Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 367–386

Association of Socioeconomic and Demographic Factors With COVID-19 Related Health Outcomes in SAARC Nations

Sangeeta Chakravarty[1], Gurprit Grover[2] and Sanya Aggarwal[2]
[1]*Institute of Economic Growth, Delhi University Enclave, Delhi*
[2]*Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi*

Received: 11 December 2020; Revised: 11 March 2021; Accepted: 13 March 2021

Abstract

COVID-19 pandemic has reshaped our world in a timescale much shorter than what we can understand and is now a major global health threat. As there was no preparedness on this virus, authorities around the world took restrictive policy measures to control the spread to ensure the wellbeing of the people. This pandemic affected both developed and underdeveloped countries equally. Moreover, existing socioeconomic and demographic characteristics of the countries may be contributing to the variation in health outcomes between countries. This study aims to analyse the influence of socioeconomic and demographic factors on COVID-19 related health outcomes in SAARC nations. The study is important as the objectives behind SAARC are regional integration and economic development of its member countries.

Panel regression analysis and Negative binomial regression are used to identify country specific factors that are associated with COVID-19 related Case Fatality Rate (CFR) and count data, such as, daily cases and active cases, respectively. The findings of the study indicate that increasing CFR are associated with countries having higher cardiovascular death rates, diabetes prevalence, health expenditure (percentage of GDP) and life expectancy. It is also found that co-morbidities such as cardiovascular disease, Tuberculosis and diabetes prevalence are associated with increased national caseloads and mortality, respectively. The study may help government to evaluate policies that can aid in managing the effects of the pandemic by utilizing resources and capabilities in an efficient way.

Key words: COVID-19; Case Fatality Rate; SAARC nations; Socioeconomic factors; Demographic factors; Negative binomial regression; Panel data analysis.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causing coronavirus disease 2019 (COVID-19), was first reported in Wuhan, Hubei Province, China in December 2019 [Yang *et al*. (2020)]. It was declared global pandemic by World Health Organisation (WHO) on 11th March 2020 [Cucinotta and Vanelli (2020)]. As on 31st January 2021, more than 100 million people were infected with COVID-19 and 2.2 million have already died [WHO (2020)]. After initial breakout of COVID-19 in China, the epicentre changed to Italy, United Kingdom (UK) and then to United States of America (USA) [Gupta and Misra (2020)]. Most infected cases were in USA followed by India and Brazil [WHO (2020)]. SARS-CoV-2 has a stronger transmission capacity as compared with the SARS-CoV that caused an outbreak

Corresponding Author: Sanya Aggarwal
Email: sanya.aggarwal.stats@gmail.com

of SARS in 2003 [Ma *et al*. (2020)]. Possible modes of transmission of virus causing COVID-19 includes animal-to-human transmission, human-to-human through casual contact, droplets, airborne, fomite, fecal-oral, bloodborne and mother-to-child transmission [World Health Organization (2020)]. Although most people infected with the SARS-COV-2 virus will experience mild to moderate respiratory illness and recover without requiring special treatment. Older people, on the other hand, with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more likely to develop serious illness [World Health Organization (2020)].

The South Asian Association for Regional Cooperation (SAARC) countries comprises 3% of the world's area and home to 21% of the world's total population and comprising of eight nations—Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka—has experienced the wave of pandemic much later than Europe and America [Shohan *et al*. (2020), WHO (2020)]. SAARC nations share a common regional space with similar geographical conditions and population, yet they differ significantly in the prevalence, severity, mortality and management of the pandemic. The first case of COVID-19 in this region was reported in Nepal on 23rd January 2020 [WHO (2020)]. As on 31st January 2021, India has the highest prevalence of COVID-19 in the region and ranks second globally, on contrary, Bhutan records one death due to the virus [WHO (2020)]. Some of these variations could be ascribed to demographic, social and economic factors, as well as health infrastructure, access to healthcare, political and public health response. Although SAARC countries had gained upper hand in demarcating the initial entry of COVID-19 into the countries, the region is much more vulnerable to its severe impacts. Infectious diseases are the major cause of mortality and morbidity in South Asia [Zaidi *et al*. (2020)]. Recently, World Bank has warned that South Asia faces its worst economic performance in ten years due to this deadly SARS-COV-2 virus. An emergency fund in response to the pandemic has been set up by these nations where each country has voluntarily contributed to secure the people of the region [Augustine (2020)]. But the region is less prepared against pandemic due to poverty, poor medical infrastructure and medical care facilities, as well as the lower number of physicians. An evidence-based study thus becomes imperative to assist policy makers and government in limiting the impact of COVID-19.

Good health improves learning, working production and income and as such health contributes to economic growth and development of the nation. For an unprecedented epidemic such as COVID-19 where individual level data is not available, frequency level estimation such as number of cases, number of deaths, number of active cases, *etc*., becomes a viable choice. Various studies have been conducted on COVID-19 related impacts on SAARC nations. Sultana and Reza (2020) studied the impact of COVID-19 from the perspective of working population of SAARC nations. Shohan *et al*. (2020) examined the onset and transmission of the virus in each SAARC country at an early stage and critically appraised their response with respect to their medical facilities for diagnosis and management. Awasthi (2020) discussed challenges faced by SAARC countries in the wake of COVID-19 pandemic and how India's endeavour is bringing all the nations together in combating the pandemic. Deo *et al*. (2020) predicted the dynamics of COVID-19 pandemic in India. Some studies have reported the effect of country specific factors on COVID-19 around the world. Chaudhry *et al*. (2020) conducted a country level exploratory analysis to assess the impact of timing and type of national health policy/actions undertaken towards COVID-19 mortality and related health outcomes. Yang *et* *al.* (2020) studied the impact of COVID-19 in Wuhan, China and suggests that older patients with co-morbidities had increased risk of death.

This paper aims to examine how country-specific socioeconomic and demographic factors effect health outcomes related to COVID-19 in SAARC nations. The importance of selecting SAARC for this study is that the geographical position of some of the member countries is such that they share their borders with China, where the cases first reported. Panel regression analysis and Negative Binomial (NB) regression modelling are utilised to model COVID-19 related health outcomes such as, Case Fatality Rate (CFR), daily infected cases and total active cases against country specific factors.

2. Materials and Methodology

2.1. Data

The study includes twelve country specific factors of eight countries. COVID-19 related health outcomes included in this study are Case fatality rate (CFR), number of reported cases and number of active cases. Publicly available information on COVID-19 related health outcomes such as number of cases, recovered cases and total deaths were extracted from various websites [Roser *et al.* (2020), COVID (2019)]. CFR is defined as the ratio of number of deaths by number of infected cases due to disease over a certain period of time. For any disease to be less severe, the CFR should be less than 1 % [Global Health Observatory (2020)]. The higher CFR suggests that the disease is severe and requires measures by government and individuals to minimise the fatalities. Daily cases are calculated by subtracting total number of cases at time t to total number of cases at time (t–1). Active cases are the number of cases which are neither dead nor recovered but are still infected. It is calculated by subtracting recovered and dead cases from the number of infected cases. Various other rates that are utilized in the study to measure the severity of COVID-19 related health outcomes are recovery rate, percentage of active cases and infection cases per capita. Recovery rate and percentage of active cases are calculated similarly as CFR with numerator changed to number of recovered cases and number of active cases, respectively. Infection per capita is another measure for understanding severity of the disease. It is calculated as the number of infections in each region to the total population in that region over a certain period of time.

Data on country level variables and indices were captured through various sources (see Appendix Table A.1). These includes total population (2019), population density, life expectancy, cardiovascular death rates, diabetes prevalence, GDP per capita and handwashing facilities [Roser *et al.* (2020)]. Other factors included were health expenditure (% of GDP), Tuberculosis (TB) prevalence, age dependency ratio, hospital beds per ten thousand population and proportion of employed population below poverty line [Asian Development Bank (ADB) (2020)]. Global health security (GHS) is another factor that is included in the study for each country [GHS Index Project Team (2019)].

