

# Early trends of socio-economic and health indicators influencing case fatality rate of COVID-19 pandemic

Shahir Asfahan,<sup>1</sup> Aneesa Shahul,<sup>1</sup> Gopal Chawla,<sup>1</sup> Naveen Dutt,<sup>1</sup> Ram Niwas,<sup>1</sup> Neeraj Gupta<sup>2</sup>

<sup>1</sup>Department of Pulmonary Medicine; <sup>2</sup>Department of Paediatrics, All India Institute of Medical Sciences, Jodhpur, India

## Abstract

Coronavirus disease 2019, *i.e.* COVID-19, started as an outbreak in a district of China and has engulfed the world in a matter of 3 months. It is posing a serious health and economic challenge worldwide. However, case fatality rates (CFRs) have varied amongst various countries ranging from 0 to 8.91%. We have evaluated the effect of selected socio-economic and health indicators to explain this variation in CFR. Countries reporting a minimum of 50 cases as on 14<sup>th</sup> March 2020, were selected for this analysis. Data about the socio-economic indicators of each country was accessed from the World bank database and data about the

health indicators were accessed from the World Health Organisation (WHO) database. Various socioeconomic indicators and health indicators were selected for this analysis. After selecting from univariate analysis, the indicators with the maximum correlation were used to build a model using multiple variable linear regression with a forward selection of variables and using adjusted R-squared score as the metric. We found univariate regression results were significant for GDP (gross domestic product) *per capita*, POD 30/70 (probability of dying between age 30 and exact age 70 from any of cardiovascular disease, cancer, diabetes or chronic respiratory disease), HCI (human capital index), GNI (gross national income) *per capita*, life expectancy, medical doctors per 10000 population, as these parameters negatively correlated with CFR ( $\rho = -0.48$  to  $-0.38$ ,  $p < 0.05$ ). Case fatality rate was regressed using ordinary least squares (OLS) against the socio-economic and health indicators. The indicators in the final model were GDP *per capita*, POD 30/70, HCI, life expectancy, medical doctors per 10,000, median age, current health expenditure *per capita*, number of confirmed cases and population in millions. The adjusted R-squared score was 0.306. Developing countries with a poor economy are especially vulnerable in terms of COVID-19 mortality and underscore the need to have a global policy to deal with this on-going pandemic. These trends largely confirm that the toll from COVID-19 will be worse in countries ill-equipped to deal with it. These analyses of epidemiological data are need of time as apart from increasing situational awareness, it guides us in taking informed interventions and helps policy-making to tackle this pandemic.

Correspondence: Dr. Gopal Chawla, Department of Pulmonary Medicine, All India Institute of Medical Sciences, Jodhpur, India. Tel. +91.9999883667. E-mail: dr.gopalchawla@gmail.com

Key words: COVID-19; pandemic; case fatality rate.

Funding: None.

Conflict of interest: The authors declare no conflict of interest.

Contributions: SA, AS, GC, ND, study concepts; SA, GC, ND, RN, study design; SA, GC, ND, definition of intellectual content; SA, GC, literature search; SA, GC, clinical studies; SA, GC, experimental studies; SA, GC, ND, data acquisition; SA, GC, data analysis; SA, GC, NG, statistical analysis; SA, GC, ND, manuscript preparation; SA, AS, GC, ND, manuscript editing; SA, AS, GC, RN, NG, manuscript review; SA, GC, Guarantors. All the authors have read and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Availability of data and materials: All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate: Not applicable.

Received for publication: 18 May 2020.

Accepted for publication: 15 July 2020.

