceptualization and supervision. JG: data collection, conceptualization, supervision, investigation, data analysis, and writing original draft and revision. LX: conceptualization, supervision, and funding acquisition, writing, and revision. All authors contributed to the article and approved the submitted version.

### Funding

This study received support from the National Key R&D Program of China grant (No. 2021ZD0114103), Karolinska Institutet Research Foundation Grants (2022–02329), and Väinö ja Laina Kiven Säätiö (No.20230006). We also extend our thanks to the Research Fund at Vanke School of Public Health, Tsinghua University, China, for their support. The funder has no role in the data collection, data analysis, preparation of manuscript and decision to submission.

### Availability of data and materials

The original contributions presented in the study are included in the method/supplementary material, further inquiries can be directed to [xu\\_lei@mail.tsinghua.edu.cn](mailto:xu_lei@mail.tsinghua.edu.cn).

### Declarations

#### Ethics approval and consent to participate

There has been no patient and/or public involvement in the study design, data collection, data analysis and writing of this research.

#### Consent for publication

Not Applicable.

#### Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### Author details

<sup>1</sup>Vanke School of Public Health, Tsinghua University, Beijing, China. <sup>2</sup>Institute for Healthy China, Tsinghua University, Beijing, China. <sup>3</sup>Respiratory Medicine Unit, Department of Medicine & Centre for Molecular Medicine, Karolinska Institutet, Karolinska Institutet, Slona, 171 65 Stockholm, Sweden. <sup>4</sup>Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. <sup>5</sup>Department of Respiratory Medicine, University of Helsinki, Helsinki, Finland. <sup>6</sup>The First School of Clinical Medicine, Lanzhou University, Lanzhou, China. <sup>7</sup>College of Natural Resources and Environment, Northwest A&F University, Yangling, China.

Received: 30 October 2023 Accepted: 27 May 2024

Published online: 03 June 2024

### References

- Ahmed F, Ahmed N, Pissarides C, Stiglitz J. Why inequality could spread COVID-19. *The Lancet Public Health*. 2020;5: e240.
- Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health*. 2020;74:964–8.
- WHO: COVID is here for the foreseeable future | United Nations in Turkey [Internet]. [cited 2022 May 15]. Available from: <https://turkey.un.org/en/169490-who-covid-here-foreseeable-future>, <https://turkey.un.org/en/169490-who-covid-here-foreseeable-future>
- WHO Coronavirus (COVID-19) Dashboard [Internet]. [cited 2022 Jun 8]. Available from: <https://covid19.who.int>
- Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect Dis*. 2021;21:793–802.
- Rossmann H, Shilo S, Meir T, Gorfine M, Shalit U, Segal E. COVID-19 dynamics after a national immunization program in Israel. *Nat Med*. 2021;27:1055–61.
- Dudel C, Riffe T, Acosta E, Raalte A, Strozza C, Myrskylä M. Monitoring trends and differences in COVID-19 case-fatality rates using decomposition methods: contributions of age structure and age-specific fatality. *PLoS ONE*. 2020;15: e0238904.
- Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci*. 2020;117:9696–8.
- Sanyaolu A, Okorie C, Marinkovic A, Patidar R, Younis K, Desai P, et al. Comorbidity and its impact on patients with COVID-19. *SN Compr Clin Med*. 2020;2:1069–76.
- Elezkurtaj S, Greuel S, Ihlow J, Michaelis EG, Bischoff P, Kunze CA, et al. Causes of death and comorbidities in hospitalized patients with COVID-19. *Sci Rep*. 2021;11:4263.
- Liang C-K, Chen L-K. National health care quality and COVID-19 case fatality rate: International comparisons of top 50 countries. *Arch Gerontol Geriatr*. 2022;98: 104587.
- Li C, Managi S. Impacts of air pollution on COVID-19 case fatality rate: a global analysis. *Environ Sci Pollut Res Int*. 2022;29(18):1–14.
- Ozkan A, Ozkan G, Yalaman A, Yildiz Y. Climate risk, culture and the Covid-19 mortality: a cross-country analysis. *World Dev*. 2021;141: 105412.
- Pana TA, Bhattacharya S, Gamble DT, Pasdar Z, Szlachetka WA, Perdomo-Lampignano JA, et al. Country-level determinants of the severity of the first global wave of the COVID-19 pandemic: an ecological study. *BMJ Open*. 2021;11: e042034.
- El Mouhayyar C, Jaber LT, Bergmann M, Tighiouart H, Jaber BL. Country-level determinants of COVID-19 case rates and death rates: An ecological study. *Transbound Emerg Dis*. 2021. <https://doi.org/10.1111/tbed.14360>.
- Onder G, Rezza G, Brusaferro S. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA*. 2020;323:1775–6.
- Hueniken K, Somé NH, Abdelhack M, Taylor G, Elton Marshall T, Wickens CM, et al. Machine learning-based predictive modeling of anxiety and depressive symptoms during 8 months of the COVID-19 global pandemic: repeated cross-sectional survey study. *JMIR Ment Health*. 2021;8: e32876.
- Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. New York, NY, USA: Association for Computing Machinery; 2016 [cited 2021 Dec 12]. p. 785–94. Available from: <https://doi.org/10.1145/2939672.2939785>
- Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2:56–67.
- Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19). *Our World in Data* [Internet]. 2020 [cited 2022 Apr 30]; Available from: <https://ourworldindata.org/covid-vaccinations>
- Song S, Ma L, Zou D, Tian D, Li C, Zhu J, et al. The global landscape of SARS-CoV-2 genomes, variants, and haplotypes in 2019nCoV. *Genomics Proteomics Bioinformatics*. 2020;18:749–59.
- Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, et al. A global database of COVID-19 vaccinations. *Nat Hum Behav*. 2021;5:947–53.
- Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman L, Haas EJ, et al. Waning immunity after the BNT162b2 vaccine in Israel. *N Engl J Med*. 2021;385: e85.
- Dolgin E. COVID vaccine immunity is waning — how much does that matter? *Nature*. 2021;597:606–7.
- Wu Y, Lin S, Shi K, Ye Z, Fang Y. Seasonal prediction of daily PM2.5 concentrations with interpretable machine learning: a case study of



