Factors Associated with Coronavirus (Covid-19) Deaths and Infections: A Cross-Country Evidence

Shafiun Nahin Shimul[1]
Muhammad Ihsan- Ul- Kabir[2]
Fariha Kadir[3]

ARTICLE INFO

Article history:
Date of Submission: 10-11-2024
Date of Acceptance: 07-03-2025
Date of Publication: 30-10-2025

Keywords:
COVID-19, deaths, infections, LASSO

ABSTRACT

While countries across the world were crippled by COVID-19 during 2020-2021, there had been substantial variations in death and infection rates. Some countries found themselves overwhelmed, while others remained relatively spared. The reasons behind such discrepancies remain largely unclear. This study aims to elucidate the determinants of death and infection rates associated with COVID-19 across various countries, utilizing multivariate regression analysis and the least absolute shrinkage and selection operator (LASSO) regression. The OLS estimates show that the ageing population and hospital beds per capita are significantly associated with the fatality rate of COVID-19, while urbanization has a positive correlation with the infection rate. The study suggests that enhancing the capacity of healthcare systems can play a crucial role in substantially reducing fatality rates associated with COVID-19.

[1] Professor, Institute of Health Economics, University of Dhaka, Bangladesh. Email: shafiun.ihe@du.ac.bd

[2] Assistant Professor, Institute of Health Economics, University of Dhaka, Bangladesh. Email: ihsan.ihe@du.ac.bd

[3] Assistant Professor, Institute of Health Economics, University of Dhaka, Bangladesh. Email: fariha.ihe@du.ac.bd (Corresponding Author)

DOI: https://doi.org/10.3329/ssr.v42i1.85319

Introduction

The outbreak of coronavirus disease 2019 (COVID-19), caused by the novel coronavirus, was first reported on 31st December 2019 in Wuhan, China. It has since been spread to 215 countries and regions globally. Till 20th October 2024, there were about 776.7 million confirmed cases (n= 776,753,259) and 7,073,249 deaths.[4] According to the World Health Organisation, the preliminary R_o estimate of COVID-19 is reported as 1.4-2.5 (WHO, 2020) indicating its larger infectiousness. However, while it has walloped many countries, many countries have been spared. In addition, there have been significant variations in deaths across countries.

Data from China, South Korea, Iran, and Italy suggest that the case fatality rate (CFR) increases drastically with age. It also indicates that the CFR is higher among the group with underlying comorbidities (The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020). Social distancing measures targeting these groups could be the most effective way to reduce morbidity and concurrent mortality. More severe COVID-19 symptoms have been linked with ageing and chronic conditions, such as obesity, diabetes, respiratory disease, kidney disease, and cardiovascular diseases (CVD), resulting in the development of acute respiratory distress syndrome (ARDS) and progression from ARDS to death (ChaominWu, et al., 2020).

The nature of jobs or occupations is likely to determine the exposure. A study finds that occupation may expose a person to risk by the nature of their job (Abdalla, Apramian, Cantley, & Cullen, 2017). The risk of infection spread through droplets is higher when a person is involved in work requiring constant human contacts, interaction with others or caring for people (Rule, et al., 2018). In the context of COVID-19, occupation likely has a direct impact on the risk of infection and an indirect influence on disease severity and mortality, given the link between occupational social class and underlying health conditions (Koh, 2020).

Low income and lower education levels are indirectly associated with factors that elevate the risk of severe COVID-19, such as higher smoking rates, poor nutrition, and substandard housing conditions, all of which can compromise the immune system. A recent systematic review of five retrospective and prospective studies confirmed a strong association between smoking and worse outcomes in COVID-19 cases (Vardavas & Nikitara, 2020).

However, to identify the factors most likely associated with COVID-19 outbreaks, high-quality data on socioeconomic conditions are urgently needed. Such data will have important implications for the development of health system measures (Khalatbari-Soltani, Cumming, & Delpierre, 2020).

---

4 https://covid19.who.int

Even though COVID-19 has spread across the globe, not all countries were equally affected. While the spread ballooned in some countries, some countries were relatively spared. Non-pharmaceutical interventions played a significant role. Even though from various sources, it was seen that the ageing population was more vulnerable, the health system's capacity could also dramatically reduce that vulnerability. While a lot of micro or patient-level data and studies are emerging and can better answer the underlying factors of case fatality rates and infection rates once COVID-19 reaches its ending phase, a cross-country comparison can elicit some demographic and health system covariates of the death and infection rates. A few such efforts, however, have been observed. In this study, our objective is to assess the influence of underlying factors using the most recent available data by employing diverse empirical estimation methods. This study attempts to examine the effect of all possible underlying factors behind COVID-19 deaths and infections with special consideration to the ageing population.

Related Literature

We have identified published studies through a rapid review and have observed that numerous social and economic requisites have been projected as potential determinants contributing to the empirical diversity in the outcomes of the coronavirus. Hardly any of the determinants, however, can robustly unravel the extent of the coronavirus pandemic. Stojkoski et al. (2020) find that population size and government health expenditure are the two determinants strongly associated with coronavirus cases (Stojkoski, Utkovski, Jolakoski, Tevdovski, & Kocarev, 2020). Their study reveals that more populous economies exhibit a higher resistance to virus infection, while countries with higher government expenditures demonstrate a greater susceptibility to the virus. Rodrigues, Prata, & Camargo (2020) investigate the regional differences in the occurrence of COVID-19 in Brazil and its relationship with climatic and demographic factors. Their results show that temperature variation maintains a relationship with the reduction in the number of COVID-19 cases. Depending on the contagion process in progress, the expected reduction in the number of cases of COVID-19 is -3.4% for each increase of 1°C (Rodrigues, Prata, & Camargo, 2020).

