Archives of Community and Family Medicine

ISSN: 2638-4787

Volume 3, Issue 1, 2020, PP: 23-27

BCG Immunization Appears to Explain 26% of Variance in Cases of Covid19 per Capita

Zack Z. Cernovsky*, M. Lakshman D. Fernando, Simon Chiu

*Department of Psychiatry, University of Western Ontario, London, Ontario, Canada.

***Corresponding Author:** Zack Z. Cernovsky, Department of Psychiatry, University of Western Ontario, London, Ontario, Canada.*

Abstract

***Background**: Biological scientists Gursel and Gursel have compiled very valuable tabular data showing that countries with national BCG immunization program have, per capita, significantly less cases and lower death toll from covid19. These authors also highlighted the importance of selecting the best BCG vaccines from among its various strains. With respect to statistical analysis, Gursel and Gursel's study reported only a significant p value: it is very important to recalculate their data to determine the actual correlational size, proportions of explained variance, and the effect sizes. This statistical information is crucial for determining if, or to what extent, the BCG vaccines could contribute to containing the covid19 pandemic.*

***Method**: We recalculated Gursel's tabular data to measure the proportion of variance explained by the underlying correlational relationships of BCG immunization to the per capita cases of covid19 and to the per capita death from the covid19.*

We also calculated the correlations of these two per capita rates to population density.

***Results and Discussion**: Countries with current national BCG immunization programs had 6 times less per capita cases of covid19: the correlation coefficient is statistically significant, and of moderate strength (Pearson r=.51, p<.001, 1-tailed), and accounts for approximately 26.0% of variance. This indicates that BCG vaccination is an important strategy of reducing cases of covid19, especially if carried out with the best strains of the vaccine. The BCG vaccination is an important potential statistical predictor of covid19 cases.*

The death toll from covid19 was also 6 times lower in countries with national BCG vaccination: the correlation coefficient is also statistically significant, but its size is weak (Pearson r=.27, p=.049, 1-tailed) thus probably accounting for only 7.3% of variance. This suggests that the BCG vaccination is a less powerful (but not a negligible) predictor of lethal outcomes. Population density was unrelated to covid19 cases (Pearson r=.07, p=.377) and to the death toll (Pearson r=.05, p=.326).

The BCG immunization programs are currently more common in low income countries than elsewhere. Death rate per capita divided by the number of cases per capita can serve as a predictor of lethal outcome of existing cases: this index was unrelated to presence or absence of the BCG programs (Pearson r=.03, p=.432).

***Conclusions**: The number of covid19 cases and the related death toll are 6 times smaller in countries which have the national BCG coverage.*

The BCG vaccination is an important statistical predictor of proportions of covid19 cases per capita. This effect is very large (Cohen's d=1.19). The BCG vaccination is also associated with lower per capita death rate from covid19, but that statistical effect is of medium size (Cohen's d=0.56). The choice of the best vaccine strain is presumably crucial, as suggested by Gursel and Gursel: this needs further statistical research.

Keywords: *covid19, BCG vaccine, case rate per capita, death rate per capita*

Introduction

Bacillus Calmette–Guérin (BCG) vaccine has been primarily used to prevent tuberculosis. The review by Curtis's team[1] published a few weeks ago in Lancet indicates that "*Randomised controlled trials have provided evidence that the BCG vaccine's immunomodulatory properties can protect against respiratory infections*" and possibly also against covid19. Very important data collection was in this respect compiled by biological scientists Mayda Gursel and Ihsan Gursel.[2] Their publication from 2020 includes tabular summaries of reported cases, per capita, of covid19 and death rate, per capita, from covid19 for 20 countries with national BCG immunization coverage and for 20 countries that did not have or have ceased their national BCG vaccination programs. Gursel and Gursel[2] found that the countries with BCG immunization programs have statistically significantly lower covid19 case and death rate ($p<.05$) than countries without current BCG immunization programs. Unfortunately, their publication includes only the p values of statistical significance without statistical measures of the size of the underlying relationships, the proportion of explained variance, or effect size indicators such as Cohen's d.