The proportion of the employed population below the international poverty line of US$1.90 per day, also referred to as the working poverty rate, reveals the proportion of the employed population living in poverty despite being employed, implying that their employment-related incomes are not enough to lift them and their families out of poverty and ensure decent living conditions [United Nations SDG indicators (2020)]. Age dependency ratio (% of working population) is the ratio of dependents of people younger than 15 or older than

64, to the working-age population, that is, between 15-64 years of age. Data are shown as the proportion of dependents per 100 working-age population [World Bank, (2019)]. Similarly, the data on stringency index was also obtained for each country at each point of time (Hale *et al.* 2020). The index is published and updated real time by a research group from Oxford university on nine response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest) [Jayatilleke *et al*. (2020), Hale *et al.* (2020)].

Another index utilised in the study is Global Health Security index (GHS). The index is the comprehensive assessment of 195 countries' health security and related capabilities cross six categories, 34 indicators, and 85 sub-indicators. The six categories are as follows: Prevention of the emergence or release of pathogens; early detection and reporting for epidemics of potential international concern; rapid response to and mitigation of the spread of an epidemic; Sufficient and robust health system to treat the sick and protect health workers; commitments to improving national capacity, financing plans to address gaps and adhering to global norms; and; overall risk environment and country vulnerability to biological threats [GHS Index Project Team (2019)].

As the data is continuously evolving, the period for the study considered is from 25th January 2020 to 31st January 2021. The data is divided into two parts and the point of partition is obtained by plotting the average stringency index per day and recording the date when stringency index was below 60. The date thus obtained is 15th September 2020. For this study, the first phase is considered from 25th January 2020-14th September 2020 and the second phase is considered from 15th September 2020-31st January 2021. We then determine the impact of the socioeconomic and demographic factors on COVID-19 health outcomes in these two periods of the pandemic.

2.2. Statistical models

The descriptive analysis was conducted on COVID-19 related health outcomes of SAARC nations. For modelling the relationship between CFR and country specific variables, Panel regression modelling technique is utilized. Panel regression modelling is used to model longitudinal data.

2.2.1. Panel regression modelling

The basic linear panel models can be described through suitable restrictions of the following general model:

$$y_{it} = \alpha_{it} + \beta'_{it}x_{it} + \mu_{it} \tag{1}$$

where, $i = 1,2,\ldots,n$ is the individual country index, and, $t = 1,2,\ldots,T$ is the time index and μ_{it} is a random disturbance term of mean 0 [Menard (2007), Croissant and Millo (2008)]. When t is same for all countries, it is called balanced data, otherwise it is unbalanced data. The data is recorded from the occurrence of first COVID-19 case in the each of the SAARC country, the data set is thus unbalanced. When the assumption of parameter homogeneity is taken, that is, $\alpha_{it} = \alpha$ for all *i, t* and $\beta_{it} = \beta$ for all *i, t*; the resulting model is standard linear pooled model, written as,

$$y_{it} = \alpha + \beta' x_{it} + \mu_{it} \tag{2}$$

To model individual heterogeneity, the error term assumes two separate components, one of which is specific to the individual and does not change over time. This is called the unobserved effects model which can be represented as:

$$y_{it} = \alpha + \beta' x_{it} + \mu_{it} + \varepsilon_{it} \tag{3}$$

The appropriate estimation method for this model depends on the properties of the two error components. If the individual component is missing altogether, pooled OLS is the most efficient estimator for β. To check 'poolability' of the data, pooling tests is conducted i.e., the hypothesis that the same coefficients apply across all individuals. It is a standard F test, based on the comparison of a model obtained for the full sample and a model based on the estimation of an equation for each country [Croissant and Millo (2008)]. Rejection of null hypothesis implies the rejection of poolability and other techniques should be utilized to analyse the data.

2.2.2. Poisson model

For studying the relationship between frequency type dependent variable and other independent variables, Poisson and Negative Binomial modelling are recommended. Poisson regression is typically used to evaluate count data in public health. It is often assumed that the number of events follows a Poisson distribution with a conditional mean μ depending upon a set of regressors x and corresponding parameters β for a participant's linear predictor. Using a log link, we can express the expected number of events for country i as $\mu_i = E(y_i|x_i) = e^{\beta' x_i}$. The Poisson probability distribution of y_i given x_i can be expressed as:

$$P(Y_i = y_i) = \frac{e^{-\mu_i} {\mu_i}^{y_i}}{y_i!} \tag{4}$$

where, y_i is a non- negative integer. The log likelihood for the model can be expressed as:

$$LL(\beta) = \sum_{i=1}^{n} y_i X' \beta - \mu_i - \log{(y_i!)} \tag{5}$$

However, this model assumes the variance is equal to mean, an assumption which is often violated [Rose *et al.* (2006)]. The most common alternative for over dispersion of dependent variable over Poisson regression is Negative Binomial (NB) model, which has a built-in dispersion parameter and can account for variance greater than mean [Agresti (2003)].

2.2.3. Negative Binomial (NB) model

The NB regression model allows for over dispersion by introducing an unobserved heterogeneity term for observation [Sheu *et al.* (2004)], *i.e.*, $\mu_i = e^{(\beta' x_i + e_i)}$. We normally assume that $\exp{(e_i)}$ has a gamma distribution with mean 1 and variance a so that the conditional mean of y_i is still μ_i but the conditional variance of y_i becomes $\mu_i(1 + a\mu_i)$. As a approaches zero, y becomes a Poisson distribution and as a becomes larger the distribution becomes more dispersed. The NB probability distribution for country i is given by:

$$P(Y_i = y_i) = \frac{\Gamma(y_i + a^{-1})}{\Gamma(y_i + 1)\Gamma(a^{-1})} \left(\frac{1}{1 + a\mu_i}\right)^{a^{-1}} \left(\frac{a\mu_i}{1 + a\mu_i}\right)^{y_i} \tag{6}$$

where, $\mu_i, a, and\ \Gamma(.)$ refer to the mean of the count distribution, the NB dispersion parameter, and the gamma function [Rose *et al.* (2006)]. The log likelihood for the model can be expressed as:

$$LL(\beta, a) = \sum_{i=1}^{n}\{\log[\Gamma(y_i + a^{-1})] - \log[\Gamma(y_i + 1)] - \log[\Gamma(a^{-1})] - a^{-1}\log(1 + a\mu_i) + y_i\log(\mu_i) + y_i\log(a) - y_i\log(1 + a\mu_i)\} \quad (7)$$

which can be maximized by iterative methods (preferably Newton–Raphson) to obtain the estimates of β and a.

2.2.4. Model comparison

To compare the predictive performance of NB regression model with that of Poisson regression, common model selection criterion, Akaike information criterion (AIC) is used. AIC is calculated as $-AIC = -2\ log\ L + k,$ where L denotes the likelihood function of the model evaluated at maximum likelihood estimates and k is the total number of parameters in the model. The models which had a higher log-likelihood, or a lower AIC value are considered to be the best. Model's goodness of fit was accessed by AIC and Cox and Snell pseudo R-squared statistic. Cox and Snell pseudo R-squared that uses likelihood ratio to assess overall fit compared to null model. It is calculated as :

$$Cox\ and\ Snell\ pseudo\ R - square = 1 - \left[\frac{LR(full\ model)}{LR(null\ model)}\right]^{2/n} \quad (8)$$

where n is the sample size and LR is the likelihood ratio of the model. NB regression does not have an equivalent to the R-squared measure found in ordinary least squares (OLS) regression, pseudo R-square measure are utilised. Its value ranges from 0 to l higher value indicates a better fitting model [Allison (2014)].

Given the limited sample size of 8 countries, the potential independent variables included in the models were identified using forward selection process. Population density was adjusted on logarithmic scale for ease of calculation. The results of the selected regression models were reported in the incidence rate ratio (IRR) where a value less than one suggests a decreased likelihood and a value of greater than one denotes an increased likelihood of the event under investigation. Similar analysis is then carried out on the two parts of the data as explained in Section 2.1.The data was managed in excel and the statistical analysis was carried out using R software.

3. Results

The situation of COVID-19 related health outcomes of 8 SAARC countries as on 31st January 2021 are presented in Table A.2 (see Appendix). India has recorded highest number of cases with 10,757,610 infected individuals, followed by Pakistan 546,428 and Bangladesh 535,139. It is evident that death toll was highest in India with 154,392 people dying due to COVID-19 followed by Pakistan (11,683) and Bangladesh (8,127). Bhutan on the other hand had only one death due to COVID-19. Highest number of recovered and active cases were seen in India. Pakistan and Bangladesh recorded recovered cases at 501,252 and 479,744, respectively.

