



# Worldwide inverse correlation between Bacille Calmette–Guérin (BCG) immunization and COVID-19 mortality

Willis X. Li<sup>1</sup>

Received: 16 September 2020 / Accepted: 10 December 2020 / Published online: 25 January 2021  
© Springer-Verlag GmbH, DE part of Springer Nature 2021

## Abstract

**Purpose** The coronavirus disease 2019 (COVID-19) pandemic has spread to all countries in the world, and different countries have been impacted differently. The study aims to understand what factors contribute to different COVID-19 impacts at the country level.

**Methods** Multivariate statistical analyses were used to evaluate COVID-19 deaths and cases relative to nine other demographic and socioeconomic factors in all countries and regions of the world using data as of August 1, 2020. The factors analyzed in the study include a country's total COVID-19 deaths and cases per million population, per capita gross domestic product (GDP), population density, virus tests per million population, median age, government response stringency index, hospital beds availability per thousand population, extreme poverty rate, Bacille Calmette–Guérin (BCG) vaccination rate, and diphtheria–tetanus–pertussis (DTP3) immunization rate.

**Results** The study reveals that COVID-19 deaths per million population in a country most significantly correlates, inversely, with the country's BCG vaccination rate ( $r = -0.50$ ,  $p = 5.3e-5$ ), and also significantly correlates a country's per capita GDP ( $r = 0.39$ ,  $p = 7.4e-3$ ) and median age ( $r = 0.30$ ,  $p = 0.042$ ), while COVID-19 cases per million population significantly correlate with per capita GDP and tests per thousand population. To control for possible confounding effects of age, the correlation was assessed in countries propensity score matched for age. The inverse correlation between BCG vaccination rates and COVID-19 case ( $r = -0.30$ ,  $p = 0.02$ ) and death ( $r = -0.42$ ,  $p = 0.0007$ ) remained significant among the top 61 countries with the highest median age.

**Conclusion** This study contributes to a growing body of evidence supporting the notion that BCG vaccination may be protective against COVID-19 mortality.

**Keywords** SARS-CoV-2 · COVID-19 · Bacille Calmette–Guérin (BCG) · Diphtheria–tetanus–pertussis (DTP3)

## Introduction

The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, has reached all countries and territories in the world. As of August 1, 2020, there have been more than 21 million total accumulated cases and > 760,000 total deaths worldwide due to COVID-19, with a fatality rate of 4% [1].

However, different countries have been impacted very differently, with about a quarter of all cases and deaths happening in the United States, while certain other countries recording much lower levels of COVID-19 cases and deaths. The differential impact of COVID-19 on different countries in the world begs the question of what factors are important in influencing COVID-19 morbidity and mortality.

As an infectious respiratory disease transmitted through human contacts, COVID-19 should have higher impact in countries with high population density, with low disease control stringency, and in countries that lack medical resources, such as physicians and hospital beds, which are often associated with low per capita GDP. In addition, there have been observations of COVID-19 association with climate, BCG immunization, and various socioeconomic conditions of countries. Some of the observations

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s15010-020-01566-6>.

✉ Willis X. Li  
wxli@health.ucsd.edu

<sup>1</sup> Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA

were made during the early stage of the disease transmission when COVID-19 data were still accumulating, hindering the ability to draw definitive conclusions. Now more than eight months have passed since COVID-19 was identified and all countries have been affected, it is time to examine its transmission again in order to identify factors that are most important for COVID-19 morbidity and mortality.

I have previously posted on the Research Square preprint server [2] that COVID-19 morbidity and mortality inversely correlate with BCG immunization by Pearson's correlation coefficient analysis with the data available for all countries in the world as of June 1, 2020, and that the correlation remains statistically significant after controlling for age, a factor important for COVID-19 susceptibility. The current study used the same methods to evaluate COVID-19 impact on countries worldwide as of August 1, 2020. During the two months, the world has witnessed a tripling of COVID-19 cases from 6,327,540 to 17,999,745, and a doubling of total COVID-19 deaths from 385,096 to 687,582. Despite the large increase in case and death numbers, the current analysis found that the previous results mostly hold true, suggesting that BCG vaccination may have protective effects against COVID-19.

## Results

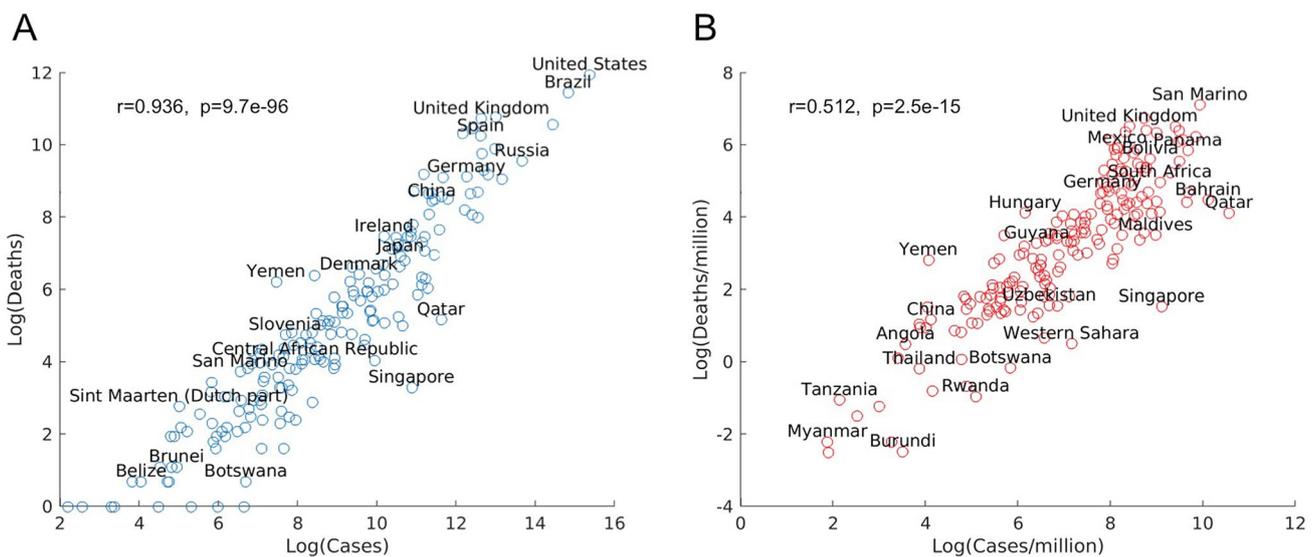
### Disproportional distribution of COVID-19 cases and deaths in world countries

Although the COVID-19 pandemic has spread to all countries in the world, different countries have been impacted

differently, based on the data as of August 1, 2020. The U.S. recorded more than 5.5 million cases and more than 170,000 deaths due to COVID-19, whereas in some other countries, these numbers are much lower, as in a scatter blot (Fig. 1a). After controlling for population size, disparity for countries still remains, although a few countries with small population sizes, such as San Marino, lead the chart in COVID-19 cases and deaths per million population (Fig. 1b). The case and death numbers are highly correlated, i.e., in general the higher the case number, the higher the death number in a country, although there are outliers with high case number but low death number, such as Qatar and Singapore (Fig. 1). The unequal distribution of COVID-19 death and case numbers in different countries raises the question of what factors are important for the susceptibility of a population to SARS-CoV-2 infection.

