

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Country-level Determinants of the Severity of the First Global Wave of the COVID-19 Pandemic: An Ecological Study
AUTHORS	Pana, Tiberiu; Bhattacharya, Sohinee; Gamble, David; Pasdar, Zahra; Szlachetka, Weronika; Perdomo-Lampignano, Jesus; Ewers, Kai; McLernon, David; Myint, Phyoo

VERSION 1 – REVIEW

REVIEWER	Chandrashekhara Sreeramareddy International Medical University
REVIEW RETURNED	01-Aug-2020

GENERAL COMMENTS	<p>Thank you for the opportunity to review this manuscript. Indeed very interesting at this point in time when Covid-19 is still continuing to surge world wide. Perhaps, the importance of the study and its implications will be become relevant future pandemics if not that much at the current phase of covid-19 pandemic. Authors have cited few other studies which did analyses on the worst-hit countries, the two studies some what different findings as they did included different variables into the analyses and regressions modelling approaches vary by study, depending on how the authors conceptualized the design of such analyses/papers.</p> <p>Well, I have no objections in the variables included albeit a number of limitations of data they had and an inherent limitations of ecological approach. I have some concerns on the data analyses approach and variables included in the models.</p> <p>Results looks like restrictions on entry into the country as indicated by arrivals was associated with mortality is too simplistic conclusions, as even this analyses is not needed for such a conclusion as per the discussed WHO pandemic preparedness. I feel this is most distal factors. Perhaps, the demographic structure, sunlight exposure (vitamin-d related immunity), BCG-vaccination coverage, may have been important variables to be included as well. Authors have discussed increasing number of cumulative cases and decreasing testing intensity (test per million population), these are more directly linked to the mortality used by the authors. % of population over 65 years rather than median age, would be more relevant, as the population structure vary by countries.</p> <p>Further, increase in cases not only linked to arrivals but also border control as mentioned by authors about Thailand, i.e. testing and isolation, mortality would also be determined by speed of increase in case loads and health care capacity, i.e. intensive care units per million population etc, since, lives can be saved in care was quality care was given in timely and of quality (this often</p>
-------------------------	--

	<p>suffers with increases occupancy in intensive care). One last item, is stage at which hospital care was sought would directly determine the clinical outcomes.</p> <p>Not withstanding these points i raised, i feel that analyses should account to random noise, i.e. models are fixed effects or random effects. Seems like authors entered all variables into the models at first step and based on their criteria on significance, they entered remaining variables in the final model. I suggest they used some conceptual framework (perhaps develop their own) which factors they tested would be most proximal to their outcome i.e deaths, and which outcomes would be most distal to the outcome.</p> <p>Perhaps, WHO health index has some healthcare quality but hidden inside a composite index, as is border closure timing with time reference to the pandemic hidden inside stringency index. Authors want include them as stand alone instead of composite indices. Try these variables in the models mentioning random or fixed effects, and step wise from distal most and proximal variables in the modelling.</p>
--	--

REVIEWER	Amalia Magaret University of Washington
REVIEW RETURNED	07-Oct-2020

GENERAL COMMENTS	<p>Review: Pana2020 BMJ Open Number of International Arrivals Is Associated with the severity of the first global wave of COVID-19</p> <p>Summary</p> <p>Summary of country-level risk factors for COVID deaths over the ramp up period while the virus began to spread. There is an impressive range of data sources collected to answer these questions.</p> <p>Main points</p> <p>Page 10, line 38, Can you give more details about what it means for the curve to reach a peak? I could imagine the analysis would be sensitive to that definition and to the time interval used to compute the outcome.</p> <p>I would be interested in a bit more information about the stringency index, whether it is thought to capture meaningfully the prevention efforts.</p> <p>Does it make sense given the changing temperature to run a model with a time point per month for each country, as a sensitivity analysis perhaps, in addition to using a single time point over the entire period? It seems that the number of international arrivals varied during that time period as well, while restrictions were put into place and then eased.</p> <p>I wonder about adjusting for population size as a covariate when the rate of deaths is on the log scale. Should there be a sensitivity analysis using incidence of death rather than absolute number of deaths over time? Perhaps adjusting for natural log of population</p>
-------------------------	--

	<p>instead of absolute population size would achieve the same goal? It seems number of deaths is so closely linked to available bodies that this might be considered directly.</p> <p>Table 2, some of the confidence intervals and p-values are not consistent. For example, the confidence interval for the coefficient for population in row 1 includes zero, while the p-value is less than 0.05. In the multivariable analysis in the same row, the confidence interval does not include zero, but the p-value is greater than 0.05. Also, the authors state that international arrivals were the only significant predictor in the multivariable model, but the p-value for GDP is 0.009 (while again the confidence interval is inconsistent b/c it includes zero).</p> <p>Minor points</p> <p>I might prefer to see the countries listed by continent rather than alphabetically.</p> <p>Page 11 lines 24 to 32, it is still unclear what the measure of smoking is. Proportion who smoke? Average cigarettes per person?</p> <p>Perhaps don't give an acronym to the stringency index.</p> <p>What is the difference between the list of predictors included in multivariable analysis found at the bottom of page 12 and those found on page 16 lines 12-20?</p> <p>Page 21, "overfitting" is not used correctly here.</p> <p>Figures 2a and 2b, please give more information about the confidence bands displayed here.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer: 1

Reviewer Name: Chandrashekhara Sreeramareddy

Institution and Country: International Medical University, Nepal.