- Beijing China. *Environ Sci Pollut Res*. 2022. <https://doi.org/10.1007/s11356-022-18913-9>.
26. Lei T, Guo J, Wang P, Zhang Z, Niu S, Zhang Q, et al. Establishment and validation of predictive model of tophus in gout patients. *J Clin Med*. 2023;12:1755.
  27. Fan Z, Jiang J, Xiao C, Chen Y, Xia Q, Wang J, et al. Construction and validation of prognostic models in critically ill patients with sepsis-associated acute kidney injury: interpretable machine learning approach. *J Transl Med*. 2023;21:406.
  28. Wang K, Tian J, Zheng C, Yang H, Ren J, Liu Y, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. *Comput Biol Med*. 2021;137: 104813.
  29. Hart WS, Miller E, Andrews NJ, Waight P, Maini PK, Funk S, et al. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *Lancet Infect Dis*. 2022;22:603–10.
  30. Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, et al. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *The Lancet Infectious Diseases* [Internet]. 2021 [cited 2021 Dec 4];0. Available from: [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(21\)00475-8/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(21)00475-8/fulltext)
  31. Fisman DN, Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario. *Canada CMAJ*. 2021;193:E1619–25.
  32. Fullman N, Yearwood J, Abay SM, Abbafati C, Abd-Allah F, Abdela J, et al. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet*. 2018;391:2236–71.
  33. Atherstone CJ, Sarah AJ, Guagliardo AH, O’Laughlin K, Wong K, Sloan ML, Henao O, Rao CY, McElroy PD, Bennett SD. COVID-19 Epidemiology during Delta Variant Dominance Period in 45 High-Income Countries, 2020–2021. *Emerg Infect Dis*. 2023. <https://doi.org/10.3201/eid2909.230142>.
  34. Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet*. 2021;398:2126–8.
  35. Nyberg T, Ferguson NM, Nash SG, Webster HH, Flaxman S, Andrews N, et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *The Lancet*. 2022;399:1303–12.
  36. Andrews N, Stowe J, Kirsebom F, Toffa S, Rickeard T, Gallagher E, et al. Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *N Engl J Med*. 2022;386:1532–46.
  37. Lenton TM, Boulton CA, Scheffer M. Resilience of countries to COVID-19 correlated with trust. *Sci Rep*. 2022;12:75.
  38. Devine D, Gaskell J, Jennings W, Stoker G. Trust and the coronavirus pandemic: what are the consequences of and for trust? An early review of the literature. *Political Studies Review*. 2021;19:274–85.
  39. Nor NSM, Yip CW, Ibrahim N, Jaafar MH, Rashid ZZ, Mustafa N, et al. Particulate matter (PM<sub>2.5</sub>) as a potential SARS-CoV-2 carrier. *Sci Rep*. 2021;11:2508.
  40. COVAX [Internet]. [cited 2022 Jun 8]. Available from: <https://www.who.int/initiatives/act-accelerator/covax>
  41. COVID-19 Vaccine Market Dashboard [Internet]. [cited 2022 Apr 30]. Available from: <https://www.unicef.org/supply/covid-19-vaccine-market-dashboard>
  42. Russell CD, Lone NI, Baillie JK. Comorbidities, multimorbidity and COVID-19. *Nat Med*. 2023;29:334–43.
  43. Thi HNN, Ou T-Y, Huy LD, Shih C-L, Chang Y-M, Phan T-P, et al. A global analysis of COVID-19 infection fatality rate and its associated factors during the Delta and Omicron variant periods: an ecological study. *Front Public Health*. 2023;11:1145138.