### Demographic and socioeconomic factors correlated with COVID-19 impact

To understand what demographic or socioeconomic factors are important for the COVID-19 impact, two different multivariate statistical analyses were employed. First, a pairwise Pearson's correlation coefficient analysis was carried out for several factors with the data available for most countries. These factors include each country's total cases and deaths per million population due to COVID-19; total numbers of SARS-CoV-2 virus tests carried out; government response stringency index, which records the strictness of 'lockdown style' policies that primarily restrict people's behavior [3]; population density (number of people per square kilometer of land); median age; per capita gross



**Fig. 1** a Scatter plots in log scale of COVID-19 cases and deaths. b Scatter plots in log scale of COVID-19 cases and deaths per million. Only a subset of countries is labeled above its marker

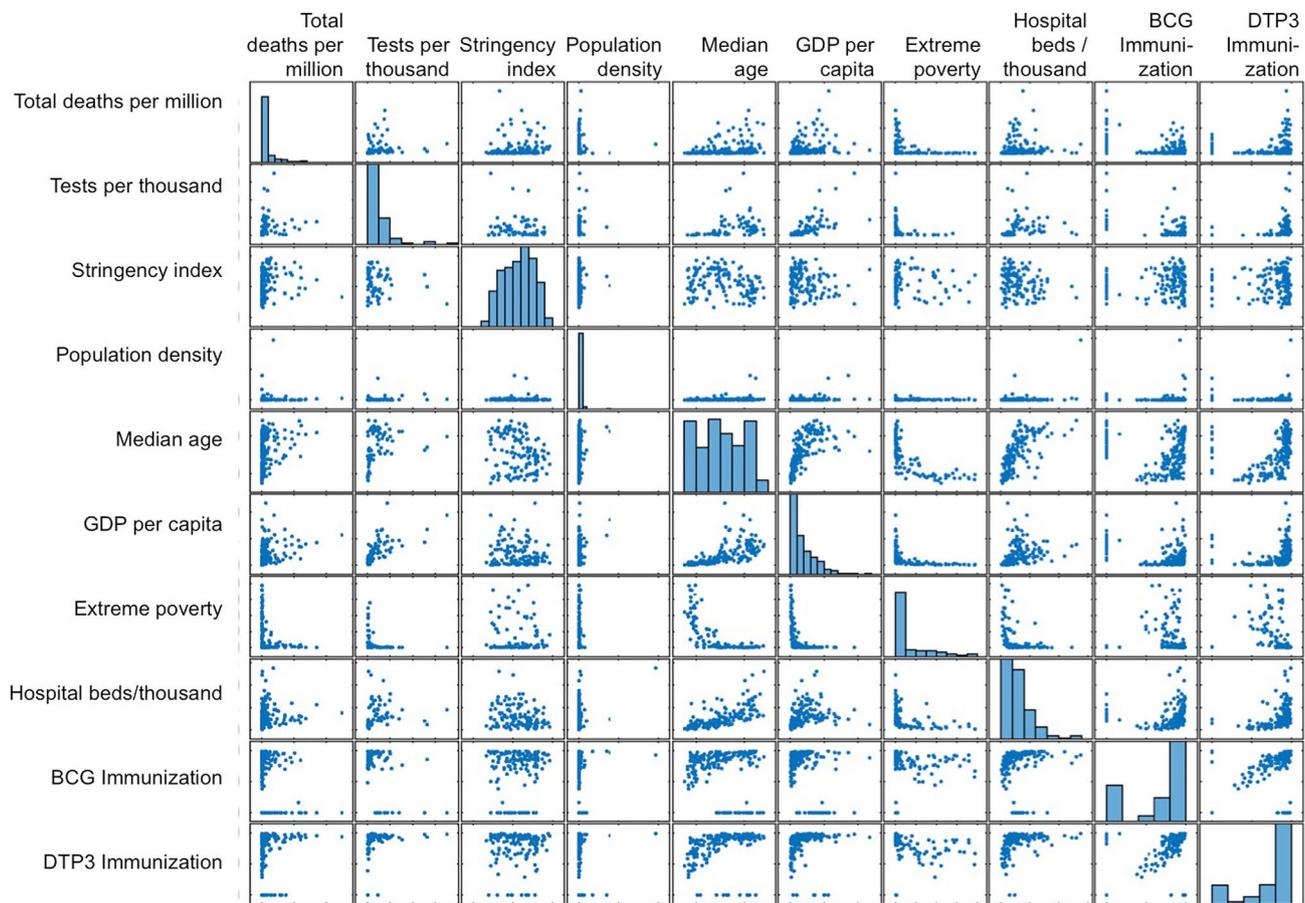
domestic product (GDP) (2019); extreme poverty index, hospital beds per thousand people; the coverage rates of Bacille Calmette–Guérin (BCG) vaccination; and diphtheria–tetanus–pertussis (DTP3) immunization rates (Table S1). Since COVID-19 death and case numbers are highly correlated, they were used separately in the following two different statistical multivariate analyses.