©Copyright: the Author(s), 2020

Licensee PAGEPress, Italy

Monaldi Archives for Chest Disease 2020; 90:1388

doi: 10.4081/monaldi.2020.1388

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Introduction

Corona virus (CoV) comprises of a large family of viruses that are common in human beings as well as animals. Other viruses of this family are known to cause severe illnesses like MERS (Middle East respiratory syndrome) and SARS (severe acute respiratory syndrome). SARS-CoV-2 is a novel virus which was first identified in December 2019 as a cause of upper and lower respiratory tract infection in Wuhan, a city in the Hubei Province of China. This infection was labelled as Coronavirus disease 2019, *i.e.* COVID-19. It rapidly spread through person-to-person transmission via droplets and fomites, resulting in an epidemic throughout China, and then gradually spreading to other parts of the world [1]. A pandemic by definition involves a wide spectrum of societies and nations [2]. Historically, the impact of pandemics have varied between countries. Socio-economic and health indicators of a country may reflect the levels of preparedness and ability to handle a pandemic. The current pandemic of COVID-19 is not different. It continues to exact a heavy toll in terms of mortality

and different countries have adopted different strategies to tackle this. Although the disease is of recent origin, the case fatality rates (CFR) are different for different countries with some countries bearing the brunt of it while others have tried to limit the impact for now [3]. Previous experience with the H1N1 pandemic has shown that levels of socio-economic and health indicators do influence case fatality rates. In previous H1N1 pandemic an economically stronger region like Europe performed better with a pooled hospital fatality rate while in some developing countries it was as high as 52% [4]. This study aims to evaluate the various factors that are involved in the determination of CFR associated with COVID-19. We have selected a wide spectrum of socio-economic and health indicators of countries dealing with COVID-19 and attempted to explain some of the variability in CFR so that this data might be useful in planning strategies in countries yet to bear the full impact of COVID-19.

## Materials and Methods

Data on the total numbers of confirmed cases and the number of deaths as of 14<sup>th</sup> March 2020 were accessed from the WHO database. CFR of all countries reporting at least 50 confirmed cases of COVID-19 were considered for analysis (Figure 1) [5]. Although the cut-off date of 14<sup>th</sup> March, 2020 appears quite early in this pandemic, however this study was envisaged to understand the initial dynamics of the spread and impact of pandemic.

Case fatality rate has been defined as the number of deaths

reported in patients confirmed with COVID-19 divided by the number of confirmed cases reported. Socio-economic and health indicators were considered to prepare an explanatory model to chart the variation in CFR of different countries so far (Table 1).

The following socioeconomic indicators were taken from the World Bank database [6]:

- (i) Gross domestic product (GDP)
- (ii) Population
- (iii) Life expectancy
- (iv) Gross national income (GNI) *Per Capita*
- (v) GDP *per capita*
- (vi) Human capital index (HCI)

GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. GDP is an indicator of the overall level of economic development and influences the level of healthcare accessible to its citizens. Population of the country was included because it can impact natural resources and social infrastructure. This can place pressure on a country's sustainability and imposes a burden on the healthcare systems. Life expectancy at birth indicates the number of years a new-born infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life. It is an important indicator of the health status of the country. GNI *per capita* is the gross national income, converted to U.S. dollars using the World Bank Atlas method, divided by the midyear population. GNI in comparison to the GDP may better reflect the economic status of countries with large foreign remittances/external aid. GDP per capita is a measure of a country's economic output divided by the population. It is a good measure of a country's standard of living which in turn affects access to affordable healthcare [6]. Human capital index measures the amount of human capital that a child born today can expect to attain by age 18, given the risks of poor health and poor education that prevail in the country where she lives. It is designed to highlight the impact of improvements in current health and education which shapes the productivity of the next generation of workers.