Only 814 patients were recovered in Bhutan with 44 patients still active. Sri Lanka reports 57,159 recovered and 6,682 active cases. Maldives stands at 14,139 recovered patients with Nepal at 266,336 and Afghanistan at 47,679. Figure A.1 (see Appendix) shows the progression of the epidemic from the first reported case in each of the SAARC nations. Table

1 enlists the percentage of COVID-19 related outcomes. Even though India records highest COVID-19 related cases, but it is observed that infection per capita (percentage of people infected with COVID-19) is highest in Maldives with 2.93 per cent of people infected, followed by Nepal at 0.93 and India at 0.78 per cent. Bangladesh has 0.32 per cent of population infected and Sri Lanka with 0.3 per cent.

Table 1: COVID-19 related events rate as on 31st January, 2021

Countries	Infection per capita	CFR	Recovery rate	Active cases rate
Afghanistan	0.14	4.36	86.65	8.99
Bangladesh	0.32	1.52	89.65	8.83
Bhutan	0.11	0.12	94.76	5.12
India	0.78	1.44	97.00	1.56
Maldives	2.93	0.33	89.26	10.42
Nepal	0.93	0.75	98.29	0.96
Pakistan	0.25	2.14	91.73	6.13
Sri Lanka	0.30	0.49	89.09	10.42

Bhutan reports lowest per capita infection of 0.11 percent, with Afghanistan at 0.14 percent and Pakistan 0.25. It can be observed that percentage of deaths out of total infected cases, also known as CFR, was highest in Afghanistan (4.36) followed by Pakistan (2.14), Bangladesh (1.52) and India (1.44). Maldives and Sri Lanka record the fatality rate at 0.33 and 0.49 percent, respectively. Bhutan, on the other hand, witnessed 0.12 percent fatalities. The highest recovery rate has been recorded in Nepal of 98.29 percent followed by India where 97 percent of COVID-19 infected patients have gained recovery. Bhutan has also witnessed a recovery rate of 94.76 percent while the rate of recuperation in the Bangladesh stands at 89.65 percent. Maldives and Sri Lanka both have 10.42 percent of active cases out of total infected cases whereas Nepal has only 0.96 percent of active cases and India with 1.56 active cases. Figure 1 illustrates these rates from the first reported cases in each of the SAARC nations.

Table 2: Descriptive Statistics of the variables

Variable	N	Mean	Standard Deviation	Minimum	Maximum
Number of cases	8	582574	1963378	1	10757610
Number of recovered cases	8	516435	1827691	0	10434983
Number of deaths	8	9108	29157.71	0	154392
Number of active cases	8	57031	158712.3	0	1017754
Number of daily cases	8	4391	14244.32	0	97894
Case Fatality Rate (CFR)	8	1.27	1.42	0	12.82
Recovery rate (RR)	8	61.47	32.21	0	100
Stringency Index	7	37.26	25.26	2.78	100

Descriptive analysis of the COVID-19 related health outcomes and stringency index recorded for each of the SAARC nations are recorded in the Table 2. On an average 582,574 infected cases, 516,435 recovered cases and 9,108 deaths were recorded for all nations. Average daily cases recorded were 4,391 with total active cases average being 57,031. Average

CFR was 1.27, reaching maximum at 12.82. Average recovery rate recorded was 61.47. The countries' stringency index average was 37.26 with minimum being 2.78 and highest recorded value being 100 (Table 2).

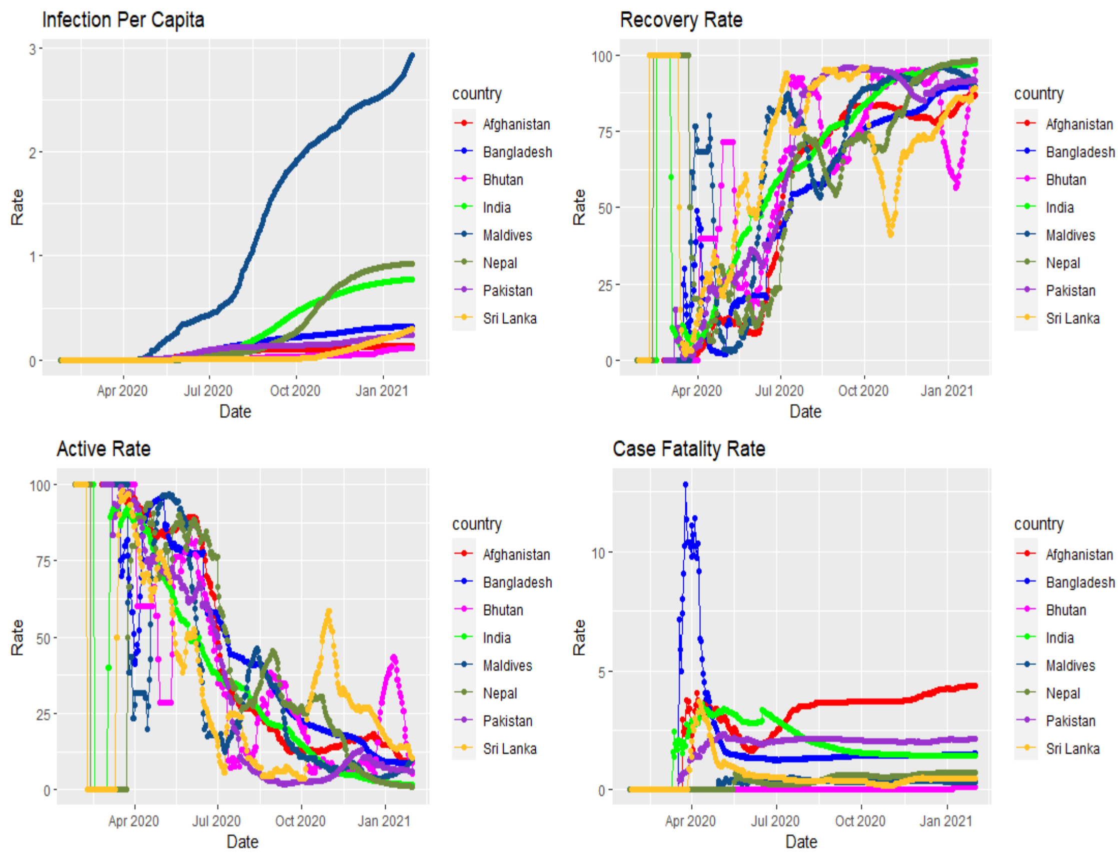

Figure 1: COVID-19 related rates as on 31st January 2021 in SAARC nations

Results of the tests to validate the choice of the regression models are presented in Table 3. The pooling tests to check 'poolability' of the data, *i.e.*, the hypothesis that the same coefficients apply across all individuals, has the normal test statistic –1.369 and *p*-value 0.9145. Null hypothesis is not rejected implying that the individual effect of coefficient is missing, and pooling technique is the most suitable for CFR. The suitable model for daily cases and active cases was assessed using AIC. For daily reported cases, NB model has smaller AIC (40376.93) than Poisson model (49386782), suggesting that NB model is better fit. Similarly, for active cases, NB model shows better fit over Poisson model.

Table 3: Tests to validate the choice of models for each dependent variable

Dependent variable 1: CFR		**Dependent variable 2 : Daily cases**		**Dependent variable 3 : Active cases**	
	Pooling test		AIC		AIC
Test statistic	–1.369	Poisson	49386782	Poisson	537267252
p-value	0.9145	NB	40376.93	NB	56942.78

Table 4: Multiple regression models for the dependent variable 'daily cases'

	Poisson distribution			NB distribution		
Variables	**Estimate**	**S.E.**	***p*-value**	**Estimate**	**S.E.**	***p*-value**
(Intercept)	–18.00	0.02131	<.0001	–27.37	0.6842	<.0001
GHS	0.3421	0.00015	<.0001	0.3344	0.0105	<.0001
Population density	1.1070	0.00163	<.0001	2.168	0.0335	<.0001
TB prevalence	0.0032	0.00003	<.0001	0.0022	0.0008	<.0001
Age dependency ratio	0.0798	0.00023	<.0001	0.1165	0.0079	<.0001
Health expenditure (% of GDP)	0.0379	0.00079	<.0001	0.2194	0.0303	<.0001
Stringency index	0.0119	0.00002	<.0001	0.02441	0.0015	<.0001
AIC	12159689			33821		
Residual deviance	12143840			2998.6		
Degrees of freedom	2452			2452		

Note: S.E. stands for Standard Error; GHS stands for Global Health security index

Poisson and NB regression models fitted to the data with daily cases as dependent variable are presented in Table 4. The AIC of NB regression (33,821) is much lower than that of Poisson regression (12,159,689). It can be observed that residual deviance of NB regression (2,998.6) is much lower suggesting that the model estimates decent amount of variation than the Poisson regression model. The same results can be verified from regression plots in Figure A.2.(see Appendix). As can be observed from Q-Q plot, the points fall relatively closer to the dashed line for NB regression model than Poisson regression. The residual vs leverage plot shows that there are several problematic points in the Poisson model and fewer in the NB regression model. In general, NB regression model provides a better fit for daily reported cases.

