BMJ Open

COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data

Yang Cao ![ORCID],[1,2] Ayako Hiyoshi,[1,3] Scott Montgomery ![ORCID] [1,4,5]

ABSTRACT

Objective To investigate the influence of demographic and socioeconomic factors on the COVID-19 case-fatality rate (CFR) globally.

Design Publicly available register-based ecological study.

Setting Two hundred and nine countries/territories in the world.

Participants Aggregated data including 10 445 656 confirmed COVID-19 cases.

Primary and secondary outcome measures COVID-19 CFR and crude cause-specific death rate were calculated using country-level data from the Our World in Data website.

Results The average of country/territory-specific COVID-19 CFR is about 2%–3% worldwide and higher than previously reported at 0.7%–1.3%. A doubling in size of a population is associated with a 0.48% (95% CI 0.25% to 0.70%) increase in COVID-19 CFR, and a doubling in the proportion of female smokers is associated with a 0.55% (95% CI 0.09% to 1.02%) increase in COVID-19 CFR. The open testing policies are associated with a 2.23% (95% CI 0.21% to 4.25%) decrease in CFR. The strictness of anti-COVID-19 measures was not statistically significantly associated with CFR overall, but the higher Stringency Index was associated with higher CFR in higher-income countries with active testing policies (regression coefficient beta=0.14, 95% CI 0.01 to 0.27). Inverse associations were found between cardiovascular disease death rate and diabetes prevalence and CFR.

Conclusion The association between population size and COVID-19 CFR may imply the healthcare strain and lower treatment efficiency in countries with large populations. The observed association between smoking in women and COVID-19 CFR might be due to the finding that the proportion of female smokers reflected broadly the income level of a country. When testing is warranted and healthcare resources are sufficient, strict quarantine and/or lockdown measures might result in excess deaths in underprivileged populations. Spatial dependence and temporal trends in the data should be taken into account in global joint strategy and/or policy making against the COVID-19 pandemic.

Strengths and limitations of this study

► This is the first study that investigated the relationship between COVID-19 case-fatality rate and demographic and socioeconomic factors globally.

► Our study addressed the question from a geospatial perspective, which may inspire new reflections to fight against the COVID-19 pandemic globally.

► Asymptomatic carriers and victims of COVID-19 were not taken into account in the analysis.

► No detailed information on time from diagnosis to death and comorbidity of the COVID-19 cases is available in the current study, which might bias the association in an unknown direction.

► Country-level analysis may conceal huge discrepancies between subnational entities in terms of both outcomes and predictors.

reported in Wuhan, a city in the Hubei province in China, in December 2019, has become a major global health concern.[1] Poor outcomes in those with COVID-19 infections correlate with clinical and laboratory features of cytokine storm syndrome, an exaggerated systemic inflammatory phenomenon due to overproduction of proinflammatory cytokines by the immune system that results in diffuse inflammatory lung disease and acute respiratory distress syndrome (ARDS).[2] It may be complicated by sepsis, respiratory failure, ARDS and subsequent multiorgan failure.[3] Although COVID-19-related deaths are not clearly defined in the international reports available so far, many governments are warning people to be particularly stringent in following the recommended prevention measures because COVID-19 may result in severe conditions that need critical care, including ventilation or death.[4] Untill the end of June 2020, the pandemic has resulted in over 10 million confirmed cases and 500

INTRODUCTION

The pandemic caused by the SARS-CoV-2 virus/COVID-19, which has been initially

000 deaths worldwide.[5] COVID-19-related fatality rates are difficult to assess with certainty, but according to the estimates based on the data from China, the UK, Italy and the Diamond Princess cruise ship, the overall death rate from the confirmed COVID-19 cases is around 0.7%–1.3%, sharply rising from less than 0.002% in children aged 9 years or younger to 8% in people aged over 80, which is much greater than seasonal influenza at about 0.1%.[4 6–10]

During the COVID-19 pandemic, numerous studies on the global public health emergency have covered a significant range of disciplines, including medicine, mathematics and social sciences. The spatial spread is one of the most important properties of COVID-19, a characteristic which mainly depends on the epidemic mechanism, human mobility and control strategy.[11] Spatial statistical methods are frequently used to uncover relationships between spatiotemporal patterns of infectious diseases and host or environmental characteristics,[12] to generate detailed maps to visualise the distribution of the diseases' morbidity or mortality[13 14] and to identify hotspots, clusters and potential risk factors.[15–18]

Clinical risk factors for mortality of adult patients with COVID-19 have been investigated in numerous studies, and the identified factors include older age[19]; male sex[20]; higher sequential organ failure assessment score[21]; obesity[22 23]; pre-existing concurrent diabetes[24]; cardiovascular, cerebrovascular[25] and kidney diseases[26]; and macroeconomic and environmental risk factors, such as socioeconomic deprivation,[27] air pollution[28 29] and diurnal temperature variation.[30] However, there is a lack of published studies on the effects of country-level demographic and socioeconomic characteristics on COVID-10 case-fatality rates (CFRs). It is an important issue for governments and regional or international non-governmental organisations to identify country characteristics that are associated with high CFR and to help develop prevention and intervention measures to fight against this global public health crisis.

Using the publicly available data from the non-governmental organisation Our World in Data,[31] we aimed to investigate the relationship of key country-level demographic and socioeconomic indices and COVID-19 case-fatality, and to explore factors associated with COVID-19 CFR, which may indicate treatment efficiency and strain in healthcare resources, while controlling for the spatial dependence of the data collected at different locations.

METHODS

Data on COVID-19 by Our World in Data

The COVID-19 dataset used in the study was downloaded from the Our World in Data website on 2 July 2020, which is a collection of the COVID-19 data maintained by the organisation Our World in Data and updated daily. The dataset includes country-level daily data on confirmed cases, deaths and testing, as well as other variables of potential interest.[31 32] The data sources of the dataset,

including the European Centre for Disease Prevention and Control, the International Organisation for Standardisation (ISO), national government reports, the Department of Economic and Social Affairs of the United Nations (UN), UN Population Division, UN Statistics Division, Oxford COVID-19 Government Response Tracker, the World Bank, the Global Burden of Disease Collaborative Network, and Eurostat of the Organisation for Economic Cooperation and Development.[31] There are in total 34 indices from 209 countries and territories in the dataset by 2 July 2020. The dataset was linked to the global geospatial vector database using the ISO 3166-1 alpha-2 codes for the spatial modelling.[33]

Case-fatality rate

CFR of COVID-19 was calculated as the number of total deaths due to COVID-19 divided by the number of total confirmed COVID-19 cases by 2 July 2020 multiplied by 100. CFR was investigated in our study because it may reflect disease severity, as well as the efficiency of treatment and healthcare response and strain. CFR is not constant. It can vary between populations and over time, depending on the interplay between the causative agent of the disease, the host and the environment, as well as available treatments and quality of patient care. For example, it can increase if the healthcare system is overwhelmed by the sudden increase of cases.[34]

We also calculated the crude cause-specific death rate (CDR) of COVID-19 in a sensitivity analysis and compared it with CFR. The CDR was calculated as the number of total deaths due to COVID-19 divided by the production of population and months of the data collected, multiplied by 1 000 000.[35]

Statistical analysis

The numbers of confirmed cases, tests and tests per thousand people were not included in the analysis because they were depending much on the population, detection policy, and quarantine and isolation policies of the countries and territories. Instead, we included the Stringency Index in the analysis, which is a composite measure based on nine response indicators, including school closures, workplace closures, testing policy and travel bans, rescaled to a value from 0 to 100 (100=strictest response).[36] The Stringency Index data were obtained from the World Intellectual Property Organisation website on 1 July 2020.[37] Because the variable 'proportion of the population with basic handwashing facilities on premises' has missing values in more than 50% of the countries/territories, it was also excluded from the analysis.