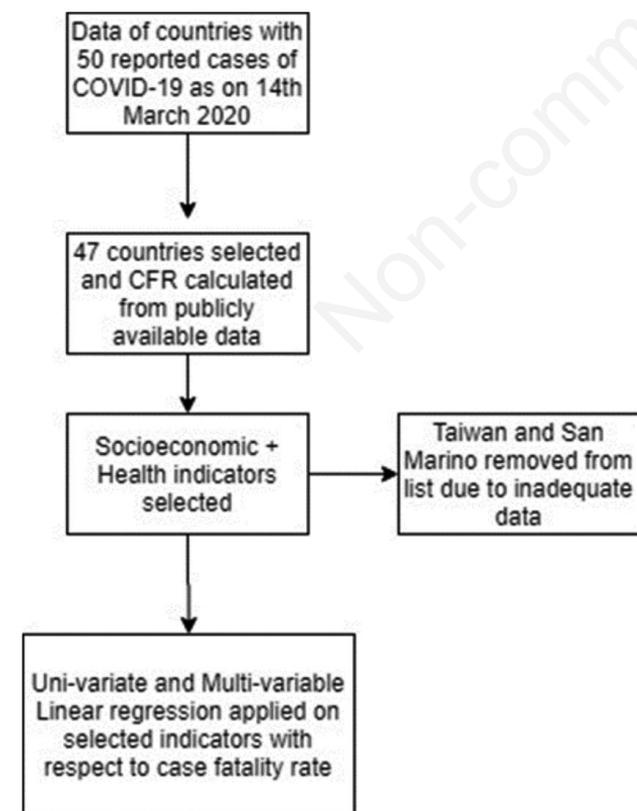


Figure 1. Flowchart of the study.

Table 1. Mean and standard deviations of the indicators selected for analysis.

Variables	Mean $\pm$ SD (n=45)
GDP in billions (US\$)	1629.96 $\pm$ 3614.9
Population in millions	99.08 $\pm$ 283.1
Life expectancy	79.06 $\pm$ 4
GNI <i>per capita</i> (US\$)	33561.33 $\pm$ 22588.75
GDP <i>per capita</i> (US\$)	38532.31 $\pm$ 21334.42
HCI	0.7 $\pm$ 0.12
Median age (years)	36.23 $\pm$ 6.66
CHE <i>per capita</i> (US\$)	2901.33 $\pm$ 2590.95
CHE as % GDP	7.72 $\pm$ 3.04
Prevalence (%) of tobacco smoking	23.41 $\pm$ 7.23
POD 30/70	14.05 $\pm$ 5.21
Hospital bed/10,000	38.62 $\pm$ 26.67
Medical doctors/10,000	28.18 $\pm$ 12.91
Confirmed cases	3185.84 $\pm$ 12307.92
Confirmed cases per million population	70.51 $\pm$ 118.13
Population density/km <sup>2</sup>	389.39 $\pm$ 1248.28
Case fatality rate	1.55 $\pm$ 2.25

The following health indicators were incorporated into the model and accessed from the WHO database [7]:

- (i) Median age
- (ii) Current health expenditure (CHE) *per capita*
- (iii) Current health expenditure (CHE) as a percentage of GDP
- (iv) Prevalence (%) of tobacco smoking
- (v) Probability (%) of dying between age 30 and exact age 70 from any of cardiovascular disease, cancer, diabetes or chronic respiratory disease (POD 30/70)
- (vi) Hospital beds per 10000 population
- (vii) Medical doctors per 10000 population
- (viii) Confirmed cases of COVID-19
- (ix) Confirmed cases per million population
- (x) Population density/Km<sup>2</sup>

Median age is a determinant of age distribution in the society. Age distribution could be a factor in the case fatality rates as COVID-19 is preferentially more severe in the older age group. Current Health Expenditure (CHE) *per capita* is the health budget divided by the population. CHE as a percentage of GDP is the percentage of GDP spent on health resources for the country. CHE per capita and CHE as a percentage of GDP are a measure of spending on public healthcare and this may influence the response of a country to a pandemic. The prevalence of tobacco smoking mirrors the respiratory ailments in a country which might predispose to respiratory complications of COVID-19. POD 30/70 may reflect the background health co-morbidities in a country which might get complicated by COVID-19 and may impact mortality as it has been shown that case fatality rises in patients with co-morbidities. Hospital beds and doctors per 10000 population are indicative of the robustness of the healthcare system in the country. We evaluated confirmed cases per million population and population density to assess whether these determinants have a role in determining CFR [7].