Qiu, Chen, and Shi (2020) explore the transmission dynamics of the novel coronavirus in 2019, focusing on both intra-city and inter-city spread within China. Using a machine learning technique called Lasso, their findings suggest that population movement from the outbreak source significantly increases the risk to other locations, more so than factors like geographic proximity or economic similarities. Additionally, the study provides evidence that public responses can interrupt the transmission chain. Variations in weather conditions also influence past infection rates, helping to identify the causal effect of previous infections on new cases (Qiu, Chen, & Shi, 2020).

A systematic review conducted by Khalatbari-Soltani, Cumming, & Delpierre (2020) shows that the impact of socioeconomic determinants on the COVID-19 outbreak fluctuates in high-income countries over low-and middle-income countries (LMICs). According to this study, the community spread of the disease COVID-19 appearing initially from abroad, will depend upon the local infrastructures and social inequalities in each context (Khalatbari-Soltani, Cumming, & Delpierre, 2020). For instance, the usual public health practices of social distancing measures including lockdowns, quarantine, and self-isolation are hardly performed in Low and Middle-Income Countries (LMICs). Moreover, in low-income settings where absolute poverty is a significant concern, access to necessities such as food, housing, water, and sanitation plays a crucial role in determining the feasibility of adhering to social distancing measures, especially when isolating older adults or other vulnerable individuals. In addition, in many parts of the world, health services are not readily available when needed, and healthcare systems, already strained even in usual times, are likely to become rapidly overwhelmed during pandemics.

Containing contact rates is a vital strategy for outbreak control which also hinges on population densities. Keeping more than one-meter distance between people coughing and sneezing, as recommended by the WHO becomes more intricate with higher population densities (WHO, 2020). Therefore, avoiding situations with high population densities will be essential to limiting the spread of COVID-19 (Rocklöv & Sjödin, 2020).

A study indicates that social distancing measures reduce the effective reproduction number by about 60%, if the intrinsic transmission potential declines in the warm summer months (Anderson, Heesterbeek, & Klinkenberg, 2020). As an added precaution, implementing broader-scale social distancing measures allows health systems the necessary time to manage cases effectively and enhance their capacity for the development of long-term treatments.

According to a study by Balmford, et al. (2020), governments all around the world have depended on preventative measures that attempt to limit the number of persons exposed to the virus and lower the average number of new cases per infection. The study also finds that age is one of the most firmly documented mortality risk factors. Even controlling for all other factors, older citizens had a higher death risk from COVID-19 infections. In addition, the amount of population concentration in major metropolitan centres may have an impact on COVID-19-related mortality. The study states that more than 75% of the cross-country variation in COVID-19 mortality differences remained unexplained even after considering the list of exogenous factors that have been hypothesized to be major drivers of cross-country variation in mortality rates (Balmford, Annan, Hargreaves, Altoè, & Bateman, 2020).

Similarly, another recent study by Bollyky, et al. (2022) discovered that factors such as age distribution, population density, smoking prevalence, and the number of hospital beds per capita affect the variation in COVID-19 infections among countries. The most important factors influencing the variation in the COVID-19 infection-fatality ratio were age structure, GDP per capita, and national mean BMI. However, standardized infection rates and pandemic preparedness indices were not significantly correlated with each other. (Bollyky, et al., 2022).

A key contribution of this paper is that it serves as a stepping stone toward a more comprehensive understanding of the relationship between COVID-19 infections and deaths and socio-economic or macroeconomic determinants.

Data and Estimation Method

Data

Two major data sources are used in this study: *Our World in Data COVID-19 Dataset*[5] by the University of Oxford, and the World Bank's *World Development Indicators* (WDI[6]).

In our empirical estimations, we use values of population ages 70 years and above (% of the total population) as aged population; population density; urban population (% of the total population); hospital bed per 1,000 population; immunization, measles (% of children ages 12-23 months); smoking prevalence, total (ages 15+). We used the mean values of these variables from 2008 to 2019. We used the population aged 70 years and above to be consistent with other similar literature. The use of same age group is evident in similar studies, For instance, Brandén, et al. (2020); Fristedt, Carlsson, Kylén, Jonsson, & Granbom (2021); Modig, Lambe, Ahlbom, & Ebeling (2021); Ningthoujam & Khomdram (2020); Tana, et al. (2023) use the similar age group to capture ageing population. While data were also available for the 60–70 age group, individuals aged 65 years and above are typically classified as the aged population. To avoid any confusion, we opted to use 70 years and above as the criterion for defining the aged populationValues of case fatality rate (CFR); total tests per thousand; total cases per million; cardiovascular diseases (CVD) death rate; and extreme poverty rate were collected from *Our World in Data COVID-19 Dataset* by the University of Oxford where extreme poverty rate was considered as a macroeconomic variable. Values for tropical country dummy (=1 for tropical countries) were collected from *World Population Review*[7]. Our cross-section data set consists of observations for each of the 186 countries of the world. The data is up to December 2022.

