medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Association of BCG vaccination policy with prevalence and mortality of COVID-19

Giovanni Sala[1] and Tsuyoshi Miyakawa[1]

[1]Division of Systems Medical Science,

Institute for Comprehensive Medical Science, Fujita Health University

Correspondence: Tsuyoshi Miyakawa email: miyakawa(at)fujita-hu.ac.jp

Abstract

There is some evidence that tuberculosis vaccine bacillus Calmette-Guérin (BCG) has non-specific beneficial effects against non-related infections. Here, we examined the possible association between BCG vaccination with prevalence and mortality by COVID-19 by using publicly available data of COVID-19 in 199 countries/regions and the BCG World Atlas. By using linear regression modeling, we found that the number of total cases and deaths per one million population were significantly associated with the country's policy concerning BCG vaccine administration. The amount of variance in cases and deaths explained by BCG vaccination policy ranged between 12.5% and 38%. Importantly, this effect remained significant after controlling for the country's life expectancy and the average temperature in February and March 2020, which themselves are significantly correlated with the cases and deaths indices, respectively. By contrast, the ratio between deaths and cases was weakly affected. This latter outcome suggested that BCG vaccination may have hindered the overall spread of the virus or progression of the disease rather than reducing mortality rates (i.e., deaths/cases ratio). Finally, by roughly dividing countries into three categories showing high, middle, or low growth rate of the cases, we found a highly significant difference between the slope categories among the BCG groups, suggesting that the time since the onset of the spread of the virus was not a major confounding factor. While this retrospective epidemiological study potentially suffers from a number of unknown confounding factors, these associations support the idea that BCG vaccination may provide protection against SARS-CoV-2, which, together with its proven safety, encourages consideration of further detailed epidemiological studies, large-scale clinical trials on the efficacy of this vaccine on COVID-19, and/or re-introduction of BCG vaccination practice in the countries which are currently devoid of the practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Introduction

The recent outbreak of coronavirus disease 2019 (COVID-19) caused by the 2019 novel coronavirus (SARS-CoV-2) is one of the most serious threats to the human race in this millennium. There appears to be a large variability among countries in the risk of infection and mortality by this virus[1]. For example, Japan, which has close physical and social distance from China, is one of the first countries that recorded the first case in the earliest phase of this pandemic. As of March 30th, while many governments have 'locked-down' numerous cities and even the entire country, the Japanese government still keeps applying mild measures to mitigate the spread of the virus. The citizens have simply been 'advised' to stay home. Nevertheless, most people in Japan, with the notable exception of Tokyo metropolitan, are currently working as usual. Surprisingly, despite such a liberal measure taken and relatively low number of PCR tests conducted, Japan is ranked at 117th and 69th among 199 countries and regions[1], as of March 28th, regarding the number of cases per 1 million population and the number of deaths per 1 million population, respectively. Also, there seem to be visible differences in COVID-19 prevalence and mortality between countries in Western Europe and Eastern Europe. There are many potential geographical, social, and biological factors, such as temperature, humidity, life expectancy, average income, social norms, and ethnical genetic background, that could potentially explain such variabilities among countries, but it remained fully explored. In case they are successfully identified, such information could potentially contribute to the mitigation of the spread of the virus.

Some vaccines (e.g., measles and Bacillus Calmette–Guérin (BCG) vaccines) are associated with a lower risk of illness and death from other disorders[2–4]. A systematic review of epidemiological studies provided evidence for the non-specific beneficial effects of BCG vaccine on all-cause mortality[5,6]. Induction of cytokines associated with trained immunity is proposed to underly such protection against non-related viral infections[7]. Researchers in four countries will soon start a clinical trial of BCG vaccine on this disease, based on the proposed beneficial effect of BCG[4]. Interestingly, a blogger noticed that there might be a correlation between the administration of BCG vaccine in each country and the extent to which the spread of COVID-19 occurs[8]. In this study, we evaluated the hypothesis that BCG vaccination has some protective effect against COVID-19 by utilizing publicly available datasets through linear regression modeling. Among the potentially confounding factors, we assessed two factors, life expectancy and the average

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

temperature of the countries. We also tried to evaluate the potential effect of BCG vaccination policy on the growth rate of COVID-19, controlled by the onset of its propagation in each country.

