**Joshua Goetz**

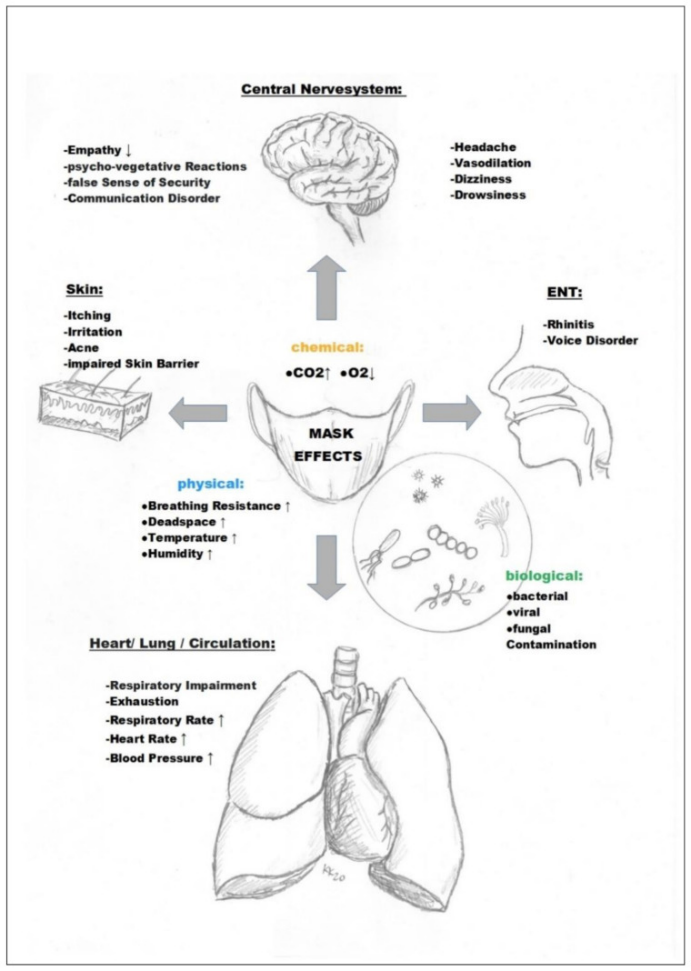
**Course** (DSC 503: Statistics Data Science)

**Cost-Benefit Analysis of Mask-Wearing Policy Stringency**

**Introduction**

In April 2020, the WHO recommended the use of masks only for symptomatic, ill, and healthcare workers. By June 2020, the WHO then advised the use of masks for the general population in some scenarios, such as crowded spaces. However, the very report advising this states that “there is no direct evidence … on the effectiveness of universal masking healthy people in the community to prevent the infection with respiratory viruses, including COVID-19” (World Health Organization, 2020). A slippery slope fallacy led to the widespread belief that masking healthy people was helpful to society and would reduce coronavirus transmission. Yet, no more than weak evidence for a reduction in viral transmission because of masks was found in other meta-analyses (Ioannidis, 2020; Chu, et al., 2020). Given that there is no strong evidence for the benefits of healthy people masking in public, some justification for wearing the mask included an adage akin to “better safe than sorry,” but was the cost of being so ‘safe’ worth the potential risk of being sorry?

There was not much review literature on the adverse effects of regular mask use prior to the pandemic emergency beginning in 2020. Published incidences of adverse effects came from a wide array of medical practices. Yet, it was not until 2021, that Kisielinski, et al., produced one of the first and most comprehensive literature reviews on the potential hazards and undesirable side effects of the everyday use of a mask. The literature reveals significant adverse effects of wearing masks from numerous disciplines, including psychology, psychiatry, gynecology, dermatology, ENT medicine, dentistry, epidemiology, pediatrics, sociology, occupational medicine, and environmental medicine with symptoms ranging from itching to fungal, viral and bacterial contamination; difficulty breathing; increased temperature; raised heart rate; and more. A visual summary of statistically significant adverse mask effects alongside their anatomical structure, as seen in the Kisielinski et al., 2021 article, is in **Figure 1** below. The wide range of disturbances in the biological system could lead to serious long-term issues such as cardiorespiratory diseases, triggering of the sympathetic stress response, and psychological deficits in positive emotion, drive, and cognition. Not only are there significant biological downsides, but the waste from used masks and mask production causes health hazards across a variety of ecological levels.



