BCG Vaccine Cross-protection from COVID-19: Statistical Study through Data Science

Dr.S.S. Sridhar[a], Rajarshi Chatterjee[b*], Saswat Pania[c]

[a]Professor, Department of CSE, SRMIST, Kattankulathur, India. E-mail: ssridhar@srmist.edu.in
[b*]B.Tech, Department of CSE, SRMIST, Kattankulathur, India. E-mail: rk6880@srmist.edu.in
[c]B.Tech, Department of CSE, SRMIST, Kattankulathur, India. E-mail: ss2501@srmist.edu.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: COVID-19, caused by SARS-COV-2 virus, has spread to almost all the countries in the world. Much like the SARS pandemic in 2002 and 2003, conventional control measures like travel restriction and patient isolation are being used to control spread of COVID-19. However, the morbidity rate of COVID-19 is still more than 1. Despite the high morbidity rate in all the countries COVID-19 has spread to, mortality rate due to COVID-19 varies greatly between different countries. Through a few recent epidemiological studies, it has been suggested that a negative association between national BCG vaccination policy and prevalence and mortality of COVID-19 exists. It has been observed that countries with BCG in their national immunisation programs have a lower mortality rate than countries without BCG in their national immunisation programs. We survey public datasets and repositories from websites like BCGAtlas, WHO, John Hopkins GitHub and World Bank, to understand and predict, with the power of data science, the implications of a plausible BCG cross-protection from severe COVID-19. We study COVID-19 related statistics to understand the spread of the disease, compare it with other countries to understand the role of cross-protective vaccines like BCG.

Keywords: Cross-protection, BCG Vaccine Policy, COVID-19, Coronavirus.

1. Introduction

The current pandemic of COVID-19 that initially originated in China in December 2019 had quickly spread over to other continents like Europe, America and other parts of globe affecting many countries. The severity of COVID 19 infection and its resultant mortality was not similar in all the countries. Rather it varied amongst them significantly. Unfortunately, we don't have very good universal tool for comparing these data. This is due to the fact that the socioeconomic status, demography, the measures available geographically to reduce physical contact and promote digital communication, type of tests used, sensitivity of tests and national guidelines for management of COVID–19 were hugely different amongst the countries. [1,2,3]

One striking observation was that the severity of COVID–19 and its mortality was lesser in the countries who had long term national Bacillus Calmette-Guerin (BCG) vaccination policy than those who do not have BCG in their national vaccination policy or those who previously used to give BCG but have discontinued later. BCG is a live attenuated vaccine derived from a bovine strain of Mycobacterium bovis used widely across the world as a vaccine for Tuberculosis. Many countries like Japan, China etc., including our country, India, have a universal national BCG vaccination policy in newborns for long term. Due to the epidemiological reasons of low incidence of disease and variable efficacy of BCG vaccine that is currently used to control Tuberculosis, countries like Spain, France, and Switzerland, have discontinued it from their national vaccine policies. Countries like United States, Italy, and the Netherlands, have not yet adopted any universal policy to vaccinate their children with BCG. [4]

Like few other vaccines, BCG vaccination has been shown to produce positive "heterologous" or non-specific immune effects leading to improved response against other non-mycobacterial pathogens. One animal study with mice in 2009 showed that previously BCG vaccinated mice when inoculated with vaccine virus produced increased IFN-Y from CD4+ cells and did not develop the disease. This is popularly named as 'trained immunity'. Immune cells, following successful BCG vaccination, undergo metabolic and epigenetic changes expressing genetic regions encoding for pro-inflammatory cytokines more leading to more cytokine release like IFN -Y, IL-1B that play vital role in prevention of viral infection against which the vaccine is not intended to. [5,6,7]

There are numerous epidemiological evidences that BCG vaccination has broad non-specific protective effects against non-M. tuberculosis infection. In one study in Sweden, as early as in 1927, showed that mortality rate in BCG vaccinated children was three-fold less than in unvaccinated children and it was not related to decrease in tuberculosis incidence due to vaccination. It was suggested that the decrease of mortality in BCG vaccinated children was due to non-specific immunity.[8]

A study from Guinea-Bissau, showed that BCG vaccinated children suffered less from neonatal sepsis, respiratory infection, and fever than those who did not receive BCG. This lower incidence of respiratory infection was not found in children who received vaccines other than BCG proving that the infection lowering effect was due to BCG itself. [9]

Another long-term study from Spain showed that children vaccinated with BCG had less hospitalization due to respiratory illnesses not attributable to Tuberculosis proving that this broad protective effect of BCG can be long lasting. [10]

Presence of accumulated evidence of non-specific protective effect from viral infections of BCG and its long-term persistence due to trained immunity offer a rational biological basis for the potential protective effect of BCG vaccination from severe coronavirus disease 2019 (COVID-19). [11]

Given our current understanding of the BCG vaccine's non-specific immunotherapeutic mechanisms and by analysing current epidemiological data, this investigation aims to identify a possible correlation between the existence of universal BCG vaccine policies and the morbidity and mortality associated to COVID-19 infections all over the world.