A subcomponent of the Stringency Index is the government policy on the access to COVID-19 test by four groups: 0, no policy; 1, only those who were both symptomatic and met specific criteria; 2, anyone symptomatic; and 3, open public testing, such as drive-through testing.[36] The testing policy indicator was used for stratification of the analysis.

Collinearity and multiple collinearity between the variables were examined using Pearson's correlation coefficient and multiple correlation coefficient, respectively.[38] Spatial autocorrelation (or spatial dependence) is defined as the relationship between spatial proximity among some observational units and similarity among their values; positive spatial autocorrelation refers to situations in which the nearer the observational units, the more similar their values (and vice versa for its negative counterpart).[39] This feature violates the assumption of independence among observations on which many regression analyses are applied. Spatial autocorrelation among the fatality rates of the countries/territories was examined using a multivariate linear regression model and the Moran's I test.[40] The autocorrelation was visualised using the Matérn correlation coefficient.[41]

The Matérn correlation model, a commonly used model for spatial correlated data, was fitted for our data to investigate the relationship between COVID-19 case-fatality and the demographic and socioeconomic variables. The latitude and longitude of the centroid of the countries/territories were used as random effects in the Matérn correlation model.[42]

Variables with skewed distribution were log-transformed before entering the regression models. The multiple imputation method was used to handle the missing values in the data. The missing values were assumed to be missing at random. A total of 10 copies of the data were created, each of which had the missing values imputed by using switching regression, an iterative multivariable regression technique. Then, each complete dataset was analysed independently. Estimates of parameters of interest were then averaged across the 10 copies to give a single estimate using Rubin's rule.[43]

Subgroup analysis was conducted by economic levels of the countries/territories according to the World Bank's newest classification.[44]

The associations of the studied variables with COVID-19 CDR (per 1 000 000 person-months) of the countries/territories from 31 December 2019 to 1 July 2020 were also evaluated using the same methodology but using the Poisson regression model because of the rare event, and the results were presented in the supplemental materials.

All the analysis were conducted in R V.4.02 (the R Foundation for Statistical Computing, Vienna, Austria) using the package *spaMM*[45] and in Python V.3.6 (Python Software Foundation)[46] using the packages *geopandas* and *geoplot*.[47]

In the study, estimates of health indicators at the global level were reported according to the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) statement.[48]

Patient and public involvement

The study is a worldwide public available register-based study; therefore, it was not required and also not possible to involve patients or the public in the design, conduct, reporting or dissemination plans of our research.

RESULTS
Descriptive characteristics of the variables

In total, 10 445 656 confirmed COVID-19 cases and 511 030 deaths from 209 countries and territories and 17 variables (including latitude and longitude) were included in the study. The descriptive statistics of the variables are shown in table 1. The pairwise relationships of the variables are shown in figure 1. The CFR, CDR, cardiovascular disease (CVD) death rate and diabetes prevalence shown in table 1 and figure 1, and/or following tables and figures were not adjusted/standardised for age and sex; therefore, they are crude rates. High multicollinearity was found between the number of confirmed cases and population size (pairwise Pearson's r=0.76, p<0.001; figure 1), and gross domestic product (GDP) per capita and life expectancy, median age, the proportion of aged 65 years or older, the proportion of extreme poverty and hospital beds per 1000 people (all pairwise Pearson's r greater than or approximate to 0.70 and p<0.001; multiple correlation coefficient=0.92, p<0.001; figure 1). Therefore, median age, the proportion of aged 65 or older and hospital beds per 1000 people were excluded in later regression analysis. Although the proportion of extreme poverty is an interesting factor to be investigated, it is highly correlated with GDP per capita (r=−0.83), and the latter is available in more countries; therefore, the proportion of extreme poverty was also excluded from the analysis.

Worldwide COVID-19 CFR distribution

Distribution of COVID-19 CFR of the 209 countries/territories is shown in figure 2. The mean and median CFR worldwide are 3.31% and 2.19%, respectively (table 1), with the highest rates found in Yemen (27%), West and North Europe (14%–19%), and North America (9%–12%, figure 2).

Spatial autocorrelation of the COVID-19 CFR

Statistically significant spatial autocorrelation was found among the countries/territories' COVID-19 CFR. The residuals from the common (non-spatial) multivariate linear regression models show apparent spatial dependence around the countries/territories with high fatality (figure 3). The p value from Moran's I test for the spatial autocorrelation of the residuals is 2.32×10^{-5}.

The estimated spatial autocorrelation coefficient of COVID-19 CFR between two locations against their distance is shown in figure 4, with a strength parameter v of 2.48 and a decay parameter ρ of 0.12. This indicates that locations more than 20° (in longitude or latitude) away have an autocorrelation coefficient below 0.5 (figure 4).

Associations of demographic and socioeconomic variables with COVID-19 CFR

Overall, after controlling for the spatial dependence, the statistically significant variables associated with COVID-19 CFR are population size and proportion of female smokers in a country/region (table 2). The multivariate adjusted results indicate that, approximately, a doubling in size of

Table 1 Descriptive statistics of the variables

	N	Mean	SD	Median	Min	Max	IQR
Confirmed COVID-19 cases	209	49 979	218 713	2221	3	2 634 432	18 420
Confirmed COVID-19 cases (log)	209	7.75	2.84	7.71	1.10	14.78	4.13
COVID-19 deaths	209	2445	11 014	51	0	127 410	346
COVID-19 CFR (%)	209	3.31	3.79	2.19	0.00	26.94	3.75
COVID-19 CDR (per 1 000 000 person-months)	209	13.98	28.75	3.31	0.00	205.84	10.94
Population size (million)	209	37.08	142.84	6.87	0.01	1439.32	25.38
Population size (log)	209	15.28	2.52	15.74	6.70	21.09	3.29
GDP per capita (US$1000)	182	19.28	19.69	13.03	0.66	116.94	23.47
GDP per capita (log)	182	9.28	1.21	9.48	6.49	11.67	1.84
Population density (per km)^2	198	328.16	1507.86	87.25	0.14	19347.50	176.45
Population density (log)	198	4.42	1.52	4.47	−1.99	9.87	1.74
CVD death rate (per 100 000)	185	256.17	116.51	242.65	79.37	724.42	153.98
CVD death rate (log)	185	5.44	0.46	5.49	4.37	6.59	0.65
Diabetes prevalence (%)	192	8.06	4.26	7.13	0.99	23.36	5.13
Diabetes prevalence (log)	192	1.94	0.56	1.96	−0.01	3.15	0.67
Stringency Index	172	44.22	9.16	45.46	7.68	65.20	11.16
Life expectancy	206	73.47	7.54	75.07	53.28	86.75	9.91
Median age	185	30.55	9.10	29.90	15.10	48.20	16.50
Age 65 years or older (%)	182	8.81	6.21	6.66	1.14	27.05	10.87
Age 65 years or older (log)	182	1.91	0.74	1.90	0.13	3.30	1.42
Proportion of extreme poverty (%)	121	13.86	20.54	2.2	0.10	77.60	20.8
Proportion of extreme poverty (log)	121	1.08	2.03	0.79	−2.30	4.35	4.57
Proportion of female smoker (%)	140	10.44	10.48	6.05	0.10	44.0	17.3
Proportion of female smoker (log)	140	1.63	1.40	1.80	−2.30	3.78	2.31
Proportion of male smoker	138	32.62	13.71	31.30	7.70	78.10	19.90
Hospital beds per 1000 people	164	3.01	2.46	2.36	0.10	13.80	2.63
Hospital beds per 1000 people (log)	164	0.77	0.88	0.86	−2.30	2.62	1.11

CDR, cause-specific death rate; CFR, case-fatality rate; CVD, cardiovascular disease; GDP, gross domestic product; ;IQR, interquartile range; SD, standard deviation.

population is associated with a 0.48% (95% CI 0.25% to 0.70%) increase in COVID-19 CFR, and a doubling in proportion of female smokers is associated with a 0.55% (95% CI 0.09% to 1.02%) increase in COVID-19 CFR. Open public testing policy is associated with decreased CFR (beta=−2.23, 95% CI −4.25 to 0.21) compared with no testing policy.