## Statistical analyses

Statistical analysis was performed using Python 3.6. Continuous variables are presented as mean  $\pm$  SD. Univariate correlation of each feature was done against the case fatality rate using

the Pearson correlation (Table 2). Factors with a p-value  $<0.2$  were considered to build a model. Features with the lowest p-values were selected in ascending order as part of the forward selection and multiple variable linear regression was applied against case fatality rate ordinary least squares (OLS). Forward selection of features was done based on the adjusted R-squared score.

Heat map of all the features was constructed to assess the correlations between different features (Figure 2). All features were standardized to unit variance based on the following formula. The standard score of sample  $x$  is calculated as  $Z = (X - U) / S$  where  $U$  is the mean of the training samples, and  $S$  is the standard deviation of the training samples. Standardization was done to maintain uniformity.

## Results

A total of 47 countries were selected for analysis based on the inclusion criteria of reporting at least 50 cases of confirmed COVID-19 (Figure 1). Taiwan and San Marino were removed from analysis due to the paucity of data from the above-mentioned sources.

A total of 143,363 were confirmed and 5379 deaths were reported in total from our dataset with a mean CFR of 3.75%. CFR exhibited a wide range (0 to 8.91%) with a standard deviation of 2.25% (Figure 3). In the univariate analysis, the parameters correlating positively with CFR were POD 30/70, number of confirmed cases, population in millions, GDP, prevalence of tobacco smoking, and confirmed cases per million population. Parameters associated with a negative correlation in univariate analysis were GDP per capita, HCI, GNI per capita, life expectancy, medical doctors per 10000 population, median age, CHE per capita, hospital beds per 10000, CHE as a percentage of GDP and population density/km<sup>2</sup>. Some of these correlations may appear counter-intuitive at first, however, we believe these changes reflect the differential impact of the pandemic on the developed and developing countries, particularly in its initial course.

In the multiple variable linear regression, a total of nine indicators were selected based on the forward selection which gave the

**Table 2. Univariate regression with Pearson correlation coefficient between indicators and case fatality rates.**

	Pearson Correlation Coefficient	p
GDP per capita(US\$)	-0.482	0.001
POD 30/70	0.459	0.002
HCI	-0.454	0.002
GNI per capita(US\$)	-0.439	0.003
Life expectancy	-0.438	0.003
Medical doctors/10,000	-0.435	0.003
Median age	-0.389	0.008
CHE per capita(US\$)	-0.291	0.053
Confirmed cases	0.259	0.086
Population in millions	0.24	0.113
Hospital bed/10,000	-0.188	0.217
CHE as % GDP	-0.167	0.274
GDP in billions (US\$)	0.159	0.295
Prevalence(%) of tobacco smoking	0.123	0.42
Population density/km <sup>2</sup>	-0.105	0.494
Confirmed cases per million population	0.035	0.817

best adjusted R-squared score of 0.306 (Table 3). The F-statistic (3.152) for the model was significant ( $p=0.00691$ ). Individually in the model, GDP per capita and the number of confirmed cases had a  $p<0.05$  (Table 3).

## Discussion

We conducted a study to assess the impact of various social and health indicators on case fatality rate of COVID-19. In 45 countries included for the study according to inclusion and exclusion criteria, we observed that various indicators had a significant effect on case fatality. These factors may be one of the reasons for varying CFR in various countries. We propose the model of selected indicators that may help in identifying the countries where CFR in the coming days of an evolving pandemic can be very high.

COVID-19 is a disease that presents in a variety of ways ranging from being asymptomatic to being severely ill. While the severity of illness contributes to the fatality of this new disease, the asymptomatic and mild cases play a major role in transmission [1].

Resources that influence pandemic response can be categorized into financial, human and physical resources. Apart from these resources leadership, intergovernmental relationships, onsite response and information sharing have also been identified as factors that influence response [8]. Here in our study we have tried to study financial and human response to the current pandemic. In the previous studies on the H1N1 pandemic, evaluating the effects of economic indicators on CFR and the effect of CFR on economic indicators were influential in tailoring the responses to future pandemics. Some of the measures like school closures and quarantines have been re-instated in the present pandemic as learned from that experience [9]. As in the previous study that explored determinants of H1N1 hospitalization, populations at the margins of society with poor social security were at higher risk even within an ecosystem of an individual first world country [10].