Please state any competing interests or state 'None declared': None declared

Dear Editor,

Thank you for the opportunity to review this manuscript. Indeed very interesting at this point in time when Covid-19 is still continuing to surge world wide. Perhaps, the importance of the

study and its implications will be become relevant future pandemics if not that much at the current phase of covid-19 pandemic. Authors have cited few other studies which did analyses on the worst-hit countries, the two studies some what different findings as they did included different variables into the analyses and regressions modelling approaches vary by study, depending on how the authors conceptualized the design of such analyses/papers.

Response: Thank you for taking the time to review our work and for highlighting that our work is of interest.

Wel, i have no objections in the variables included albeit a number of limitations of data they had and an inherent limitations of ecological approach. I have some concerns on the data analyses approach and variables included in the models.

Response: Thank you for your comment. We have discussed the inherent limitations of our ecological study in the limitations paragraph of the discussion section of our originally submitted manuscript. Furthermore, we have only drawn cautious conclusions from our analyses, accounting for the study design.

Results looks like restrictions on entry into the country as indicated by arrivals was associated with mortality is too simplistic conclusions, as even this analyses is not needed for such a conclusion as per the discussed WHO pandemic preparedness.

Response: Thank you for highlighting the above issues. As noted in the discussion section of our originally submitted manuscript, the WHO recommendations for pandemic preparedness and resilience suggest that points of entry into the country should be monitored by focussing on surveillance and risk communication to travellers but falls short of closing down international travel. The results of our study highlight the strength of association between international travel and the severity of the first wave of the COVID-19 pandemic, which remained highly significant despite adjustments for restrictions on international travel, suggest that at least in the particular case of SARS-CoV-2, early restrictions on international travel should be considered in order to limit the severity of the pandemic.

I feel this is most distal factors. Perhaps, the demographic structure, sunlight exposure (vitamin-d related immunity), BCG-vaccination coverage, may have been important variables to be included as well. Authors have discussed increasing number of cumulative cases and decreasing testing intensity (test per million population), these are more directly linked to the mortality used by the authors. % of population over 65 years rather than median age, would be more relevant, as the population structure vary by countries. Further, increase in cases not only linked to arrivals but also border control as mentioned by authors about Thailand, i.e. testing and isolation, mortality would also be determined by speed of increase in case loads and health care capacity, i.e. intensive care units per million population etc, since, lives can be saved in care was quality care was given in timely and of quality (this often suffers with increases occupancy in intensive care).

Response: Thank you for your kind and helpful suggestions. We have adjusted our variables for the country-level proportion of population aged 65 and over instead of median age. Furthermore, we have added the following country-level determinants into our analyses: ambient UV radiation level (as a proxy for sunlight exposure), BCG vaccination coverage and testing capacity, quantified as the total number of COVID-19 tests per 1000 population performed until the 8th of June 2020. Of note, a few countries included in our analyses (Algeria, Brazil, Egypt, France, Germany, the Netherlands, Spain and Sweden), did not report COVID-19 testing numbers. We have thus performed a secondary analysis (results detailed in the newly added Table 3) excluding these countries but including testing

capacity as an additional variable. We have revised the methods, results and discussion sections of manuscript to reflect the results of the updated analyses.

Under **METHODS**, page 12, lines 1-6:

“Country-level exposure to UV radiation was quantified as the population-weighted average daily ambient ultraviolet radiation level measured in J/m² for the years 1997-2003³⁵. BCG vaccination coverage was quantified as the average percentage of 1 year-old children having received the BCG vaccine between 1980 and 2019 in each country. Testing capacity was quantified as the total number of COVID-19 tests per 1000 population performed until the 8th of June 2020.”

Under **METHODS**, page 13, lines 3-13:

“Given that testing capacity data for 8 (Algeria, Brazil, Egypt, France, Germany, the Netherlands, Spain and Sweden) of the 37 included countries were not available, a secondary analysis also including testing capacity as a predictor was performed considering only the remaining 29 countries. Linear regressions were performed to assess the univariable relationship between each country-level predictor and the calculated mean mortality rate for each country. Predictors reaching a P-value <0.3 at univariable level were then included in a multivariable logistic regression model with the natural logarithm of the mean mortality rate as outcome: the logarithm of the total population in 2018, percentage of population aged 65 and over, international arrivals, population density, prevalent neoplasms, prevalent hypertension, GDP per capita, UV radiation exposure, mean BCG coverage, the stringency index and testing capacity.”

Under **RESULTS**, page 17, lines 1-17:

“Table 2 details the results of the linear regression analyses. The following country-level predictors showed a statistically significant relationship with log mean mortality rate at univariable level: international arrivals in 2018 (coefficient (95% confidence interval) = 0.049 (0.033, 0.064), $P < 0.001$), prevalent neoplasms (0.614 (0.209, 1.019), $P = 0.005$) and prevalent hypertension (-0.150 (-0.254, -0.045), $P = 0.008$)- natural logarithm of population, international arrivals, prevalent neoplasms, prevalent hypertension, GDP per capita and BCG vaccination coverage. ~~The multivariable model included the following predictors, which were selected from univariable models: median age, pollution levels, mean temperature, international arrivals, prevalent neoplasms, prevalent hypertension, WHO health index, percentage of population living in urban areas, GDP per capita and the stringency index.~~ Upon multivariable adjustment, International arrivals in 2018, as a marker of global connection, was the ~~only~~-main statistically significant predictor of log mean mortality rate (0.040 (0.017-, 0.063) for 1 million increase in international arrivals, $P = 0.002$) along with mean BCG vaccination coverage (-0.018 (-0.034, -0.002) for 1% increase in BCG vaccination coverage, $P = 0.031$). Figures 2 and 3 detail the relationship between the country-level log mean mortality rate (predicted and observed) and each country-level predictor included in the multivariable regression model.”