2. Literature Survey

II.1 Papers Surveyed

In another study done by the researchers in Virginia Polytechnic Institute and Sate University suggests that BCG index, which is the degree of BCG vaccine deployment in a country, among European countries with similarity in social conditions found that COVID-19 mortality rates dropped by 10.4% for every 10% increase in BCG index. They also corroborated that the correlation between BCG in national immunization policy and a decrease in mortality due to COVID-19. [3]

A study done by researchers in NYIT, New York in March 28, 2020 found through epidemiological studies carried in the institution that the impact of COVID-19 was differed according to the country that was being studied. They proposed that this difference in impact of COVID-19 could be attributed to the BCG vaccination policy present in the country at the start of the pandemic. Countries were divided into three groups according to their income group and the COVID-19 data generated in these countries were compared, which showed a correlation between BCG vaccination policy present in these countries and their corresponding COVID-19 mortality. [4]

Scientists belonging to ICMR from India found that BCG vaccine gives innate immune memory among the elderly population aged between 60-80 years in COVID hotspots in India. They also suggest the existence of correlation between BCG vaccination and lowering of COVID-19 mortality rates in the country based on the 86 individuals the vaccine was initially tested on. They have started clinical trials sponsored by Tuberculosis Research Centre, India.[12]

The BCGAtlas website is used to store the BCG vaccination policy information about all the countries in the world along with the information of their economic condition, human development index and other related information about all the countries in the world. It contains the other BCG related information like the strain of BCG that is used in these countries and also the start of BCG deployment in countries and also about the cessation of BCG inoculation in countries where it is not in national immunization policy any more. [13]

Doctors from Royal Children’s Hospital and researchers from Murdoch Children’s Research Institute are collaborating in a clinical trial with two groups with randomized control method with 10078 healthcare workers to access the possible correlation of reduction in COVID-19 severity and the immunity gained from BCG vaccination.[14]

II.2 Inference from the Survey

All the papers involved in this study helped us in understanding the cross-protection from BCG vaccine against COVID-19 and other viral diseases. The papers also gave very important insights about the social conditions that may affect this correlation between BCG vaccine policy and COVID-19 mortality. The eleventh paper in this survey helped us understand the BCG vaccine coverage in each country, the status of BCG in their national immunization programs and the region these countries fall in. The papers helped us understand the current situation of BCG vaccine policy and also helped in understanding the social conditions that affect it.

3. Proposed Work

The proposed study in the paper can understood using the following UML diagram, in figure1. which shows the workflow of the whole study.