Univariate regression results were significant for GDP *per capita*, POD 30/70, HCI, GNI *per capita*, life expectancy, medical doctors per 10000 population and median age. Predictably all the above-mentioned factors were negatively correlated with CFR except POD 30/70. However, GNI *per capita* was removed from the analysis as it correlated strongly with GDP *per capita*. These

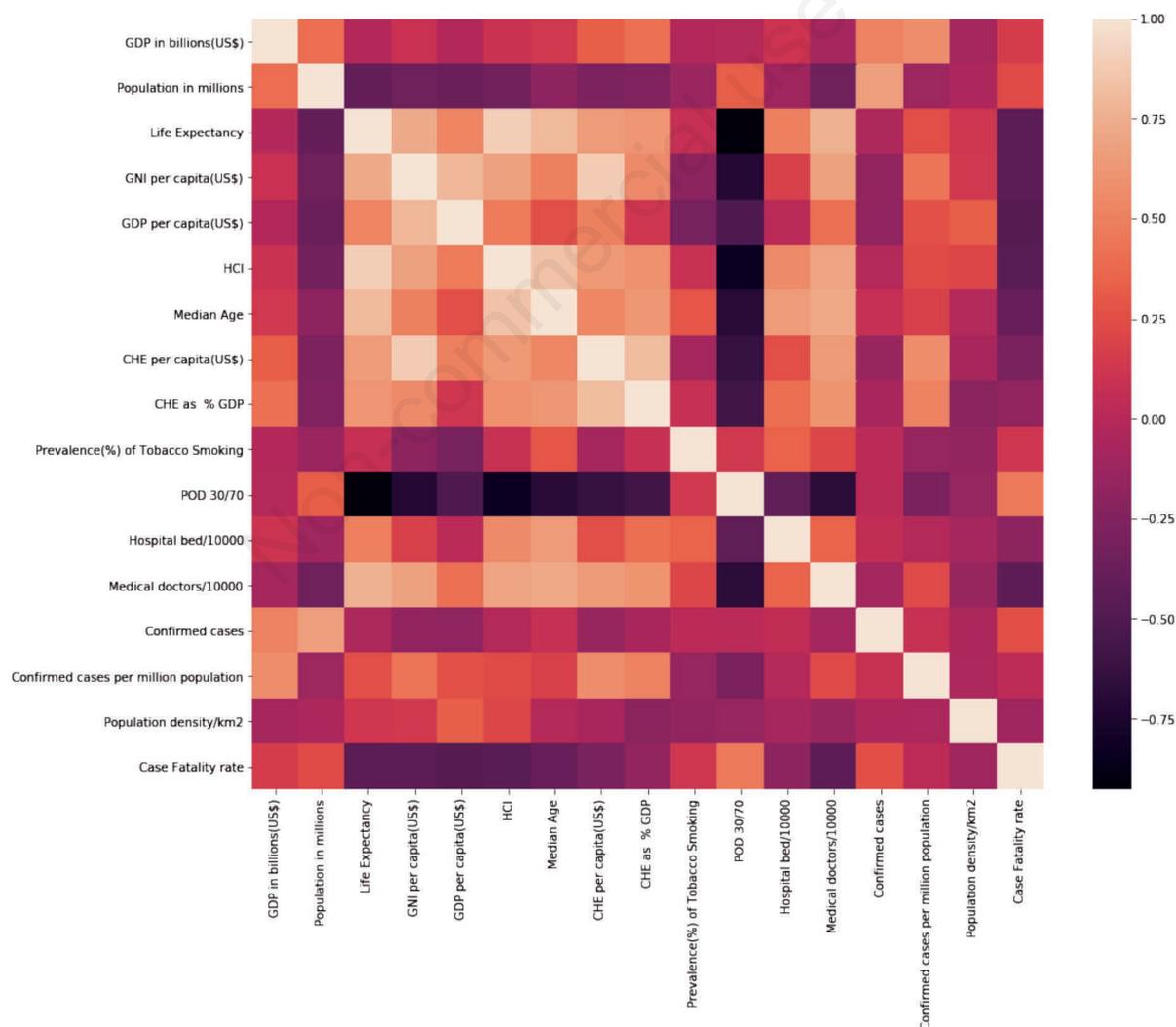


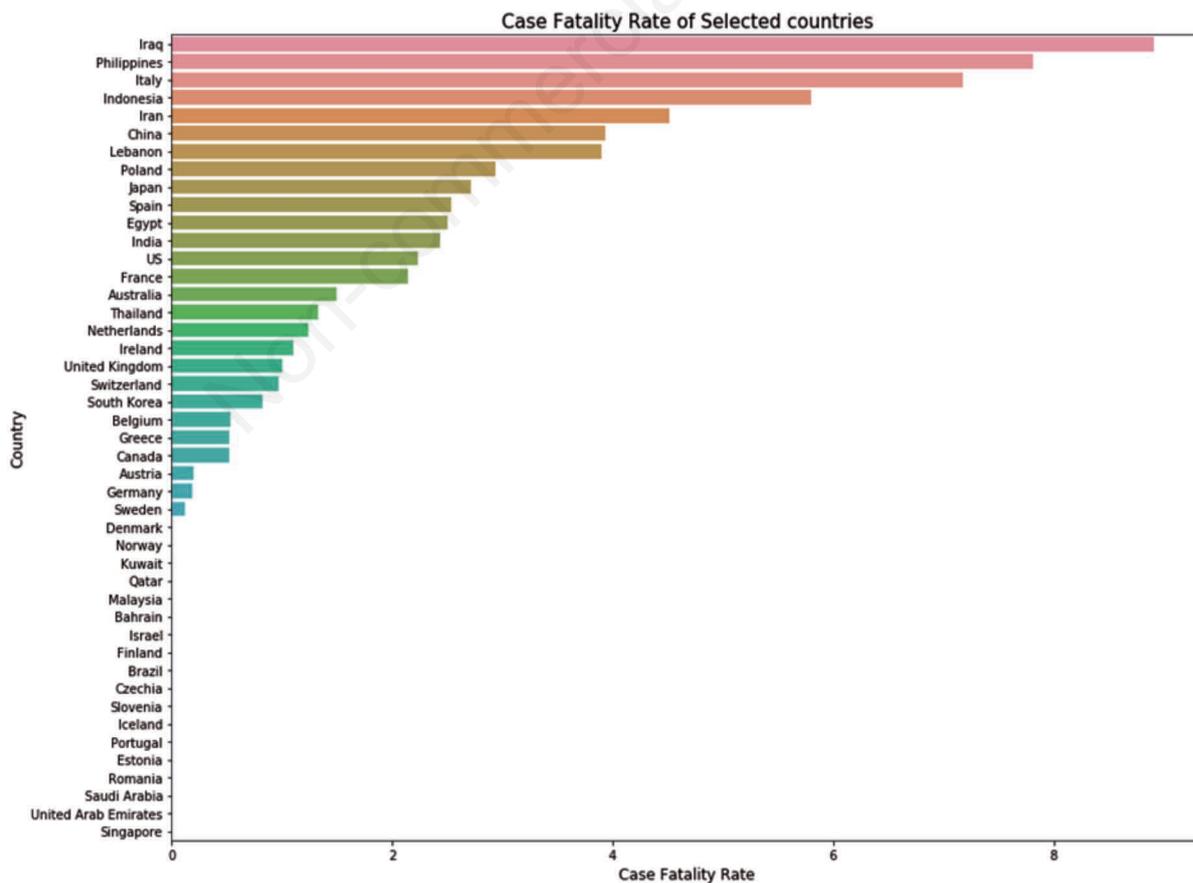
Figure 2. Heat map of the correlation between all the features under study.

factors are strongly related to the overall development status of a country and reflect the strength of the medical systems in place. In multivariate analysis, in addition to the significant univariate factors, CHE *per capita*, the number of confirmed cases and popula-

tion of the country helped improve the explanatory model for CFR. The overall model was significant as determined by the F-statistic, however, in the adjusted model, GDP *per capita* and numbers of confirmed cases were significant variables underlying the impor-