Table 5: Multiple regression model for dependent variable ‘active cases'

	Poisson distribution			NB distribution		
Variables	**Estimate**	**S.E.**	***p*-value**	**Estimate**	**S.E.**	***p*-value**
(Intercept)	–3.771	0.03597	<0.0001	–3.7711	0.6587	<0.0001
Population density	1.981	0.00290	<0.0001	1.9810	0.0394	<0.0001
Cardiovascular death rate	0.009	0.00003	<0.0001	0.0094	0.0012	<0.0001
Age dependency ratio	–0.095	0.00043	<0.0001	–0.0947	0.0127	<0.0001
Employed people BPL	0.093	0.00010	<0.0001	0.0938	0.0051	<0.0001
Hospital beds(per 10,000 people)	–0.330	0.00005	<0.0001	–0.3301	0.0067	<0.0001
Handwashing facilities	0.121	0.00007	<0.0001	0.1211	0.0043	<0.0001
AIC	147431462			47480		
Residual deviance	147408287			3042.4		
Degrees of freedom	2411			2411		

Note: BPL stands for Below Poverty Line, S.E. stands for Standard Error

Similarly, the results from Poisson regression model and NB regression model for ‘Active cases’ as the dependent variable are recorded in Table 5. It can be observed that the NB regression model has comparatively smaller residual deviance (3,042.4) and AIC (47,480) value than Poisson regression model implying that NB regression model has better fit. The results can be verified from the regression plots (Figure A.3, see Appendix). The Q-Q plot shows that the NB regression model should be preferred over the Poisson model. The analysis thus suggests that NB regression model has better than Poisson regression model for active cases variable.

Table 6: Panel regression analysis on COVID-19 Case fatality rate (CFR)

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Case Fatality Rate (CFR)	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
Cardiovascular death rate	1.025 (1.023-1.027)	1.029 (1.026-1.032)	1.019 (1.018-1.019)
Diabetes prevalence	2.735 (2.469-3.029)	3.551 (3.024-4.157)	1.792 (1.757-1.829)
Hospital beds (per 10,000 people)	0.835 (0.821-0.849)	0.792 (0.771-0.813)	0.905 (0.902-0.908)
Employed people BPL	0.905 (0.891-0.920)	0.882 (0.861-0.905)	0.943 (0.940-0.946)
Health expenditure (% GDP)	1.148 (1.110-1.187)	1.143 (1.085-1.205)	1.145 (1.137-1.152)
Life expectancy	2.249 (2.101-2.409)	2.885 (2.582-3.201)	1.555 (1.535-1.576)
R-squared	0.53	0.42	0.99

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, BPL stands for Below Poverty Line

The findings for association between CFR and country specific factors from Panel regression modelling are presented in Table 6. The significant factors associated with the CFR are cardiovascular death rates, prevalence of diabetes, hospital beds per ten thousand people, employed persons below poverty line, health expenditure (% of GDP) and life expectancy. There was negative association between hospital bed per ten thousand people (IRR = 0.835; 95% CI: 0.821-0.849) and CFR. People employed below poverty line, earning less than US$1.99 per day was also negatively associated with CFR (IRR = 0.905; 95% CI: 0.891-0.920).

Table 7: Negative Binomial regression analysis on COVID-19 daily reported cases

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Daily Cases	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
GHS	1.397 (1.366-1.430)	1.436 (1.398-1.477)	1.472 (1.435-1.510)
Population density	8.731 (8.082-9.409)	8.633 (7.788-9.544)	8.967 (8.301-9.665)
TB prevalence	1.002 (1.0003-1.004)	1.009 (1.006-1.012)	0.991 (0.989-0.994)
Age dependency ratio	1.123 (1.103-1.143)	1.171 (1.144-1.199)	1.218 (1.187-1.251)
Health expenditure (% GDP)	1.249 (1.184-1.318)	1.037 (0.952-1.141)	1.212 (1.147-1.281)
Stringency index	1.024 (1.019-1.029)	1.057 (1.050-1.063)	1.057 (1.048-1.065)
Cox and Snell pseudo R-square	0.65	0.70	0.84

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, GHS stands for Global Health security index

In contrast, countries with higher cardiovascular death rates (IRR = 1.025; 95% CI: 1.023-1.027), higher diabetes prevalence (IRR= 2.735; 95% CI: 2.469-3.029), spends higher percentage of GDP on healthcare (IRR= 1.148; 95% CI:1.110-1.187) and have higher life expectancy (IRR = 2.249; 95% CI: 2.101-2.409) had significantly higher CFR. The R-squared value is 0.53 reveals that model explains 53% of the variation in the response variable CFR. This implies that the model has decent fit.

The results of NB regression for daily reported cases are presented in Table 7 in terms of IRR. Socioeconomic and demographic factors positively associated with the increasing daily cases are Global health score (GHS) (IRR = 1.397; 95% CI: 1.366 -1.430), population density (IRR = 8.731; 95% CI: 8.082-9.409), higher prevalence of Tuberculosis (IRR = 1.002; 95% CI: 1.0003-1.004), higher age dependency ratio (% of working population) (IRR = 1.123; 95% CI: 1.103-1.143) and higher stringency index (IRR = 1.024; 95% CI: 1.019-1.029). Higher healthcare expenditure as percentage of GDP (IRR = 1.249; 95% CI: 1.184-1.318) was associated also with higher number of daily reported infected cases. Cox & Snell's pseudo R-squared value is 0.65 which implies that the model has decent fit.

The findings of NB regression analysis of the total active cases on each day (Table 8) suggests that factors significantly associated with increased active cases are: population density (IRR = 7.254 95% CI: 6.710- 7.832), higher cardiovascular death rates (IRR = 1.009; 95% CI: 1.007- 1.012), higher employed people working below poverty line (with less than US$1.99 per day) (IRR = 1.099; 95% CI: 1.088- 1.109) and higher handwashing facilities in the country (IRR = 1.129; 95% CI: 1.120- 1.138). In contrast, higher age dependency ratio (IRR = 0.910; 95% CI: 0.887- 0.933) and more hospital bed available per ten thousand people (IRR = 0.719; 95% CI: 0.709- 0.728) were associated with lower number of active cases in the country. Cox & Snell's pseudo R squared value reported as 0.72 implies that the model has decent fit.

Table 8: Negative Binomial regression analysis on COVID-19 active cases

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Active Cases	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
Population density	7.254 (6.710-7.832)	7.749 (6.862-8.722)	7.064 (6.727-7.414)
Cardiovascular death rate	1.009 (1.007-1.012)	1.017 (1.013-1.021)	1.003 (1.002-1.005)
Age dependency ratio	0.910 (0.887-0.933)	0.869 (0.837-0.904)	0.934 (0.919-0.949)
Employed people BPL	1.099 (1.088-1.109)	1.086 (1.069-1.103)	1.114 (1.106-1.120)
Hospital beds(per 10,000 people)	0.719 (0.709-0.728)	0.724 (0.709-0.738)	0.708 (0.702-0.714)
Handwashing facilities	1.129 (1.120-1.138)	1.129 (1.114-1.143)	1.132 (1.126-1.138)
Cox & Snell pseudo R-square	0.72	0.63	0.95

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, BPL stands for Below Poverty Line

Further, the results of partitioned data to ascertain any differential change in impact of socioeconomic and demographic variables on the health outcomes are presented in Table 6-8. The point of partition is fixed on 15th September 2020, when the average stringency index was below 60 for the first time (see Appendix Figure A.4) Notable difference appears in the association of Tuberculosis prevalence and daily reported cases. In the phase from 15th September 2020 to 31st January 2021 there is negative association between TB prevalence and daily reported cases (Table 7). The result contrasts with the positive association during the first phase and full data analysis. All the other results are in line with the complete data analysis (Table 6-Table 8).