Methods and Results

The data on COVID-19 and life expectancy were retrieved on March the 28th from "COVID-19 CORONAVIRUS PANDEMIC" page[1] and "Life Expectancy of the World Population" page[9] in worldmeter website. The average temperature data were retrieved from "List of cities by average temperature" page[10] in Wikipedia. Average temperature of each country was obtained by averaging those of the cities in each country. The raw data on which our analyses are based is attached as Supplementary Table 1. Also, the STROBE Statement for cohort studies is in the supplementary materials[11].

Three sets of linear regressions were run. The dependent variables (DVs) were the rate of total cases per one million population (Model 1), the rate of death per one million population (Model 2), and the ratio between these two variables (Model 3). The country's BCG policy was the independent variable (IV) of interest. In accordance with the BCG world atlas[12], this variable had three levels: Group A, the country currently administers the BCG; Group B, the country used to administer the BCG and currently does not; and Group C, the country has never administered the vaccine. The country's life expectancy and mean temperature (°C) in February and March 2020 were added as control variables in all the models in order to rule out potentially confounding effects. In fact, infection rates and mortality rates may have been influenced by the subject's age (i.e., the older, the more vulnerable) and climate conditions (i.e., the higher the temperature, the less contagious the virus). Finally, the Order Quantile (ORQ) normalization transformation was used to standardize the continuous variables. Beyond meeting the statistical assumption of linear modeling (normalization), this procedure is also useful to mitigate potentially confounding effects of outliers.

This study included 136 countries or regions for which complete data about total cases per one million population, life expectancy, and temperature were available (Model 1). The data regarding total deaths per one million population and deaths/cases ratio were available for 91 countries (Models 2 and 3). All the analyses were run with the R software[13], and the graphs were produced by using the package ggplot2[14]. The ORQ transformation was performed with the

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

bestNormalize R package[15]. Finally, the rsq R package was employed to calculate partial R^2 statistics (i.e., unique variances)[16]. Descriptive statistics of each variable (e.g., means, medians, and quantiles) are retrievable from the R codes provided online.

Regarding total cases per one million population, the linear model showed a statistically significant effect for both Group B ($b = 0.6483$, $p = .0002$) and Group C ($b = 0.8666$, $p = .0025$; as compared to Group A). Temperature was barely significant (Figure 1a; $b = –0.1232$, $p = .0424$). As anticipated, life expectancy was significant (Figure 1c; $b = 0.5544$, $p = 6.33\times10^{-14}$). The model's adjusted R^2 was .6409 (i.e., approximatively 64% of the observed variance was explained). The amount of unique variance – i.e., the variance explained by a certain variable by controlling for all the other variables in the model – accounted for by the BCG variables was 12.50%. The amount of unique variance accounted for by life expectancy and the country's mean temperature was 34.52% and 2.37%, respectively.

Since life expectancy was the strongest predictor of cases, we reran the analyses only with those countries with a life expectancy higher than 78 ($n = 45$). This way, the mean life expectancy was homogeneous across BCG groups (A: 80.92±2.31; B: 82.41±1.27; C: 81.33±2.44). The pattern of results was essentially the same. There was a statistically significant effect for both Group B (Figure 2a; $b = 0.6122$, $p = .0024$) and Group C (Figure 2a; $b = 0.6511$, $p = .0326$). The temperature had no significant effect. The amount of unique variance accounted for by the BCG variables was 19.78%.

Regarding total deaths per one million population, the linear model again showed a statistically significant effect for both Group B ($b = 0.7262$, $p = .0007$) and Group C ($b = 1.495$, $p = 7.11\times10^{-5}$). While life expectancy was significant (Figure 1d; $b = 0.4251$, $p = 6.20\times10^{-6}$), the country's mean temperature was not (Figure 1b; $b = –0.0761$, $p = .3453$). The model's adjusted R^2 was .5473. The amount of unique variance accounted for by the BCG variables was 20.37%. The amount of unique variance accounted for by life expectancy and the country's mean temperature was 20.33% and 0%, respectively.