Because associations might differ by the proportion of the population aged 65 years or older (65+), we produced stratified estimates by the proportion of people aged 65+ years (see online supplemental table S1). Briefly, population size and testing policy were found to be associated with CFR in the countries with a proportion of people aged 65+ years between 5% and 10%; and GDP per capita, population size, population density and the proportion of smokers were associated with CFR in the countries with a proportion of people aged 65+ years larger than 15%.

The estimated contour of COVID-19 CFRs worldwide is shown in figure 5. The areas with the higher risks are mainly around North America and West Europe.

The subgroup analysis by economic level indicates that population size, CVD death rate and diabetes prevalence are statistically significantly associated with COVID-19 CFR in the lower-income to middle-income countries; population size, diabetes prevalence, and testing only symptomatic and specified policy and testing anyone symptomatic policy are statistically significantly associated with COVID-10 CFR in the upper-income to middle-income countries (table 2).

However, the subgroup analysis in upper-income to middle-income and high-income countries by testing policy indicates that, if testing was ensured (testing policy=2 or 3), increment in Stringency Index is associated with increased COVID-19 CFR (beta=0.14, 95% CI 0.01 to 0.27) (table 3). The finding might imply an open
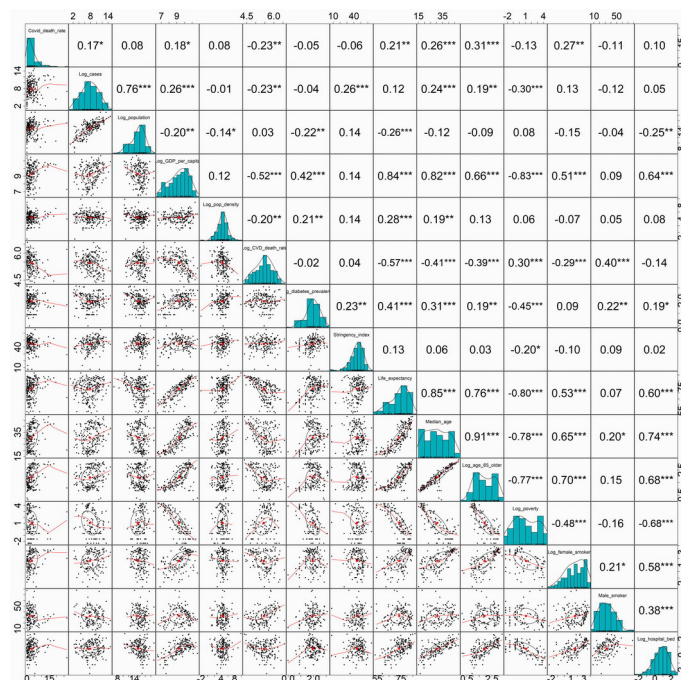
Figure 1 Pairwise scatter plots and Pearson's correlation coefficients of the variables. *P<0.05, **P<0.01, ***P<0.001.

society policy does not necessarily lead to a great CFR when testing is ensured. Meanwhile, diabetes prevalence is inversely associated with CFR (beta=−3.30, 95% CI −5.85 to 0.74).

DISCUSSION
Geospatial analysis of the COVID-19 pandemic
The COVID-19 pandemic is still full of unknowns, and many of them have a spatial dimension that leads to understanding the emergency geographically. Analysis of the COVID-19 data requires an interdisciplinary approach, including spatial statistics, that may provide important implications to policies addressing the spatial issues in the pandemic.[49] A recently published review summarised studies by 1 May 2020 on geospatial and spatial–statistical analysis of the COVID-19 pandemic. In total, 63 studies were categorised into five subjects: spatiotemporal analysis, health and social geography, environmental variables, data mining and web-based mapping.[50] Although

15 of the studies address the question globally, none of them investigated the association of COVID-19-related deaths with country-level demographic and/or socioeconomic factors. From a global health perspective, there is a knowledge gap in the research field. Our geospatial analysis fills this gap and shows the utility of the analysis to improve the understanding of the consequences of COVID-19 and their related factors from a global level, and contributes to the predictive modelling and decision-making to combat the pandemic.

Proportion of smokers and COVID-19 CFR
The proportion of female smokers was positively associated with COVID-19 CFR in the overall analysis, but the association diminished when the analysis was stratified by the economical level of the countries/territories. This is counterintuitive, given that severer COVID-19 was associated with male sex due to possibly immune system and hormone levels[51] and smoking.[52 53] The observed association between smoking in women and COVID-19 CFR
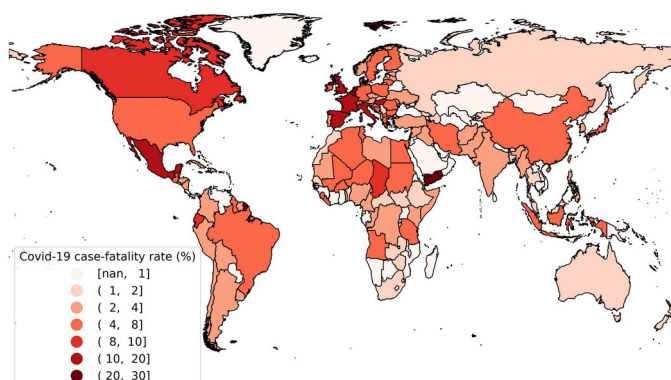

Figure 2 Global COVID-19 case-fatality rates (%). nan, no data available.
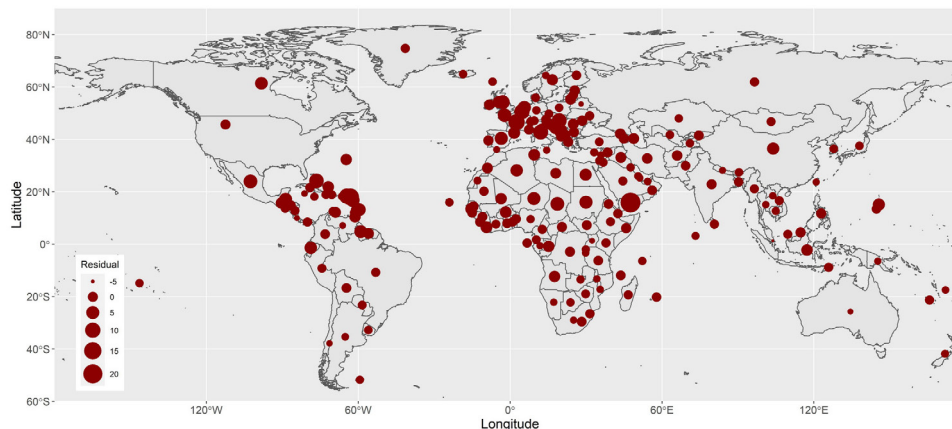
Figure 3 Residuals of the common (non-spatial) multivariate linear regression.