[5] https://ourworldindata.org/coronavirus-data

[6] https://databank.worldbank.org/source/world-development-indicators#

[7] https://worldpopulationreview.com/countries/tropical-countries/?fbclid=IwAR3lkkH3tseEu3WB7aI ZBYjwAXj_ERrpbGHzTTl8MtjiymK0lG8y4dDL6lQ

Estimation method

In our empirical estimations, we use a *t*-test, correlation matrix, and Least Absolute Shrinkage and Selection Operator (LASSO) to find out the correlation of the determinants of COVID-19. We employed the difference in means of seven-factor variables. Considering ageing population (70 years and above) of 10%, the variable is valued as 1 if a country has more than or equal to 10% of its population aged 70 years and above, and 0 otherwise. Similarly, we used the other six factor (or dummy) variables including 3.5 hospital beds per 1,000 population, 80% urban population, tropical country, CVD death rate of 300, 80% measles immunization, and smoking prevalence of 25%. We also reported pairwise correlation coefficients between the variables. The choice of a cut-point was the mean/median values, unless otherwise stipulated.

OLS & LASSO

While t-test and correlation test can provide some insights on the relationships between variables, they cannot separate the effects of confounding variables, and therefore, the conclusion made from these could be spurious. Regression analysis can solve this problem of omission of control variables. To understand the effects of demographic, health system and other correlates of death rates and infection, we used the following regression equation:

$$d_i = \beta_{1d} + \beta_2 Demographic_{id} + \beta_3 Health\_Sytem_{id} + \beta_4 Other\_Controls_{id} + \varepsilon\varepsilon_{id} \ldots (1)$$

$$I_i = \beta_{1I} + \beta_2 Demographic_{iI} + \beta_{3I} Health\_Sytem_{iI} + \beta_{4I} Country\_Characteristics_{\,iI} + \beta_5 Other\_Controls_{iI} + \varepsilon_{iI} \ldots (2)$$

Here, d is the death rates, I is the infection rate per million population, and on the right-hand side we have all relevant determinants. $Demographic_{id}$ is the set of demographic variables for country i, $Health\_Sytem_{id}$ is the set of health system variables in country i¸ $Country\_Characteristics_{\,iI}$ is the other characteristics of country i.

First, we applied Ordinary Least Square (OLS) regression to understand the determinants of both infection and death rates. However, right-hand co-variates are likely correlated. If the determining variables are correlated, we might not be able to precisely estimate the coefficients resulting in non-rejection of the null hypothesis. Another issue with COVID-19 is that many things are still unclear to researchers and so one common or typical modeling framework would be less reliable. Thankfully, a good alternative appears to exist. LASSO has been enormously popular among data scientists for prediction and model selection and it has been widely used in such a situation. We also applied LASSO, a machine learning tool, that accepts some level of bias to minimize the variance and select an appropriate model from the data. Though traditional LASSO measures are not meant for inference, rather they are better suited for predictions, with some modification, we can apply LASSO for inference. We applied a double-selection lasso linear regression model for inference. In this process, instead of running a single regression, we employed double-selection LASSO linear regression following (Belloni, 2016) & (StataCorp., 2019). The following process is followed:

Step 1: running LASSO of key variables on other controls.

Step 2: running LASSO of left-hand side variable on other controls.

Step 3: finding the union of the selected variables from both step 1 and step 2, and estimate the predicted value.

Step 3: regressing left-hand side variable on key variables and predicted value from step 3.

Step 5: making inference.

LASSO estimates coefficients, standard errors, and confidence intervals and performs tests for variables of interest while selecting from among potential control variables. Therefore, the linear regression model is as follows:

$E\,[y \mid d, x] = d\alpha' + x\beta'$ … *(3)*

Where d is the variables for which we wish to make inferences and x is the from which the lassos select. Potential control variables include case fatality rate, the proportion of population aged 70 or older as aged population, a dummy variable for tropical countries, population density, proportion of urban population, death rate from cardiovascular disease (CVD), hospital beds per 1,000 people, total tests per thousand, immunization rate, smoking prevalence, extreme poverty rate, obesity prevalence rate, etc.

Findings

As shown in Table 1, the average case fatality rate (CFR) is 0.024 with a maximum of 0.066. Average tests per thousand and average cases per million are 36.25 and 1,776, respectively, with a maximum of 184.86 and 8,414.

Table 1: Descriptive statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
Case fatality rate (CFR)	54	0.024	0.013	0.007	0.066
Total tests per thousand	54	36.25	44.304	.808	184.859
Total cases per million	54	1776.168	2118.736	4.797	8414.116

Variables	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Population ages 70 years and above (% of total population)	54	7.679	4.504	1.308	16.24
Tropical country dummy (=1 for tropical countries)	54	0.315	0.469	0	1
Population density	54	123.336	187.847	3.202	1265.036
Urban population (% of total population)	54	64.721	21.088	17.859	97.785
Cardiovascular diseases (CVD) death rate	54	221.793	113.518	85.755	539.849
Hospital bed per 1,000 population	54	3.2	2.373	0.3	9.033
Immunization, measles (% of children ages 12-23 months)	54	91.148	7.805	61.182	99
Smoking prevalence, total (ages 15+)	54	22.379	10.235	4.157	44.929
Extreme poverty rate	54	5.433	10.059	0.1	41.6
Obesity prevalence rate	54	20.424	8.533	3.4	31.3

About 7.6% of the population aged 70 years and above and 31.5% of countries are in the tropics. However, there is a large variation in the proportion of the ageing population ranging from 1.3 to 16.24. So far evidence suggests that the ageing population is more vulnerable to COVID-19 and therefore the variation in the ageing population across countries can significantly predict the death rates of the countries.