**Figure 1:** Visual representation of mask-induced physiological, somatic, psychological, and pathologic changes found in numerous scientific literature papers from a myriad of disciplines.

The burden of proof to freely utilize the respiratory system without obstruction should not be on the citizen. It should be of concern to national policymakers to evaluate whether the increase of strictness for policies regarding mask use has a benefit or not. Since we now have the data on mortality rates for different countries over the two years that various masking policies were implemented, we should be able to compare the mortality rates with the level of masking within and between countries. By drawing correlations between the mask-policy stringency (MPS) and the excess mortality rates in the given countries using regression analysis, we can deduce the overall effectiveness of the MPS. I do not expect to be able to accept the alternative hypothesis that more MPS caused fewer cumulative excess deaths.

**Data Sourcing**

The data source used to index the levels of MPS, our primary dependent variable, comes from the Oxford COVID-19 Government Response Tracker (Hale, Angrist, Goldszmidt, et al., 2021.) The paper links to their GitHub repository, which includes all datasets used. For our purposes, we will use the ‘OxCGRT\_compact\_national\_v1.csv’ file in the data folder (Permalink - <https://github.com/OxCGRT/covid-policy-dataset/blob/e7f66ee39654293b5c068efd2f195bd591dc27f6/data/OxCGRT_compact_national_v1.csv>). Beside the health system policies of MPS, we will also include levels of governmental economic income support (EIS), and school closing stringency (SCS), for a broader perspective and evaluation of potential confounding variables. Our dependent variables with their official name [current-use name], descriptions, measurement, and coding parameters, as seen in the ‘documentation\_and\_codebook.md’ from the OxCGRT covid-policy-dataset repository, are available in Table 1 below.

**Table 1:** Personalized Naming Convention and Description/Coding of Variables

Our study uses the 'Cumulative excess mortality per million people' as the independent variable, sourced from 'Our World in Data'. This data, drawing from the Human Mortality Database (2023), World Mortality Dataset (2023), and Karlinsky and Kobak (2021), measures total deaths against expected figures from past trends. We compare actual deaths to projected ones instead of direct country comparisons. This approach accounts for varying factors affecting death rates in different countries, providing a more accurate reflection of the impact of our study's dependent variable. The cumulative count starts from January 1st, 2020 and ends on December 31st 2022 or the most recent occurrence after October 1st 2022.

**Data Cleaning**

Processing for our mask-wearing data begins by reindexing the data to include only the columns we need: 'CountryName', 'CountryCode', 'Date', 'H6M\_Facial coverings', ‘C1M\_School closing’, ‘E1\_Income support’. Next, remove all records with a date before April 2020 because this was prior to a common consideration of mask-use policies and is most likely just ‘0’. Our data contains the level of MPS, SCS, IS on every given day. To get a single, average value for each country, the mean stringency level will be calculated for each country.

For our data with cumulative excess deaths count, we will take the cumulative count of excess deaths from the last day of 2022, December 31st, 2022, because this is where the facial coverings data cuts off, and many would consider the pandemic to be essentially over at this point. If the country did not have a data point for that specific day, the cumulative excess death count from the first date prior will be used for the country; the cumulative excess death count over multiple years is unlikely to have a dramatic difference over a small time period, however, some countries have there last cumulative count more than 3 months before the end of 22; assuming those three months might have a significant effect on the cumulative count, the record for these countries will be removed. The variable distributions can be seen from the boxplots in **A comparison of a diagram

Description automatically generated with medium confidence**Figure 2 below.

**Figure 2:** Box Plots for our Target and Feature Variables

Seeing what appears to be outliers within Excess Mortality and Mask Policy Stringency, these records will be removed from our data frame using the inter quartile function. The resulting data table is sorted by Cumulative Excess Mortality per Million and shown in Table 2 below with the head and tail end of excess mortality values. There are a total of 81 records.