might be due to the finding that the proportion of female smoker reflected broadly the income level of a country (figure 6A). Linking to the theory of diffusion of innovation, the spread of smoking habit has been illustrated to take several stages of a rise, levelling and decline, from rich to poor, men to women, and young to old.[54–56] In the early phrase, the prevalence of smoking has increased in men, and women take up smoking about a few decades later. Subsequently, male smoking starts to decline and female follows later on. This pattern has been found to spread from rich to poor countries. In general, Asian and African countries tend to have low female smoking but high male smoking, while in European countries, prevalence is similar between men and women.[57 58] Therefore, female smoking in the overall analysis is a marker of the development of a country and it diminishes when the analysis is stratified by it.
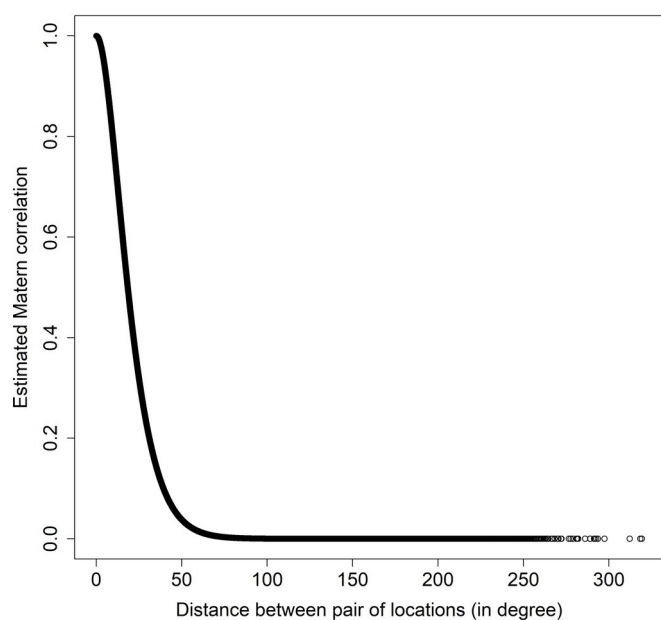

Figure 4 Strength and decay of the spatial autocorrelation between pair of locations.

Population size and density and COVID-19 CFR

Our results indicate that a larger population size is the most consistently associated with higher COVID-19 CFR (figure 6B), but population density is not associated with the outcome, controlled for some demographic and socioeconomic variables and spatial dependence worldwide. The association of population size with CFR can be interpreted in at least two ways. It could be because larger countries have experienced a greater number of deaths and have conducted relatively less test compared with countries with smaller population. Alternatively, in larger countries, healthcare might have been strained and resulted in a relatively higher number of deaths among confirmed cases than smaller countries.[59] We are unable to disentangle the mechanisms of the association. Therefore, it is recommended that the analysis is replicated by a study with more detailed healthcare information on both individual and country levels. However, in national studies, higher density has been shown to be associated with higher COVID-19 prevalence in Japan,[60] Italy[61] and Iran.[62] The lack of an association between COVID-19 CFR and population density globally might be due to the confounding by testing strategies and economic levels. In countries like Germany and South Korea, which took a more active testing strategy than, for example, the UK, where PCR test for COVID-19 was only performed among those who were with severe symptoms and were admitted to the hospital at the beginning,[63] CFR naturally showed lower values. There are weak but statistically significant negative associations between population size and population density (r=−0.15, p=0.042) and GDP per capita (r=−0.20, p=0.006). To minimise the confounding, we conducted a stratified analysis by economic level (table 2) and testing policy (only within upper-income and middle-income and high-income countries, table 3). Furthermore, we conducted a sensitivity analysis for CDR and the results were similar (see online supplemental tables S2, S3 and figures S1–S4). The results suggest that, globally, healthcare strain should be first relieved and treatment efficiency should be improved in countries with large populations.

Table 2 Estimation for variables' regression coefficients (beta) for CFR

Variables	Unadjusted		Adjusted	
	Beta (95% CI)	P value	Beta (95% CI)	P value
Overall (n=209)				
GDP per capita (log)	0.14 (−0.38 to 0.65)	0.616	0.03 (−0.68 to 0.73)	0.777
Population (log)	0.42 (0.20 to 0.63)	<0.001	0.48 (0.25 to 0.70)	<0.001
Population density (log)	0.01 (−0.33 to 0.35)	0.797	0.10 (−0.25 to 0.46)	0.595
Stringency Index	−0.02 (−0.08 to 0.03)	0.460	−0.02 (−0.08 to 0.03)	0.434
Proportion of female smokers (log)	0.47 (0.06 to 0.89)	0.040	0.55 (0.09 to 1.02)	0.034
Proportion of male smokers	−0.02 (−0.06 to 0.02)	0.339	−0.02 (−0.06 to 0.03)	0.478
CVD death rate (log)	−1.48 (−2.62 to -0.34)	0.017	−0.79 (−2.43 to 0.85)	0.390
Diabetes prevalence (log)	−0.87 (−1.93 to 0.18)	0.139	−0.37 (−1.52 to 0.78)	0.480
Testing policy				
0 (ref)				
1	−0.72 (−1.94 to 0.49)	0.266	−0.97 (−2.12 to 0.18)	0.126
2	−1.08 (−2.47 to 0.32)	0.168	−1.23 (−2.56 to 0.10)	0.109
3	−1.37 (−3.47 to 0.73)	0.240	−2.23 (−4.25 to −0.21)	0.043
Low-income countries (n=29)				
GDP per capita (log)	−0.22 (−2.92 to 2.48)	0.874	0.88 (−3.73 to 5.49)	0.592
Population (log)	0.49 (−0.98 to 1.96)	0.516	0.02 (−1.82 to 1.86)	0.428
Population density (log)	0.31 (−1.03 to 1.65)	0.651	−0.33 (−1.61 to 0.95)	0.417
Stringency Index	−0.01 (−0.11 to 0.08)	0.724	0.01 (−0.15 to 0.16)	0.354
Proportion of female smokers (log)	0.53 (−0.50 to 1.55)	0.295	0.34 (−1.52 to 2.19)	0.357
Proportion of male smokers	0.09 (0.00 to 0.19)	0.234	−0.01 (−0.22 to 0.21)	0.375
CVD death rate (log)	1.63 (−3.80 to 7.06)	0.557	2.58 (−6.40 to 11.56)	0.396
Diabetes prevalence (log)	0.00 (−2.01 to 2.01)	0.989	−1.01 (−2.76 to 0.74)	0.292
Testing policy				
0 (ref)				
1	−2.52 (−4.68 to −0.36)	0.152	−0.18 (−2.82 to 2.46)	0.354
2	−3.50 (−5.48 to −1.52)	0.018	−2.23 (−5.59 to 1.13)	0.233
3	−2.75 (−5.48 to −0.01)	0.163	−1.00 (−6.10 to 4.10)	0.577
Lower-indome to middle-income countries (n=44)				
GDP per capita (log)	0.74 (−0.02 to 1.5)	0.072	0.59 (−0.17 to 1.35)	0.152
Population (log)	0.27 (0.08 to 0.46)	0.005	0.44 (0.25 to 0.62)	<0.001
Population density (log)	0.06 (−0.23 to 0.35)	0.695	−0.08 (−0.36 to 0.20)	0.551
Stringency Index	0.00 (−0.03 to 0.02)	0.804	0.03 (−0.01 to 0.06)	0.234
Proportion of female smokers (log)	−0.12 (−0.48 to 0.24)	0.538	0.10 (−0.24 to 0.44)	0.418
Proportion of male smokers	0.01 (−0.01 to 0.03)	0.436	0.02 (0.00 to 0.04)	0.184
CVD death rate (log)	−0.61 (−1.72 to 0.50)	0.297	−1.78 (−2.98 to −0.58)	0.008
Diabetes prevalence (log)	0.71 (−0.06 to 1.47)	0.089	1.00 (0.34 to 1.66)	0.008
Testing policy				
0 (ref)				
1	−0.09 (−0.85 to 0.68)	0.757	−0.11 (−0.95 to 0.72)	0.644
2	0.24 (−0.92 to 1.39)	0.691	−0.22 (−1.16 to 0.72)	0.606
3	0.10 (−1.26 to 1.45)	0.856	−1.66 (−3.22 to −0.09)	0.051
Upper-income to middle-income countries (n=54)				
GDP per capita (log)	0.73 (−0.58 to 2.05)	0.275	−0.24 (−1.67 to 1.20)	0.549