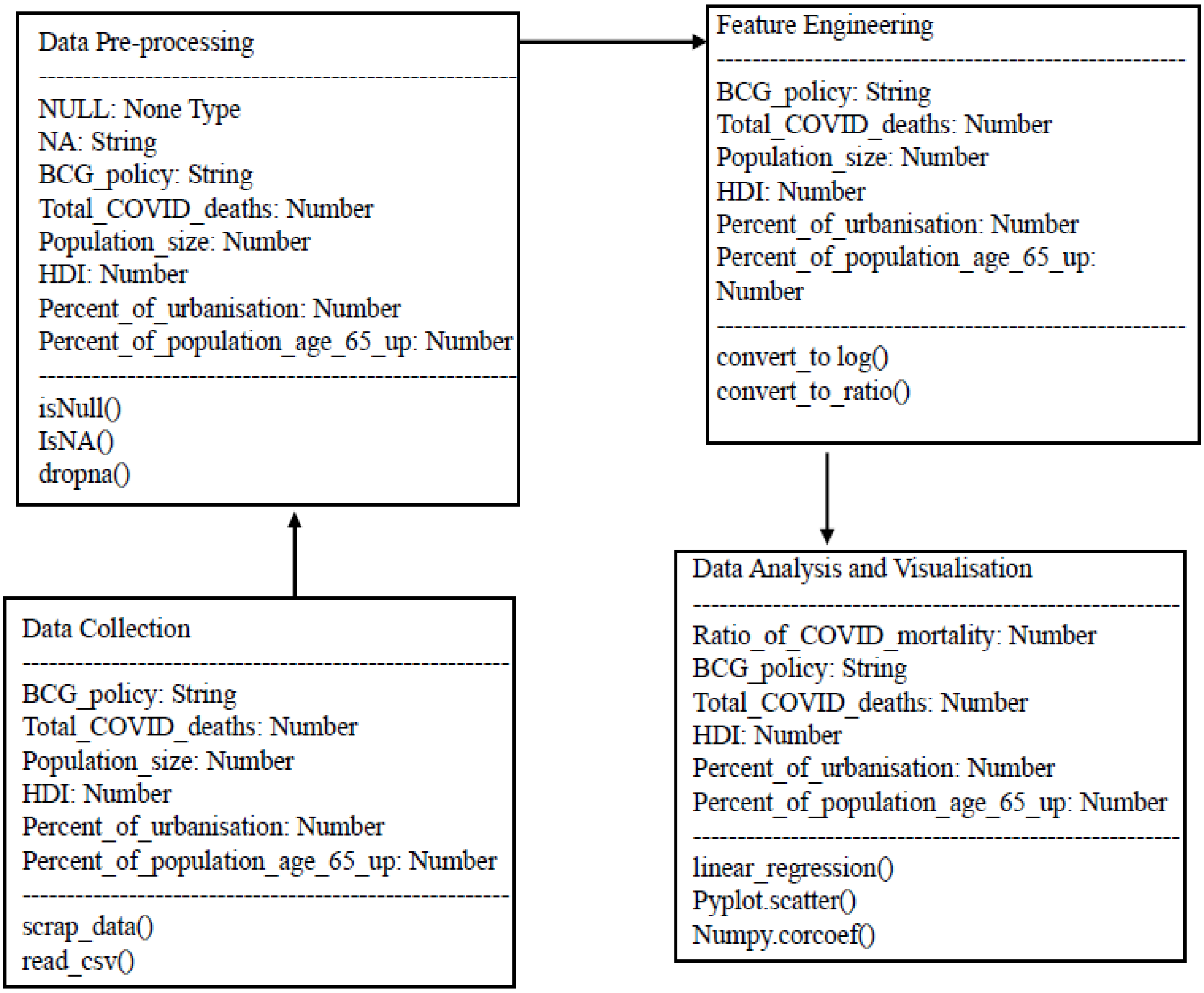

Figure 1. UML Diagram of Study Process

Abbreviations and Acronyms

The abbreviations and acronyms used in this paper are: BCG: Bacillus Calmette-Guerin

COVID-19: Corona Virus Disease 2019, HDI: Human Development Index

4. Implementation

This study on BCG vaccine policy and COVID-19 mortality was conducted in Python using Jupyter notebooks. The implementation can be understood with the help of the architecture diagram as follows.