### Factors correlate with COVID-19 death and case in pairwise analysis

Pearson's correlation coefficients were calculated from an input table assembled from publicly available data as of August 1, 2020 (Table S1). The table rows are observations (210 countries and regions) and columns are independent variables (e.g., deaths per million, BCG rates). When these 10 factors were thus analyzed as independent variables, it is found that COVID-19 deaths per million most significantly (negatively) correlates with the country's BCG vaccination coverage rates ( $r = -0.50$ ,  $p = 5.3e-5$ ). COVID-19

deaths per million also significantly (positively) correlates with a country's per capita GDP ( $r = 0.39$ ,  $p = 0.0074$ ), and with median age ( $r = 0.30$ ,  $p = 0.042$ ) (Fig. 2, Table 1). COVID-19 death is found not significantly correlated, however, with a country's stringency index, population density, virus tests, extreme poverty rates, hospital bed availability, and DTP3 immunization coverage (Table 1). Thus, fewer COVID-19 deaths were found in countries with higher rates of BCG vaccination, whereas more COVID-19 deaths are found in countries with higher per capita GDP or more elderly people.

When COVID-19 cases per million were analyzed with the same method, however, it is found that it significantly correlates only with tests per thousand and GDP per capita, suggesting that high-income countries that carried more tests have more COVID-19 cases (Table 2). The difference between COVID-19 case and death numbers in terms of their relationships with other demographic and socioeconomic factors begs further analysis with different methods.



**Fig. 2** Pairwise scatter plots of the indicated independent variables. Green shaded regions indicate  $p < 0.05$  by Pearson's correlation coefficient (see Table 1 for values)

**Table 1** Pearson's correlation coefficients and *p* values for COVID deaths and other factors

	Total deaths per million	Tests per thousand	Stringency index	Population density	Median age	GDP per capita	Extreme poverty	Hospital beds/ thousand	BCG Immunization	DTP3 Immunization	<i>r/p</i>
Total deaths per million	1	0.24006	0.12363	-0.020814	0.30459	0.39375	-0.24	0.02363	-0.49612	0.23477	<i>r</i>
Tests per thousand	0.24006	1	0.41846	0.89205	0.041914	0.0074484	0.11229	0.87755	5.285e-4	0.12057	<i>p</i>
Stringency index	0.12363	0.41846	1	-0.33173	0.38069	0.76633	-0.29337	0.28211	-0.47359	0.34804	<i>r</i>
Population density	0.89205	0.81189	0.026004	1	0.0098853	8.4839e-10	0.050481	0.060442	0.0010164	0.019138	<i>p</i>
Median age	0.41846	0.1241	0.1241	0.41669	1	0.00013723	0.24135	-0.44037	0.045849	-0.35718	<i>r</i>
GDP per capita	-0.020814	-0.33173	0.1241	0.00013723	0.00097169	1	0.11023	0.0024671	0.76489	0.016012	<i>p</i>
Extreme poverty	0.89205	0.81189	0.41669	1	-0.13849	-0.13351	0.17546	-0.077743	-0.069474	-0.16258	<i>r</i>
Hospital beds/ thousand	0.30459	0.38069	-0.53818	0.41669	0.36428	0.38192	0.24896	0.61173	0.65021	0.28594	<i>p</i>
BCG Immunization	0.041914	0.0098853	0.00013723	-0.13849	1	0.64273	-0.68238	0.70813	-0.24357	0.65846	<i>r</i>
DTP3 Immunization	0.39375	0.76633	-0.47519	0.36428	1	1.9316e-06	2.436e-07	5.2981e-08	0.10689	8.8096e-07	<i>p</i>
	0.0074484	0.00097169	0.00097169	-0.13351	0.64273	1	-0.49649	0.44154	-0.62533	0.48072	<i>r</i>
	-0.24	-0.29337	0.24135	0.38192	1.9316e-06	1	0.00052269	0.0023952	4.3777e-06	0.00083048	<i>p</i>
	0.11229	0.050481	0.11023	0.17546	-0.68238	-0.49649	1	-0.44738	0.12645	-0.64954	<i>r</i>
	0.02363	0.28211	-0.44037	0.24896	2.436e-07	0.00052269	1	0.0020614	0.40782	1.383e-06	<i>p</i>
	0.87755	0.060442	0.0024671	-0.077743	0.70813	0.44154	-0.44738	1	0.042477	0.51754	<i>r</i>
	-0.49612	-0.47359	0.045849	0.61173	5.2981e-08	0.0023952	0.0020614	1	0.78174	0.00027185	<i>p</i>
	5.285e-4	0.0010164	0.76489	-0.069474	-0.24357	-0.62533	0.12645	0.042477	1	-0.09481	<i>r</i>
	0.23477	0.34804	-0.35718	0.65021	0.10689	4.3777e-06	0.40782	0.78174	1	0.53558	<i>p</i>
	0.12057	0.019138	0.016012	-0.16258	0.65846	0.48072	-0.64954	0.51754	-0.09481	1	<i>r</i>
				0.28594	8.8096e-07	0.00083048	1.383e-06	0.00027185	0.53558	1	<i>p</i>

Pearson's correlation coefficient and *p* value for the indicated independent variables compared in pairwise. Two-sided significance threshold was set at  $p < .05$  (shaded in green)