**Table 2:** Cleaned Data Table

**Modeling**

After the proper conditions are met for each data set, they will inner merge on the country code. We now have a data frame of all countries with the independent values (MPS, SCS, EIS) and our independent value ‘Excess Mortality’ in one data set. Now, to fit this data on an Ordinary Least Squares (OLS) regression model to see the relationship between our IVs and ‘Excess Mortality.’ This will show us our coefficients for the dependent variables, indicating the number of excess deaths per million changes given one unit increase in the IV. By evaluating our t-score and p-value for the coefficient of each variable, we can tell whether it is significant.

**Results**

The r-squared score of our OLS model was a modest 6.47. An F-statistic of 47.61 and the p-value associated with the f-statistic is well below 0.05, indicating that our model is highly statistically significant. The coefficients for our IVs, along with their standard error, t-score, and p-value are presented in Table 3 below.

**Table 3:** Results of the OLS model



Here, we see that MPS was the only feature to have a statistically significant effect on excess mortality with a t-score of 2.381 and a p-value <0.05. The coefficient of 1079.78 tells us that for every one-unit increase in MPS, we expect to find an increase of 1,080 cumulative excess deaths over the course of the pandemic when holding all other features constant. The other features show positive coefficients. However, they are not nearly statistically significant, with p-values from 0.89 to .45. Graphic Representation of our models predicted excess mortality values, and the actual excess mortality values can be seen in Figure 3 below.

A diagram of a mask policy

Description automatically generated

**Figure 3:** Predicted (red) and Actual (blue) Excess Mortality Values in Relation to MPS

**Discussion**

Our final analysis provides evidence that an increase in mask policy stringency is significantly correlated with an increase in cumulative excess mortality. We also saw what appears to be a positive correlation between school closing policy stringency and income support, but these were not significant. Drawing features from separate categories of government policy interventions (health policies for MPS; containment policies for SCS; economic policies for IS) we were able to hold all variables constant while avoiding collinearity in our model.

After removing records with outliers in our excess mortality and mask policy stringency, it appears there may have been new outliers from the updated data frame; more scaling could be done to remedy this; however, initially analyzing the data with respect to the meaning behind the original ordinal scaling variable is the best start.

Further, analysis should include breaking down the data into different time buckets. This way we can see the features’ correlations with cumulative excess mortality over different periods of the pandemic. Also, the original source data contains other variables that could be analyzed and run through a feature selection model to find which feature combinations have the most accuracy in predicting excess mortality.

Given the evidence of the adverse health effects of mask use, it’s of little surprise to see increased mask-wearing policies likely increased cumulative excess mortality within nations. Further evaluation should be done before recommendation of everyday mask use.

**Work Cited/ References**

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., & Schünemann, H. J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet*, *395*(10242). <https://doi.org/10.1016/S0140-6736(20)31142-9>

Ioannidis, J. P. A. (2020). Infection fatality rate of COVID-19 inferred from seroprevalence data. *Bulletin of the World Health Organization*, *99*(1), 19–33F. https://doi.org/10.2471/blt.20.265892

Kisielinski, K., Giboni, P., Prescher, A., Klosterhalfen, B., Graessel, D., Funken, S., Kempski, O., & Hirsch, O. (2021). Is a Mask That Covers the Mouth and Nose Free from Undesirable Side Effects in Everyday Use and Free of Potential Hazards? *International Journal of Environmental Research and Public Health*, *18*(8), 4344. https://doi.org/10.3390/ijerph18084344

‌World Health Organization. (‎2020)‎. Advice on the use of masks in the context of COVID-19: interim guidance, 5 June 2020. World Health Organization. <https://iris.who.int/handle/10665/332293>. License: CC BY-NC-SA 3.0 IGO‌

Hale, T., Angrist, N., Goldszmidt, R. et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 5, 529–538 (2021). <https://doi.org/10.1038/s41562-021-01079-8>‌