Under **RESULTS**, page 20, lines 1-11:

“Table 3 details the results of the secondary linear regression analyses, including only countries having reported COVID-19 testing data up to the 8th of June 2020. The following country-level predictors showed a statistically significant relationship with log mean mortality rate at univariable level: natural logarithm of population, international arrivals, prevalent neoplasms, prevalent hypertension, BCG vaccination coverage and total COVID-19 tests per 1000 population performed until the 8th of June 2020. Upon multivariable adjustment, the statistically significant predictors of log mean mortality rate were: international arrivals in 2018 (0.036 (0.008, 0.063) for 1 million increase in international arrivals, $P = 0.013$), prevalent hypertension (-0.129 (-0.246, -0.012) for 1% increase in country-level hypertension prevalence, $P = 0.032$) and testing capacity (0.018 (0.001, 0.034) for 1 per 1000 population increase in the number of total COVID-19 tests performed until the 8th of June 2020, $P = 0.039$).”

Under **DISCUSSION**, page 23, lines 3-14:

“In this ecological study including data from 37 countries which were most severely affected by COVID-19 in the first wave of current Global pandemic, we assessed 169 country-level socioeconomic, environmental, health and healthcare system, and globalisation parameters as potential predictors of variation in death rates from COVID 19 infection. In the multivariable linear regression model, the ~~only~~ main predictor that reached statistical significance was international arrivals, a proxy of global connection: increases in international arrivals were associated with higher mean mortality rate. Furthermore, country-level BCG vaccination coverage was associated with decreases in the COVID-19 mean mortality rate during the first wave of the pandemic. Finally, in our secondary analyses including only country with available testing capacity data, the total number of COVID-19 tests performed per 1000 population until the 8th of June 2020 was associated with increases in the COVID-19 mean mortality rate.”

Under **DISCUSSION**, page 25, lines 15-24 and page 25, lines 1-8:

“Our multivariable model also suggests an inverse relationship between BCG vaccination coverage the mean mortality rate, in which increasing BCG vaccination coverage was associated with decreased mean mortality rate. The relationship between BCG vaccination and the evolution of the COVID-19 transmission and disease severity remains controversial^{43,44}. While the BCG vaccine has been postulated to exhibit non-specific immunomodulatory properties, which may reduce SARS-CoV-2 viraemia after exposure⁴³, current epidemiological evidence is derived from ecological studies⁴⁵ and needs to be interpreted in the light of the inherent limitations of this study design. Further ongoing studies (NCT04327206⁴⁶, NCT04328441⁴⁷) may provide more robust evidence regarding the association between BCG vaccination and COVID-19.

Our analyses also revealed a few surprising findings: the intensity of COVID-19 testing was apparently associated with mean mortality rate increases while the country-level prevalence of hypertension was apparently associated with mean mortality rate decreases. These findings appear to be contradictory to previous evidence suggesting that testing intensity may be associated with decreased COVID-19 mortality⁴⁸, while hypertension was clearly associated with increased mortality⁴⁹. These surprising findings need to be interpreted in the light of our ecological study design in which residual confounders may influence these associations.”

Under **CONCLUSION**, page 28, lines 12-16:

“The associations between other predictors, such as BCG vaccination coverage, prevalent hypertension and COVID-19 testing capacity, and the outcome were weaker and need to be interpreted in the light of our ecological study design. Further studies are required to determine the relationship between previous BCG vaccination and COVID-19 disease progression.”

With regards to intensive care (ICU) capacity, as also highlighted in our response to the editorial board, we have considered adding this variable to our analyses, as per your suggestion. Nevertheless, in the absence of a centralised database of country-level ICU capacity, the reporting of this variable remains poor as it relies on countries' individual reporting, leading to potential between-country heterogeneity. Furthermore, we were unable to find publicly available data on ICU capacity for several important countries included in our analyses, such as Algeria, Argentina, Chile, the Dominican Republic, Ecuador, Egypt, India, Indonesia, Peru, the Philippines, Saudi Arabia and Ukraine. Under these circumstances, we have decided not to include ICU capacity in our analyses. Nevertheless, it is important to note that we have already included country-level hospital beds per 10,000 population in our analyses as an indicator of health systems' capacity to cope with the pandemic.

Under **METHODS**, page 11, lines 12-15:

“Data on hospital beds per 10,000 population were defined by the World Bank as including ‘inpatient beds available in public, private, general, and specialized hospitals and rehabilitation centres’. The published data for countries included was from 2000 to 2017. In most cases beds for both acute and chronic care are included.²²”

One last item, is stage at which hospital care was sought would directly determine the clinical outcomes.