**Table 2** Pearson's correlation coefficients and *p* values for COVID cases and other factors

	Total cases per million	Tests per thousand	Stringency index	Population density	Median age	GDP per capita	Extreme poverty	Hospital beds/ thousand	BCG Immunization	DTP3 Immunization	<i>r/p</i>
Total cases per million	1	0.42886	0.2582	-0.099699	0.11561	0.36497	-0.22875	-0.021876	-0.22236	0.18427	<i>r</i>
Tests per thousand	0.42886	1	0.086803	0.51465	0.4495	0.013698	0.13068	0.88658	0.14207	0.22561	<i>p</i>
Stringency index	0.0032881	0.0032881	1	-0.036493	0.38069	0.76633	-0.29337	0.28211	-0.47359	0.34804	<i>r</i>
Population density	0.2582	-0.33173	0.026004	0.81189	0.0098853	8.4839e-10	0.050481	0.060442	0.0010164	0.019138	<i>p</i>
Median age	0.086803	0.026004	1	0.1241	-0.53818	-0.47519	0.24135	-0.44037	0.045849	-0.35718	<i>r</i>
GDP per capita	-0.099699	0.036493	0.1241	0.41669	0.00013723	0.00097169	0.11023	0.0024671	0.76489	0.016012	<i>p</i>
Extreme poverty	0.51465	0.81189	0.41669	1	-0.13849	-0.13351	0.17546	-0.077743	-0.069474	-0.16258	<i>r</i>
Hospital beds/ thousand	0.11561	0.38069	-0.53818	1	0.36428	0.38192	0.24896	0.61173	0.65021	0.28594	<i>p</i>
BCG Immunization	0.4495	0.0098853	0.00013723	0.36428	1	0.64273	-0.68238	0.70813	-0.24357	0.65846	<i>r</i>
DTP3 Immunization	0.36497	0.76633	-0.47519	-0.13351	0.64273	1	2.436e-07	5.2981e-08	0.10689	8.8096e-07	<i>p</i>
	0.013698	8.4839e-10	0.00097169	0.38192	1.9316e-06	1	0.00052269	0.0023952	4.3777e-06	0.00083048	<i>r</i>
	-0.22875	-0.29337	0.24135	0.17546	-0.68238	-0.49649	1	-0.44738	0.12645	-0.64954	<i>p</i>
	0.13068	0.050481	0.11023	0.24896	2.436e-07	0.00052269	1	0.0020614	0.40782	1.383e-06	<i>r</i>
	-0.021876	0.28211	-0.44037	-0.077743	0.70813	0.44154	-0.44738	1	0.042477	0.51754	<i>p</i>
	0.88658	0.060442	0.0024671	0.61173	5.2981e-08	0.0023952	0.0020614	1	0.78174	0.00027185	<i>r</i>
	-0.22236	-0.47359	0.045849	-0.069474	-0.24357	-0.62533	0.12645	0.042477	1	-0.09481	<i>p</i>
	0.14207	0.0010164	0.76489	0.65021	0.10689	4.3777e-06	0.40782	0.78174	1	0.53558	<i>r</i>
	0.18427	0.34804	-0.35718	-0.16258	0.65846	0.48072	-0.64954	0.51754	-0.09481	1	<i>p</i>
	0.22561	0.019138	0.016012	0.28594	8.8096e-07	0.00083048	1.383e-06	0.00027185	0.53558	1	<i>p</i>

Pearson's correlation coefficient and *p* value for the indicated independent variables compared in pairwise. Two-sided significance threshold was set at *p* < .05 (shaded in green)

## Multivariate analysis for COVID-19 cases and deaths

A second method, generalized linear regression, was used for a multivariate statistical analysis of COVID-19 cases and deaths. Here COVID-19 deaths and cases per million, respectively, in each country was used as observed responses, and all other factors were used as predictors. The results from multivariate analysis with generalized linear model regression are similar to those from the above Pearson's correlation coefficient analysis, although with interesting differences. Specifically, the most significant factor predicting COVID-19 deaths is, again, BCG immunization rates ( $p=3.7e-7$ ; Table 3), consistent with Pearson's correlation coefficient analysis. Also consistent with the Pearson's correlation coefficient is that per capita GDP significantly correlates with COVID-19 deaths ( $p=0.016$ ). Unlike Pearson's correlation coefficient, the multivariate linear regression found that DTP3 immunization rates ( $p=0.0001$ ), but not median age ( $p=0.0647$ ), significantly correlates with COVID-19 deaths (Table 3). However, DTP3 immunization rates positively correlates with COVID-19, the significant result is less informative, because it would suggest that DTP3 immunization promoted COVID-19 death. Thus, two different statistical methods both suggest BCG vaccination reduces COVID-19 deaths.

When COVID-19 case numbers were analyzed with multivariate linear regression, it is found that per capita GDP ( $p=1.75e-7$ ) and tests per 1000 population ( $p=0.0093$ ) significantly correlates with COVID-19 cases (Table 4), consistent with Pearson's correlation coefficient. In addition, multivariate linear regression indicates that median age, government response stringency, and poverty levels are also correlated with COVID-19 cases (Table 4), unlike the results from the Pearson's correlation coefficient analysis.

**Table 3** Multivariate analysis of COVID deaths

Predictors	Estimate (b)	SE	<i>t</i> Stat	<i>p</i> value
(Intercept)	85.1258	30.0668	2.8312	0.0051
Tests per thousand	-0.2288	0.1477	-1.5492	0.1229
Stringency index	0.5202	0.3479	1.4952	0.1364
Population density	0.0011	0.0073	0.1472	0.8831
Median age	-2.0553	1.1066	-1.8574	0.0647
GDP per capita	0.0018	0.0007	2.4233	0.0163
Extreme poverty	-1.0002	0.6836	-1.4631	0.1450
Hospital beds/thousand	1.2075	5.5085	0.2192	0.8267
BCG Immunization	-1.8233	0.3467	-5.2596	3.70e-07
DTP3 Immunization	1.8905	0.4680	4.0397	0.0001

Coefficient estimates (b) for a generalized linear regression of COVID deaths as responses on multiple predictors calculated using MATLAB glmfit function, with standard errors (SE), *t* statistics, and *p* values