  44. ISO - ISO 3166 — Country Codes [Internet]. ISO. [cited 2022 Jun 15]. Available from: <https://www.iso.org/iso-3166-country-codes.html>
  45. World Population Prospects - Population Division - United Nations [Internet]. [cited 2022 Apr 8]. Available from: <https://population.un.org/wpp/Download/Standard/Population/>
  46. World Development Indicators [Data Catalog [Internet]. [cited 2022 Apr 8]. Available from: <https://datacatalog.worldbank.org/search/dataset/0037712>
  47. Roser M, Ortiz-Ospina E. Global Education. Our World in Data [Internet]. 2016 [cited 2022 Apr 8]; Available from: <https://ourworldindata.org/global-education>
  48. GDP per capita (current US\$) | Data [Internet]. [cited 2022 Apr 8]. Available from: <https://data.worldbank.org/indicator/NY.GDPPCAP.CD>
  49. GBD Results Tool | GHDX [Internet]. [cited 2022 May 10]. Available from: <https://ghdx.healthdata.org/gbd-results-tool>
  50. Total NCD mortality rate (per 100 000 population) , age-standardized [Internet]. [cited 2022 Apr 8]. Available from: <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-ghc-ncd-mortality-rate>
  51. Prevalence of overweight among adults, BMI  $\geq$  25 (age-standardized estimate) (%) [Internet]. [cited 2022 Apr 8]. Available from: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-overweight-among-adults-bmi-25-\(age-standardized-estimate\)-\(-\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-overweight-among-adults-bmi-25-(age-standardized-estimate)-(-))
  52. Crowther TW, Glick HB, Covey KR, Bettigole C, Maynard DS, Thomas SM, et al. Mapping tree density at a global scale. *Nature*. 2015;525:201–5.
  53. World Bank Climate Change Knowledge Portal [Internet]. [cited 2022 Apr 8]. Available from: <https://climateknowledgeportal.worldbank.org/>
  54. Population density (people per sq. km of land area) | Data [Internet]. [cited 2022 Jun 8]. Available from: <https://data.worldbank.org/indicator/EN.POPDNST>
  55. National Health Emergency Framework (IHR SPAR) [Internet]. [cited 2022 Apr 8]. Available from: <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/national-health-emergency-framework>
  56. Hospital beds (per 10 000 population) [Internet]. [cited 2022 Apr 8]. Available from: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-\(per-10-000-population\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-(per-10-000-population))
  57. Current health expenditure (CHE) per capita in US\$ [Internet]. [cited 2022 Jun 8]. Available from: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/current-health-expenditure-\(che\)-per-capita-in-us\\$](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/current-health-expenditure-(che)-per-capita-in-us$)
  58. Total density per 100 000 population: Hospitals [Internet]. [cited 2022 Jun 8]. Available from: <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/total-density-per-100-000-population-hospitals>
  59. Wellcome Global Monitor 2020: Covid-19 [Internet]. Wellcome. [cited 2022 Apr 8]. Available from: <https://wellcome.org/reports/wellcome-global-monitor-covid-19/2020>

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.