From the statistical perspective, Welkowitz, Ewen, and Cohen[3] emphasize that "*statistical significance does not imply that the results have practical utility as well. Instead, it is necessary to convert statistically significant results into measures which express the strength of the relationships between the variables in question, or otherwise get some notion as to how large an effect is*." For these reasons, we re-calculated tabular data from Gursel and Gursel[2] to measures the size of the relationship between BCG vaccination and per capita case and death rates from covid19, to determine the proportion of variance in these outcome variables explained by BCG vaccination.

Population density might have some impact on the per capita case and death rates of covid19. We used worldometer data for population density in 2019[4] to calculate these correlations.

Method

As already mentioned, the excellent epidemiological study by Gursel and Gursel[2] compared 20 countries with BCG vaccination coverage to 20 countries without the coverage. The 20 countries with current BCG coverage in Gursel's study were China, Iran, S. Korea, Portugal, Brazil, Turkey, Malaysia, Japan, Ireland, Ecuador, Pakistan, Poland, Chile, Thailand, Greece, Indonesia, Romania, Saudi Arabia, Singapore, and Qatar. The 20 countries without the coverage were Italy, USA, Spain, Germany, France, Switzerland, UK, Netherlands, Austria, Belgium, Norway, Sweden, Canada, Australia, Denmark, Israel, Czechia, Luxembourg, Finland, and Iceland.

Tabular data in Gursel's study[2] include covid19 case and death rates, per capita, for those two sets of countries. In countries with current BCG coverage, the highest per capita rates of covid19 cases in tabular data presented by Gursel and Gursel[2] were those for Iran (274 per million) and Ireland (228 per million) and the highest per capita death rates were for Iran (21.54 per million), China (2.26 per million), and Portugal (2.26 per million).

In countries without current BCG coverage, the highest per capita rates of covid19 cases in tabular data presented by Gursel and Gursel were those for Iceland (1,723 per million), Luxemburg (1,398 per million) and Italy (1,057 per million) and highest per capita death rates were for Italy (100.48 per million) and Spain (47.22 per million).

In countries with current BCG coverage, the lowest reported per capita rates of covid19 cases in tabular data presented by Gursel and Gursel were those for Indonesia (2 per million), Pakistan (4 per million), and Brazil (8 per million), and the lowest reported per capita death rates were reported for Thailand (0.01 per million), Pakistan (0.03 per million), and Chile (0.10 per million).

In countries without current BCG coverage, the lowest reported per capita rates of covid19 cases in tabular data presented by Gursel and Gursel were those for Canada (54 per million) and Australia (67 per million) and lowest reported per capita death rates were for Czechia (0.09 per million), Israel (0.12 per million), and Finland (0.18 per million).

For obvious reasons, the transmission rates of covid19 cases might depend on population density. For this reason, we obtained data on population density per km^2 for all 40 countries from the worldometer website as listed for 2019.[4]

The size of correlational relationships was measured via Pearson product moment coefficients. Some statisticians would prefer nonparametric correlational

procedures by evoking the possibility of violations of assumptions of normal distribution and the type of measurements scale. It has, however, been adequately determined in statistical investigations by Havlicek and Peterson that "*the Pearson r is insensitive to extreme violations of the basic assumptions of normality and of the type of measurement scale.*"[5]

Some readers might also wonder why bivariate Pearson r is calculated on variables one of which is a dichotomy (countries with BCG coverage versus those without it), assuming it is an incorrect statistical choice. This is a frequent misconception. The results obtained by using the formula for regular Pearson r are identical to those obtained with those for point-biserial coefficients on variables one of which is a dichotomy and the same is also true for the phi coefficients on variables both of which are only dichotomies.[6]

Results and Discussion

Average Covid19 Case and Death Rates with and without BCG Coverage

Using tabular data compiled by Gursel and Gursel in their excellent study,[2] we calculated the average case and death rate of covid19 for countries with and those without BCG coverage. The results are listed in Table 1. The data show that the number of covid19 cases per capita is about 6 times higher in countries without current BCG coverage. It is of interest that per capita death rate from covid19 is also about 6 times higher in countries without current BCG coverage.