Response: Thank you for highlighting this important issue. We acknowledge that the delay between COVID-19 symptom onset and hospitalisation may be an important factor in the overall clinical prognosis of patients with severe COVID-19 disease. Nevertheless, given that our analyses rely on country-level determinants and in the absence of individual patient data, it is impossible to ascertain the country-level trends of delay to hospital admission. Notwithstanding, it is important to highlight that some other country-level parameters pertaining to the accessibility of healthcare included in our analyses such as the number of hospital beds per 10,000 population, proportion of population living in urban areas as well as the WHO health index may account for such differences. We have included a discussion of these aspects in the limitations’ paragraph of our discussion section.

Under **DISCUSSION**, page **27**, lines **22-25** and page **28**, lines **1-4**:

“Furthermore, the delay between COVID-19 symptom onset and hospitalisation may be an important factor in the overall clinical prognosis of patients with severe COVID-19 disease. Nevertheless, given that our analyses rely on country-level determinants and in the absence of individual patient data, it is impossible to ascertain the country-level trends of delay to hospital admission. Notwithstanding, some other country-level parameters pertaining to the accessibility of healthcare included in our analyses such as the number of hospital beds per 10,000 population, proportion of population living in urban areas as well as the WHO health index may account for such differences.”

Notwithstanding these points i raised, i feel that analyses should account to random noise, i.e. models are fixed effects or random effects. Seems like authors entered all variables into the models at first step and based on their criteria on significance, they entered remaining variables in the final model. I suggest they used some conceptual framework (perhaps develop their own) which factors they tested would be most proximal to their outcome i.e deaths, and which outcomes would be most distal to the outcome. Perhaps, WHO health index has some healthcare quality but hidden inside a composite index, as is border closure timing with time reference to the pandemic hidden inside stringency index. Authors want include them as stand alone instead of composite indices. Try these variables in the models mentioning random or fixed effects, and step wise from distal most and proximal variables in the modelling.

Response: Thank you for your comment. With regards to the point you raised about adopting linear mixed models accounting for between-country random effects, we think that whilst this would be an appropriate analytical strategy for individual patient data clustered by country, such models would not be appropriate for our country-level dataset. We have created a dataset in which each included country is a separate entry, for which we have populated the country-level variables of interest, such as population, international arrivals, etc. with data extracted from publicly available datasets, such as the WHO Global Health Observatory. Supplementary File 1 is a Microsoft Excel sheet containing the dataset which was utilised for the statistical analyses described in our manuscript, which will exhibit in more detail the structure of our data. Thus, given that we analysed each country as a separate datapoint, as opposed to analysing individual patient data with patients clustering in different countries, we have decided not to adopt linear mixed models with random effects by country as this would not be suitable for our study design.

With regards to your second point about hierarchical selection of country-level predictors, we think that adopting such an approach instead of our pre-specified univariate predictor selection would make little difference to the overall results, since the predictors in question appear to have no statistically significant relationship with the outcome. Furthermore, such an approach may further hinder the clinical implications drawn from our results, as there remains considerable divergence in opinion regarding colliders and over-adjustment. Some experts have argued that the possibility of colliders has only a minor impact on the results, and that one should aim to comprehensively adjust as far as possible for confounders (Clin Epidemiol. 2018 Jul 6;10:771-788). Thus, we have chosen to keep our original approach.

Reviewer: 2

Reviewer Name: Amalia Magaret

Institution and Country: University of Washington, USA

Please state any competing interests or state 'None declared': none

Summary

Summary of country-level risk factors for COVID deaths over the ramp up period while the virus began to spread. There is an impressive range of data sources collected to answer these questions.

Response: Thank you for taking the time to review our work and for highlighting that our work is of interest.

Main points

Page 10, line 38, Can you give more details about what it means for the curve to reach a peak? I could imagine the analysis would be sensitive to that definition and to the time interval used to compute the outcome.

Response: Thank you for highlighting this issue. We have defined the peak of each mortality curve as the first point at which the first derivative of the COVID-19 mortality as a function of the pandemic timeline became zero. We have now detailed this in the methods section of our manuscript:

Under **METHODS**, page **10**, lines **2-3**:

“The peak of each mortality curve was defined as the first point at which the first derivative of the COVID-19 mortality as a function of the pandemic timeline became zero.”

I would be interested in a bit more information about the stringency index, whether it is thought to capture meaningfully the prevention efforts.

Response: We apologise for not having provided more comprehensive details about the stringency index in the originally submitted version of our manuscript. The stringency index contains comprehensive information regarding country-level public health measures aimed at limiting the spread of COVID-19, such as school and workplace closures, cancelling public events, restrictions on

gatherings, public transport closures, stay-at-home requirements, restrictions on internal movements and international travel controls. We have provided further details regarding the sub-components of the stringency index in the methods section of our manuscript

Under **METHODS**, page 11, lines 15-23:

“The ~~S~~Stringency Index is an overall indicator of public health measures adopted by each country in response to the COVID-19 pandemic and includes containment and closure indicators (school closures, workplace closures, cancelling public events, restrictions on gatherings, public transport closures, stay-at-home requirements, restrictions on internal movements, international travel controls), economic response indicators (income support, debt/contract relief, fiscal measures, international support) as well as health systems indicators (public information campaigns, testing policy, contact tracing, emergency investment in healthcare, investment in vaccines).¹¹”

Does it make sense given the changing temperature to run a model with a time point per month for each country, as a sensitivity analysis perhaps, in addition to using a single time point over the entire period? It seems that the number of international arrivals varied during that time period as well, while restrictions were put into place and then eased.