Again, since life expectancy was a strong predictor in the model, we reran the analyses only with those countries with a life expectancy higher than 78 ($n = 40$). The pattern of results was the same. There was a statistically significant effect for

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

both Group B ($b = 0.7918$, $p = .0011$) and Group C ($b = 1.498$, $p = 8.05\times10^{-5}$). The temperature had no significant effect. The amount of unique variance accounted for by the BCG variables was 38.39%.

Regarding the ratio between deaths and cases, the linear model showed a statistically significant effect for Group C ($b = 1.140$, $p = .0267$) but not Group B ($b = 0.1356$, $p = .6442$). This time, life expectancy was not significant ($b = –0.1294$, $p = .3019$). By contrast, the country's mean temperature was significant ($b = 0.2629$, $p = .0226$). The model's adjusted R^2 was .1019. The amount of variance accounted for by the BCG variables was 3.39%.

Finally, a potential issue of the above analyses is that we used data that were not corrected by the onset of the spread of the virus. We are now in the middle of the rapid development of a pandemic and the time since the onset significantly differs across different countries, which may serve as a potentially confounding factor. To address this issue, we utilized the curves of the cumulative number of confirmed cases by the number of days since 100th case (Figure 3a), which was reported on a website[17]. Since we did not obtain the numerical data behind the figure in the website, we classified the 70 countries into three groups, that is, high (h), middle(m), and low(l) growth rate (Figure 3a; Boundary between h and m was approximately 2500 cases and the one between m and l was approximately 1400 cases at Day 15 since 100 cases; The countries in which 15 days was not reached, the value of Day 15 was estimated by projecting the growth curve to 15th day). While based on a rough classification, this analysis was useful to explore the potential impact of the country's onset of the virus on our findings.
The Group A countries significantly tend to be classified in lower growth group compared to Groups B and C countries (as a single group) with all the 68 countries with known BCG policy (Figure 3b; Chi-squared test: $p = 0.0002$; B and C groups were combined for this test). As in the linear regression models, given that life expectancy may affect growth rate, we also conducted a Chi-squared test after excluding the countries that had a shorter life expectancy (≤78 years old, $n = 40$), and found that the difference remained significant after the exclusion (Figure 3c; Chi-squared test: $p < 0.0001$; B and C groups were combined for this test), where p-value became even smaller adjusting life expectancy despite the decrease of sample number. Both the analyses thus showed that Group A countries tended to exhibit a

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

lower growth rate compared to Groups B and C countries regardless of their life expectancy.

Discussion

Associations of BCG policy with COVID-19 after controlling two major confounding factors

In this study, we have shown that the countries that currently adopt universal BCG vaccination programs (Group A) have, compared to the other countries, a lower number of cases and deaths per one million people. The countries that no longer recommend BCG vaccination for everyone (Group B) and those that have never had universal BCG vaccination programs (Group C) reports more cases and deaths. We have also evaluated the impact of two potentially confounding factors, the country's life expectancy and the average temperature in February and March 2020. Crucially, the effects of BCG groups on the two dependent variables remain significant, even when the country's life expectancy and temperature are controlled for. The amount of variance explained by BCG vaccination is about 12.50% and 20% for cases and deaths per one million population, respectively. The percentage of explained variance is greatly increased when only countries with a life expectancy above 78 years are considered (about 20% and 38% for cases and deaths, respectively). Also, only Group C appears to play a role in the deaths/cases ratio and explains little more than 3% of the observed variance. This latter result may suggest that, unlike cases and deaths per one million population, the death rate is weakly affected by the type of BCG vaccination policy adopted by the country. We thus hypothesize that the protective effect of the vaccine, if any, may consist of a significant reduction of the spread of the virus rather than a reduced mortality rate. It is also possible that the vaccine prevents progression of the disease after infection, since only the persons with severe symptoms tend to be able to receive PCR tests in many of the countries at present. This hypothesis requires additional empirical corroboration. Finally, in line with the results of the linear regression analysis, the Chi-squared analyses highlight that BCG vaccination policy appears to affect how quickly this virus spreads. Taken together, these results suggest that the overall impact of BCG vaccination on the COVID-19-related cases and deaths is, at the very least, non-negligible and worth of further empirical investigation.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