Table1. *Per capita covid19 cases and death rates versus BCG coverage*

	BCG coverage	No current BCG coverage
Mean **case rate** per million	76.1 (SD=85.0)	465.0 (SD=472.9)
Mean **death rate** per million	1.9 (SD=5.0)	11.5 (SD=23.5)
Pearson correlation coefficient involving covid19 case rates per capita	Pearson r=.51 p<.001 (1-tailed)	
Pearson correlation coefficient involving covid19 death rates per capita	Pearson r=.27 p=.049 (1-tailed)	
Variance % in covid19 case rates explained by BCG coverage	26.0%	
Variance % in covid19 death rates explained by BCG coverage	7.3%	
Cohen's effect size involving covid19 case rates per capita	Cohen's d=1.19	
Cohen's effect size involving covid19 death rates per capita	Cohen's d=0.56	

Variance % in Covid19 Cases and Death Rates Explained by BCG Coverage

As shown in Table 1, the difference in BCG coverage explains 26.0% of variance in covid19 case rates per capita and 7.3% of variance in death rate.

The statistics of 26% of explained variance in per capita case rates shows how well the covid19 case rate could be predicted solely by knowing if the particular country has or does not have the BCG coverage. Conversely, the remaining 74% of variance in covid19 case rates per capita is assumed to be caused by some other hypothetical or also some yet unknown factors.

The 7.3% of explained variance in per capita death rates shows how well the covid19 death rate could be predicted by knowing if the particular country has or does not have the BCG coverage. Conversely, the remaining 92.7% of variance in covid19 death rates per capita is assumed to be caused by some other hypothetical or also some yet unknown factors.

A major hypothetical factor that potentially influences the per capita case and death rates of covid19 is suggested in the review by Curtis's team and also in the recent article by Mayda Gursel and Ihsan Gursel, namely the strain of the BCG vaccine. As explained by these authors, some of the strains are assumed to be more efficient against covid19. The strain of BCG vaccines differs from country to country. Occasionally, several different strains of the vaccine are used within the same country, thus making the related statistical analyses complicated.

The proportion of explained variance is sometimes referred to as "coefficient of determination."

The estimate of the % of explained variance is obtained simply by squaring the correlation coefficient (here, the Pearson r). The Pearson r is a measure of the size of relationship between two variables: its absolute size ranges from 0.00 (indicative of no linear relationship)

to 1.00 (the number 1 indicates perfect relationship that allows precise predictions). The interpretation of size of Pearson r in scientific studies is often as follows: Pearson rs lower than 0.15 are usually of too little or no importance for statistical predictions because the % of explained variance is excessively low (only 2.3% or less). Pearson rs below 0.40 are usually referred as "low," those from 0.40 up to almost 0.60 are often described as "moderate," and those of 0.60 or above are often referred as "high." However, these "interpretative categories" may vary from author to author. For instance, Downie and Heath[7] consider only an r of .8 and above as "high."

Cohen's d is a widely used measure of effect size, i.e., of the magnitude of the phenomenon. According to Sawilowsky, Cohen's d of 0.20 could be considered as "small, the one of 0.50 as "medium," the one of 0.80 as "large," and the one of 1.20 as "very large." Thus, using this terminology, the effect of BCG vaccine on rates of covid19 cases per capita can be labelled as **very large** and the vaccine's effect on related per capita death rates as **medium**, see data in Table 1.

For physicians or various administrators with limited familiarity with statistics, it may be more easy to comprehend the practical value of BCG vaccine when it is more simply stated that, per capita, the number of covid19 cases and the related death toll are 6 times smaller in countries which have the national BCG coverage.

Proportion of the Death Rate Corrected for Case Rate

The fraction with the per capita death rate as the numerator and the case rate in the denominator provides an index of the lethal outcome in infected cases. The highest proportion was found in Italy (lethal outcomes in approximately 10% of covid19 cases), i.e., in a country which, as pointed out by Gursel and Gursel[2] historically never had a national BCG vaccination policy for all.