**Table 4** Multivariate analysis of COVID cases

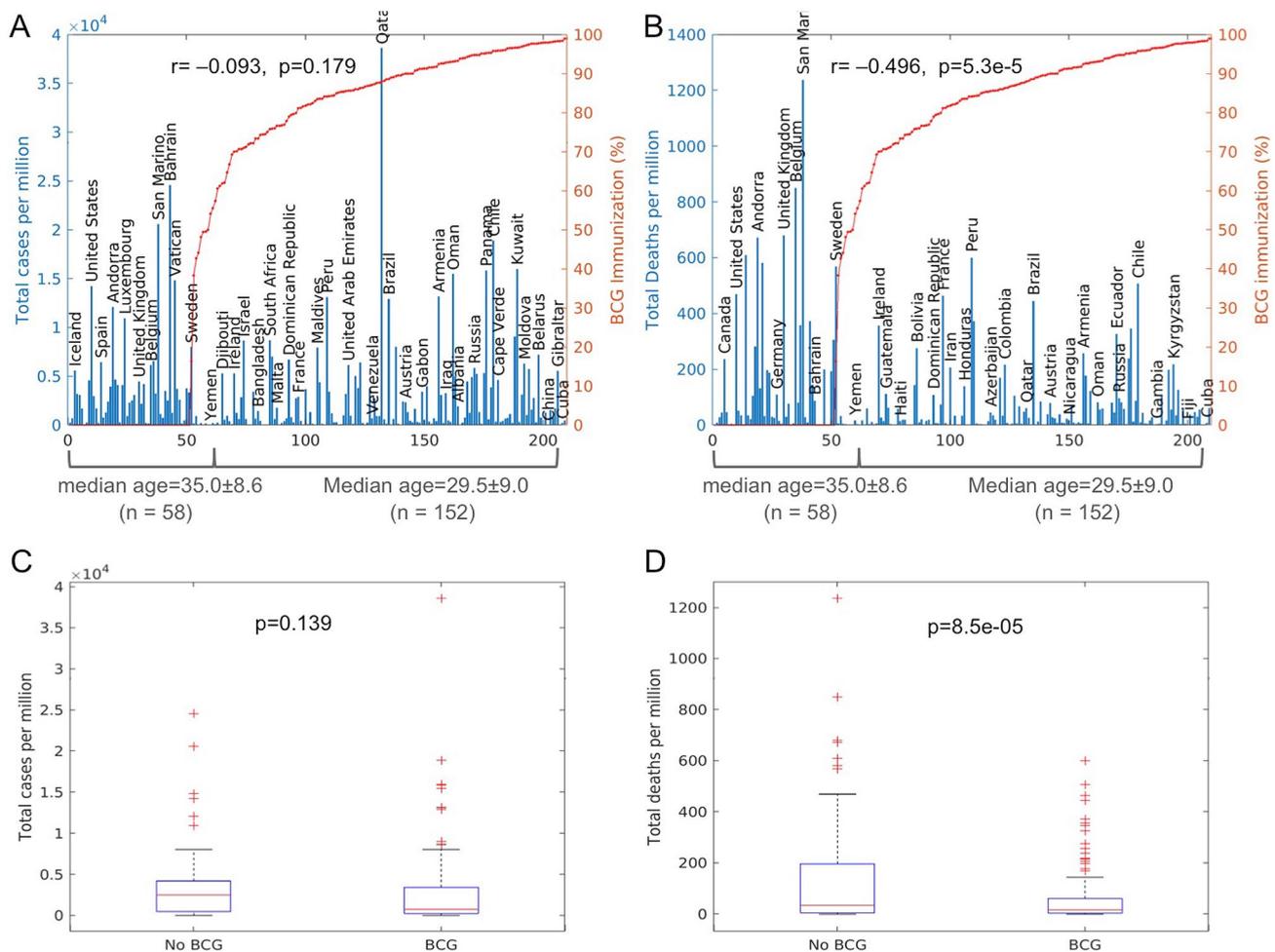
Predictors	Estimate	SE	<i>t</i> Stat	<i>p</i> value
(Intercept)	2260.4548	795.8747	2.8402	0.0050
Tests per thousand	10.2672	3.9101	2.6258	0.0093
Stringency index	25.7003	9.2095	2.7906	0.0058
Population density	0.0222	0.1920	0.1156	0.9081
Median age	-105.1955	29.2909	-3.5914	0.0004
GDP per capita	0.1074	0.0198	5.4144	1.75e-07
Extreme poverty	-39.1161	18.0948	-2.1617	0.0318
Hospital beds/thousand	-186.4294	145.8114	-1.2786	0.2025
BCG Immunization	1.0827	9.1760	0.1180	0.9062
DTP3 Immunization	15.1585	12.3878	1.2237	0.2225

Coefficient estimates (b) for a generalized linear regression of COVID cases as responses on multiple predictors calculated using MATLAB glmfit function, with standard errors (SE), *t* statistics, and *p* values

BCG is a tuberculosis (TB) vaccine used in many countries with high TB prevalence [4]. However, BCG is not generally used in countries with low risks of TB infection, such as the US and most Western European countries. There have been reports of COVID-19 and BCG association with both positive and negative results [5–7], and currently there are ongoing clinical trials using BCG as vaccine against COVID-19 [8, 9].

This study used BCG vaccination rates from Our World in Data [10], with the rates averaged for all the years with available data (1980–2019) for each country, and these were further verified with data from The BCG World Atlas [11]. Although the world health organization (WHO) started recording BCG coverage data only after 1980, most countries with BCG policy started immunization before 1980 [11]. The average BCG vaccination rates over all the years show highly significant inverse correlation with COVID-19 deaths by two statistics methods (Tables 1, 3). Indeed, when countries are sorted by mean BCG immunization rate, it can be seen that the high COVID-19 deaths occurred more in countries with low BCG vaccination rates (Fig. 3).

The world countries' BCG coverage rates show a dichotomous pattern, and there is a steep drop from high coverage rates (nearly 100%) to no coverage (0%) (Fig. 3a, b). Very few counties have a BCG coverage rate around 50% (a total of 9 countries have BCG rates from 40 to 60%). When countries were separated by BCG coverage rates into two groups of BCG rates  $\geq 50\%$  (denoted as "BCG") and BCG rates  $< 50\%$  (denoted as "No BCG"), the two groups of countries show significantly different mean deaths, but not significantly different mean cases (Fig. 3c, d). This is in agreement with the results from both Pearson's correlation coefficient analysis and multivariate linear regression.



**Fig. 3** COVID-19 morbidity and mortality vs BCG immunization rates. **a, b** Total COVID-19 cases and deaths per million against BCG immunization rates for all countries and regions ( $n=210$ ) in the world. Only selected countries are labeled. Values of  $r$ ,  $p$  indicate Pearson correlation coefficient and  $p$  value. **c, d** Boxplots of

total COVID-19 cases and deaths per million for all countries and regions ( $n=210$ ) in the world, grouped by BCG vaccination rate of either  $< 50\%$  (“No BCG”) or  $\geq 50\%$  (BCG).  $p$  values indicate the levels of difference between the two groups by two-tailed Student’s  $t$  test

In summary, factors affecting COVID-19 cases and deaths were analyzed using Pearson’s correlation coefficient and multivariate linear regression. Both statistical methods indicate that COVID-19 deaths most significantly correlate with a country’s BCG immunization rate, while the number of coronavirus positive cases correlate with the country’s resourcefulness and testing capabilities.

### Correlation between COVID-19 death and GDP may be due to age

A single factor identified by two statistical methods that significantly correlates with both COVID-19 cases and deaths is per capita GDP (Tables 1, 2, 3, 4). It has been reported that COVID-19 affects older people more than younger people [12]. Indeed, worldwide COVID-19 deaths significantly

correlate with the median age of nations ( $r=0.30$ ,  $p=0.042$ ) by Pearson’s correlation (Fig. 2; Table 1). On the other hand, the median age of a country highly significantly correlates with its per capita GDP ( $r=0.64$ ,  $p=1.93 \times 10^{-6}$ ). Thus, a plausible explanation is that high per capita GDP leads to longer lifespan; thereby, raising the median age of a country. In other words, the higher case and death numbers in wealthy countries were likely due to the higher age as a confounding factor in those countries.