4. Discussion

It is important to analyse the significant association between country specific socioeconomic and demographic factors and COVID-19 related health outcomes. The three most affected SAARC countries are India, Bangladesh and Pakistan, respectively, which are also the densely populated nations of SAARC (see Appendix Table A.1). It is evident that the death toll was highest in India followed by Pakistan and Bangladesh. Bhutan on the other hand records only one COVID-19 related death till date and has least number of COVID-19 confirmed cases. On contrary, it is observed that the infection per capita was highest in Maldives, followed by Nepal and India. This is possibly because Maldives has the highest population density with 1,454 people living per square km, which might have resulted in higher transmission rates of infections. It was found that CFR is highest in Afghanistan, being 4.36% deaths of the confirmed cases. The reason of high fatality may be limited health resources and poor health knowledge. Poverty is another issue contributing to worsening the situation of COVID-19 in a war-torn Afghanistan, which is in line other studies, such as, Husseini and Kamil (2020), Sultana and Reza (2020), among others.

The findings of impact of country specific factors on CFR suggests that countries with higher diabetes prevalence and cardiovascular death rates are associated with higher CFR. Also, countries with higher cardiovascular death rate have higher number active cases. Complications are more common in patients with cardiac complication and diabetes with higher mortality than those without it. The results are supported by various studies such as Yang *et al*. (2020) and Zheng *et al*. (2020). Another finding of this analysis is that countries with higher life expectancy have higher CFR. Reports by Onder *et al*. (2020) and Hussain *et al*. (2020) showed similar results that older patients with chronic diseases were at higher risk for severe COVID-19 related mortality.

Health infrastructure is an important factor that affects cases and fatalities. It is found that higher number of hospital beds per ten thousand people was associated with lower CFR and reduced active cases each day. Hospital beds are crucial during an outbreak such as COVID-19 to assess the health facility, as critical cases need medical care in hospital settings for a longer time compared to ordinary patients, thus reducing the active cases and mortality rates. The result is line with the findings of Khan *et al*. (2020) and Blondel and Vranceanu (2020) reporting that COVID-19 fatalities are lower in countries with significant resources dedicated to health care such as hospital beds. Further we have also found that the countries which spend higher percentage of GDP on healthcare witnessed higher number of daily reported cases and CFR. This means that nations that spent more percentage of GDP on healthcare were not insulated from COVID-19 related deaths. This trend was also seen among the wealthy nations such as North America and Europe. There are a few possible explanations for this result between healthcare investment and CFR related to COVID-19 among SAARC nations. Higher healthcare expenditure (% of GDP) was not associated with higher GDP per capita (see Appendix Table A.1). For example, Afghanistan has the lowest GDP per capita but spends more that 10% of GDP on healthcare, which also reports highest CFR. With higher underlying disease burden and higher population, these nations have much lower number of hospital beds and advanced equipment per population, and fewer medical staff to respond to this unprecedented threat from the pandemic. Similar results have been reported by Sorci *et al*. (2020) and Khan *et al*. (2020).

The findings shows that countries with higher TB prevalence have more reported cases per day. SAARC has 37% of the global burden of TB [STAC (2018)]. It may be the case that chronic respiratory diseases such as active TB could lead to increase in susceptibility to the COVID virus in this region. Other studies have also arrived on similar results [Liu, Bi *et al*. (2020), Maciel *et al*.(2020)].

Our analysis also suggests that countries with higher population density are associated with increased daily cases and active cases. Higher population density may potentially facilitate interactions between susceptible and infectious individuals, which sustains continued transmission and spread of COVID-19. This has been observed in case of Maldives which has the highest population density and has the highest infection per capita. Higher age-dependency ratio is associated with have higher number of daily cases and lower active cases. Younger individuals tend to have a higher proportion of asymptomatic or mild symptoms which are less likely to be detected in testing [Cortis (2020)]. On the contrary, elderly family members with underlying comorbidities are more susceptible to the disease [Liu *et al*. (2020)]. Further, demography of Asia has a lower proportion of elderly individuals than Western nations, with about 85% of the population in India, Pakistan and Bangladesh is younger than 50 years [Sultana and Reza (2020), Gupta and Misra (2020)]. Younger population has shorter disease course than elderly, hence lower active cases [Yang, Hung *et al*. (2020)].

The results also show that the countries with higher proportion of employed people earning below poverty line are associated with increased number of active cases but witnessed lower CFR. About 33.4% of the population in South Asia is living on less than US$1.99 per day income [Sultana and Reza (2020)]. This large population needs to go out to earn living, which increases the chances of infection spread. Also, poor housing facilities and overcrowded accommodation with limited access to personal outdoor space reduces compliance with social distancing thus increasing the overall active cases. The potential reason for lower CFR might include low testing and poor quality of data [Sannigrahi *et al.* (2020), Schellekens and Sourrouille (2020)]. As the huge proportion of this population is underprivileged, illiterate with poor health knowledge and have poor access to healthcare services due to limited income may contribute to the reason behind poor death records.

It is found that even countries that were in relatively more prepared condition according to the GHS index witnessed higher number of daily cases. This may point to the issue that health security is essentially weak in the SAARC nations. The average GHS index of SAARC nations is 36.55 and the global average is just 40.2 [Index Project Team (2019)]. The results also show that countries with higher handwashing facilities have higher reported active cases. The potential reason for such relationship is that high population density in SAARC nations makes it very difficult, if not impossible, to follow basic handwashing, hygiene and physical distancing practices advised during the COVID-19 outbreak, increasing transmission risks, and leading to increased precariousness in living conditions.

Another important finding is that the countries with higher stringency index witnessed increase in daily reported cases. This suggests that the stringent measures did not insulate the nations from spread of infection. Imposing stringent measures is very resource intensive which requires widespread testing and scrupulous contact tracing. The weak healthcare system with low testing rate in SAARC nations coupled with other factors such as, economically unprivileged population and high population density made maintaining social distance and lockdown challenging in these countries [Niazi *et al.* (2020)].

There are important limitations of the study. First, this study does not consider post COVID-19 factors such as increase in isolation camps, hospital beds, handwashing facilities and number of tests performed in each country. There was missing data for stringency index (Maldives) and handwashing facilities (Sri Lanka). This may have introduced important unintended bias. Missing values were not treated. Various other assumptions are also considered in the study. The basic definition of CFR is utilized for global comparison. The other popular definition, such as, the ratio of number of deaths to the sum of recovered and death cases was not used. Secondly, the asymptomatic COVID-19 population is not considered in the study. Thirdly, the population is assumed to be constant, i.e., it is closed for birth, death and migration. Also, the basic definition of population density is used that is the number of people living per square kilometre. This definition in denominator includes non-habitable lands where no or very little population resides. These assumptions might have led to underestimation of the results.

5. Conclusion

The COVID-19 pandemic has had a significant impact all over the world. During this time, a high number of deaths, public stress and economic damage was witnessed. This study addresses the association of various socio-economic and demographic factors with pandemic related health outcomes in the countries of the SAARC region. The results reveal that diseases, such as tuberculosis, cardiovascular diseases and diabetes, are related with increased mortality and national caseloads. Higher life expectancy is associated with increased mortality. Healthcare infrastructure such as higher number of hospital beds are associated with reduced active cases and mortality. Countries with higher GHS index witnessed higher number of caseloads. Increasing active cases and daily reported caseloads have a positive association with high population density. The findings from the data also suggest that during the later phase of the study period, socioeconomic factors such as, health expenditure (% of GDP), proportion of employed people earning below poverty line, hospital bed, age dependency ratio became more prominent in describing the path of pandemic.