The average population density per square kilometer is 123, with a maximum of 1,265. Countries have an average of 65% urban population. Average cardiovascular disease (CVD) death rate was 222 per 100,000 population and there were 3 beds per 1,000 population in the countries. About 91% of children ages 12-23 months are immunized. Average rate for smoking prevalence is 22%. Besides, the average extreme poverty rate was 5.4% with a maximum of 41.6%. In addition, later tables report estimates of the correlation matrix and regression results. This study also conducted more relations among variables which are reported in Figure 1- Figure 7 of the appendix.

Table 2 shows the pairwise correlation matrix of the variables. Results show that the ageing population has a significant negative correlation and the CVD death rate has a significant negative correlation with case fatality rate (CFR). On the contrary, population density is positively correlated with urban population and immunization. Extreme poverty rate has a significant negative correlation with an ageing population and an urban population. However, all the significant correlations are at a significance level of 1%.

Table 2: Pairwise correlation matrix

	CFR	Aging population	Tropical country	Population density	Urban population	CVD death rate	Hospital bed per 1,000 population	Total cases per million	Total tests per thousand	Immunization	Smoking prevalence	Extreme poverty rate	Obesity prevalence
CFR	1.000												
Ageing population	0.386*	1.000											
Tropical country	-0.045	-0.409*	1.000										
Population density	-0.058	0.031	0.065	1.000									
Urban population	0.157	0.495*	-0.285*	0.229*	1.000								
CVD death rate	-0.219*	-0.340*	-0.111	-0.178	-0.433*	1.000							
Hospital bed per 1,000 population	0.116	0.692*	-0.430*	0.008	0.338*	-0.002	1.000						
Total cases per million	0.173	0.717*	-0.409*	0.158	0.529*	-0.401*	0.471*	1.000					
Total tests per thousand	0.005	0.314*	-0.443*	0.090	0.442*	-0.179	0.227	0.669*	1.000				
Immunization	0.040	0.386*	-0.282*	0.098	0.360*	-0.257*	0.327*	0.422*	0.304*	1.000			
Smoking prevalence	0.096	0.491*	-0.508*	-0.033	0.216	0.180	0.527*	0.468*	0.240	0.177	1.000		
Extreme poverty rate	-0.180	-0.558*	0.344*	-0.033	-0.619*	0.189	-0.447*	-0.533*	-0.350*	-0.473*	-0.385*	1.000	
Obesity prevalence	0.219*	0.476*	-0.413*	-0.106	0.728*	-0.181	0.340*	0.525*	0.403*	0.481*	0.347*	-0.682*	1.000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

From the correlational analysis, we find some expected results such as a positive correlation between ageing and death rate; a positive correlation between the proportion of urban population and infection; a negative correlation of tropical country status with both infections and death rates, and other interesting results have emerged. It is worth mentioning that even though regression is mostly used for understanding causation, regression cannot show causation with a clear identification strategy such as a randomized experiment. Otherwise, regression will show most of the correlation.

Table 3 presents the regression estimates. The first two columns show the OLS and LASSO estimates for case fatality rate and the third column shows OLS estimates for total cases per million.

Table 3: Regression estimates

	(1) CFR_OLS	(2) CFR_LASSO	(3) Total Cases Per Million
Population aged 70 years and above (% of total population)	0.561**	0.836***	-115.549
	(0.221)	(0.323)	(115.971)
Tropical country dummy	0.515	0.814	-352.492
	(1.448)	(1.525)	(760.233)
Population density	0.002	0.002	-0.101
	(0.003)	(0.002)	(1.513)
Urban Population (% of total population)	0.025	0.061	34.739
	(0.049)	(0.067)	(25.856)
CVD death Rate	-0.005	0.001	-7.589**
	(0.007)	(0.007)	(3.460)
Hospital bed per 1,000 population	-0.031	0.165	62.566
	(0.384)	(0.593)	(201.740)
Total tests per 1,000 population	-0.029*	-0.033**	15.544*
	(0.015)	(0.014)	(7.859)
Immunization, measles (% of children ages 12-23 months)	-0.005	0.006	-16.644
	(0.077)	(0.059)	(40.267)
Smoking prevalence, total (ages 15+)	-0.006	-0.034	26.572
	(0.071)	(0.067)	(37.520)
Extreme poverty rate	0.020	0.072	-6.491
	(0.075)	(0.077)	(39.276)
Obesity prevalence	0.031	0.014	-19.609
	(0.140)	(0.153)	(73.529)
Constant	-0.238		2816.312
	(7.394)		(3882.358)
Observations†	54	54	54

Adjusted R^2	0.206	0.310

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

From the OLS estimates we see that the ageing population has a significant effect on case fatality rate (p<0.05). CFR increases by 0.561 with the ageing population increasing by 1% of the total population. Total tests per and total test per 1,000 population has also significant effect on case fatality rate (P<0.10). It also reveals that one unit increase in tests per 1,000 population results in a 0.029 unit reduction in CFR. For the LASSO estimates, consistent results are detected for both the estimates. Coefficients are 0.836 for the ageing population (P<0.01) and -0.033 for tests per 1,000 population.

On another note, total cases per million is determined by total tests per thousand. It features that one unit increment in total tests per thousand results in an increase of 15.5 units in total cases per million (p<0.10). In particular, it symbolizes the more a country tests COVID-19, the more cases are found. At the same time, a 1% increase in urban population results in an increase of 34.8 cases per million. Again, OLS estimates report that extreme poverty and obesity result in increased case fatality rate which is consistent with LASSO estimates. Most other covariates are found to be less relevant (insignificant) in explaining infection and death rates.