We have reported that a large part of the variance in the two dependent variables is explained by life expectancy, while the effect of temperature seems to be marginal at best. This is not surprising, considering the fact that age is a major risk factor in COVID-19[18,19]. Life expectancy is highly correlated with the income level of individuals[20] and countries[21], which would also be related to the quality of medical care that people can receive. If life expectancy is controlled, higher income and quality of medical care may *decrease* the risk and mortality of diseases in general, which is expected to decrease the positive effect of life expectancy on the risks in COVID-19. Considering the highly significant contribution of life expectancy on the risks, we could consider that the income level or the quality of medical care cannot be the confounding factor that explains the association between BCG vaccination and the risks in COVID-19. Regarding the tight relationships between age and mortality in COVID-19[18,19], BCG protection may wane with time since vaccination[22], and it could be potentially contributing to the higher mortality in aged populations to some extent. These possibilities need to be examined by further analyses.

As seen, the average temperature in countries exhibits a small and sometimes even non-significant negative correlation with the risks of COVID-19 infection and mortality. That said, we have not assessed the effect of humidity, which we plan to do in the near future, since absolute humidity is considered to be associated with influenza infection[23,24]. However, it is unlikely that absolute humidity significantly affects the results. In fact, absolute humidity does not seem to play any major role in the transmission rates of the COVID-19[25]. Moreover, absolute humidity and temperature are usually highly correlated, and it is thus unlikely that humidity may play a major role in explaining the spread of the COVID-19 if the temperature does not.

BCG vaccination policy is associated with a lower COVID-19 growth rate

We are still in the midst of the pandemic, during which the virus keeps spreading in most of the countries (except for China, where the rise has nearly been stopped). Also, the role of the observed between-country differences concerning the date of the spread of the COVID-19 is still unclear. In brief, the current state of affairs might change quickly and, therefore, our findings should be interpreted with caution. That being said, we have found a highly significant difference among the BCG groups of the slope of increase of the cases in log scale, that is, adjusted by the

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

day when 100th cases have been achieved. While this analysis is based on a somewhat subjective classification of the slope steepness in each country, and the number of countries assessed is only 68 here, the result indicates that the difference of the onset cannot be a major contributing factor underlying the association between COVID-19 indices and BCG vaccination.

The Group B countries are heterogenous regarding 1) the years in which the vaccination started and stopped[12] and 2) the strain of the vaccine[26]. Also, in the A groups, there are differences in coverage levels among countries[27]. For example, many of A countries reside in Africa, where BCG vaccination is mostly mandatory[12], but coverage is not necessarily high[27], which could potentially have distorted the results. However, in this regard, analyzing those countries with a life expectancy above 78 years, which excluded most of the African countries, did not substantially affect the results. Since the initiation of the spread of COVID-19 in African countries were relatively late compared to the ones in other continents, it might be necessary to conduct similar analyses by taking the coverage rate into consideration after further spreading of the disease.

Potential mechanisms

BCG, originally developed against tuberculosis, is hypothesized to develop 'frontline' immunity, training it to respond non-specifically to certain viruses with greater intensity[7,3,4]. This idea is supported by clinical and epidemiological studies, which showed that BCG appeared to lower overall mortality in children[5,2,6]. BCG, which can remain alive in the human skin for several months, triggers not only specific memory B and T cells, but also stimulates the innate blood cells for a prolonged period[7]. In a randomized placebo-controlled study, it was shown that BCG vaccination protects against experimental infection with a weakened form of the yellow fever virus[7,4]. By showing a strong association of BCG vaccination program with COVID-19-related risks, our results are in line with the idea that BCG vaccination provides non-specific protection against COVID-19.

Recommended future research

The retrospective nature of the study requires extra caution in interpreting the results presented here because they may be affected by unknown confounding factors. Examples include the country's overall hygienic conditions, number of PCR testing per population, ethnical genetic backgrounds, prevalence of SARS-CoV-2 strains, strength of the mitigation/suppression policy, people's compliance to such

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

policies, and culture/life style (e.g., shaking hands and hugging, wearing masks, and washing hands). Some of these elements might affect the observed relationships between COVID-19 indices and BCG vaccination policy by affecting both the susceptibility to SARS-CoV-2 and BCG vaccination policy in an indirect manner. These and other factors need to be examined in future studies (or even before formal publication of this manuscript in a journal with peer review; see "production note" below).