There are many challenges before the SAARC nations, especially in the health sector. Due to this pandemic, healthcare has become a central point of economic and social well-being of all, even more so than before. It has made us realise how important it is to work on all dimensions jointly to save the mankind's present and future. This study will be helpful for evaluation of public health policies in SAARC countries.

Acknowledgement

The authors are very grateful to the reviewers and the editors for their invaluable comments and suggestions, which have helped us to improve the paper substantially.

References

Agresti, A. (2003). *Categorical Data Analysis* (Vol. 482). John Wiley & Sons. (ISBN: 0-471-36093-7).

Allison, P. D. (2014, March). Measures of fit for logistic regression. In *Proceedings of the SAS Global Forum* 2014 *Conference* (pp. 1-13).

Asian Development Bank (ADB). (2020). *Basic Statistics* 2020. Accessed on: 1st September 2020. Website URL: https://www.adb.org/publications/basic-statistics-2020

Augustine P., VID, C. Covid-19 (2020). A New Hope for Global Multilateralism. Website URL:http://ris.org.in/newsletter/diary/2020/Covid%2019%20II/pdf/AUGUSTINE%20 PETER.pdf

Awasthi, S. (May,2020). *India and SAARC Combating COVID Pandemic*. BIPSS commentary.

Blondel, S. and Vranceanu, R. (2020). COVID-19 mortality and health expenditures across European countries: the positive correlation puzzle. Available at: SSRN 3679972.

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J. and Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinical Medicine*, 25, 100464.

Cortis, D. (2020). On determining the age distribution of COVID-19 pandemic. *Frontiers in Public Health*, **8**, 202.

COVID, C. (2019). *Dashboard by the Center for Systems Science and Engineering* (*CSSE*) at Johns Hopkins University (JHU), 2020.

Croissant, Y. and Millo, G. (2008). Panel data econometrics in R: The plm package. Journal of statistical software, **27(2)**, 1-43.

Cucinotta, D. and Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. Acta bio-medica: Atenei Parmensis, 91(1), 157–160. DOI: https://doi.org/10.23750/abm.v91i1.9397

Deo, V., Chetiya, A. R., Deka, B. and Grover, G. (2020). Forecasting transmission dynamics of COVID-19 in India under containment measures - A time-dependent state-space SIR approach. *Statistics and Applications*, **18**(**1**), 157-180.

Global Health Observatory, WHO. (2020). Cholera case fatality ratio (%). Accessed on: 30th September 2020. Website URL: https://www.who.int/gho/epidemic_diseases/cholera/situation_trends_case_fatality_ratio/en/

Global Health Security (GHS) Index Project Team. (2019). GHS index: building collective action and accountability [Online]. Website URL: https://www.ghsindex.org/wp-content/uploads/2020/04/2019-Global-Health-Security-Index.pdf

Gupta, R. and Misra, A. (2020). COVID19 in South Asians/Asian Indians: Heterogeneity of data and implications for pathophysiology and research. *Diabetes Research and Clinical Practice*, **165**, 108267.

Hale, T., Webster, S., Petherick, A., Phillips, T. and Kira, B. (2020). Oxford covid-19 government response tracker. Blavatnik School of Government, **25**.

Hussain, A., Bhowmik, B. and do Vale Moreira, N. C. (2020). COVID-19 and diabetes: Knowledge in progress. Diabetes Research and Clinical Practice, 108142. DOI: https://doi.org/10.1016/j.diabres.2020.108142

Husseini, A. A. and Kamil, A. A. (2020). Estimating COVID-19 Dynamics in Afghanistan. *Erciyes Med J*, **42**(**4**), 468-73.

Indicators, S. D. G. Metadata repository. United Nations. The Sustainable Development Goal indicators. (2020). Website URL: https://unstats.un.org/sdgs/metadata/files/Metadata-01-01-01b.pdf

Jayatilleke, A. U., Dayarathne, S., de Silva, P., Siribaddana, P., Abeygunawardana, R. A. B., Nieveras, O., de Silva, N. and de Silva, J. (2020). COVID-19 case forecasting model for Sri Lanka based on Stringency Index. medRxiv. DOI: https://doi.org/10.1101/2020.05.20.20103887

Khan, J. R., Awan, N., Islam, M. and Muurlink, O. (2020). Healthcare Capacity, Health Expenditure, and Civil Society as Predictors of COVID-19 Case Fatalities: A Global Analysis. Frontiers in Public Health, **8**, 347.

Liu, K., Chen, Y., Lin, R. and Han, K. (2020). Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *Journal of Infection*, **80**(**6**), e14-e18.

Liu, Y., Bi, L., Chen, Y., Wang, Y., Fleming, J., Yu, Y., Gu, Y., Liu,C., Fan, L., Wang,X. and Cheng, M. (2020). Active or latent tuberculosis increases susceptibility to COVID-19 and disease severity. *Medrxiv*. DOI: https://doi.org/10.1101/2020.03.10.20033795

Ma, C., Su, S., Wang, J., Wei, L., Du, L. and Jiang, S. (2020). From SARS-CoV to SARS-CoV-2: safety and broad-spectrum are important for coronavirus vaccine development. *Microbes and Infection*, **22**(**6-7**), 245–253. DOI: https://doi.org/10.1016/j.micinf.2020.05.004

Maciel, E. L. N., Gonçalves Júnior, E. and Dalcolmo, M. M. P. (2020). Tuberculosis and coronavirus: what do we know? *Epidemiologia e Serviços de Saúde*, **29**, e2020128.

Menard, S. (Ed.). (2007). *Handbook of Longitudinal Research*: *Design, Measurement, and Analysis*. Elsevier. (ISBN: 978-0-12-370481-8).

Niazi, A., Kifayat, S., Javed, N. and Khan, M. S. (2020). Assessing the causes for a relatively lower caseload of Covid-19 in South Asia. DOI: https://doi.org/10.31219/osf.io/sf59y

Onder, G., Rezza, G. and Brusaferro, S. (2020). Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *Jama*, **323**(**18**), 1775-1776.

Rose, C. E., Martin, S. W., Wannemuehler, K. A. and Plikaytis, B. D. (2006). On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. Journal of biopharmaceutical statistics, **16**(**4**), 463-481.

Roser, M., Ritchie, H., Ortiz-Ospina, E. and Hasell, J. (2020). Coronavirus pandemic (COVID-19). Our world in data. Website URL: https://ourworldindata.org/coronavirus?fbclid=IwAR10byAszhXJ7hAN6VITOfYv8llw5BHhWfa-BsFKW2D_3TqWvg9uv7FnoXY

SAARC Tuberculosis and HIV/AIDS Centre (STAC) (2018). SAARC Epidemiological Response On Tuberculosis, SAARC Tuberculosis & HIV/AIDS Centre (STAC), (www.saarctb.org)

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S. and Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*, **62**, p.102418.

Schellekens, P. and Sourrouille, D. M. (2020). Covid-19 mortality in rich and poor countries: a tale of two pandemics? . *World Bank Policy Research Working Paper*, (9260). Available at SSRN: https://ssrn.com/abstract=3614141

Shohan, M.U.S., Alam, A.R.U., Rakhi, N.N., Kabir, M., Siam, M.K.S., Hasan, M.M., Sheriff, R., Karim, M.A., Rahman, A., Moni, M.A. and Arif, M. (2020). Onset, Transmission, Impact, and Management of COVID-19 Epidemic at Early Stage in SAARC Countries. Authorea Preprints. DOI: 10.22541/au.159775079.90952648

Sorci, G., Faivre, B. and Morand, S. (2020). Explaining among-country variation in COVID-19 case fatality rate. *Scientific Reports*, **10**(**1**), 1-11.

Sultana, F. and Reza, H. M. (2020). Are SAARC countries prepared to combat COVID-19 to save young, working-age population? *AIMS Public Health*, **7**(**3**), 440.

WHO (2020). WHO Coronavirus Disease (COVID-19) Dashboard [Online]. Website URL: https://covid19.who.int/?gclid=CjwKCAjwnef6BRAgEiwAgv8mQezLbX7QG5rMuuWdkPVAjZlngItN69DBphoPXo4L0mn17SlE0NUqChoClxMQAvD_BwE

World Bank. (2019). Age dependency ratio (% of working-age population). Website URL: https://data.worldbank.org/indicator/sp.pop.dpnd

World Health Organization. (2020). Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations: scientific brief, 27 March 2020 (No. WHO/2019-nCoV/Sci_Brief/Transmission_modes/2020.1). World Health Organization.