Response: Thank you for pointing out this important issue. As described in the methods section of our originally submitted manuscript, both the temperature levels were calculated as monthly mean temperatures for the months January to May between 2010 and 2016. Furthermore, the total stringency index introduced in our model was calculated as the daily mean stringency index recorded between the 31st of December 2019 and the 8th of June 2020 for each country considered in our analyses. Thus, our original analytic approach allows us to consider values for temperature data as well as the stringency index that are representative of each country for the length of the exposure period (31/12/19-08/06/20). Finally, we have used the total number of country-level international arrivals in 2018 as a proxy for country-level international connections rather than considering international arrival data occurring during the exposure period. We have thus decided not to perform any sensitivity analyses splitting our exposure period by month, since such an approach would only reduce the power of our analyses by decreasing the number of mortality data points which would be used to calculate the mean mortality rate and would subsequently hinder the interpretation of our results.

I wonder about adjusting for population size as a covariate when the rate of deaths is on the log scale. Should there be a sensitivity analysis using incidence of death rather than absolute number of deaths over time? Perhaps adjusting for natural log of population instead of absolute population size would achieve the same goal? It seems number of deaths is so closely linked to available bodies that this might be considered directly.

Response: Thank you for pointing out this issue. With regard to your point about looking at incident mortality as opposed to the absolute number of deaths over time, we agree that such an approach would be desirable in the context of individual patient data, where a time-to-event analysis framework would allow a better interpretation of the mortality data. Nevertheless, this ecological study did not consider individual patient data, but country-level parameters and thus we are unable to pursue time-to-event analytic approach of our dataset.

With regard to your second point about the natural logarithm of the population data, we have adjusted our analyses in the originally submitted manuscript for the natural logarithm of the population, as we have deemed that only a linear regression considering the logarithm of the country-level population data as an independent variable would satisfy the assumption of homoscedasticity. We have inadvertently described this predictor in the originally submitted manuscript “population” instead of “log population”. We would like to apologise for this error, which has now been corrected throughout the revised version.

Under **METHODS**, page 12, lines 8-24:

“All analyses were performed in Stata 15.1SE, Stata Statistical Software. A 5% threshold of statistical significance was utilised for all analyses ($P < 0.05$). Linear regressions ~~was~~ were performed to assess the univariable relationship between each country-level predictor and the calculated mean mortality rate for each country. The following predictors were included in the univariable analyses: the natural logarithm of the population in 2018 (10 million increase) (~~natural logarithm~~), percentage of population aged 65 and over (~~median age~~), pollution levels, mean temperature (January-May), international arrivals, population density, prevalent diabetes, prevalent neoplasms, median BMI, prevalent hypertension, smoking prevalence, hospital beds (per 10,000 population), WHO health index, percentage population living in urban areas, GDP per capita (PPP), UV radiation exposure, mean BCG coverage and the stringency index (~~and the Stringency Index~~). Predictors reaching a P -value < 0.3 at univariable level were then included in a multivariable logistic regression model with the natural logarithm of the mean mortality rate as outcome: the logarithm of the total population in 2018, percentage of population aged 65 and over, pollution, mean temperature (January-May), international arrivals, population density, prevalent neoplasms, prevalent hypertension, the WHO health index, population living in urban areas, GDP per capita, UV radiation exposure, mean BCG coverage and the stringency index.”

Table 2, some of the confidence intervals and p-values are not consistent. For example, the confidence interval for the coefficient for population in row 1 includes zero, while the p-value is less than 0.05. In the multivariable analysis in the same row, the confidence interval does not include zero, but the p-value is greater than 0.05. Also, the authors state that international arrivals were the only significant predictor in the multivariable model, but the p-value for GDP is 0.009 (while again the confidence interval is inconsistent b/c it includes zero).

Response: Thank you for pointing out this issue. Both in our original as well as in our revised analyses (please note that the analyses have changed in response to comments from reviewer 1 asking for the inclusion of additional country-level determinants in the analyses), GDP per capita was only a significant predictor in univariable, but not multivariable models. Thus, we have not considered it to be a significant predictor of the severity of the first wave of the COVID-19 pandemic according to our pre-specified protocol. We apologise for the minor errors which appeared in Table 2 in the originally submitted manuscript which rendered non-significant results appear significant. Those have now been revised, with all confidence intervals being consistent with their respective P values.

Under **RESULTS**, pages 18-19:

Table 2. Results of the linear regression assessing the country-level predictors of the daily increase in deaths. The predictors achieving a 30% statistical significance level at univariable levels ($P < 0.3$) were included in the multivariable model.