### BCG vaccination and COVID-19 death inversely correlate in countries with high median age

Since age may be a significant confounder in the correlation between GDP and COVID-19, is age similarly a confounding factor in the correlation between BCG vaccination and

COVID-19? Correlation coefficient analysis shows that BCG vaccination does not significantly correlate with a country's median age ( $r = -0.24$ ,  $p = 0.107$ ; Table 1), although there appears to be a trend that countries with lower BCG vaccination rates may have a higher median age. This raised the question of whether the higher COVID-19 case and death numbers in countries lacking BCG vaccination can be confounded by their having more elderly population.

To reduce the possible confounding effects of age, propensity score matching for age was used to further evaluate the correlation between BCG vaccination and COVID-19. To this end, all countries were divided into three groups by median age. The median ages of the world's 210 countries and regions range from 15.1 (Niger) to 48.2 (Japan) (Table 5). These countries were discretized into "young", "medium", and "old" groups based on their median age, and their correlation with COVID-19 were analyzed separately in age-matched subgroups. It is found that BCG vaccination is not associated with COVID-19 cases and deaths for "young" and "medium" aged countries, but BCG vaccination remains significantly negatively associated with both COVID-19 case and death only in "old" countries (Table 5). Importantly, in these 61 "old" countries, BCG immunization rate and median age are no longer inversely correlated ( $r = 0.116$ ,  $p = 0.373$ ) (Table 5).

When separated by BCG coverage rates ( $\geq 50\%$  vs  $< 50\%$ ), among the 61 "old" countries, 36 countries with  $\geq 50\%$  BCG coverage have an average median age of  $41.8 \pm 2.7$ , which is almost the same as the average median age of  $41.1 \pm 3.025$  for the countries with  $< 50\%$  BCG coverage (Fig. 4). Thus,

after controlling for the confounding effects of old age by propensity score matching, the inverse correlation remains statistically significant between BCG vaccination and COVID-19 case ( $r = -0.30$ ,  $p = 0.019$ ), and between BCG vaccination and COVID-19 death ( $r = -0.42$ ,  $p = 0.0007$ ), in high median-age ("old") countries. Thus, BCG vaccination may reduce COVID-19 mortality in high median age countries.

## Discussion

Analyses for correlation between COVID-19 deaths and nine socioeconomic factors was carried out in order to assess which factors are important for the COVID-19 impact on world countries. The study identified BCG immunization as most significantly correlated, inversely, with COVID-19 mortality eight months into the pandemic, and this correlation remains significant after controlling for the confounding effect of age. Other factors significantly correlate with COVID-19 deaths a country's per capita GDP and median age. These two factors highly significantly correlate with each other, and age, which is known to affect COVID-19 susceptibility, is likely a confounder for the effects of per capita GDP.

As an infectious respiratory disease transmitted through human contacts, it can be assumed that the higher the human density should result in more COVID-19 cases and deaths. Social distancing works by reducing effective human density. However, this study found that COVID-19 cases and deaths do not significantly correlate with population density. Therefore, most densely populated counties or regions, such as Monaco, Singapore, and Hong Kong, do not have the highest number of COVID-19 cases or deaths, even based on per million population. Likewise, it was surprising that COVID-19 deaths did not significantly correlate with factors, such as government disease control stringency and the availability of hospital beds. However, there are caveats in this interpretation as certain data are incomplete, affecting statistical analysis. In particular, the policy stringency index data are incomplete (119 countries are missing data) and do not necessarily reflect the real timely responses and control in the complex situations of different countries.

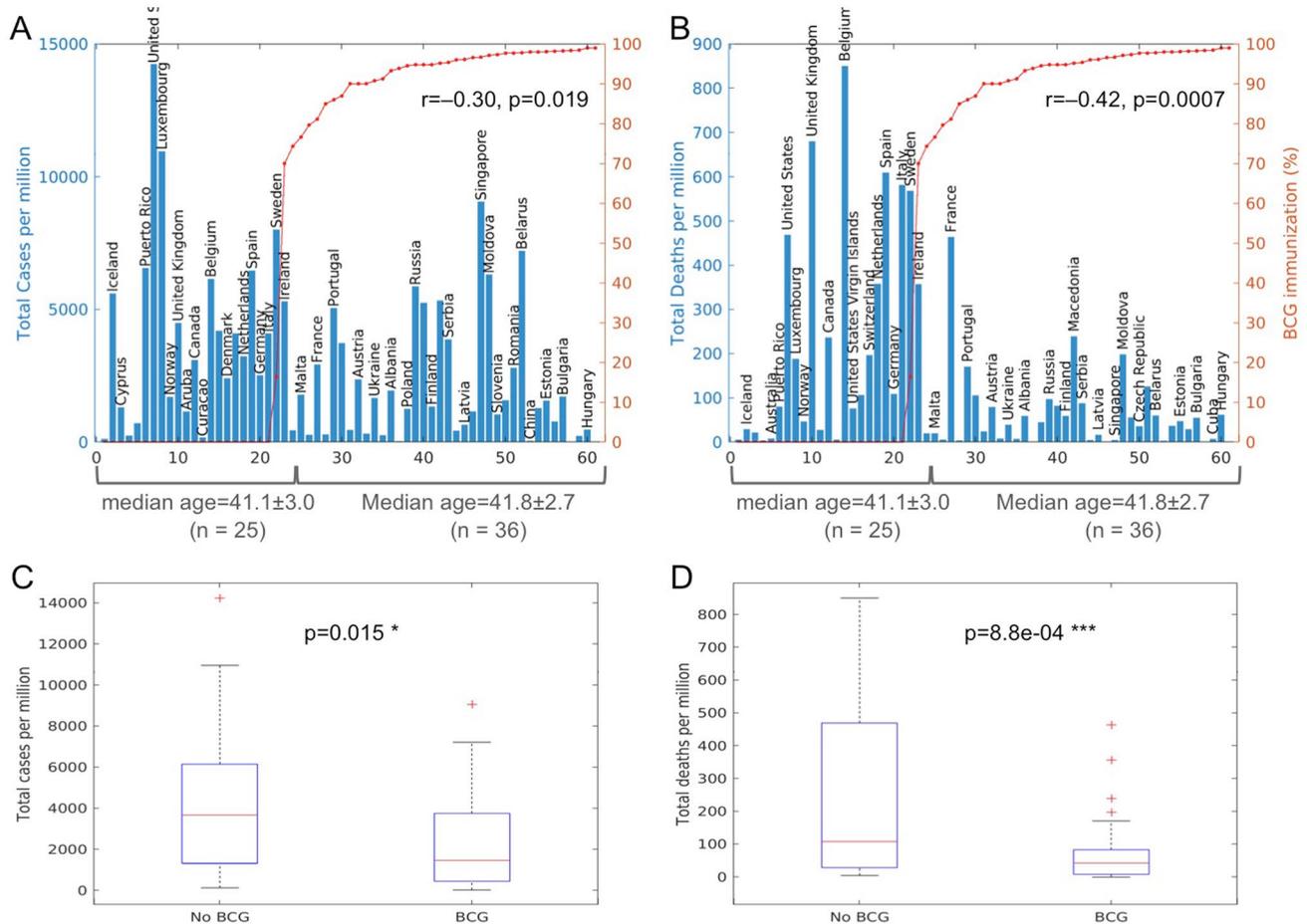
Contrary to the perception that low living standard and poverty may incur high number of COVID-19 casualty, it is found that the high numbers of per million COVID-19 cases and deaths correlate with high per capita GDP, a general measure of a country's wealth. Indeed, the top 20 countries in both total number of deaths and deaths per million include many of the world's wealthiest nations, including the United States and most Western European countries. However, as a country's median age is significantly correlated with its per capita GDP. It is believed that the reason the elderly