Yang, M. C., Hung, P. P., Wu, Y. K., Peng, M. Y., Chao, Y. C. and Su, W. L. (2020). A three-generation family cluster with COVID-19 infection: should quarantine be prolonged?. Public Health, 185, 31-33.

Yang, X., Yu, Y., Xu, J., Shu, H., Xia, J., Liu, H., Wu, Y., Zang, L., Yu, Z., Fang, M., Yu, T., Wang, Y., Pan, S., Zou, X., Yuan, S. and Shang, Y. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, ***8*(5)**, 475-481.

Zaidi, A. K., Awasthi, S. and de Silva, H. J. (2004). Burden of infectious diseases in South Asia. *BMJ*, **328(7443)**, 811-815.

Zheng, Y. Y., Ma, Y. T., Zhang, J. Y. and Xie, X. (2020). COVID-19 and the cardiovascular system. *Nature Reviews Cardiology*, **17**(**5**), 259-260.

Appendix

Table A.1: Socioeconomic and demographic factors of SAARC nations

Factors	Afghanistan	Bangladesh	Bhutan	India	Maldives	Nepal	Pakistan	Sri Lanka
Population	38928341	164689383	771612	1380004385	540542	29136808	220892331	21413250
Population density	54.42	1265.04	21.19	450.42	1454.43	204.43	255.57	341.96
GDP per capita	1803.99	3523.98	8708.60	6426.67	15183.62	2442.80	5034.71	11669.08
Cardiovascular death rate	597.03	298.00	217.07	282.28	164.91	260.80	423.03	197.09
Diabetes prevalence	9.59	8.38	9.75	10.39	9.19	7.26	8.35	10.68
Handwashing facilities	37.75	34.81	79.81	59.55	95.80	47.78	59.61	n.a
Life expectancy	64.83	72.59	71.78	69.66	78.92	70.78	67.27	76.98
Health expenditure (% of GDP)	10.20	2.40	3.50	3.70	10.60	6.30	2.80	3.90
Employed persons below poverty line	40.10	9.20	1.30	10.70	1.70	6.10	2.30	0.30
Hospital beds (per 10,000)	3.90	7.95	17.40	5.30	43.00	3.00	6.30	41.50
GHS	32.30	35.00	40.30	46.50	33.80	35.10	35.50	33.90
TB prevalence	189	221	149	199	33	151	265	64
Age dependency ratio	84	49	47	50	31	57	66	53

Source: Roser et al. (2020), Asian Development Bank (ADB) (2020), GHS Index Project Team (2019)

Table A.2: COVID-19 situation as on 31st January 2021

Country	Confirmed	Deaths	Recovered	Active
Afghanistan	55023	2400	47679	4944
Bangladesh	535139	8127	479744	47268
Bhutan	859	1	814	44
India	10757610	154392	10434983	168235
Maldives	15841	52	14139	1650
Nepal	270959	2029	266336	2594
Pakistan	546428	11683	501252	33493
Sri Lanka	64157	316	57159	6682

Source: Our World in Data (Roser *et al.*(2020))