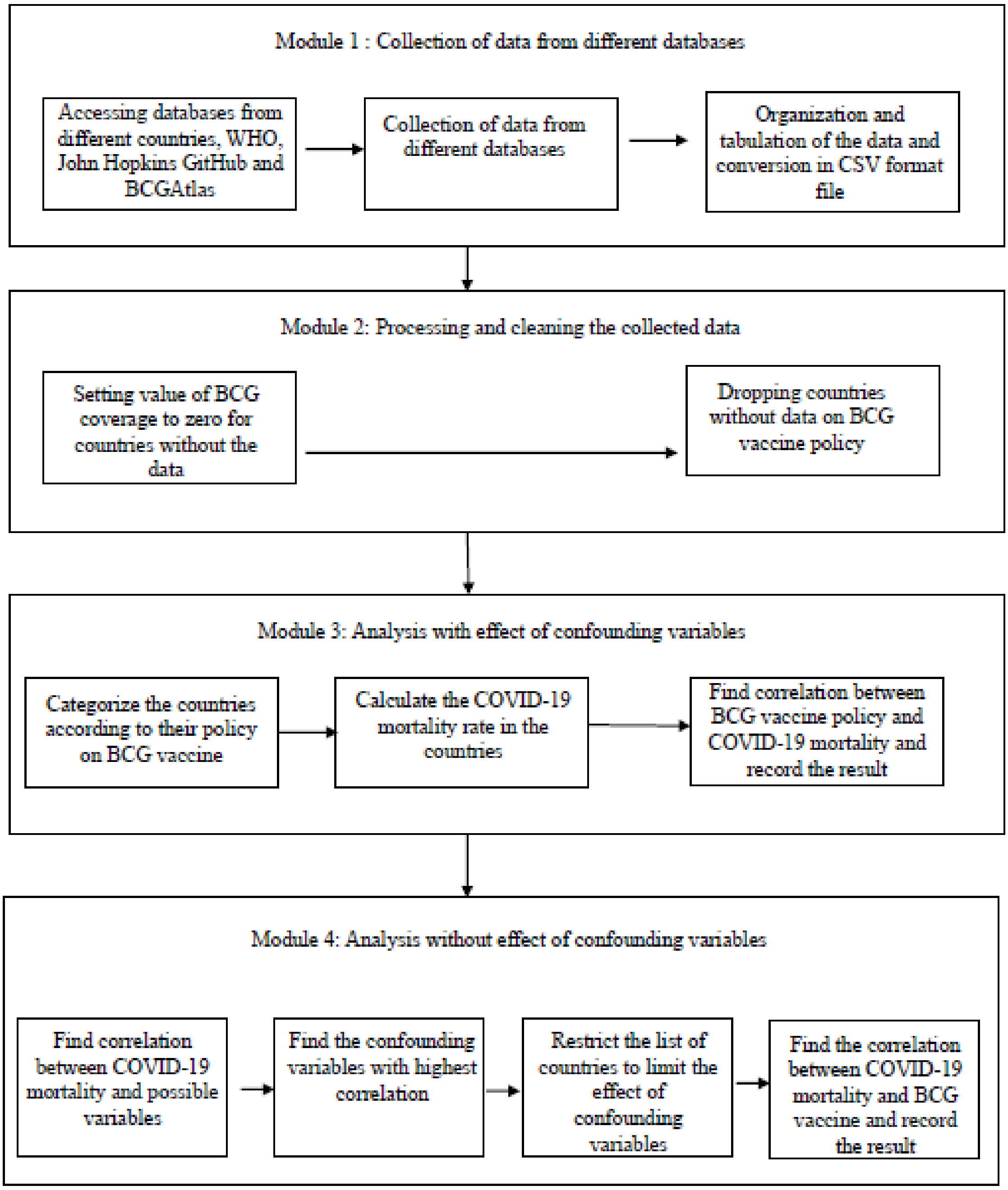

Figure 2. Architecture Diagram of the Study Process

IV.1 Module 1: Collection of data from different sources.

The data has been collected from BCG World Atlas, which is publicly available on the website, https://www.bcgatlas.org/, from WHO on https://covid19.who.in/, and from John Hopkins GitHub page and from World Bank. The data on COVID-19 cases and deaths was obtained on April 22, 2020. The data is then put together and tabulated into an excel file which was then converted into a CSV file and uploaded to Google Drive for easy access of all participants.

IV.2 Module 2: Processing and Cleaning of collected data.

The dataset collected had a lot of missing or NA values which we dealt with by removing them or assigning to the value of zero. The missing or NA values of BCG vaccine coverage, which was collected country-wise, to zero. [3] The BCG coverage values were in floating point format and missing or NA values belonged to countries where BCG vaccination drives were never conducted.

The missing or NA values for BCG vaccine policy belonged to countries never publicly stated a policy. So, these countries had to be removed from the study.

IV.3 Module 3: Analysis with effect of confounding variables.

In this part of the study, we divide the countries among different group according to their BCG vaccination policy – current, never, and interrupted. The countries which have BCG in their national immunization policy fall under current category. The countries which do not have BCG in their national immunization policy fall under never category. The countries which had BCG in their national immunization policy and later discontinued it fall under interrupted category.

COVID-19 mortality is calculated as number of total COVID-19 divided by the population of the country.

We used t-test to access the relationship between BCG vaccine policy and COVID-19 mortality among the different categories of BCG vaccine policy and find the correlation.

We do not take into consideration the effect of confounding variables in this part of the study.

IV.4 Module 4: Analysis without effect of confounding variables.

In this part of the study, we take into account the effect of confounding variables that may affect this correlation between BCG vaccine and COVID-19 mortality. These confounding variables are social conditions that were present in the country before the pandemic had started.

We filtered the list of countries in this study to consist of countries which have more than zero deaths due to COVID-19 and have publicly available data on the present status of BCG in their national vaccine policy and BCG vaccination coverage.

We didn’t consider climate conditions of a country as a confounding variable as there is no sufficient evidence on the temperature dependency on the spread or death due to COVID-19.[15]

We considered an individual’s ability to access healthcare facilities, access to proper education and income as the potential confounding variables. The confounding variables considered were Human Development Index or HDI [16], population density [17], percentage of urbanized region [18] and the percentage of people who are above 65 in age [19].

We then calculated the correlation of these confounding variables with COVID-19 mortality in all countries. We evaluated the associations and then restricted our study to those confounding variables with the highest value for coefficient of correlation. After, the finding the appropriate confounding variables we restricted our list of countries such that the confounding variables have the least effect on BCG vaccine policy and COVID-19 mortality. The relationship between COVID-19 and BCG vaccine policy was studied again and the results were recorded.

5. Results and Discussion

V.1 Analysis with Effect of Confounding Variables

We performed a t-test between the COVID-19 mortality rates of countries with different BCG vaccine policy – current, never, interrupted.

The t-test for current vs never is -5.390 and the p-value is 2.635e-07 which lower than the significant point of 0.05. So, the COVID-19 mortality for countries with current and never values for BCG policy are significantly different.

The t-test for current vs interrupted is -2.625 and the p-value is 0.009 which lower than the significant point of 0.05. So, the COVID-19 mortality for countries with current and never values for BCG policy are significantly different.

V.2 Analysis without Effect of Confounding Variables

We found that total number of COVID-19 deaths in a country has a positive relation with three variables – HDI, the percentage of population with age above or equal to 65, and the percentage of urbanized region. COVID-19 mortality has the highest coefficient of correlation value, of 0.78, with HDI.