**Table 3. Multiple variable linear regression results of our model OLS (ordinary least squares) regression results.**

Dependent variable	Case fatality rate	R squared	0.448		
No. of observations	45	Adj R squared	0.306		
Df residuals	35	F- statistic	3.152		
Df model	9	Prob (F statistic)	0.00691		
	Coefficient	Standard error	p	[0.025	0.975]
	0	0.126	1.000	-0.255	0.255
GDP <i>per capita</i> (US\$)	-0.4010	0.164	0.020	-0.735	-0.067
POD 30/70	0.4755	0.357	0.191	-0.248	1.199
HCI	-0.2728	0.331	0.415	-0.945	0.399
Life expectancy	0.5541	0.505	0.280	-0.471	1.579
Medical doctors per 10,000	-0.2548	0.222	0.260	-0.706	0.197
Median age	-0.2165	0.283	0.449	-0.790	0.357
CHE <i>per capita</i> (US\$)	0.2922	0.186	0.125	-0.086	0.670
Confirmed cases	0.3871	0.183	0.042	0.015	0.759
Population in millions	-0.2318	0.206	0.268	-0.650	0.187



**Figure 3. Case fatality rates by countries reporting at least 50 confirmed cases of COVID-19.**

tance of these factors in the overall explanation of CFR in these countries. Our model of selected socio-economic and health indicators could explain 30.6% of the variability in CFR.

GDP *per capita* is an overall metric of a nation and it can be extrapolated to the kind of health care systems which exists in the country. It comes as no surprise that increasing GDP per capita is associated with decreasing mortality. The total number of confirmed cases was directly related to an increase in CFR. It may be possible that a higher number of confirmed cases may result in overwhelming healthcare systems leading to higher mortality. POD 30/70 reflects the background co-morbidities in the population and an increase in this metric was associated with an increase in the CFR. This implies that increased protection and care need to be put in place for people with significant co-morbidities. An increase in HCI was associated with lower CFR indicating that long term development correlates with better management of pandemics. Increasing life expectancy was associated with higher CFR which may reflect the higher burden of the older population in the country. An increase in medical doctors per 10,000 population led to a decrease in CFR which is understandable as it is a metric of better health care systems. This measure can be improved to tackle future pandemics and in the short term, doctors could be deputed to underserved areas to tackle this disease.

Counter-intuitively, we can see that increasing median age led to lower CFR and higher CHE *per capita* led to higher CFR. It could be possible that lower median age is a characteristic of developing countries which due to the nature of their relatively under-developed healthcare systems could have led to higher CFR seen so far [11]. As for higher CHE per capita leading to higher CFR, we believe that higher CHE *per capita* is a characteristic of a developed country which implies a better testing for COVID-19, better reporting of data and lesser number of missed deaths attributable to other causes when in fact it could have been caused by COVID-19 [4].

Although countries with a fledgling economy like India have been relatively unaffected so far with fewer cases compared to others in the region and beyond, however, going by past experience of handling H1N1 pandemic, it portends an ominous outcome if unchecked and un-helped. Glaring health in-equalities in a developing country may allow for unchecked local transmission with potentially disastrous results [12]. Digging early into indicators will help look into other parts of response management. Because model which can be useful in China or United states may not be feasible in countries like India or Pakistan. Since economic and health care response could prove to be deficient in these countries we may need to look at other parts of response management which are leadership, intergovernmental relationships and innovative measures.

As we go forward in this yet evolving pandemic there are some short-term and long-term measures that can be taken. Short term measures are those where intervention is anticipated for 12 weeks. Countries with low GDP *per capita* like those in sub-Saharan Africa, Southeast Asia, and Latin America may need help with resources to augment their healthcare systems to tide over this crisis. Measures like improving hygiene, encouraging handwashing, better droplet etiquettes, social distancing, closing of schools postponing public gatherings and partial to complete lockdown including public transport ban can prove helpful in halting the spread of infection. All effective measures need to be put in place to limit the total number of confirmed cases so that the health systems can respond better and do not impact the mortality of the other background diseases [13]. Indeed, long term measures should be focussed towards improving healthcare systems including the

number of medical doctors per 10000 population, HCI, health care budget spent *per capita* and instituting population control measures which will enable us to effectively tackle future pandemics.