Table 2 Continued

Variables	Unadjusted		Adjusted	
	Beta (95% CI)	P value	Beta (95% CI)	P value
Population (log)	0.36 (0.04 to 0.68)	0.030	0.61 (0.24 to 0.97)	0.011
Population density (log)	−0.26 (−0.71 to 0.20)	0.275	−0.48 (−0.88 to −0.07)	0.057
Stringency Index	0.00 (−0.06 to 0.06)	0.586	0.03 (−0.03 to 0.10)	0.321
Proportion of female smokers (log)	0.03 (−0.39 to 0.44)	0.211	−0.09 (−0.53 to 0.34)	0.389
Proportion of male smokers	−0.04 (−0.09 to 0.01)	0.177	−0.01 (−0.07 to 0.04)	0.588
CVD death rate (log)	−1.12 (−2.79 to 0.56)	0.213	−0.06 (−1.73 to 1.61)	0.783
Diabetes prevalence (log)	1.74 (−0.02 to 3.50)	0.108	2.37 (0.76 to 3.98)	0.009
Testing policy				
0 (ref)				
1	−0.89 (−1.76 to −0.02)	0.086	−1.70 (−2.73 to −0.68)	0.005
2	−0.64 (−1.90 to 0.61)	0.32	−2.10 (−3.41 to −0.78)	0.007
3	0.39 (−3.30 to 4.08)	0.836	−0.99 (−4.62 to 2.64)	0.578
High-income countries (n=82)				
GDP per capita (log)	−1.05 (−3.56 to 1.46)	0.453	−1.08 (−3.49 to 1.32)	0.417
Population (log)	0.56 (0.22 to 0.90)	0.001	0.43 (0.04 to 0.81)	0.097
Population density (log)	0.16 (−0.40 to 0.72)	0.596	0.65 (0.01 to 1.28)	0.084
Stringency Index	0.07 (−0.05 to 0.20)	0.285	0.06 (−0.07 to 0.19)	0.515
Proportion of female smokers (log)	0.85 (−0.10 to 1.79)	0.177	0.53 (−0.84 to 1.91)	0.300
Proportion of male smokers	−0.04 (−0.14 to 0.06)	0.451	−0.11 (−0.22 to 0.00)	0.161
CVD death rate (log)	−2.14 (−4.49 to 0.20)	0.097	0.31 (−2.60 to 3.22)	0.567
Diabetes prevalence (log)	−2.63 (−4.77 to −0.48)	0.035	−2.26 (−5.03 to 0.50)	0.214
Testing policy				
0 (ref)				
1	1.14 (−1.59 to 3.87)	0.365	0.02 (−2.51 to 2.56)	0.541
2	0.35 (−2.52 to 3.23)	0.524	−0.11 (−2.75 to 2.54)	0.707
3	−0.49 (−4.01 to 3.04)	0.654	−0.67 (−3.98 to 2.64)	0.642

Testing policy: 0, no policy; 1, tested only symptomatic and specified; 2, tested anyone symptomatic; 3, tested open public.
CFR, case-fatality rate; CVD, cardiovascular disease; GDP, gross domestic product; ref, reference.

Economic level and COVID-19 CFR

In our analysis, high COVID-19 CFR was found mainly around North America and West Europe (figure 1). One of the possible reasons might be that these countries
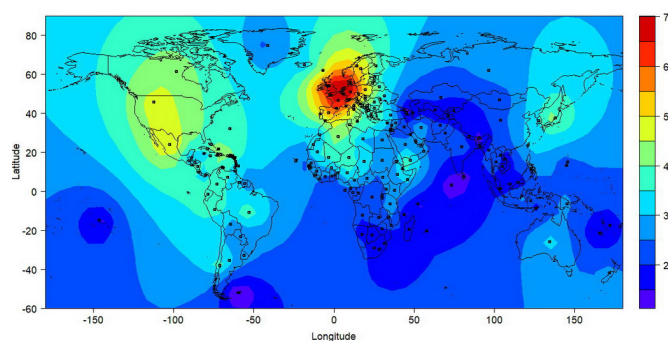

Figure 5 Contour plot of estimated COVID-19 case-fatality rate (%).

counted COVID-19 deaths by including those who died with it, not only from it.[8 64] Determination of COVID-19 deaths also differed by country. Some countries recorded a COVID-19 death as any death once the patient became a confirmed case, even the death happened after 2 months possibly by other reasons (such as an accident), while in some other countries, a COVID-19 death was recorded as the death that occurred within a certain period (ranging from 2 to 8 weeks) after COVID-19 symptom onset.[65] Furthermore, the extent that the counting covered home, institutions and hospitals in high-income countries is different from that in low-income countries.[64]

It has been reported that CFR was more favourable in low-income countries.[64 66] There are three possibilities to explain this unusual pattern: it may be because of younger population, poor data quality or it was still the early stage (at the time of writing this paper) of COVID-19 infection.[64] There is a tight relationship between the income

Table 3 Estimation for variables' regression coefficients (beta) for CFR by testing policy in upper-income to middle-income and high-income countries