Again, we conducted a similar analysis for the first six months of the COVID-19 pandemic as the impact of COVID-19 was vast at the beginning of the pandemic and limited strategies were in place for combatting the pandemic then. Moreover, the long-term dynamics of transmission and deaths could be strikingly different due to the existing knowledge base about the pandemic. The results are presented in Appendix Table A1. It was found from the analysis that the ageing population has a significant (p<0.01) positive effect and hospital beds per 1,000 population has a significant (p<0.01) negative effect on case fatality rate. Total cases per million was significantly (p<0.01) determined by the number of tests per thousand population. These findings are consistent with the findings from all three years' data. However, tropical country status and population density have a negative effect on case fatality rate during the early six months of the pandemic but a positive effect on case fatality rate when all three years' data is considered. Again, during the first six months, hospital beds per 10,000 population had a significant (p<0.01) negative effect on case fatality rate.

Discussion

Based on the cross-sectional analysis of COVID-19 cases, we observed that demographic characteristics and the country's health systems capacity have a strong correlation with COVID-19 deaths. On the other hand, very few predictors of infection was also observed. However, several key findings from this study merit further discussion.

First, the country with a larger number of ageing population had to face higher death rates. It is not surprising as various studies found that the ageing population is the most vulnerable, and so countries with a higher proportion of the ageing population have higher death rates.

Second, a proxy of health system variable—hospital beds per capita—appears to be highly important in COVID-19. It is not unexpected since more hospital beds mean the patients will not overwhelm the health systems, and thus treatment can be ensured for many, and therefore, people will not die just due to lack of basic treatment such as oxygen. In addition, more hospitals are also likely to come with more ICU beds which again can save more lives.

Third, whenever it comes to infection, the story seems to be still unknown; the puzzle still remains. However, we find urbanization has a strong correlation with infection. It is not shocking since urban areas have higher densities and so the individuals may have higher interactions and transmissions. Most of the worst-affected countries also had the worst-affected cities such as New York, London.

Some limitations of this study are worth noting. This study used mainly empirical exercise without any strong theoretical basis. Another limitation of the study is that infection dynamics is complex so all dimensions of these dynamics may not be explained by the variables considered here. Besides, a few other variables are not included as many of them are not available for a lot of countries used in our studies, e.g., air pollution, and race distribution. Moreover, the under-reporting of deaths and infections is prevalent in many countries. Excess mortality for death estimates and seroprevalence for infection could potentially be used in the studies. However, in this study, we only stick to the reported cases. Hence, the findings of the study should be interpreted with caution. Again, since the aggregate data were used, regional variations and countries with large geographical variations could not be captured. COVID-19 infection was reported by confirmed cases, but there may be variations across countries regarding reporting of cases.

Conclusion

Using cross-country data, the study attempted to understand the variation of case fatality rates and infection rates of COVID-19 across countries. The study finds that countries' health systems and demographic structures play a significant role in determining death rates. The countries with a higher ageing population will have higher death rates, and countries with a good health system will have lower death rates. In addition, countries with higher urbanization experience more severe outbreaks. Other variables such as weather, vaccination coverage, etc. do not have an association with either death rates or with the level of infections. The study suggests improving the system at a maximum capacity to reduce the death rates, while special attention to social distancing measures could be enforced for the entire country especially in the highly dense urban areas. Even though pandemic appears to be over now, we need more research on this issue to better prepare for the future pandemics.

Statements and Declarations

We, hereby, declare that we are the authors of this research work. We have not submitted this article elsewhere. The authors have no competing interests to declare that are relevant to the content of this article.

Ethical Approval

Ethical approval is not required as publicly available secondary data were used.

Acknowledgement

None

Funding

None

Authors' contribution

SNS conceptualized the study, and analysed the data and finalized the manuscript, MIUK analysed the data and was a major contributor in writing the draft. FK reviewed the related literatures and was a major contributor in writing the draft. All authors read and approved the final manuscript.

References

Abdalla, S., Apramian, S. S., Cantley, L. F., & Cullen, M. R. (2017). Occupation and Risk for Injuries. In C. N. Mock, R. Nugent, O. Kobusingye, & K. R. Smith, *Injury Prevention and Environmental Health* (p. 97). Washington (DC): The International Bank for Reconstruction and Development / The World Bank. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK525209

Afifi, A. A., May, S., & Clark, V. A. (2012). *Practical Multivariate Analysis* (5th ed. ed.). Boca Raton, FL: CRC Press.

Anderson, R., Heesterbeek, H., & Klinkenberg, D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet, 395*. doi:https://doi.org/10.1016/S0140-6736(20)30567-5

Balmford, B., Annan, J. D., Hargreaves, J. C., Altoè, M., & Bateman, I. J. (2020). Cross-Country Comparisons of Covid-19: Policy, Politics and the Price of Life. *Environmental and Resource Economics, 76*, 525–551. doi:https://doi.org/10.1007/s10640-020-00466-5

Belloni, A. V. (2016). Post-selection inference for generalized linear models with many controls. *Journal of Business & Economic Statistics*, 34: 606–619. doi:10.1080/07350015.2016.1166116

Bollyky, T., Hulland, E. N., Barber, R. M., Collins, J. K., Kiernan, S., Moses, M., . . . Castellano, R. (2022). Pandemic preparedness and COVID-19: an exploratory analysis of infection and fatality rates, and contextual factors associated with preparedness in 177 countries, from Jan 1, 2020, to Sept 30, 2021. *The Lancet, 399*(10334), 1489-1512. doi:https://doi.org/10.1016/S0140-6736(22)00172-6

Brandén, M., Aradhya, S., Kolk, M., Härkönen, J., Drefahl, S., Malmberg, B., . . . Gunnar Andersson, G. (2020). Residential context and COVID-19 mortality among adults aged 70 years and older in Stockholm: a population-based, observational study using individual-level data. *The Lancet Healthy Longevity, 1*(2), e80 - e88.