It is pertinent to note that our model despite an extensive repertoire of indicators could only explain up to 30.6% of the variability in the case fatality rates. This indicates that the COVID-19 pandemic is difficult to predict and model at present. Either more and different indicators need to be evaluated or more data is needed or more time is needed to understand the dynamics of this disease.

The limitation of this study is that with approximately 3 months of data so far into this pandemic we were only able to catch the early trends and we believe with the passage of time and more data at our disposal, we may be able to refine the model.

---

## Conclusions

Each pandemic brings its own set of problems, but analysing early trends during the pandemic may help in shaping the response to the pandemic itself. COVID-19 appears to be difficult to control because of this virus' high infectivity. The globalisation of economy and integration of various industrial sectors across the globe has made this pandemic spread fast, unlike any other pandemic. Countries with better economic infrastructure in place might fare well in this pandemic compared to lesser privileged countries. Policymakers at international, regional and governmental levels need to formulate policies to address this imbalance as a breakdown of the healthcare system in one place will put all connected countries at risk. A rapidly increasing number of confirmed cases may overwhelm any healthcare systems; hence all possible measures should be instituted to keep it down so that effective care can be provided to affected patients. Measures like quarantine, lockdowns, travel restrictions may help in keeping the mortality rates low. Overall, this pandemic like any pandemic exposes the lacunae in existing systems and also provides an opportunity to learn what measures work best and how to be prepared for newer challenges and what we learn today, may shape a better tomorrow.

---

## References

1. Guo YR, Cao QD, Hong ZS, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil Med Res* 2020;7:11.
2. WHO. What is a pandemic? Accessed on: 19 March 2020. Available from: [http://www.who.int/csr/disease/swineflu/frequently\\_asked\\_questions/pandemic/en/](http://www.who.int/csr/disease/swineflu/frequently_asked_questions/pandemic/en/)
3. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2020;20:553-8.
4. Wong JY, Kelly H, Cheung C-MM, et al. Hospitalization fatality risk of influenza A(H1N1)pdm09: A systematic review and meta-analysis. *Am J Epidemiol* 2015;182:294-301.
5. WHO. Novel Coronavirus (2019-nCoV) situation reports. Accessed on: 19 March 2020. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
6. World Bank. Open Data. Accessed on: 19 March 2020. Available from: <https://data.worldbank.org/>
7. WHO. MoNITOR: Mother and Newborn information for

- tracking and results. Accessed on: 19 March 2020. Available from: <https://www.who.int/data/maternal-newborn-child-adolescent/monitor>
8. Lee K-M, Jung K. Factors influencing the response to infectious diseases: Focusing on the case of SARS and MERS in South Korea. *Int J Environ Res Public Health* 2019 22;16:1432.
  9. Pasquini-Descomps H, Brender N, Maradan D. Value for money in H1N1 influenza: A systematic review of the cost-effectiveness of pandemic interventions. *Value Health* 2017;20:819–27.
  10. Lowcock EC, Rosella LC, Foisy J, et al. The social determinants of health and pandemic H1N1 2009 influenza severity. *Am J Public Health* 2012;102:e51–8.
  11. Our World in Data [Internet]. Age structure. 20 September 2019. Accessed on: 19 March 2020. Available from: <https://ourworldindata.org/age-structure>
  12. Kumar S, Quinn SC. Existing health inequalities in India: informing preparedness planning for an influenza pandemic. *Health Policy Plan* 2012;27:516-26.
  13. Class RJ, Glass LM, Bayeler WE, Min HJ. Targeted social distancing designs for pandemic influenza. *Emerg Infect Dis* 2006;12:1671-81.

Non-commercial use only