Variables	Unadjusted		Adjusted	
	Beta (95% CI)	P value	Beta (95% CI)	P value
Testing policy=1 (tested only symptomatic and specified, n=45)				
GDP per capita (log)	0.51 (−1.05 to 2.08)	0.522	−0.22 (−2.04 to 1.61)	0.805
Population (log)	0.34 (−0.29 to 0.97)	0.290	0.32 (−0.34 to 0.99)	0.393
Population density (log)	0.49 (−0.15 to 1.13)	0.132	0.58 (−0.15 to 1.30)	0.131
Stringency Index	−0.07 (−0.22 to 0.08)	0.349	−0.06 (−0.23 to 0.10)	0.469
Proportion of female smokers (log)	0.20 (−0.96 to 1.36)	0.667	0.13 (−1.28 to 1.55)	0.409
Proportion of male smokers	−0.08 (−0.19 to 0.03)	0.186	−0.04 (−0.20 to 0.11)	0.555
CVD death rate (log)	−1.69 (−3.91 to 0.52)	0.137	−0.20 (−3.70 to 3.30)	0.798
Diabetes prevalence (log)	−0.65 (−3.52 to 2.22)	0.659	−0.82 (−4.37 to 2.73)	0.672
Testing policy=2 or 3 (tested anyone symptomatic or anyone public, n=37)				
GDP per capita (log)	0.35 (−1.42 to 2.11)	0.706	0.45 (−1.19 to 2.09)	0.590
Population (log)	0.81 (0.38 to 1.23)	0.000	0.49 (0.03 to 0.96)	0.047
Population density (log)	0.08 (−0.55 to 0.7)	0.812	0.02 (−0.61 to 0.65)	0.913
Stringency Index	0.17 (0.05 to 0.28)	0.005	0.14 (0.01 to 0.27)	0.044
Proportion of female smokers (log)	0.70 (−0.22 to 1.61)	0.151	0.40 (−0.52 to 1.33)	0.400
Proportion of male smokers	0.00 (−0.10 to 0.10)	0.890	0.00 (−0.11 to 0.11)	0.845
CVD death rate (log)	−2.34 (−4.66 to −0.02)	0.049	−2.00 (−4.95 to 0.94)	0.189
Diabetes prevalence (log)	−2.45 (−5.21 to 0.32)	0.089	−3.30 (−5.85 to −0.74)	0.013

CFR, case-fatality rate; CVD, cardiovascular disease; GDP, gross domestic product.

level of a country and demographic structure. In 2015, many of African countries were classified as low-income, and the median age of the population was less than 20 years. About 61% of the population was 24 years or younger, and merely 3% was equal to or older than 65 years.[67] It has been shown that younger age is associated with a lower likelihood of severe COVID-19.[4 68 69] However, the prevalence of risk factors such as lack of hygiene facilities, handwashing soap and water is greater,[64] and higher viral load has been suggested to be linked to more severe disease.[70] Healthcare resources are usually low in low-income countries.[71] Therefore, our finding of favourable CFR for low-income countries may be in part due to the pandemic being at an early stage (figure 6C). Finally, even before the pandemic, developing countries had challenges to collect, verify and aggregate reliable data
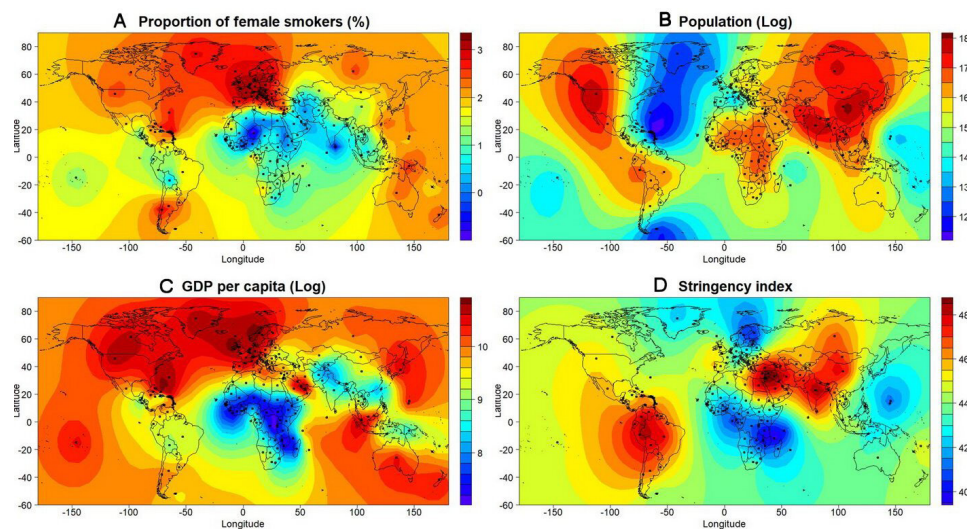

Figure 6 Contour plot of projected (A) proportion of female smoker, (B) population, (C) GDP per capita and (D) Stringency Index. GDP, gross domestic product.

Table 4 Stringency of measures against COVID-19 epidemic and COVID-19 CFR by continent

Continent	Stringency Index		CFR (%)	
	Mean	Median	Mean	Median
Africa	41.57	43.55	2.33	1.76
Asia	46.79	48.85	2.10	1.24
Europe	42.57	43.70	5.15	4.25
North America	45.71	46.42	4.28	3.10
Oceania	42.62	40.85	1.47	0.66
South America	49.35	51.28	2.96	2.88

CFR, case-fatality rate.

in a timely manner due to lack of resource, communication and technological development.[72] Moreover, the pandemic might have accentuated the pre-existing challenges.[64] The extent of bias is difficult to know, including whether it is still in the early stage of infection in developing countries. Further monitoring and investigation are necessary in the future.

Stringency of measures against the COVID-19 epidemic and COVID-19 CFR

According to the currently available data from the Oxford COVID-19 Government Response Tracker,[36] South Americans and Asians took the strictest measures (table 4 and figure 6D), and they also had relatively lower COVID-19 CFR (table 4). However, in our multivariable analysis, which controlled for other variables and spatial dependence, we did not observe a statistically significant association between Stringency Index and COVID-19 CFR. In contrast, stricter measures were even found being associated with higher CFR in the high-income countries with active testing policies (table 3), which seems to support the current argument that lockdown measures might result in excess deaths in underprivileged populations, and those in need are hit harder by the crisis.[73] So far, the evidence that stricter response reduced healthcare strain or treatment efficiency reflected by COVID-19 CFR is lacking. However, the findings need to be further examined by comparing the all-cause mortalities in previous years. Meanwhile, the reliability of the Stringency Index also needs to be further investigated. The relationship between socioeconomic measures against the pandemic and COVID-19 CFR is a complicated issue which needs deeper spatiotemporal analysis with more detailed and reliable information in the future.

Noticeably, we also observed negative associations between COVID-19 CFR and CVD death rate and diabetes prevalence in some analyses, which might be partially explained by the competing risk between the deaths and/or comorbidities,[74] because most of COVID-19 deaths are among the elderly and have one or more comorbidities.[75–77] Therefore, the COVID-19 CFR worldwide deserves deeper investigation with more detailed comorbidity information.

Strengths and limitations

To our knowledge, this is the first study that investigated relationship between COVID-19 CFR and demographic and socioeconomic factors globally. Although numerous studies have investigated the aforementioned factors related to the COVID-19 CFR, either they investigated the question locally or they did not approach this issue from a geospatial perspective. Our study may inspire new reflections from healthcare workers to work together against the COVID-19 pandemic geographically and globally. International comparison of CFR may be challenging when the ascertainment of COVID-19 cases differed by country. To tackle this, we performed a sensitivity analysis using CDR. Although some risk factors, such as CVD and diabetes, showed different patterns of association, the population showed consistent and positive associations with COVID-19 CDR (see online supplemental tables S1 and S2).