As outlined above, we have employed a rough classification for the statistical test for the relationship between BCG groups and the growth curve of COVID-19 in each country with a limited number of observations (Figure 3a). Obviously, it would be preferable to use continuous data to produce a more accurate estimation of the impact of BCG vaccination policy on how quickly the virus spreads.
Also, examining variables such as BCG strain[26] and the timing and the number of vaccinations (including the ones for other diseases) would be worth some attention. These variables might be easier to be examined in those countries where mandatory BCG vaccination was terminated by investigating the history of vaccination of COVID-19 patients and their symptoms and condition severity.
In our study, we used the data that are publicly available from several websites (see Method and results section), which have not been validated by any official institutions (to the best of our knowledge). Therefore, further analyses are recommendable once the data's reliability has been certified.
It may also be worth looking into the relationship between the age of BCG vaccination and the risk of infection and severity /mortality in COVID-19 patients[18,19]. An SNP in IL-1b (IL1B; rs16944) was found to affect the trained immunity response induced by BCG[7], and it is of interest to evaluate the association of this SNP and the protective effect of BCG vaccine against SARS-CoV-2.

Conclusion

The present study reports significant associations of BCG vaccination with prevalence and mortality from COVID-19 in 136 countries, even when the country's life expectancy and average temperature are controlled for. Together with the previous studies showing potential non-specific protection of BCG, we believe that further studies examining this hypothesis are warranted. In countries with no mandatory BCG vaccination at present, policymakers involved in health service

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

may consider, with careful evaluation of its safety, increasing/implementing BCG vaccination during this outbreak of COVID-19, at least for the categories of people who are currently receiving the vaccination in other countries.

Acknowledgement

We thank Yoko Kagami, and Harumi Mitsuya for their assistance in inputting the data and making Figures, Mamoru Tejima for letting us know the existence of potential correlation, and Hideo Hagihara, Tomoyuki Murano, Hironori Funabiki and Shinichi Nakagawa for useful discussions. We also thank Tomoko Tatsumi for helping us drafting the R codes.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Production note

We welcome the contributions from readers of any data related to variables of interest (including potentially confounding factors), which might be considered as a collaboration. It is preferable that data contributed be sorted by country name and in excel format (as in our dataset; See Supplementary Table 1). Using the same naming convention for the name of the countries as those in supplementary table 1 is also helpful for us to carry out analyses quickly.

The data we'd like to add to the excel file includes, but not limited to,

1. Humidity in February and March for each country,
2. Average income for each country,
3. Any data related to social distance (population density, culture of shaking hands, etc.) for each country
4. Administration policy of other vaccines for each country, and
5. Prevalence of infectious diseases (such as H1N1 influenza) in each country.

When you provide us data, specify where you obtained data so that we can confirm the data.

We look forward to collaborating with you in this war against COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

References

1. Coronavirus Update (Live): 686,047 Cases and 32,269 Deaths from COVID-19 Virus Outbreak - Worldometer. https://www.worldometers.info/coronavirus/. Accessed March 30, 2020.

2. Aaby P, Kollmann TR, Benn CS. Nonspecific effects of neonatal and infant vaccination: public-health, immunological and conceptual challenges. *Nature Immunology*. 2014;15(10):895-899. doi:10.1038/ni.2961

3. Netea MG, Joosten LAB, Latz E, et al. Trained immunity: A program of innate immune memory in health and disease. *Science*. 2016;352(6284). doi:10.1126/science.aaf1098

4. Vrieze J de, 2020, Am 6:25. Can a century-old TB vaccine steel the immune system against the new coronavirus? Science | AAAS. https://www.sciencemag.org/news/2020/03/can-century-old-tb-vaccine-steel-immune-system-against-new-coronavirus. Published March 23, 2020. Accessed March 28, 2020.

5. Higgins J, Soares-Weiser, K, Reingold, K. Systematic review of the non-specific effects of BCG, DTP and measles containing vaccines. *https://www.who.int/immunization/sage/meetings/2014/april/3_NSE_Epidemiology_review_Report_to_SAGE_14_Mar_FINAL.pdf?ua=1*. 2014:34.