**Table 5** Correlation coefficient in countries matched by median age

Group		"Young"	"Medium"	"Old"	
N		50	75	61	
Min age		15.1	23.1	36.2	
Country		Niger	Gabon	Trinidad and Tobago	
Max age		22.9	35.7	48.2	
Country		Guatemala	Armenia	Japan	
BCG vs COVID cases	<i>r</i>	0.095	0.065	-0.299	
	<i>p</i>	0.514	0.58	0.0192	*
BCG vs COVID deaths	<i>r</i>	0.082	0.12	-0.424	
	<i>p</i>	0.57	0.304	0.0007	***
BCG vs median Age	<i>r</i>	0.201	-0.128	0.116	
	<i>p</i>	0.161	0.274	0.373	

210 countries/regions were propensity score matched for median age into three groups. 24 countries were excluded due to the lack of median age information. Pearson's correlation coefficient (*r*) and *p* value (*p*) are shown

\*, \*\*\*Indicate significant levels,  $p < 0.5$ ,  $p < 0.001$



**Fig. 4** COVID-19 morbidity and mortality vs BCG immunization in “old” countries. **a, b** Total COVID-19 cases and deaths per million against BCG immunization rates for high median age (“old”) countries and regions ( $n = 61$ ).  $r$  and  $p$  are from Pearson’s correlation coefficient analysis. **c, d** Boxplots of total COVID-19 cases and deaths

per million “old” all countries in the world ( $n = 61$ ), grouped by BCG vaccination rate of either  $< 50\%$  (“No BCG”) or  $\geq 50\%$  (BCG).  $p$  values indicate the levels of difference between the two groups by two-tailed Student’s  $t$  test

are more susceptible to COVID-19 is because their immune systems have declined [13]. Thus, the correlation between GDP and COVID-19 is likely confounded by age.

Although it is also argued that low GDP countries may lack testing capabilities, thus underestimating COVID-19 impacts. Although this may be true in the early days or months of the pandemic, testing ability has increased to levels comparable to high-income countries [14]. This study found COVID-19 deaths not significantly correlated with total tests carried out by countries, although understandably COVID-19 cases significantly correlates with testing.

To assess if other types of vaccination also influence COVID-19 deaths, immunization against diphtheria tetanus toxoid and pertussis (DTP3) was analyzed. DTP3 is a combined vaccine offered to young children in many, but not all countries with historical data available similarly to those of BCG immunization [15]. DTP3 immunization is another long running vaccine program adopted nonuniformly by

different countries, and broad immune benefits have been reported for both BCG and DTP3 [16, 17]. However, unlike BCG, no significant correlation was found between DTP3 immunization and COVID-19 death by Pearson’s correlation ( $r = 0.24, p = 0.121$ ; Table 1). Although DTP3 did show significant correlation with COVID-19 when analyzed by multivariate linear regression, this correlation would indicate DTP3 vaccination increases COVID-19 deaths, because the coefficient is positive, as mentioned earlier. Thus, BCG vaccination might have specific effects on reducing COVID-19 deaths.

There have been other studies recently on BCG and COVID-19 correlation, with different conclusions. Miller et al. analyzed data of several countries as of March 21, 2020, and found a significant correlation between countries with universal BCG vaccination policy and reduced COVID-19 morbidity and mortality [7]. Berg et al. have published a biweekly updated report since April 2020 analyzing the

daily rates of COVID-19 case and death increase, and have shown that countries with mandated BCG vaccination policy consistently exhibit flattened curves (slower increases) [6]. A recent study by Escobar et al. carried out detailed population comparisons, e.g., between former East and West Germany, and have found significant correlation BCG immunization and reduced COVID-19 deaths especially in European countries [18]. These studies are consistent with the current study. There are also reports, however, showing the opposite. For example, Hamiel et al. found no significant difference in COVID-19 positive test rates between two groups of young adults of about 3000 each in Israel [5]. It is noted that the BCG vaccinated group (aged 39–41 years) is slightly older than the unvaccinated group (aged 35–37), and the study examined only SARS-CoV2 infection not mortality. Asahara analyzed COVID-19 cases among passengers boarding the *Diamond Princess* cruise ship and found no significant correlation between citizens from countries with mandatory BCG immunization and those with no such policy [19]. Thus, further investigation is clearly needed to establish a causal relationship, if any, between BCG vaccine and COVID-19 mortality. Nonetheless, a growing consensus seems to suggest that BCG vaccination may help reduce COVID-19 mortality.

BCG vaccine was initially developed in 1921 in France as attenuated live strains of *Mycobacterium bovis*, but the strains used in different countries have since shown difference in eliciting human immune reactions [20]. This study did not attempt to distinguish among different strains of BCG vaccines. Future studies may compare different BCG strains toward mechanistic understanding if and how BCG vaccine might be used protect against diseases other than TB. Indeed, it has been shown that BCG immunization can have long-lasting protective effects not only against TB, but also other respiratory diseases, including even lung cancer [8, 9, 16, 17]. Thus, BCG might confer broad nonspecific immunity against respiratory infections. The mechanism awaits further investigations.