ChaominWu, M., Xiaoyan Chen, M., Yanping Cai, M., Jia'an Xia, M., Xing Zhou, M., Sha Xu, M., . . . Zeyong Yang, M. (2020). Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. Retrieved from https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2763184

Fristedt, S., Carlsson, G., Kylén, M., Jonsson, O., & Granbom, M. (2021). Changes in daily life and wellbeing in adults, 70 years and older, in the beginning of the COVID-19 pandemic. *Scandinavian Journal of Occupational Therapy, 29(6), 511–521., 29*(6), 511-521. doi:https://doi.org/10.1080/11038128.2021.1933171

Hamilton, L. C. (2013). *Statistics with Stata: Updated for Version 12.* (8th ed. ed.). Boston: Brooks/Cole.

Keeling, M. J., Hollingsworth, T. D., & Read, J. M. (2020, February 17). The efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *medRxiv.org - the preprint server for Health Sciences*. doi:https://doi.org/10.1101/2020.02.14.20023036

Khalatbari-Soltani, S., Cumming, R., & Delpierre, C. (2020). Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *Journal of Epidemiology and Community Health*, 1-4. doi:10.1136/jech-2020-214297

Koh, D. (2020). Occupational Risks for COVID-19 Infection. *Occup Med (Lond), 70*(1), 3-5. doi:10.1093/occmed/kqaa036

Modig, K., Lambe, M., Ahlbom, A., & Ebeling, M. (2021). Excess mortality for men and women above age 70 according to level of care during the first wave of COVID-19 pandemic in Sweden: A population-based study. *The Lancet Regional Health - Europe, 4*(May 2021), 100072. doi:https://doi.org/10.1016/j.lanepe.2021.100072

Ningthoujam, R., & Khomdram, D. (2020). WHO statement – "Older people are at highest risk from COVID-19": Should the hypothesis be corroborated or rejected? *Med Hypotheses, 144*, 109896. doi:10.1016/j.mehy.2020.109896

Qiu, Y., Chen, X., & Shi, W. (2020). Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China. doi:https://doi.org/10.1101/2020.03.13.20035238

Rocklöv, J., & Sjödin, H. (2020). High population densities catalyse the spread of COVID-19. *Journal of Travel Medicine*. doi:10.1093/jtm/taaa038

Rodrigues, W., Prata, D., & Camargo, W. (2020). Regional Determinants of the Expansion of COVID-19 in Brazil. doi:https://doi.org/10.1101/2020.04.13.20063925.

Rule, A., Apau, O., Ahrenholz, S., Brueck, S., Lindsley, W., de Perio, M., . . . Fisher, E. (2018). Healthcare personnel exposure in an emergency department during influenza season. *PLoS One 2018*. doi: https://doi.org/10.1371/journal.pone.0203223

Spearman, C. E. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 72-101.

StataCorp. (2019). *Stata LASSO Reference Manual: Release 18*. College Station, TX: StataCorp LLC. Retrieved from https://www.stata.com/manuals/lasso.pdf

Stojkoski, V., Utkovski, Z., Jolakoski, P., Tevdovski, D., & Kocarev, L. (2020). The socio-economic determinants of the coronavirus disease (COVID-19) pandemic. doi:10.1101/2020.04.15.20066068

Tana, C., Moffa, L., Falasca, K., Vecchiet, J., Tana, M., Mantini, C., . . . Giamberardino, M. A. (2023). Approach to COVID-19 in older adults and indications for improving the outcomes. *Ann Med, 55*(2), 2265298. doi:10.1080/07853890.2023.2265298

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus disease (COVID-19). *China CDC Weekly, 2*(x), pp. 113-22.

Uyanik, G., & Guler, N. (2013). A Study on Multiple Linear Regression Analysis. *Social and Behavioral Sciences 106*, 234-240. doi:doi: 10.1016/j.sbspro.2013.12.027

van Belle, G., Fisher, L. D., Heagerty, P. J., & Lumley, T. S. (2004). *Biostatistics: A Methodology for the Health* (2nd ed. ed.). New York: Wiley.

Vardavas, C., & Nikitara, K. (2020). COVID-19 and smoking: A systematic review of the evidence. *Tob Induc Dis., 18*(20). doi:10.18332/tid/119324

WHO. (2020, June 01). *Coronavirus disease (COVID-19) advice for the public.* Retrieved from www.who.int: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

WHO. (2020, January 23). Statement on the meeting ofStatement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV. Geneva.