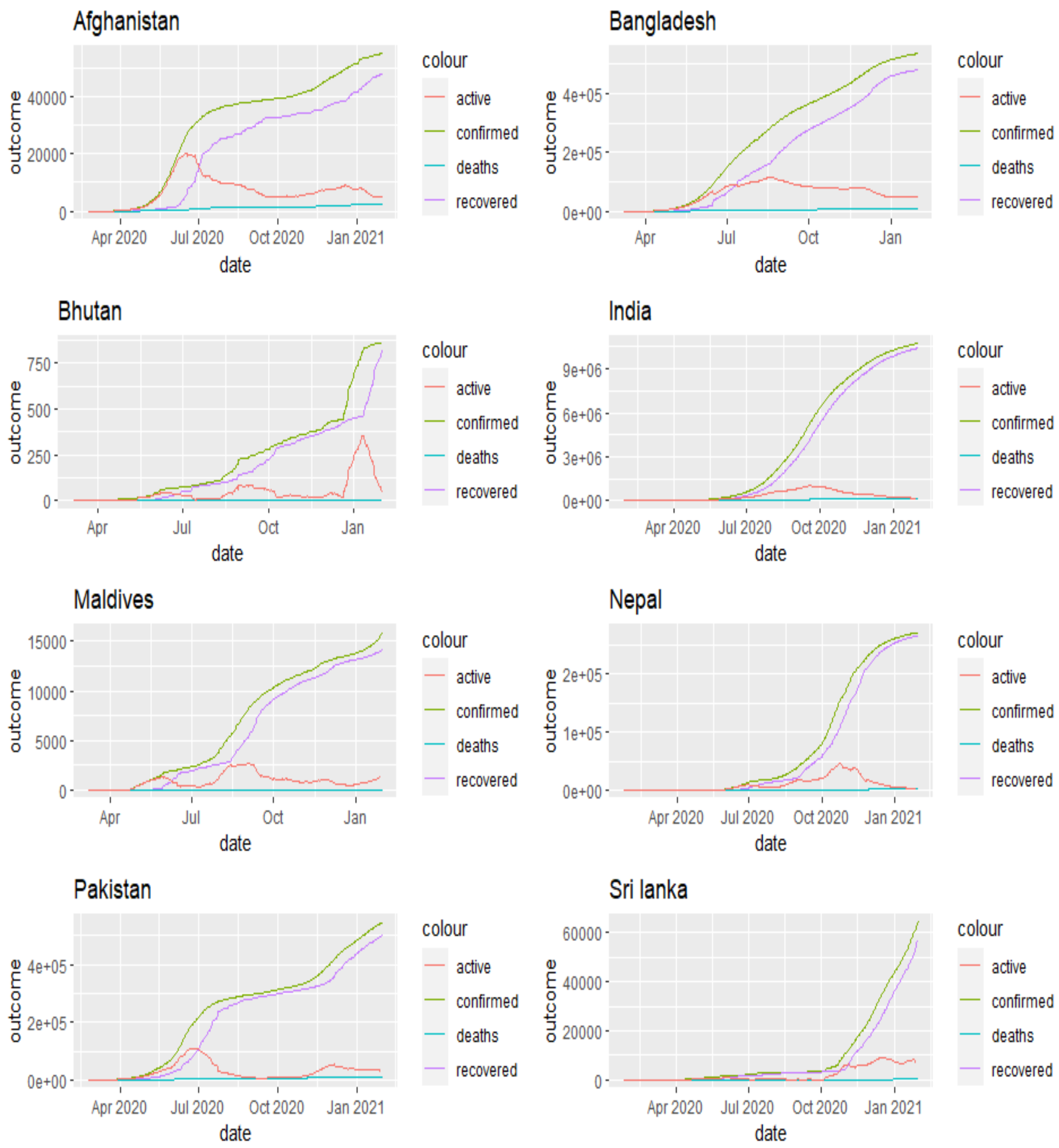

Figure A.1: COVID-19 related outcomes in SAARC nations

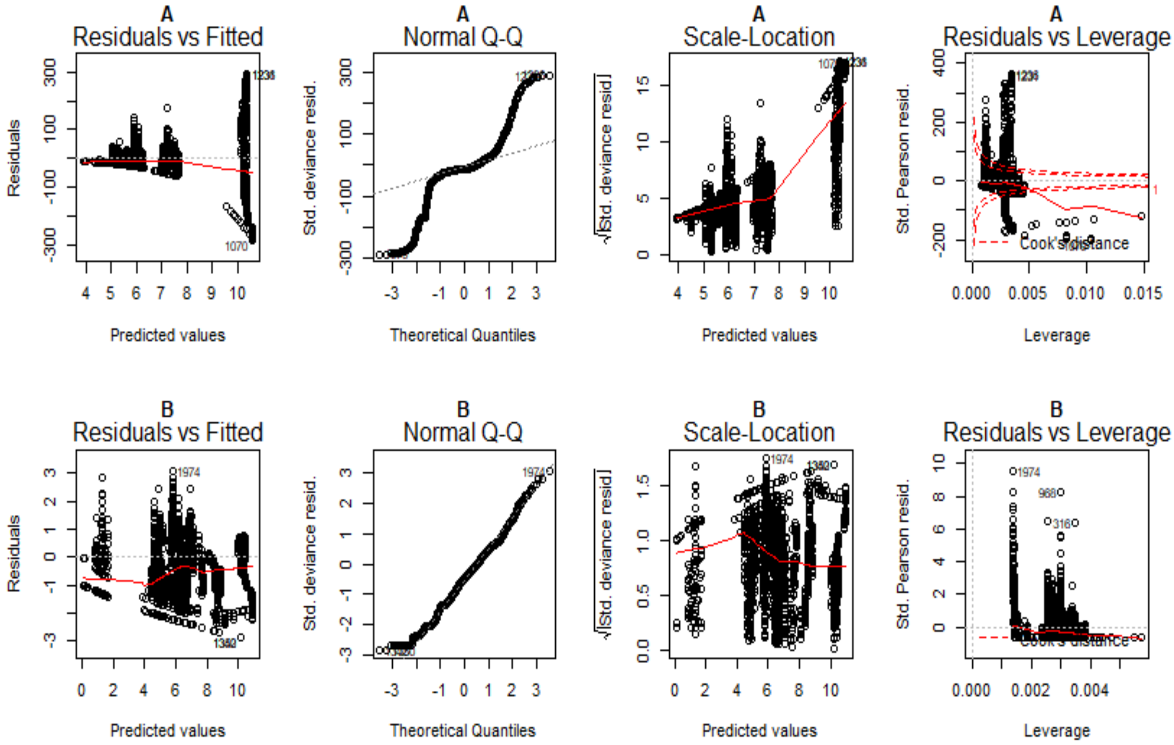

Figure A.2: Poisson regression plot (A) and Negative Binomial regression plot (B) for 'Daily cases'

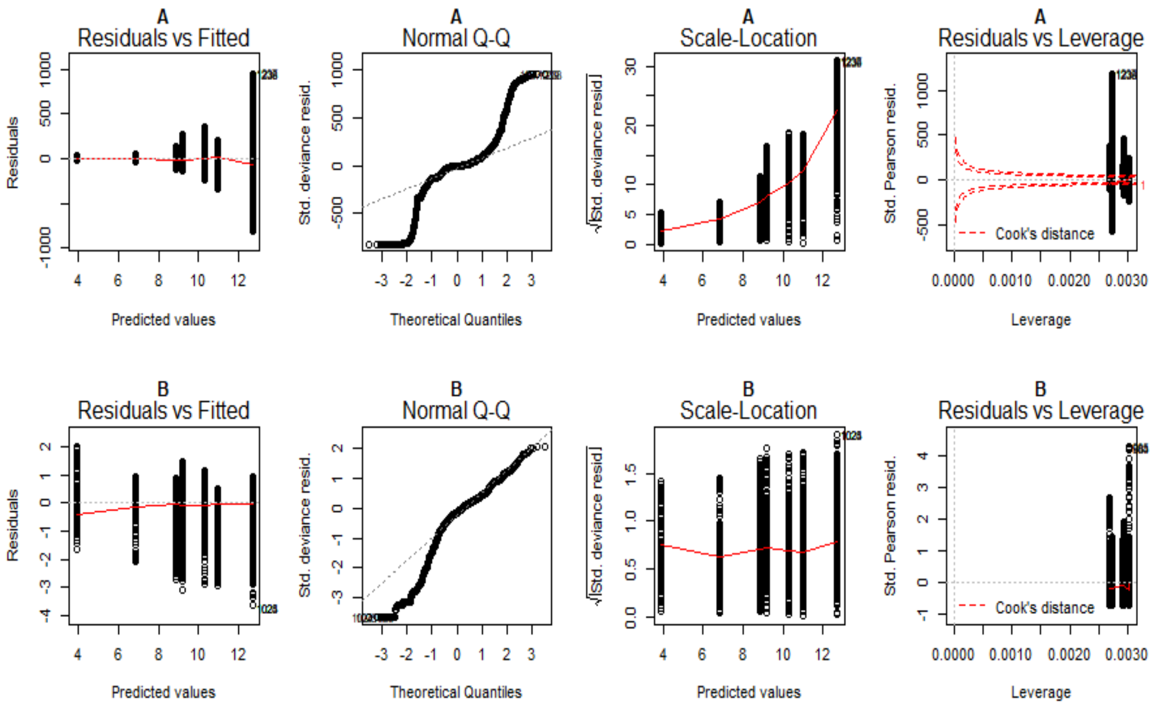

Figure A.3: Poisson regression plot (A) and Negative binomial regression plot (B) for 'Active cases'

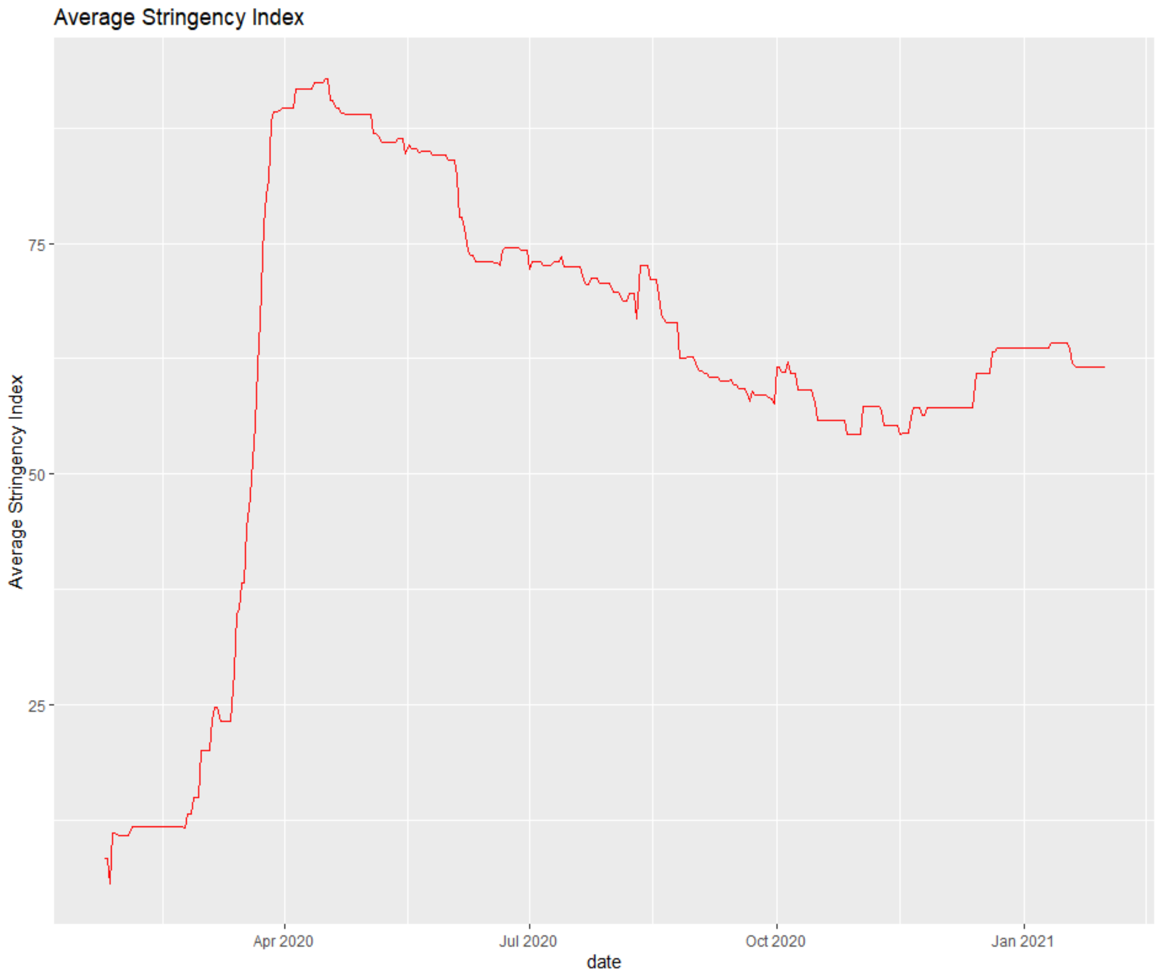

Figure A.4: Average stringency index per day among SAARC nations