There were no proven vaccines yet against SARS-CoV-2 as of August 1, 2020. If the proposed BCG protection hypothesis holds true, BCG could be used as a substitute if COVID-19 vaccines cannot be established. Ongoing trials of COVID-19 vaccines may benefit by comparing with BCG for effects against SARS-CoV-2.

As with all epidemiological studies, the current study has limitations. Notably, COVID-19 data at the country levels may not fully reflect local provincial situation and may thus represent a sampling bias in statistical studies. Future studies should be at individual levels to alleviate sampling bias. Nonetheless, the inverse correlation between BCG vaccination and COVID-19 mortality found in this and other mostly epidemiological studies suggest that BCG vaccines may offer protection against COVID-19.

## Methods

All data were from Our World in Data (OWID) based at the University of Oxford (<https://ourworldindata.org/>) [1]. The data as of 08-01-2020 were used for COVID-19 total deaths per million, stringency index, population density per thousand kilometer, median age, per capita gross domestic product (GDP), extreme poverty index, hospital beds per thousand people. Average rates of BCG vaccination coverage for 1 year olds for each country from 1980 to 2019, and for share of children immunized with DTP3 from 1980 to 2018 were also from Our World in Data (OWID) [10]. BCG coverage rates were verified using data from The World BCG Atlas (<http://www.bcgatlas.org/index.php>) [11].

For statistical analyses, The MATLAB<sup>®</sup> function ‘*corrcoef*’ was used for Pearson’s correlation coefficients for statistical relationships between independent variables. Pearson’s correlation coefficient was used for measuring statistical relationships between two independent variables that are continuous and approximately normally distributed. The MATLAB<sup>®</sup> function ‘*corrcoef*’ returns a matrix of correlation coefficients calculated from an input matrix whose rows are observations (countries) and whose columns are variables (e.g., deaths per million, BCG rates). A generalized linear model regression with the MATLAB<sup>®</sup> function ‘*glmfit*’ was used for multivariate analysis. Two-sided significance threshold was set at  $p < 0.05$ . Data processing, visualization, and statistical analyses were carried out using MATLAB<sup>®</sup> Online (R2020a; The MathWorks Inc. 2020).

**Acknowledgement** I thank Drs. A. Catanzaro, M. Seifert, J. Mandel, and A. Malhotra for helpful discussions and suggestions.

**Author contributions** WL conducted literature search, analyzed data, prepared figures and tables, and wrote the manuscript.

**Funding** W.L is supported in part by a research grant from the National Institutes of Health (R01GM131044 to W.X.L), which plays no role in the design and execution of the study.

## Compliance with ethical standards

**Conflicts of interest** The author declares no conflicts of interest.

**Availability of data and material** All data are in the text and supplementary file.

## References

1. Roser M et al. Coronavirus Pandemic (COVID-19). Our World in Data, 2020. Published online at OurWorldInData.org (Retrieved from: ‘<https://ourworldindata.org/coronavirus>’ [Online Resource]).

2. Li WX. Worldwide inverse correlation between Bacille Calmette–Guérin immunization and COVID-19 morbidity and mortality. *Res Sq*. 2020. <https://doi.org/10.21203/rs.3.rs-42927/v1>.
3. Hale T, Angrist N, Cameron-Blake E, Hallas L, Kira B, Majumdar S, Petherick A, Phillips T, Tatlow H, Webster S. Oxford COVID-19 Government Response Tracker. Blavatnik School of Government.
4. World Health Organization. BCG vaccines: WHO position paper – February 2018. *Vaccins BCG: Note de synthèse de l’OMS – Février 2018*. *Wkly Epidemiol Rec*. 2018;93:73–96.
5. Hamiel U, Kozar E, Youngster I. SARS-CoV-2 rates in BCG-vaccinated and unvaccinated young adults. *JAMA*. 2020;323:2340–1. <https://doi.org/10.1001/jama.2020.8189>.
6. Berg MK, et al. Mandated Bacillus Calmette–Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19. *Sci Adv*. 2020;6:eabc1463.
7. Miller A, et al. Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. *medRxiv*. 2020. <https://doi.org/10.1101/2020.03.24.20042937>.
8. O’Neill LAJ, Netea MG. BCG-induced trained immunity: can it offer protection against COVID-19? *Nat Rev Immunol*. 2020;20:335–7.
9. Curtis N, et al. Considering BCG vaccination to reduce the impact of COVID-19. *Lancet*. 2020;395:1545–6.
10. Vanderslott S, B. Dadonaite. Vaccination. *Our World in Data*, 2013. Published online at [OurWorldInData.org](http://OurWorldInData.org). (Retrieved from: <https://ourworldindata.org/vaccination> [Online Resource]).
11. Zwerling A, et al. The BCG World atlas: a database of global bcg vaccination policies and practices. *PLoS Medicine*. 2011;8:e1001012 ([Epub 2011 Mar 22]).
12. Zhang L, et al. Blockade of JAK2 protects mice against hypoxia-induced pulmonary arterial hypertension by repressing pulmonary arterial smooth muscle cell proliferation. *Cell Prolif*. 2020;53:e12742.
13. Nikolich-Zugich J. The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol*. 2018;19:10–9.
14. Gupta M, et al. The need for COVID-19 research in low- and middle-income countries. *Glob Health Res Policy*. 2020;5:33.
15. World Health Organization. Diphtheria vaccine: WHO position paper, August 2017 - Recommendations. *Vaccine*. 2018;36:199–201. <https://doi.org/10.1016/j.vaccine.2017.08.024>.
16. Higgins JP, et al. Association of BCG, DTP, and measles containing vaccines with childhood mortality: systematic review. *BMJ*. 2016;355:i5170.
17. Usher NT, et al. Association of BCG vaccination in childhood with subsequent cancer diagnoses: a 60 year follow-up of a clinical trial. *JAMA Netw Open*. 2019;2:e1912014.
18. Escobar LE, Molina-Cruz A, Barillas-Mury C. BCG vaccine protection from severe coronavirus disease 2019 (COVID-19). *Proc Natl Acad Sci USA*. 2020;117:17720–6.
19. Masakazu A. The effect of BCG vaccination on COVID-19 examined by a statistical approach: no positive results from the Diamond Princess and cross-national differences previously reported by world-wide comparisons are flawed in several ways. *medRxiv*. 2020. <https://doi.org/10.1101/2020.04.17.20068601>.
20. Miyasaka M. Is BCG vaccination causally related to reduced COVID-19 mortality? *EMBO Mol Med*. 2020;12:e12661.