Predictor	Univariable		Multivariable	
	Coefficient (95% CI)	P value	Coefficient (95% CI)	P value
Natural logarithm of population (10 million increase) [2018]	0.432 (-0.050, 0.814)	0.033	0.393 (-0.087, 0.873) 0.347 (-0.163, 0.798)	0.103 0.186
% population aged 65 and older Median age	0.065 (-0.010, 0.139) 0.063 (-0.006, 0.132)	0.097 0.82	-0.020 (-0.143, 0.103) 0.037 (-0.079, 0.154)	0.741 0.512

Pollution levels	-0.017 (-0.044, 0.011)	0.247	-0.005 (-0.031, 0.020) -0.007 (-0.036, 0.021)	0.6590.605
Mean Temperature (January-April-May) [2010-2016]	-0.031 (-0.078, 0.017)	0.218	0.052 (-0.025, 0.128) 0.005 (-0.061, 0.072)	0.1750.869
International Arrivals (1 million increase) [2018]	0.049 (0.033, 0.064)	<0.001	0.033 (0.012, 0.054) 0.040 (0.017, 0.063)	0.0030.002
Population Density	-0.002 (-0.006, 0.002)	0.268	-0.001 (-0.004, 0.002) 0.001 (-0.002, 0.005)	0.5600.377
Diabetes prevalence (% of population ages 20 to 79) [2019]	-0.0031 (-0.189, 0.126)	0.700	--	--
Prevalence - Neoplasms - Sex: Both - Age: Age-standardized (Percent) (%) [2017]	0.614 (0.209, 1.019)	0.005	-0.404 (-1.079, 0.271) 0.301 (-1.062, 0.461)	0.2270.423
Median BMI	0.010 (-0.297, 0.318)	0.947	--	--
Prevalent Hypertension (%), [2015]	-0.150 (-0.254, -0.045)	0.008	-0.107 (-0.249, 0.035) 0.118 (-0.250, 0.014)	0.1320.078
Smoking prevalence, 2016 total (ages 15+)	0.002 (-0.058, 0.062)	0.952	--	--
Hospital beds (per 10, 000 population)	-0.004 (-0.022, 0.014)	0.632	--	--
WHO health index, [2000]	2.259 (-0.920, 5.439)	0.173	-2.616 (-6.157, 0.925) -3.465 (-7.455, 0.526)	0.1400.086
Population living in urban areas (%)	0.023 (-0.011, 0.580)	0.193	0.010 (-0.019, 0.039) 0.004 (-0.029, 0.038)	0.4680.804
GDP per capita, PPP (\$1000 increase), [2018]	0.280 (0.037, 0.524)	0.030	0.154 (-0.174, 0.482) 0.303 (-0.051, 0.657)	0.3400.009

Country-level average daily ambient ultraviolet radiation (UVR) level - 2004	-0.000 (-0.001, 0.000)	0.133	-0.001 (-0.001, 0.000)	0.109
Mean % of BCG vaccination coverage among 1 year old children (1980-2019)	-0.027 (-0.037, -0.016)	<0.001	-0.018 (-0.034, -0.002)	0.031

Mean Daily Stringency Index	-0.036 (-0.072, -0.0000.001)	0.057	0.004 (-0.028, 0.037)0.000 (-0.035, -0.035)	0.790 0.99
-----------------------------	------------------------------	-------	---	-------------------

R2 for multivariable linear regression = 0.75658031

BMI – body mass index; WHO – world health organisation; GDP – gross domestic product; PPP – purchasing power parity; BCG – Bacille-Calmette-Guerin

Minor points

I might prefer to see the countries listed by continent rather than alphabetically.

Response: Thank you for your suggestion. Countries included in our analyses are now listed by continent in the methods section of our manuscript.

Under **METHODS**, page 7, lines 22-25 and page 8, lines 1-11:

“Countries reporting at least 25 daily deaths up to the 8th of June 2020 with available data for all chosen determinants were included. A total of 37 countries from 4 continents were included in the analysis: Africa (Algeria, Egypt, South Africa), America (Argentina, Brazil, Canada, Chile, Colombia, the Dominican Republic, Ecuador, Mexico, Peru and the United States of America), Asia (India, Indonesia, Japan, the Philippines, Saudi Arabia, Turkey) and Europe (Austria, Belgium, Finland, France, Germany, Hungary, Ireland, Italy, the Netherlands, Poland, Portugal, Romania, the Russian Federation, Spain, Sweden, Switzerland, Ukraine, the United Kingdom). ~~Algeria, Argentina, Austria, Belgium, Brazil, Canada, Chile, Colombia, the Dominican Republic, Ecuador, Egypt, Finland, France, Germany, Hungary, India, Indonesia, Ireland, Italy, Japan, Mexico, the Netherlands, Peru, the Philippines, Poland, Portugal, Romania, the Russian Federation, Saudi Arabia, South Africa, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.~~ China was not included in the analysis due to potential inaccuracies in the number of daily reported deaths which may have occurred subsequent to 1290 deaths which were retrospectively reported on the 17th of April.¹²”

Page 11 lines 24 to 32, it is still unclear what the measure of smoking is. Proportion who smoke? Average cigarettes per person?

Response: We apologise for this oversight. Smoking data was extracted as the age-standardised smoking rate across both sexes amongst adults (≥18 years) in 2013. This has now been more clearly explained in the methods section of our manuscript.

Under **METHODS**, page 10, lines 22-23:

“Data on daily cigarette smoking were extracted as the age-standardised smoking rate ~~on~~ across both sexes amongst adults (≥18 years) in 2013.³²”

Perhaps don't give an acronym to the stringency index.

Response: Thank you for your suggestion. We have now described the stringency index without using the acronym “SI” throughout the revised manuscript.

What is the difference between the list of predictors included in multivariable analysis found at the bottom of page 12 and those found on page 16 lines 12-20?