There are many limitations in our study. First, the case fatality analysed here was based on the reported COVID-19 cases and deaths by countries/territories. According to the recent estimations, asymptomatic carriers of COVID-19 could be as high as 10%–80% in a population.[78–83] However, this fraction was not taken into account in our analysis. Therefore, the CFRs presented in the study might be significantly higher than the real ones. In addition, there is no single globally accepted definition of COVID-19-related death; therefore, the variation in the reported values of CFR could not be fully explained, and the bias derived from the difference in the definition of COVID-19-related death between the countries could not be excluded using the data available so far. Second, the age structure of the population influences both prevalence and mortality of COVID-19; although we adjusted our analysis using the proportion of age over 65 years in populations, residual confounding largely remains. Third, no individual-level data are available in the current study; thus, results should not be extrapolated to individual-level association. Fourth, because no diagnostic date was available in the Our World in Data, the time between diagnosis and death was not known, which could lead to variation in patient follow-up time among countries and, therefore, potential differences in CDR (because the CDR is calculated using person-time). Fifth, country-level analysis may conceal huge discrepancies between subnational entities in terms of both outcomes and predictors. The case of Northern and Southern Italy is an epitome of this. In-depth geospatial studies conducted at subnational levels are expected to provide less biassed and more actionable results. Furthermore, during an ongoing pandemic, delayed reporting occurs for both the number of cases and deaths, and strategies against the crisis also change by time. Although the analysis using data from two different time points obtained on 17 June and on 2 July produced the same results, suggesting the bias due to delayed report might be negligible; the dynamic of the problem needs to be addressed, incorporating with temporal statistics methods.

CONCLUSION

The average of country/territory-specific COVID-19 CFR is about 2%–3% worldwide, which is higher than previously reported at 0.7%–1.3% and possibly due to the unreported asymptomatic cases. The COVID-19 CFR is statistically significantly associated with population size, especially in middle-income and high-income countries, which may indicate the healthcare strain and/or lower treatment efficiency in the countries with large populations, and secondary to higher transmission risk and generally poorer health. When testing is warranted and healthcare resources are sufficient, strict quarantine and/or lockdown measures might result in excess deaths in countries with high-income level. No statistically significant findings were found in low-income countries, which might be due to the challenges in data collection, communication and verification in the countries and need to be further investigated in follow-up studies. To make global joint strategy and/or policy against the COVID-19 pandemic, spatial dependence and temporal trends must be considered in data analysis and decision making.

Author affiliations
[1]Clinical Epidemiology and Biostatistics, Universitetssjukhuset Örebro, School of Medical Sciences, Örebro University, Örebro, Sweden
[2]Unit of Integrative Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[3]Department of Public Health Sciences, Stockholm University, Stockholm, Sweden
[4]Clinical Epidemiology Division, Department of Medicine, Karolinska Institutet, Stockholm, Sweden
[5]Department of Epidemiology and Public Health, University College London, London, United Kingdom

Twitter Scott Montgomery @clin_epi

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The data used in the study are freely available in the open source for research. There are no individual data available in the dataset. The owner of the data permits the use, distribution and reproduction of the data in any medium. Therefore, no private and confidential information could be disclosed in the study, and ethical approval is not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data used in this study are publicly available at https://ourworldindata.org/ and are referenced in the article.

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

ORCID iDs
Yang Cao http://orcid.org/0000-0002-3552-9153
Scott Montgomery http://orcid.org/0000-0001-6328-5494

REFERENCES

1. Margallo II LN, Diaz M, Lim PP. 2019 novel coronavirus pandemic: what do we know? *South Dakota Med* 2020;73.
2. Henderson LA, Canna SW, Schulert GS, *et al*. On the alert for cytokine storm: immunopathology in COVID-19. *Arthritis Rheumatol* 2020;72:1059–63.
3. Zaim S, Chong JH, Sankaranarayanan V, *et al*. COVID-19 and multiorgan response. *Curr Probl Cardiol* 2020;45:100618.
4. Jordan RE, Adab P, Cheng KK. Covid-19: risk factors for severe disease and death. *BMJ* 2020;368:m1198.
5. Worldometer. *COVID-19 coronavirus pandemic*. Worldometer, 2020. https://www.worldometers.info
6. Mahase E. Covid-19: death rate is 0.66% and increases with age, study estimates. *Bmj-Brit Med J* 2020;369.
7. Baud D, Qi X, Nielsen-Saines K, *et al*. Real estimates of mortality following COVID-19 infection. *Lancet Infect Dis* 2020;20:773.
8. 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.
9. Rajgor DD, Lee MH, Archuleta S, *et al*. The many estimates of the COVID-19 case fatality rate. *Lancet Infect Dis* 2020;20:776–7.
10. Russell TW, Hellewell J, Jarvis CI, *et al*. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the diamond Princess cruise SHIP, February 2020. *Euro Surveill* 2020;25.
11. Gross B, Zheng Z, Liu S, *et al*. Spatio-Temporal propagation of COVID-19 pandemics. *medRxiv* 2020;3:20041517.
12. Lawson AB. *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. CRC press, 2013.
13. Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994-2010. *BMC Infect Dis* 2014;14:285.
14. Palk L, Blower S. Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the HIV epidemic in Malawi. *BMC Med* 2018;16:22.
15. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. *Geogr Anal* 2006;38:5–22.
16. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Stat Med* 1995;14:799–810.
17. Sasaki S, Suzuki H, Igarashi K, *et al*. Spatial analysis of risk factor of cholera outbreak for 2003-2004 in a peri-urban area of Lusaka, Zambia. *Am J Trop Med Hyg* 2008;79:414–21.
18. Pfeiffer D, Robinson TP, Stevenson M, *et al*. *Spatial analysis in epidemiology*. Oxford University Press: Oxford, 2008.
19. Brooke J, Jackson D, people O. Older people and COVID-19: isolation, risk and ageism. *J Clin Nurs* 2020;29:2044–6.
20. Jin J-M, Bai P, He W, *et al*. Gender differences in patients with COVID-19: focus on severity and mortality. *Front Public Health* 2020;8:152.
21. Zhou F, Yu T, Du R, *et al*. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
22. Klang E, Kassim G, Soffer S, *et al*. Severe obesity as an independent risk factor for COVID-19 mortality in hospitalized patients younger than 50. *Obesity* 2020;28:1595–9.
23. Zhang F, Xiong Y, Wei Y, *et al*. Obesity predisposes to the risk of higher mortality in young COVID-19 patients. *J Med Virol* 2020:2542.
24. Guo W, Li M, Dong Y, *et al*. Diabetes is a risk factor for the progression and prognosis of COVID -19. *Diabetes Metab Res Rev* 2020;36.
25. RH D, Liang LR, Yang CQ, *et al*. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *Eur Respir J* 2020;55.
26. Cheng Y, Luo R, Wang K, *et al*. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int* 2020;97:829–38.

27 Bibbins-Domingo K. This time must be different: disparities during the COVID-19 pandemic. *Ann Intern Med* 2020.

28 Wu X, Nethery RC, Sabath BM, *et al*. Exposure to air pollution and COVID-19 mortality in the United States: a nationwide cross-sectional study. *medRxiv* 2020

29 Ogen Y. Assessing nitrogen dioxide (NO_2) levels as a contributing factor to coronavirus (COVID-19) fatality. *Sci Total Environ* 2020;726:138605.

30 Ma Y, Zhao Y, Liu J, *et al*. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci Total Environ* 2020;724:138226.

31 Our World in Data, 2020. Available: https://ourworldindata.org/ [Accessed 23 Aug 2020].

32 Roser M, Ritchie H, Ortiz-Ospina E, *et al*. *Coronavirus disease (COVID-19)–Statistics and research*. Our World in Data, 2020.

33 International Organization for Standardization (ISO). *Iso 3166-1: codes for the representation of names of countries and their subdivisions–Part 1: country codes*. Switzerland: ISO Geneva, 2006.