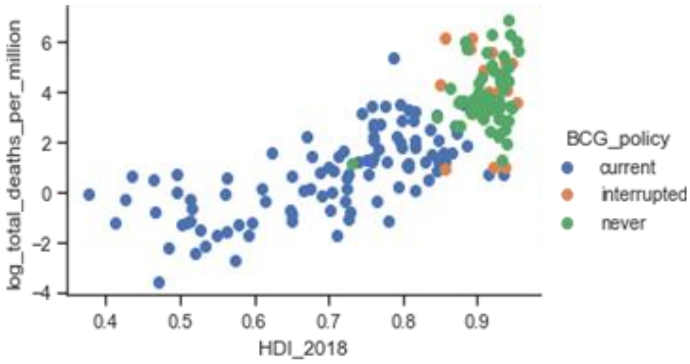

Figure 3A. Scatter Plot between HDI and Total COVID Deaths Per Million (log)

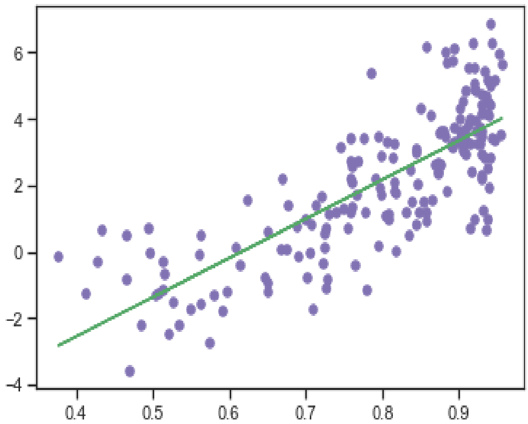

Figure 3B. Linear Regression Line, on Scatter Plot of fig. 3A, of y = (-7.271) + (11.796)x

The figure 3A and 3B, shows the positive association of HDI and rate of COVID-19 mortality. This shows that COVID-19 mortality is not lower with better access to healthcare facilities, or better education.

COVID-19 mortality has a coefficient of correlation value of 0.72 with the percentage of population of age equal to or above 65 years.

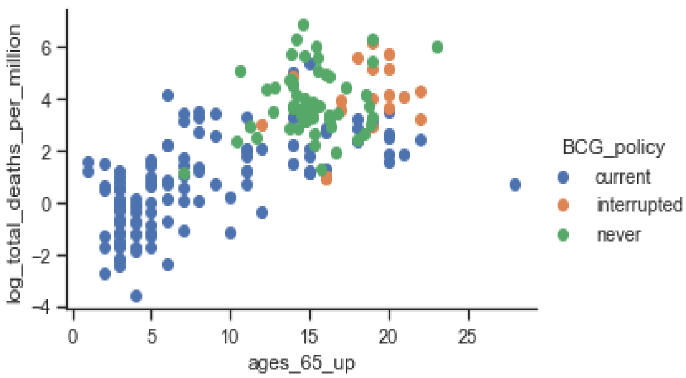

Figure 4A. Scatter Plot between Population of Age 65 Up and Total COVID Deaths Per Million (log)

The figure 4A and 4B, show the positive association of percentage of population of age equal to or above 65 and rate of COVID-19 mortality. This shows that people older than 65 will be more affected by COVID-19 than others.

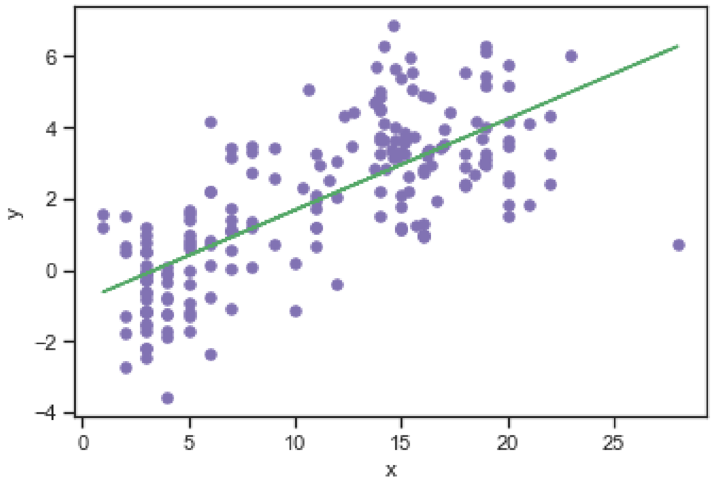

Figure 4B. Linear Regression Line, on Scatter Plot of fig. 4A, of y = (-0.875) + (0.255)x

COVID-19 mortality has a coefficient of correlation value of 0.50 with the percentage of urbanized region.

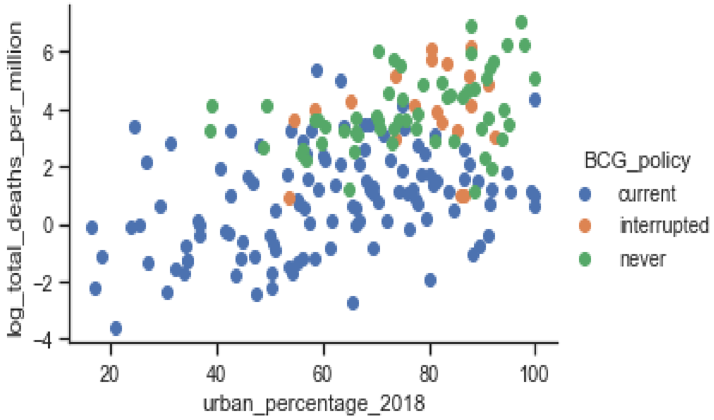

Figure 5A. Scatter Plot between the Percentage of Urbanized Area and Total COVID Deaths Per Million (log)