6. Higgins JPT, Soares-Weiser K, López-López JA, et al. Association of BCG, DTP, and measles containing vaccines with childhood mortality: systematic review. *BMJ*. 2016;355. doi:10.1136/bmj.i5170

7. Arts RJW, Moorlag SJCFM, Novakovic B, et al. BCG Vaccination Protects against Experimental Viral Infection in Humans through the Induction of Cytokines Associated with Trained Immunity. *Cell Host & Microbe*. 2018;23(1):89-100.e5. doi:10.1016/j.chom.2017.12.010

8. Sato, J. If I were North American/West European/Australian, I will take BCG vaccination now against the novel coronavirus pandemic. https://www.jsatonotes.com/2020/03/if-i-were-north-americaneuropeanaustral.html. Accessed March 28, 2020.

9. Life Expectancy by Country and in the World (2020) - Worldometer. https://www.worldometers.info/demographics/life-expectancy/. Accessed March 30, 2020.

10. List of cities by average temperature. In: *Wikipedia*. ; 2020. https://en.wikipedia.org/w/index.php?title=List_of_cities_by_average_temperature&oldid=944505832. Accessed March 30, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

11. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD

12. Zwerling A, Behr MA, Verma A, Brewer TF, Menzies D, Pai M. The BCG World Atlas: a database of global BCG vaccination policies and practices. *PLoS Med*. 2011;8(3):e1001012. doi:10.1371/journal.pmed.1001012

13. R Core Team. R: The R Project for Statistical Computing. https://www.r-project.org/. Published 2014. Accessed March 30, 2020.

14. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*.; 2009. https://ggplot2-book.org/. Accessed March 30, 2020.

15. Peterson RA, Cavanaugh JE. Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*. June 2019:1-16. doi:10.1080/02664763.2019.1630372

16. Zhang D. *Rsq R Package*.; 2018. https://cran.r-project.org/web/packages/rsq/rsq.pdf.

17. Burn-Murdoch J, Tilford C, Fray K, Bernard S. Coronavirus tracked: the latest figures as the pandemic spreads | Free to read. https://www.ft.com/content/a26fbf7e-48f8-11ea-aeb3-955839e06441. Published February 8, 2020. Accessed March 29, 2020.

18. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*. 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3

19. Dowd JB, Rotondi V, Adriano L, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *medRxiv*. March 2020. doi:10.1101/2020.03.15.20036293

20. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality, and mortality. *Am J Public Health*. 1997;87(9):1491-1498. doi:10.2105/AJPH.87.9.1491

21. Wilkinson RG. Income distribution and life expectancy. *BMJ*. 1992;304(6820):165-168.

22. Sterne JAC, Rodrigues LC, Guedes IN. Does the efficacy of BCG decline with time since vaccination? https://www.ingentaconnect.com/content/iuatld/ijtld/1998/00000002/00000003/art00005;jsessionid=gq9mh4938cdnk.x-ic-live-03. Published March 1998. Accessed March 28, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

23. Kudo E, Song E, Yockey LJ, et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. *PNAS*. 2019;116(22):10905-10910. doi:10.1073/pnas.1902840116

24. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol*. 2010;8(2):e1000316. doi:10.1371/journal.pbio.1000316

25. Luo W, Majumder MS, Liu D, et al. The role of absolute humidity on transmission rates of the COVID-19 outbreak. February 2020. doi:10.1101/2020.02.12.20022467

26. Ritz N, Curtis N. Mapping the global use of different BCG vaccine strains. *Tuberculosis*. 2009;89(4):248-251. doi:10.1016/j.tube.2009.03.002

27. World Health Organization. *National Immunization Coverage Scorecards Estimate for 2016*.; 2017.

28. 6_GVAP2017_Coverage_scorecards.pdf. https://www.who.int/immunization/sage/meetings/2017/october/6_GVAP2017_Coverage_scorecards.pdf. Accessed March 28, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figure 1