Appendix

Table A1: Regression estimates for first six months of COVID-19 pandemic

	(1) CFR_OLS	(2) CFR_LASSO	(3) Total Cases Per Million
Population ages 70 years and above (% of total population)	0.586***	0.705***	-49.147
	(0.124)	(0.184)	(96.336)
Tropical country dummy	-0.041	0.303	-231.955
	(1.026)	(0.829)	(680.691)
Population density	-0.000	-0.001*	-0.406
	(0.000)	(0.000)	(1.269)
Urban Population (% of total population)	0.027	0.035	34.324*
	(0.023)	(0.028)	(18.187)
CVD death Rate	-0.002	0.001	-5.655*
	(0.004)	(0.003)	(2.897)
Hospital bed per 1,000 population	-0.435**	-0.500***	-146.355
	(0.191)	(0.185)	(135.223)
Total tests per thousand	-0.016*	-0.018*	19.264***
	(0.009)	(0.010)	(6.979)
Immunization, measles (% of children ages 12-23 months)	-0.025	-0.017	-26.514
	(0.060)	(0.041)	(36.325)
Smoking prevalence, total (ages 15+)			28.026
			(35.725)
Extreme poverty rate			-0.855
			(36.072)
Constant	2.644		2906.861
	(5.538)		(3673.821)
Observations†	78	75	60
Adjusted R^2	0.320		0.359

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

† Number of observations varies from descriptive statistics based on the availability of data for observations

Figure 1: Correlation among variables

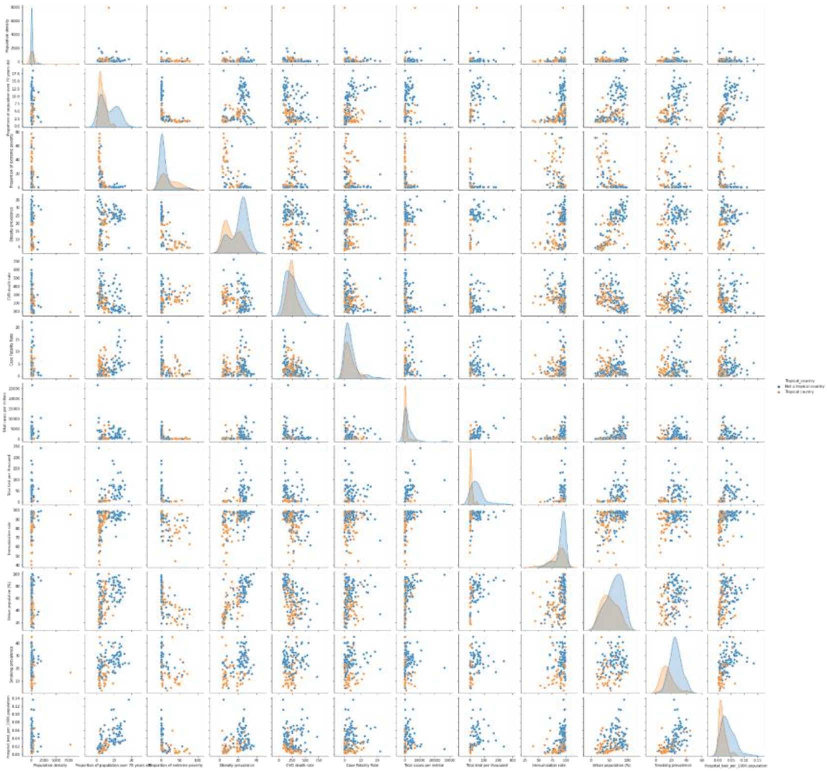

Figure 2: Ageing and case fatality by hospital beds

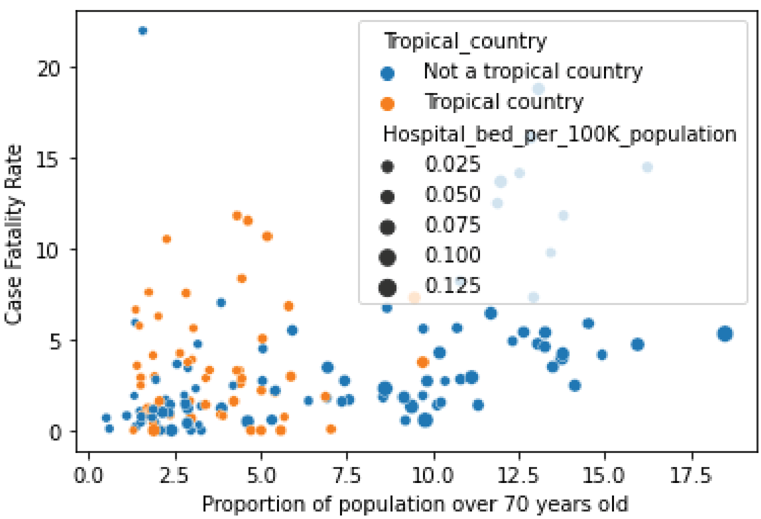

Figure 3: Urban population, case fatality by ageing population

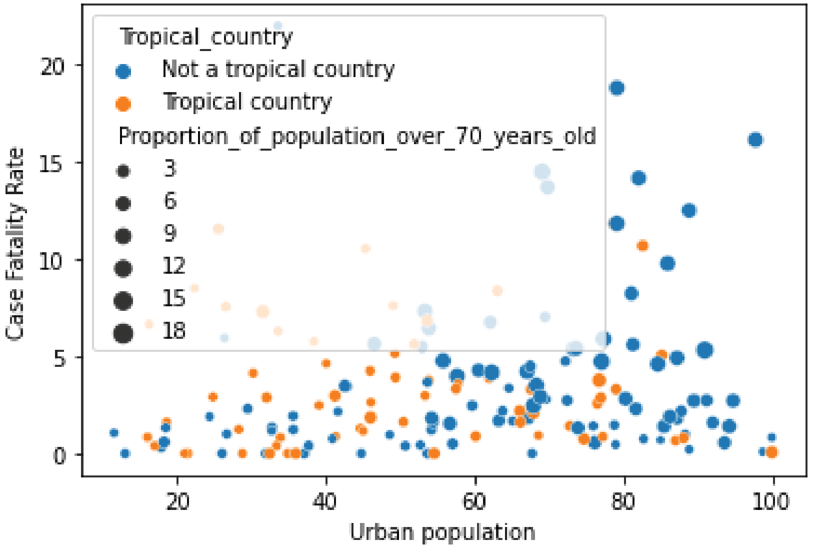

Figure 4: Ageing and infection rate by tests (initial stages)