34 Harrington RA. Case fatality rate in encyclopedia britannica, 2020. Available: https://www.britannica.com/science/case-fatality-rate [Accessed 15 Jul 2020].

35 Centers for Disease Control Prevention. *Principles of epidemiology in public health practice*. Atlanta, US: US Department of Health and Human Services, 2012.

36 Hale T, Webster S, Petherick A, *et al*. Oxford covid-19 government response tracker. *Blavatnik School of Government*. 25, 2020.

37 World Intellectual Property Organization. *WIPO's COVID-19 Response*. New York: World Intellectual Property Organization, 2020. https://www.wipo.int/covid19-policy-tracker/#/covid19-policy-tracker/ipo-operations

38 Mason CH, Perreault WD, Collinearity PWD. Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research* 1991;28:268–80.

39 Lee S-I. Correlation and spatial autocorrelation. In: Shekhar S, Xiong H, Zhou X, eds. *Encyclopedia of GIS*. Heidelberg, Germany: Springer Nature, 2017: 360–8.

40 Ren T, Long Z, Zhang R, *et al*. Moran's I test of spatial panel data model — based on bootstrap method. *Econ Model* 2014;41:9–14.

41 Matérn B. *Spatial variation*. Springer Science & Business Media, 2013.

42 Ward MD, Gleditsch KS. *Spatial regression models*. Sage Publications, 2018.

43 Royston P. Multiple imputation of missing values. *Stata J* 2004;4:227–41.

44 World Bank. World bank country and lending groups, 2020. Available: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [Accessed 30 Jun 2020].

45 Rousset F, Ferdy J-B. Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography* 2014;37:781–90.

46 van Rossum G. *Python tutorial, technical report CS-R9526, Centrum voor Wiskunde en Informatica (CWI)*. Amsterdam, 1995.

47 Jordahl K, Jd B, Wasserman J, *et al*. *Geopandas/geopandas: V0. 6.1: Zenodo*, 2019.

48 Stevens GA, Alkema L, Black RE, *et al*. Guidelines for accurate and transparent health estimates reporting: the gather statement. *PLoS Med* 2016;13:e1002056.

49 Meade MS. *Medical geography. The Wiley Blackwell encyclopedia of health, illness, behavior, and society*. Hoboken, NJ: John Wiley & Sons, Inc, 2014: 1375–81.

50 Franch-Pardo I, Napoletano BM, Rosete-Verges F, *et al*. Spatial analysis and GIS in the study of COVID-19. A review. *Sci Total Environ* 2020;739:140033.

51 Channappanavar R, Fett C, Mack M, *et al*. Sex-Based differences in susceptibility to severe acute respiratory syndrome coronavirus infection. *J Immunol* 2017;198:4046–53.

52 Vardavas CI, Nikitara K. COVID-19 and smoking: a systematic review of the evidence. *Tob Induc Dis* 2020;18:20.

53 Cai H. Sex difference and smoking predisposition in patients with COVID-19. *Lancet Respir Med* 2020;8:e20.

54 Pampel FC. Global patterns and determinants of sex differences in smoking. *Int J Comp Sociol* 2006;47:466–87.

55 Khlat M, Pampel F, Bricard D, *et al*. Disadvantaged social groups and the cigarette epidemic: limits of the diffusion of innovations vision. *Int J Environ Res Public Health* 2016;13

56 Huisman M, Kunst AE, Mackenbach JP. Inequalities in the prevalence of smoking in the European Union: comparing education and income. *Prev Med* 2005;40:756–64.

57 Hitchman SC, Fong GT. Gender empowerment and female-to-male smoking prevalence ratios. *Bull World Health Organ* 2011;89:195–202.

58 Gallus S, Lugo A, Liu X, *et al*. Who smokes in Europe? data from 12 European countries in the TackSHS survey (2017-2018). *J Epidemiol* 2020

59 World Health Organization. *COVID-19 strategy update*. Geneva, Switzerland, 2020.

60 Bassino J-P, Ladmiral G. *Socio-economic factors influencing COVID-19 spread in Japan; evidence from the second wave (March-May 2020). Evidence from the second wave (March-May 2020) (June 30, 2020)*, 2020. http://dx.doi.org/10.2139/ssrn.3637994

61 Sjödin H, Wilder-Smith A, Osman S, *et al*. Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in Italy, 2020. *Eurosurveillance* 2020;25:7–12.

62 Ahmadi M, Sharifi A, Dorosti S, *et al*. Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Sci Total Environ* 2020;729:138705.

63 Cheng H-Y, Li S-Y, Yang C-H. Initial rapid and proactive response for the COVID-19 outbreak - Taiwan's experience. *J Formos Med Assoc* 2020;119:771–3.

64 Schellekens P, Sourrouille DM. COVID-19 mortality in rich and poor countries: a tale of two pandemics? world bank policy research working paper, 2020. Available: https://ssrn.com/abstract=3614141

65 Our World in Data. Coronavirus (COVID-19) deaths, 2020. Available: https://ourworldindata.org/covid-deaths [Accessed 22 Jul 2020].

66 Cash R, Patel V. Has COVID-19 subverted global health? *Lancet* 2020;395:1687–8.

67 United Nations. *World population prospects 2019: highlights*. New York City: United nations, 2019.

68 Li X, Xu S, Yu M, *et al*. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol* 2020;146:110–8.

69 CDC COVID-19 Response Team. Severe outcomes among patients with coronavirus disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:343–6.

70 Zheng S, Fan J, Yu F, *et al*. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang Province, China, January-March 2020: retrospective cohort study. *BMJ*;15:m1443.

71 Fryatt R, Mills A, Nordstrom A. Financing of health systems to achieve the health millennium development goals in low-income countries. *Lancet* 2010;375:419–26.

72 Bram JT, Warwick-Clark B, Obeysekare E, *et al*. Utilization and Monetization of healthcare data in developing countries. *Big Data* 2015;3:59–66.

73 Melnick ER, Ioannidis JPA. Should governments continue lockdown to slow the spread of covid-19? *BMJ*;2020:m1924.

74 Berry SD, Ngo L, Samelson EJ, *et al*. Competing risk of death: an important consideration in studies of older adults. *J Am Geriatr Soc* 2010;58:783–7.

75 Garnier-Crussard A, Forestier E, Gilbert T, *et al*. Novel coronavirus (COVID-19) epidemic: what are the risks for older patients? *J Am Geriatr Soc* 2020;68:939–40.

76 Yang J, Zheng Y, Gou X, *et al*. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis* 2020;94:91–5.

77 Richardson S, Hirsch JS, Narasimhan M, *et al*. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA* 2020;323:2052.

78 Nishiura H, Kobayashi T, Miyama T, *et al*. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis* 2020;94:154–5.

79 Mizumoto K, Kagaya K, Zarebski A, *et al*. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the diamond Princess cruise SHIP, Yokohama, Japan, 2020. *Euro Surveill* 2020;25.

80 World Health Organization. Q&A: Similarities and differences – COVID-19 and influenza, 2020. Available: https://www.who.int/news-room/q-a-detail/q-a-similarities-and-differences-covid-19-and-influenza [Accessed 26 Jun 2020].

81 Kimball A, Hatfield KM, Arons M, *et al*. Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:377–81.

82 Anastassopoulou C, Russo L, Tsakris A, *et al*. Data-Based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* 2020;15:e0230405.

83 Day M. Covid-19: four fifths of cases are asymptomatic, China figures indicate. *BMJ* 2020;369:m1375.