

CDS DS310: Data Mechanics

Professor: Chris Seferlis

Mentor: Emma Steel

7th May, 2024



Report RQ: What are the most unrestrictive policies they can implement to keep the growth rate of deaths below 1% and the growth rate of new cases below 3% on a 30-day rolling average?

Team 20: Anika Bhati, Gayoung (Ellena) Ko, Alicia Cao Belen Karakullukcu

Caladan, a midsize commonwealth with a population of 3.2 million, faces the threat of a potential surge in COVID-19 cases and deaths as it approaches the spring season. Our goal is to implement some minimally restrictive policies that effectively keep the growth rate of new cases below 3% and the growth rate of deaths below 1% over a 30-day rolling average. The paper below elucidates on the policies that we've recommended as justified by our data exploration and statistics, followed by scope for further research.

Our analysis began with collecting and reviewing data on COVID-19 response strategies from 10 countries that showed effective management of the pandemic with minimal restrictions. Firstly, our exploration involved comparing timelines of policy implementation by creating a correlation matrix heatmap along with a slider to understand strong correlations between policies and affected individuals. Based on the heatmap, we found the facial covering, testing and vaccination policies to have the strongest positive correlations with cases and deaths, which led us to explore these variables further through power BI visualizations. The correlation between facial coverings with cases and death was 0.43 and 0.53 respectively. And, the correlation between testing policy with cases was 0.33 and deaths was 0.13. Since we divided the data into pre-vaccine and post vaccine, we noticed that the correlations for vaccines (in the post vaccine period) was 0.60 with cases and 0.61 with deaths. We also noticed a strong correlation between cases and deaths in the initial phases of covid which eventually tapered off as more preventative policies were introduced.

Prior to implementing any public health policies, we observed that the average death growth rate due to COVID-19 was 3.45%. However, the introduction of a level 2 facial covering policy markedly reduced this rate by 4.16%, effectively inverting it to a -0.71% growth rate. This reduction demonstrates a significant positive impact on controlling the virus's lethality, a finding corroborated by a similar reduction in the case growth rate.

In regards to testing policies, our strategy involved advancing to a level 2 policy, which mandates testing for anyone exhibiting symptoms of COVID-19. Before this policy was enacted, the death growth rate stood at 3.92%. Post-implementation, it decreased by 3.02%, reducing to 0.90%. This substantial decrease is mirrored in the case growth rate, underscoring the critical role of comprehensive testing in disease management.

Additionally, our examination of vaccine policies revealed that implementing an initial vaccine policy at an estimated level of 0.5 led to an average death growth rate across all countries of -2.92% when a level 1 policy was employed. The corresponding rate for cases was -2.35% post-policy. Given the effectiveness of these measures at level 1, our group explored even less restrictive options, settling on a potential level 0.5 stringency. This approach reflects our aim to balance public health safety while minimizing restrictions.

Based on our analysis, we recommend the following minimally invasive strategies:

1. **Policy 1:** Facial Covering (Policy level 2)
2. **Policy 2:** Vaccine Policy (Policy level 0.5), prioritizing medical workers and immunity compromised individuals
3. **Policy 3:** Testing policy(Policy level 2), testing of anyone showing covid symptoms (preferably with at home testing kits to reduce contact)

In terms of other research, our group also decided to look into machine learning models using python(click here to access [ipynb](#) file), as discussed in class, to further reinforce our results. We first looked into decision trees due to its high interpretability, but chose not to select it as it would be most effective to solve classification problems. Our main aim was to understand the importance of each policy for affected individuals over time. Due to this, we chose to run an exponential regression analysis on python, due to covid's rampant rise from the year 2020 to 2021. Through this, we were able to gauge statistically significant policy variables, for example:facial coverings, restriction on gathering, etc. We also looked into country wise populations to compare countries with a similar population as caladan while developing policies. The approval and introduction of vaccines was also studied to internally divide the data into pre-vaccine and post vaccine time periods (december 2020)¹, isolating the effects of policies when vaccines were not an influencing factor which provided a clear picture of their true efficacy.

However, there are still limitations to implementing these policies. Face coverings, while essential for public health, may pose challenges to the environment and, if not replaced regularly, several pathogens (including SARS-CoV-2) could accumulate in the mask. Next, implementing vaccine policies would require the economy of Caladan to have substantial investments in health. Lastly, enforcing a testing policy could lead to false negatives if the individual is tested too early or if the nasal swab is inserted incorrectly. To navigate these challenges, Caladan should aim to find a balance between public health and economic considerations.

In terms of further research, it would be interesting to explore the long-term effects of these policies on the economy as well as the mental health of the population through a dataset with a longer time period. Additionally, continuous monitoring of vaccine efficacy and virus mutation patterns will be crucial to adapt policies as the situation evolves. It would be useful to have more demographic variables such as age, gender, and health indexes for each country to understand individuals and country profiles better. We also noticed that policy effect and implementation differs from country to country due to confounding variables, so basing conclusions about

¹ Fortner, Andra, and David Schumacher. "First COVID-19 Vaccines Receiving the US FDA and EMA Emergency Use Authorization." *Discoveries (Craiova, Romania)*, U.S. National Library of Medicine, 5 Mar. 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC8101362/#:~:text=The%20first%20vaccine%20receiving%20emergency,Comirnaty%20from%20Pfizer%20and%20BioNTech Accessed 04 May 2024.

caladan from generalized country trends might not be the most accurate approach. To combat this, past trends about Caladan's health index, investment in medical care, and individual profiles of residents could help make our research more personalized to the country.

To conclude, our evaluation underscores the pivotal role of vaccines, testing and facial coverings as the three most significant policies to keep deaths growth rate and new cases growth rate under 1% and 3% in the commonwealth respectively. As a result, through the implementation of our suggested policies, will aim to keep the community safe while also maintaining normalcy in residents' lives.

Contributions to the project

Anika	Alicia	Belen	Gayoung (Ellena)	Everyone
Project Manager	Data Engineer	Data Analyst	Data Architect	Did everything
Whenever we did any work for the project, we made sure to do it together. However, Alicia took the lead in visiting the TA for doubts/help with the project. Belen took the lead in fabricating PowerBI visualizations. Ellena was responsible for establishing a working architecture to run our project on. And Anika was in charge for further Python exploration and creating the report and presentation.				