Response: Thank you for pointing out this issue. We apologise for any discrepancies that have appeared between the listed predictors, those should be the same and represent the predictors included in the multivariable models based on their significance level ($P < 0.3$) and univariable level. We have now corrected the multivariable predictor list from the methods section and have removed the predictor list from the results section to avoid any further confusion.

Under **METHODS**, page 12, lines 11-24:

“The following predictors were included in the univariable analyses: the natural logarithm of the population in 2018 (10 million increase) (~~natural logarithm~~), percentage of population aged 65 and over (~~median age~~), pollution levels, mean temperature (January-May), international arrivals, population density, prevalent diabetes, prevalent neoplasms, median BMI, prevalent hypertension, smoking prevalence, hospital beds (per 10,000 population), WHO health index, percentage population living in urban areas, GDP per capita (PPP), UV radiation exposure, mean BCG coverage and the stringency index ~~and the Stringency Index~~. Predictors reaching a P -value < 0.3 at univariable level were then included in a multivariable logistic regression model with the natural logarithm of the mean mortality rate as outcome: the logarithm of the total population in 2018, percentage of population aged 65 and over, pollution, mean temperature (January-May), international arrivals, population density, prevalent neoplasms, prevalent hypertension, the WHO health index, population living in urban areas, GDP per capita, UV radiation exposure, mean BCG coverage and the stringency index.”

Under **RESULTS**, page 17, lines 1-17:

“Table 2 details the results of the linear regression analyses. The following country-level predictors showed a statistically significant relationship with log mean mortality rate at univariable level: international arrivals in 2018 (coefficient (95% confidence interval) = 0.049 (0.033, 0.064), $P < 0.001$); prevalent neoplasms (0.614 (0.209, 1.019), $P = 0.005$) and prevalent hypertension (-0.150 (-0.254, -0.045), $P = 0.008$). natural logarithm of population, international arrivals, prevalent neoplasms, prevalent hypertension, GDP per capita and BCG vaccination coverage. ~~The multivariable model included the following predictors, which were selected from univariable models: median age, pollution levels, mean temperature, international arrivals, prevalent neoplasms, prevalent hypertension, WHO health index, percentage of population living in urban areas, GDP per capita and the stringency index.~~ Upon multivariable adjustment, International arrivals in 2018, as a marker of global connection, was the ~~only~~ main statistically significant predictor of log mean mortality rate (0.040 (0.017-, 0.063) for 1 million increase in international arrivals, $P = 0.002$) along with mean BCG vaccination coverage (-0.018 (-0.034, -0.002) for 1% increase in BCG vaccination coverage, $P = 0.031$). Figures 2 and 3 detail the relationship between the country-level log mean mortality rate (predicted and observed) and each country-level predictor included in the multivariable regression model.”

Page 21, “overfitting” is not used correctly here.

Response: Thank you for your comment. Overfitting is a phenomenon in which the regression model describes the random error in the data as opposed to the relationships between the variables of interest (Psychosom Med. May-Jun 2004;66(3):411-21.). Given the large ratio between the number of predictors and the number of countries included in our analyses and the very high R-squared value $> 80\%$ obtained from our regression models, we deemed that our models are at risk of overfitting. We

have now rephrased the description of this phenomenon in our limitations paragraph in order to avoid any confusion.

Under **DISCUSSION**, page 27, lines 9-13:

“Only including countries that had reported at least 25 deaths reduced our sample and consequently the power. ~~This may also result in the regression model overfitting the data.~~ Furthermore, the reasonably large number of country level predictors relative to the number of countries means that we cannot rule out the potential for overfitting in the multivariable model. This may lead to spurious associations between predictors and the outcome.”

Figures 2a and 2b, please give more information about the confidence bands displayed here.

Response: Thank you for pointing out this issue. We apologise for our oversight. We have now fully described all the elements of figures 2 and 3, including the confidence intervals.

Under **FIGURE LEGENDS**, page 37:

“**Figure 2.** Predicted (based on the results of the multivariable linear regression) and observed country-level mortality rate (mean daily increase in deaths until the peak in mortality) as a function of the recorded country-level number of international arrivals in 2018 (millions). The solid red line represents the point estimate of the predicted log daily increase in deaths, while the blue-grey area represents the corresponding 95% confidence interval. The crosses represent the observed values of the log daily increase in deaths.

Figure 3. Predicted (based on the results of the multivariable linear regression) and observed country-level mortality rate (mean daily increase in deaths until the peak in mortality) as a function of each country-level predictor included in the multivariable model. The solid red lines represent the point estimates of the predicted log daily increase in deaths, while the blue-grey areas represent the corresponding 95% confidence intervals. The crosses represent the observed values of the log daily increase in deaths.”

VERSION 2 – REVIEW

REVIEWER	Chandrashekar T Sreeramareddy International Medical University Kuala Lumpur Malaysia
REVIEW RETURNED	01-Jan-2021
GENERAL COMMENTS	Albeit the major limitations inherent in the design and data limitation; candidly admitted by the authors the manuscript has an important message for policy making. Some of the issues about analyses i.e. modelling should perhaps be explained in the analyses section by adding the same explanation provided by them here. Otherwise, i have no further comments or suggestions on this version of the manuscript.