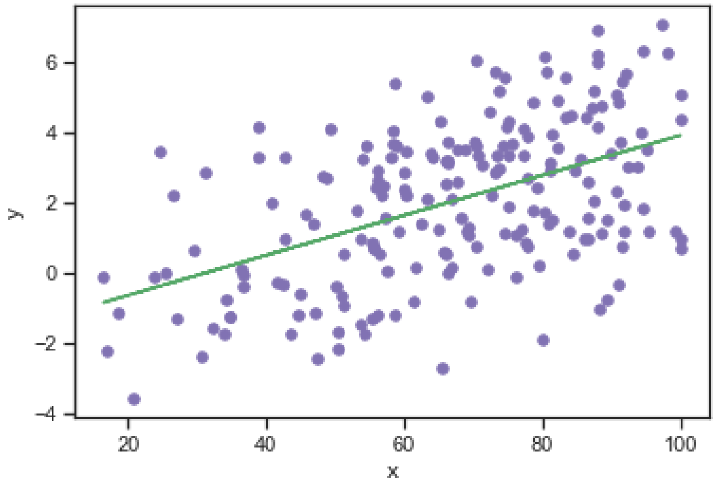

Figure 5B. Linear Regression Line, on Scatter Plot of fig. 5A, of y = (-1.781) + (0.057)x

The figure 5A and 5B, show the positive association of percentage of urbanized region and COVID-19 mortality. This shows that countries with higher percentage of rural or uninhabited region would have lower COVID-19 mortality rate.

Also, COVID-19 mortality has a negative association with BCG coverage with a coefficient of correlation value of -0.02.

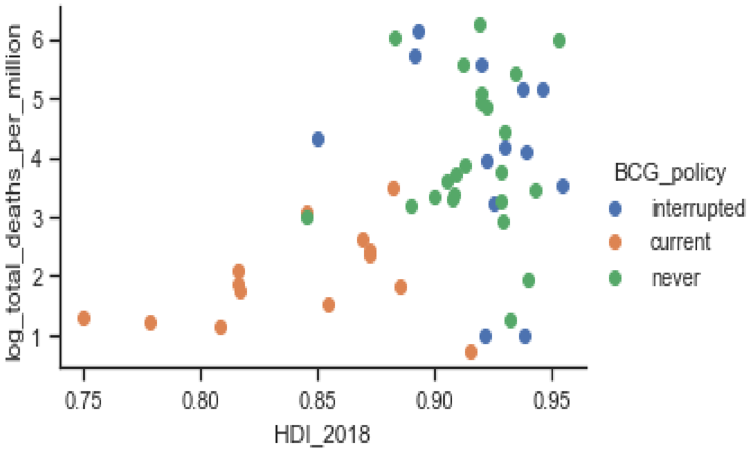

Figure 6A. Scatter Plot between HDI and Total COVID Deaths Per Million (log) for 49 Socially Similar Countries

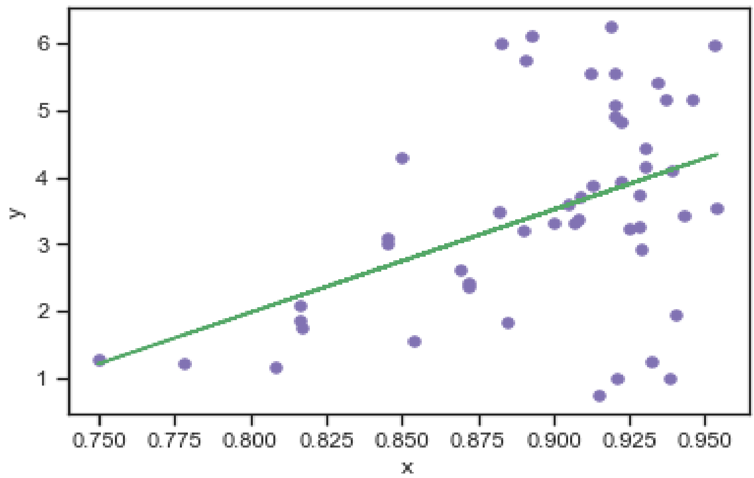

Figure 6B. Linear Regression Line, on Scatter Plot of fig. 6A, of y = (-10.305) + (15.355)x for 49 Socially Similar Countries

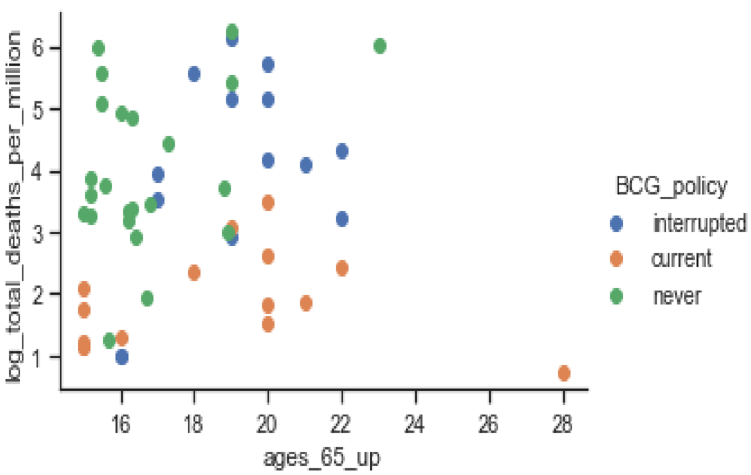

Figure 7A. Scatter Plot between Population of Age 65 Up and Total COVID Deaths Per Million (log) for 49 Socially Similar countries