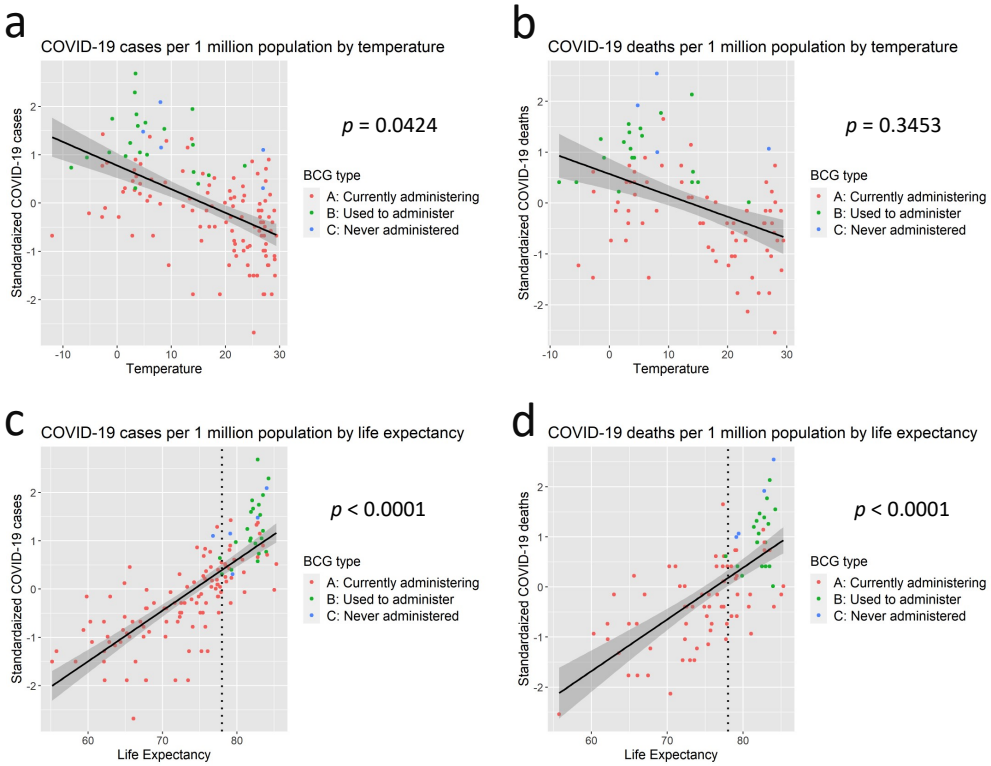


Figure 1: Correlations between temperature or life expectancy and COVID-19
a. Scatter plot of total cases per one million population by temperature. The dots represent the countries. The black line is the regression line with standard errors. Temperature (°C) is barely significantly correlated ($p = .0424$) with rate of cases of COVID-19 per one million population. b. Scatter plot of total deaths per one million population by temperature. The dots represent the countries. The black line is the regression line with standard errors (in grey). Temperature is not significantly correlated ($p = .3453$) with rate of deaths from COVID-19 per one million population. c. Scatter plot of total cases per one million population by life expectancy. The vertical dotted line indicates the threshold of 78 years, that is subsequently used for the analyses shown in Figure 2. The dots represent the countries. The black line is the regression line with standard errors (in grey). Life expectancy is significantly correlated ($p < .0001$) with rate of cases of COVID-19 per one million population. d. Scatter plot of total deaths per one million population by life expectancy. The vertical dotted line indicates the threshold of 78 years, that is subsequently used for the analyses shown in Figure 2. The dots represent the countries. The black line is the regression line with standard errors (in grey). Life expectancy is significantly correlated ($p < .0001$) with rate of deaths from COVID-19 per one million population.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figure 2