REVIEWER	Amalia Magaret University of Washington, Seattle, WA, USA
REVIEW RETURNED	06-Jan-2021

GENERAL COMMENTS	<p>Review: Pana2020 BMJ Open Number of International Arrivals Is Associated with the severity of the first global wave of COVID-19</p> <p>Summary</p> <p>Summary of country-level risk factors for COVID deaths over the ramp up period while the virus began to spread. There is an impressive range of data sources collected to answer these questions.</p> <p>Main points</p> <p>The paper is much improved and the vast majority of my concerns have been addressed. I am disappointed not to see a time-varying analysis, and I think it would increase interpretability rather than reduce power, but I believe the work contributes substantial information as is.</p> <p>Minor points</p> <p>The below portion of the discussion should be removed. The <i>timing</i> of the implementation of stringency measures is not included in the analyses, and so a claim to a lack of association when all time points are averaged together is unfounded.</p> <p>“Furthermore, the mean stringency index, which also accounts for international travel restrictions amongst other measures, was not associated with the mean mortality rate in the multivariable model. This suggest that international travel restrictions and other containment measures may have been imposed too late to influence the steepness of the mortality curve and that the level of global connectivity of each country may influence the course of the epidemic mortality curve before the number of COVID-19 related cases and deaths reaches worrying levels.”</p>
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Dr. Chandrashekhara Sreeramareddy, Manipal College of Medical Sciences

Comments to the Author:

Dear Editor,

Albeit the major limitations inherent in the design and data limitation; candidly admitted by the authors the manuscript has an important message for policy making. Some of the issues about analyses i.e. modelling should perhaps be explained in the analyses section by adding the same explanation provided by them here. Otherwise, i have no further comments or suggestions on this version of the manuscript.

Response: Thank you for taking the time to consider our work. Also, thank you very much for your kind words and highlighting the fact that our work is of interest. Also, thank you for pointing out that further changes to our methods section in line with the responses provided by us in the previous response letter will improve the quality of our finally published paper. We have subsequently added further explanatory sections in the methods section and limitations paragraph of our manuscript detailing our choice of not including country-level ICU capacity data in our analyses.

Under METHODS, page 12, lines 1-5:

“Country-level intensive care unit (ICU) capacity was not included in the analyses, given the absence of a database centralising this information and the resulting poor reporting. Furthermore, ICU capacity data were unavailable for several important countries included in our analyses, such as Algeria, Argentina, Chile, the Dominican Republic, Ecuador, Egypt, India, Indonesia, Peru, the Philippines, Saudi Arabia and Ukraine.”

Under DISCUSSION, page 28, lines 5-8:

“Finally, we did not include ICU capacity data in our analyses due to a lack of a reliable data source centralising this variable. Nevertheless, our analyses account for country-level hospital beds per 10,000 population as an indicator of health systems’ coping capacity with increased pressures related to the pandemic.”

We have also added a further explanation regarding our choice of the predictor selection method:

Under METHODS, page 12, lines 23-24 and page 13, lines 1-2:

“Such a determinant selection process was chosen in order to lessen the likelihood of excluding factors which may be important but would not reach statistical significance due to the relatively small sample size of the study.”

Reviewer: 2

Prof. A Magaret, University of Washington

Comments to the Author:

Summary: Summary of country-level risk factors for COVID deaths over the ramp up period while the virus began to spread. There is an impressive range of data sources collected to answer these questions.

Main points: The paper is much improved and the vast majority of my concerns have been addressed. I am disappointed not to see a time-varying analysis, and I think it would increase interpretability rather than reduce power, but I believe the work contributes substantial information as is.

Response: Thank you for taking the time to consider our work and we are pleased that our previous revision has addressed your main concerns. With regards to employing a time-varying analyses, whilst we agree with the reviewer that such an approach may provide more granular information, we have decided to keep our original approach for a number of reasons. Firstly, it is important to note that most of the country-level determinants considered in our study (e.g. international arrivals) are not time-dependent as these data have been collected in years prior to the pandemic. Furthermore, our chosen outcome is calculated as the mean slope of the new daily deaths curve over the whole exposure period, which we feel best describes the impact of the pandemic in each county when calculated for longer periods of time, as this approach allows more countries to reach their peak levels of new daily deaths.

Minor points: The below portion of the discussion should be removed. The timing of the implementation of stringency measures is not included in the analyses, and so a claim to a lack of association when all time points are averaged together is unfounded. "Furthermore, the mean stringency index, which also accounts for international travel restrictions amongst other measures, was not associated with the mean mortality rate in the multivariable model. This suggest that international travel restrictions and other containment measures may have been imposed too late to influence the steepness of the mortality curve and that the level of global connectivity of each country may influence the course of the epidemic mortality curve before the number of COVID-19 related cases and deaths reaches worrying levels."

Response: Thank you for pointing out this issue. We agree with the reviewer that such relationships are best described by analyses which account for the stringency index as a time-dependent variable. As we have chosen not to undertake such an approach (also see response above). we have removed this paragraph from our discussion section.

Under METHODS, page 25, lines 10-16:

“Furthermore, the mean stringency index, which also accounts for international travel restrictions amongst other measures, was not associated with the mean mortality rate in the multivariable model. This suggests that international travel restrictions and other containment measures may have been imposed too late to influence the steepness of the mortality curve and that the level of global connectivity of each country may influence the course of the epidemic mortality curve before the number of COVID-19 related cases and deaths reaches worrying levels. ”