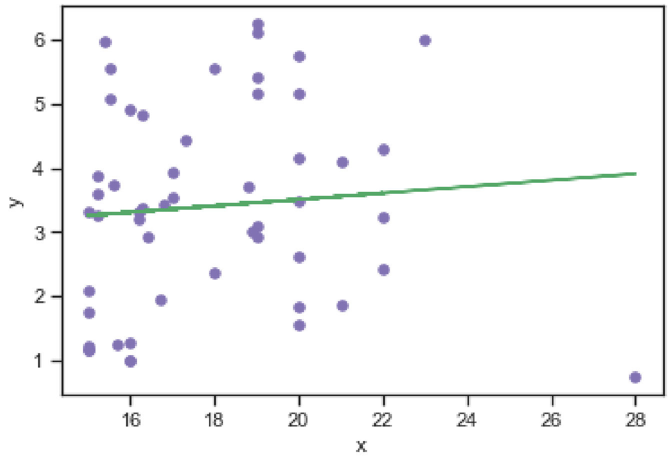

Figure 7B. Linear Regression Line, on Scatter Plot of fig. 7A, of y = (2.508) + (0.050)x for 49 Socially Similar Countries

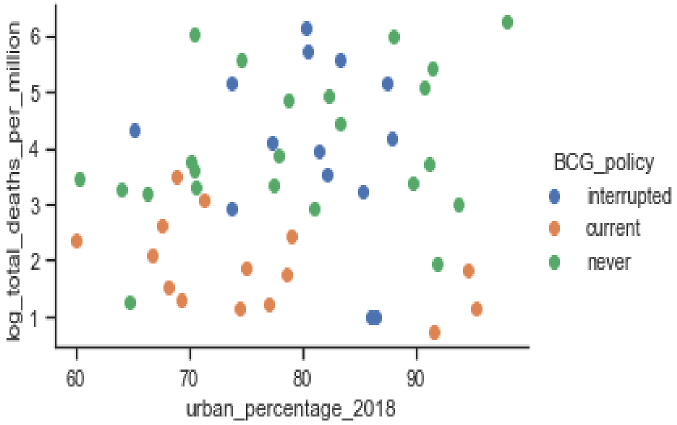

Figure 8A. Scatter Plot between the Percentage of Urbanized Area and Total COVID Deaths Per Million (log) for 49 Socially Similar Countries

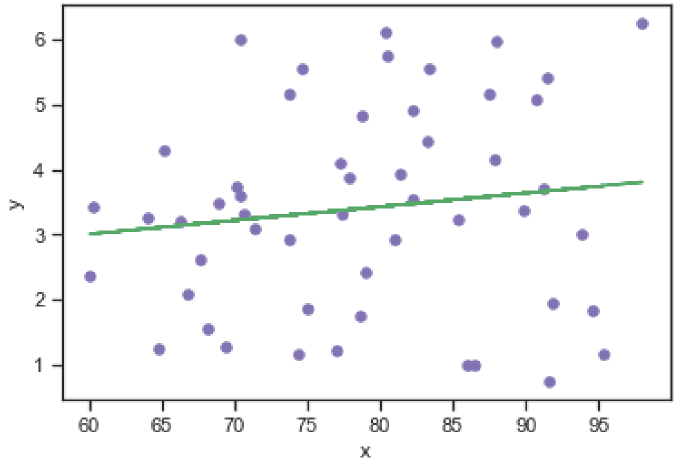

Figure 8B. Linear Regression Line, on Scatter Plot of fig. 8A, of y = (1.745) + (0.021)x for 49 Socially Similar Countries

To limit the effect of these confounding variables we restricted the list of 49 socially similar countries. The countries in this list had HDI greater than 0.7, percentage of population with age equal to or more than 65 of greater than or equal to 15% and percentage of urbanized region greater than 60%. This reduced the effect of the confounding variables which also reduced the values of coefficient of correlation between COVID-19 mortality and these variables – HDI (coeff. of corr. = 0.46) in figure 6A and 6B, percentage of population with age greater than or equal to 65(coeff. of corr. = 0.09) in figure 7A and 7B, and percentage of urbanized region (coeff. of corr.

= 0.13) in figure 8A and 8B. However, the p-value between COVID-19 mortality and these variables is still significant as their value is lower than 0.05.