The overall mean % of lethal outcomes of covid19 cases, however, was similar in countries with BCG coverage (average of 2.2%, SD=2.4) and in those without current coverage (average of 2.1%, SD=2.6). Many factors might account for this lack of difference in outcomes, including perhaps the prevalent type of medical care in the given country and of diagnosing, or also the type of BCG vaccine.

Covid19 per Capita Cases and Death Rates Versus Population Density

The transmission of the virus is presumably facilitated by social proximity such as in crowded urban centers. The Pearson correlations calculated between the worldometer data for population density and covid19 case and death rates were not statistically significant ($p>.05$, 1-tailed) and are very low ($r=.07$ for case rate and $r=.05$ for death rate), thus suggesting the lack of important relationships.

A major confounding variable undermining the value of such calculations may be the extremely uneven geographic distribution of population density over certain countries such as Canada. More than 80% of Canadians live within 300 km of the US border and more Northern areas are less populated. For instance, in Northern Canada, the Baffin Island is larger than Germany, but has density of only 0.02 persons per km^2.

It is noteworthy in this context that, among countries without BCG coverage, Canada (with population density of 4 persons per km^2) had 54 covid19 cases per million, but Iceland (population density of 3 persons per km^2) had 1,723 covid19 cases per million of inhabitants.

Weaknesses of Our Statistical Study

The rates of covid19 cases per capita and the related death rates per capita are from Gursel and Gursel:[2] unfortunately, these are only the "officially reported rates." Such officially reported rates may, for various reasons, differ from the real events. For instance, the lack of accurate and widely available diagnosing of covid19, especially in remote, medically underserviced rural areas probably causes major underestimates of covid19 cases and related death rates. In general, mild or asymptomatic cases of covid19 can remain undiagnosed, but are still relevant because such persons are carriers of the illness.

Ideally, future studies should also include the data on strains of BCG vaccines as this is presumably a potentially important explanatory variable.

Conclusions

In the 20 countries with current BCG coverage, the average covid19 case rate was 76.1 and the death rate 1.9, compared to case rate of 465.0 and death rate of 11.5 in countries without current BCG coverage: the two rates are 6 times lower in countries with BCG

coverage. The size of this effect is very large (Cohen's d=1.19) for the covid19 case rate and medium (Cohen's d=.56) for covid19 death rate. These statistical findings indicate that BCG vaccine could help to reduce or contain the impact of covid19. Further research on the BCG role in (hopefully) reducing the number of covid19 case is needed very urgently.

References

[1] Curtis N, Sparrow A, Ghebreyesus TA, Netea MG. Considering BCG vaccination to reduce the impact of COVID-19. *Lancet*. Online April 30, 2020. https://doi.org/10.1016/S0140-6736(20)31025-4

[2] Gursel M, Gursel I, Cernovsky ZZ. Global BCG Vaccination Coverage versus Cases and Mortality in the SARS-CoV-2 Pandemic. Archives of Community and Family Medicine. 2020; 3(2): 24-26.

[3] Welkowitz J, Ewen RB, and Cohen J. *Introductory Statistics for the Behavioral Sciences*. 4th edition. New York, NY: Harcourt Brace Jovanovich, 1991.

[4] Worldometer data on population density for 2019, retrieved May 8, 2020. https://www.worldometers.info/world-population/population-by-country/

[5] Havlicek LL, & Peterson NL. Effect of the violation of assumptions upon significance levels of the Pearson r. *Psychological Bulletin*. 1977; 84(2): 373–377. https://doi.org/10.1037/0033-2909.84.2.373

[6] Cernovsky Z. A frequent misunderstanding associated with the point-biserial coefficient and the phi. *Psychological Reports*. 2002; 90:65-66

[7] Downie NM, Heath RW. *Basic Statistical Methods*. New York, NY: Harper & Row, 1983.

[8] Sawilowsky S. New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*. 2009; 8(2): 467-474. doi: 10.22237/jmasm/1257035100

Citation: **Zack Z. Cernovsky, M. Lakshman D. Fernando, Simon Chiu**. *BCG Immunization Appears to Explain 26% of Variance in Cases of Covid19 per Capita. Archives of Community and Family Medicine. 2020; 3(1): 23-27.*