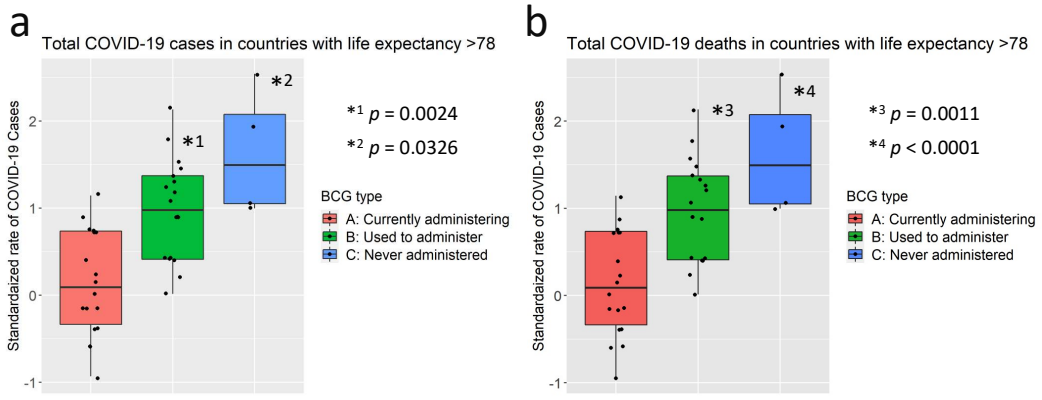
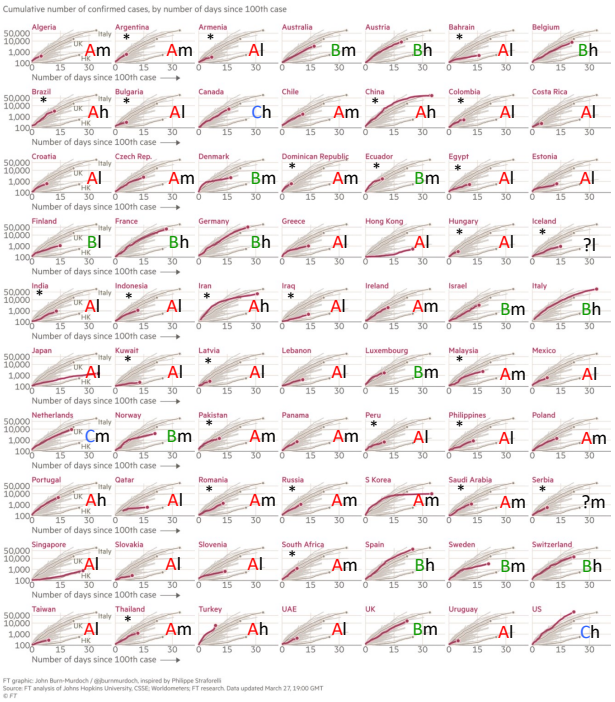


Figure 2: Effect of BCG vaccination policy on COVID-19
a. The boxplot of total cases per one million population sorted by BCG Group in countries with life expectancy higher than 78 years. Groups B and C (no current BCG vaccination) show a significantly higher rate of cases of COVID-19 compared to Group A (countries currently implementing BCG vaccination). Groups B and C (no current BCG vaccination) show a significantly (p = .0024 and p = .0326) higher rate of cases of COVID-19 compared to Group A (countries currently implementing BCG vaccination). b. The boxplot of total deaths per one million population sorted by BCG Group in countries with life expectancy higher than 78 years. Groups B and C (no current BCG vaccination) show a significantly higher rate of cases of COVID-19 compared to Group A (countries currently implementing BCG vaccination). Groups B and C (no current BCG vaccination) show a significantly (p = .0011 and p < .0001) higher rate of deaths from COVID-19 compared to Group A (countries currently implementing BCG vaccination).

medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20048165. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figure 3

a



Modified from https://www.ft.com/coronavirus-latest

*Countries with life expectancy shorter than 78 years

b

BCG groups and growth rate
(68 countries with all life exp.)

	Low	Middle	High
A	28	16	5
B	1	8	7
C	0	1	2

$p = 0.0002$ (chi square test)

c

BCG groups and growth rate
(40 countries with 78 years or longer life expectancy)

	Low	Middle	High
A	15	6	1
B	1	7	7
C	0	1	2

$p < 0.0001$ (chi square test)

Figure 3: BCG vaccination policy and COVID-19 growth rate

a. Categorial classification of growth rate of cases in 70 countries. A, B, or C designates BCG group and l, m, or h designates low, middle or high growth rate. "?" indicates that information is lacking. Countries marked by asterisk have average life expectancy below 78 years. b. Table for 68 countries on BCG vaccination groups and growth rate. c. Table for 40 countries on BCG vaccination groups and growth rate, by excluding the countries whose life expectancy is below 78 years.