6. Conclusions

From this study, we learn that there is positive correlation between BCG vaccine policy and COVID-19 mortality rates. Countries with BCG in their national immunization policy consistently show lower mortality rates due to COVID-19. However, the causality between BCG vaccine policy and COVID-19 mortality cannot be affirmed. Only randomized clinical trials, as organized in the Netherlands [20], in India [21], and in Australia [22] can help to confirm the causality between BCG vaccine and COVID-19. However, the results of the data analysis can help in making a case regarding vaccinating BCG to keep COVID-19 mortality in control.

References

1. Arulprakash M., Dynamic Evolutionary Information Diffusion Over Mobile Social Networks. http://www.arpnjournals.org/jeas/research_papers/rp_2016/jeas_1016_5098.pdf
2. Arulprakash M., Kamal A, Manisha A, QR-code scanner based vehicle sharing. http://www.arpnjournals.org/jeas/research_papers/rp_2018/jeas_0518_708.pdf
3. Escobara LE, Molina-Cruz A, and Barillas-Mury C, BCG vaccine protection from severe corona virus disease 2019 (COVID-19). 17720–17726 | *Proceedings of National Academy of Science.* 2020: 117 (30) www.pnas.org/cgi/doi/10.1073/pnas.2008410117
4. Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. Miller A, Reandelar MJ, Fasciglione K, Roumenova V, Li Y, Otazu GH. medRxiv preprint March 28, (2020). doi: https://doi.org/10.1101/2020.03.24.20042937
5. Mathurin, K. S., Martens, G. W., Kornfeld, H. & Welsh, R. M. CD4 T-Cell-Mediated C Heterologous Immunity between Mycobacteria and Poxviruses. *J. Virol., 83*, 3528–3539(2009)
6. Netea, M. G. et al. Trained immunity: A program of innate immune memory in health and disease. *Science* (80). 352, aaf1098–aaf1098 (2016)
7. Kleinnijenhuis, J. et al. Long-lasting effects of bcg vaccination on both heterologous th1/th17 responses and innate trained immunity. *J. Innate Immun., 6*, 152–158 (2014)
8. C. Näslund, Results of BCG vaccination experiments carried out in Norrbotten (Sweden) (September 1927 – December 1931). Preventative Tuberculosis Vaccination, Reports and Documents, (Institut Pasteur, Paris, France, 1932)
9. Kristensen, I., Aaby, P. & Jensen, H. Routine vaccinations and child survival: Follow up study in Guinea-Bissau, West Africa. *Br. Med. J.*, 321, 1435–1439 (2000).
10. M. J. de Castro, J. Pardo-Seco, F. Martinón-Torres, Nonspecific (heterologous) protection of neonatal BCG vaccination against hospitalization due to respiratory infection and sepsis. *Clin. Infect. Dis.*, 60, 1611–1619 (2015)
11. L. A. J. O’Neill, M. G. Netea, BCG-induced trained immunity: Can it offer protection against COVID-19? *Nat. Rev. Immunol.*, 20, 335–337 (2020).
12. BCG vaccination induces enhanced frequencies of memory T and B cells and dendritic cell subsets in elderly individuals. Kumar PN, Padmapriyadarsini C, Rajamanickam A, Bhavani PK, Nancy A, Jayadeepa B, Selvaraj N, Kumar D, Renji RM, Venkataramani V, Tripathy S, Babu S. doi: https://doi.org/10.1101/2020.10.22.20217471
13. The BCG World Atlas: A Database of Global BCG Vaccination Policies and Practices. Zwerling A, Behr MA, Verma A, Brewer TF, Menzies D, Pai M doi: https://doi.org/10.1371/journal.pmed.1001012
14. Tuberculosis and COVID-19: Lessons from the Past Viral Outbreaks and Possible Future Outcomes. Crisan-Dabija R, Grigorescu C, Pavel C, Artene B, Popa IV, Cernomaz A, Burlacu A. doi: https://doi.org/10.1155/2020/1401053
15. No Evidence for Temperature-Dependence of the COVID-19 epidemic. Jamil T, Alam I, Gojobori T, Duarte CM. doi: https://doi.org/10.3389/fpubh.2020.00436
16. United Nations Development Program, Human development data (1990-2018). Human Development Reports (2020). http://hdr.undp.org/en/data.
17. World Bank, Population density (people per sq. km of land area). https://data.worldbank.org/indicator/EN.POP.DNST
18. World Bank, Urban population (% of total population). https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS
19. World Bank, World Development Indicators 2.1: Population dynamics. http://wdi.worldbank.org/table/2.1

20. US National Library of Medicine, Reducing health care workers absenteeism in covid-19 pandemic through BCG vaccine (BCG-CORONA). ClinicalTrials.gov (2020).
21. https://clinicaltrials.gov/ct2/show/NCT04328441
22. BCG Vaccine in Reducing Morbidity and Mortality in Elderly Individuals in COVID-19 Hotspots. Clinical Trials.gov (2020). https://clinicaltrials.gov/ct2/show/NCT04475302
23. US National Library of Medicine, BCG vaccination to protect healthcare workers against COVID-19 (BRACE). Clinical Trials.gov (2020). https://clinicaltrials.gov/ct2/show/NCT04327206