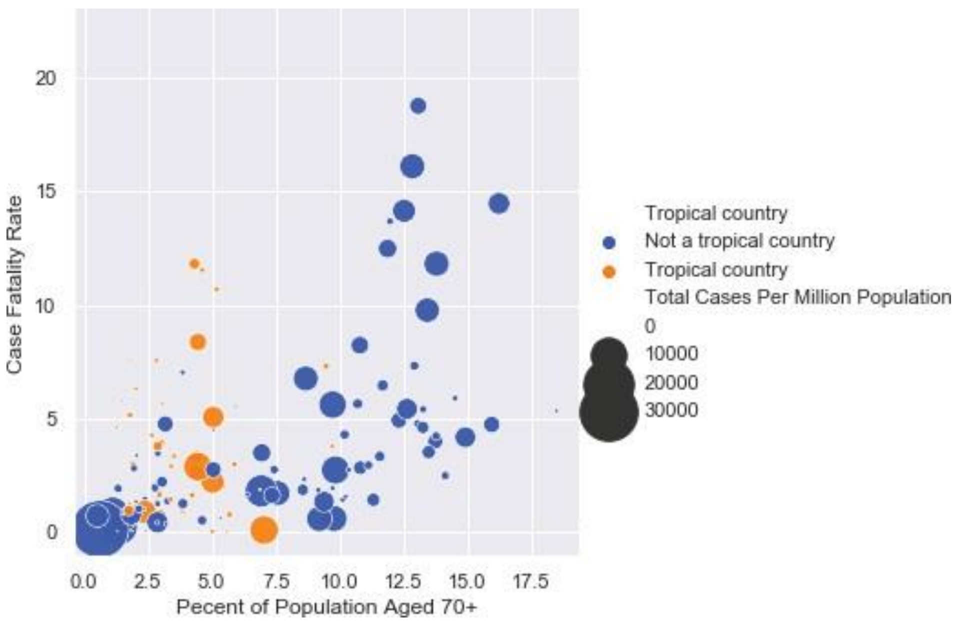

Figure 5: Case fatality rate by tropical status of the country

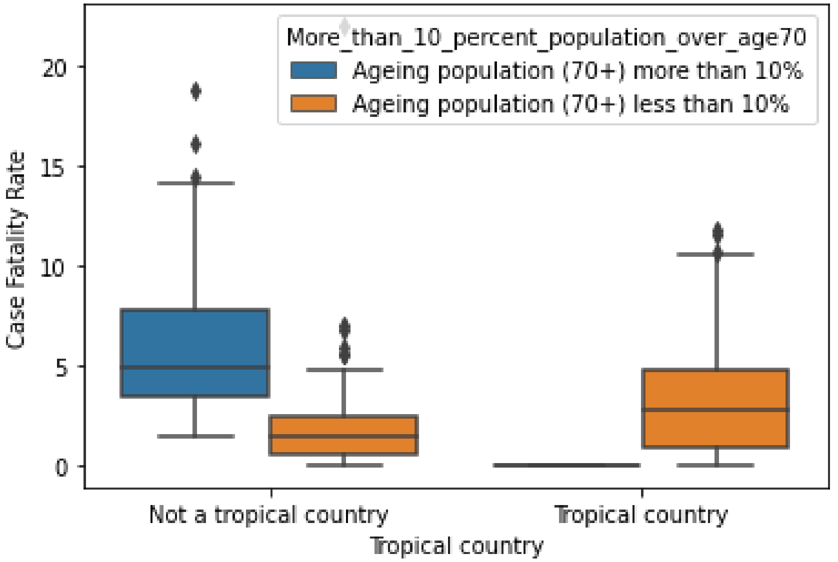

Figure 6: Case fatality rate by age (initial stages)

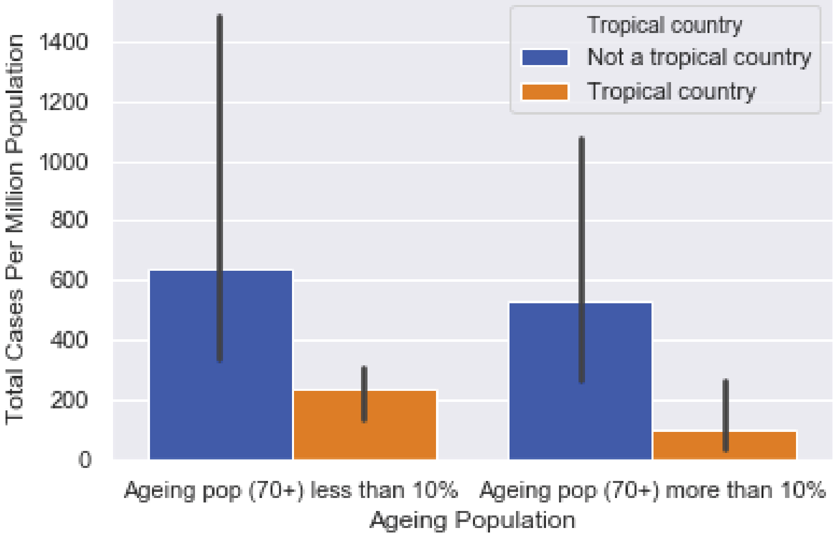

Figure 7: Total infection by country type and ageing population