The Covid-19 pandemic has resulted in over 1.3 million infections and 300,000 deaths between January and the 8th April 2020. Recently published papers suggest vaccination with Bacillus Calmette-Guérin (BCG) could have protective effects against viral infection. This work uses existing data and visualisation methods with a simple linear regression model to validate and illustrate previously reported observations that a COVID-19-attributable mortality among BCG-using countries have a lower trend than in non BCG-using countries using only the raw daily cases and daily death statistics and comparing for countries with different BCG vaccination policies.

1 Introduction

The Covid-19 pandemic has resulted in over 1.3 million infections and 300,000 deaths between January and the 8th April 2020. Recently published papers suggest vaccination with Bacillus Calmette-Guérin (BCG) could have protective effects against viral infection. BCG has been in use since 1921 and is a safe vaccine that has been shown in some trials to provide some level of protection for other a wide range of pathogens including viruses. A number of randomised clinical trials have begun to investigate whether the BCG vaccine is playing a contributory role in supressing the COVID-19 pandemic spread. There has been speculation that a higher blanket immunity among the young could be providing a shield for the elderly in some countries, reducing the pandemic mortality rates. In the interim, studies of the rates of mortality in the population could allow a better population based understanding of potential contributory factors. Hegarty et al. examined whether a relationship could be identified between national BCG vaccination programs and the mortality levels reported by countries. Inspired by Shet et al. who reported a COVID-19-attributable mortality among BCG-using countries as 5.8 times lower [95% CI 1.8- 19.0] than in non BCG-using countries, this work sought to use existing data and visualisations with a simple linear regression model to illustrate this finding using only the raw daily cases and death statistics and comparing for countries with different BCG vaccination policies.

2 Data Sources and Preparation

Daily cases and death statistics attributed to Covid-19 was sourced from the European Centre for Disease Prevention and Control. Per country BCG vaccination policy classifications was retrieved from the World Atlas of BCG.

---

*Andrew Hines*  
School of Computer Science  
University College Dublin  
Dublin, Ireland  
andrew.hines@ucd.ie

April 8, 2020

ABSTRACT

The Covid-19 pandemic has resulted in over 1.3 million infections and 300,000 deaths 80,000 deaths between January and the 8th April 2020. Recently published papers suggest vaccination with Bacillus Calmette-Guérin (BCG) could have protective effects against viral infection. This work uses existing data and visualisation methods with a simple linear regression model to validate and illustrate previously reported observations that a COVID-19-attributable mortality among BCG-using countries have a lower trend than in non BCG-using countries using only the raw daily cases and daily death statistics and comparing for countries with different BCG vaccination policies.
data was processed using scripts developed to plot visualisations of daily cumulative cases and deaths inspired by the excellent daily update visualisations produced by John Burn-Murdoch of the Financial Times.

The BCG World Atlas categorised the database of policies and practices of the Bacillus Calmette-Guérin (BCG) vaccine from countries all over the world with details on the changes in policy over time. At a high level there are three main country policies: (a) Current national BCG vaccination policy for all; (b) Past national BCG vaccination policy for all; and (c) BCG recommendation only for specific groups or none at all.

More stratified categories from [7] based on the policy commencement year, subset of populations vaccinated, etc. are not considered here. Many other factors such as those considered in the modelling of [6] or other factors such as the levels of immigration, wealth, vaccination take-up and healthcare systems could also play a factor but are not considered.

---

3 Rate of increase in Cases and Deaths

The cumulative cases and deaths per country are plotted, colour-coded by BCG policy class in Fig. 1 and Fig. 2. Both graphs plot the cumulative number of cases per day with a log scaled y-axis. The grey dashed lines show what a trend following an exponential doubling of cases per day, two days etc. would follow. They are plotted from day zero when 100 cases identified in Fig. 1 and 10 deaths occurs in Fig. 2. It should be noted that the reporting on cases varies by

---

*https://github.com/barrysmyth/data_in_the_time_of_corona*
*twitter.com/jburnmurdoch*
Figure 3: Maximum slope of 5-day sampled linear regression of cumulative daily cases and deaths by country. The figure to the right for deaths shows the red and blue lines have a steeper slope.

Figure 4: Histograms showing the maximum slope distributions for countries cumulative cases and deaths. The red lines show the individual country maximim slope rates for countries with None/Selective BCG vaccination policies; blue dashed lines are countries with past polices.

country and even within countries as the testing criteria, protocol and capacity change. Hence the daily cases should be treated as unreliable.

In Fig. 2, the cumulative deaths reported are plotted. This data has some of the same quality issues, e.g. deaths outside hospitals are delayed in reporting, are mis-classified or are reported in bulk on a lagging timescale. However, it can be considered better than the cases data in terms of quality.

Looking at both graphs there is a sense that there is a trend that the red and blue lines (i.e. classes with less or no vaccination) are following steeper trends than the grey lines but it is difficult interpret.

4 Rate of increase in Cases and Deaths

In order to compare across countries in a consistent manner, the maximum slope of the cumulative cases and deaths per country are computed as a linear regression fit over 5 day windows. This gives an indication of the peak rise in cases and deaths. These are plotted in Fig. 3. For both a trend emerges in the red and blue lines (None/Targeted BCG
and Past Policy) having steeper trajectories. The outliers are countries like Lebanon and Andorra with relatively low numbers of cases and deaths reported to date.

This data is also visualised in Fig. 4 where the full distribution of slopes are plotted in a 12 bin histogram and dashed lines indicate the slopes of the countries outside of full population vaccination class (a). Again, the cases are noisy but do illustrate the skew. The deaths highlight that all the dashed red lines except one outlier (Lebanon) lie in the top third of the peak slope distributions.

5 Conclusions

This paper presents a simple linear regression maximum slope over 5 days analysis of the daily COVID-19 case and death statistics reported and shows a correlation with national BCG vaccination policies that is similar to [6]. A followup analysis when the countries that are still in the earlier stages of the outbreak may clarify the outliers.

References


