Evaluation of non-pharmaceutical intervention effectiveness in Covid-19 pandemic by using excess mortality metric

By: Migle Urbonaite

Supervisor: Raphaela Mayerhofer

Södertörn University | School of Natural Sciences, Technology and Environmental

Master's dissertation 15 credits

Degree project, Spring 2021

SÖDERTÖRN UNIVERSITY | STOCKHOLM
sh.se

Popular summary

Due to the recent pandemic, more focus has been put on the preventative measures used in infectious disease control. One type of these measures is called nonpharmaceutical interventions. The unique aspect of these interventions is that they do not involve any medication or vaccinations and rely on the community to adhere to them. These measures include face masks in public, keeping distance, increase in handwashing frequency, etc.

In the face of the current pandemic, where a new virus was introduced, nonpharmaceutical interventions were the only available tool to stop or slow down the pandemic. However, questions arose whether there is a correct intervention or set of interventions to assist in the pandemic management.

For this reason, scientists started to look for ways to evaluate how effective interventions are in reducing the spread of the disease or deaths caused by it. In most cases, the case or death reports were used to evaluate the interventions. Sadly, the information gathered from these reports can be misleading or incorrect. First, all countries follow different rules on which deaths are assigned to the disease. Secondly, not all disease cases will be reported because not all sick people will report it to the healthcare offices. This mismatch between the data gathered and the situation in the real world can cause the conclusions of the scientific studies not to represent the real world. Thus provide usable insights or misleading conclusions. An excellent example of this is the latest WHO report in this matter, indicating that there were 1.8 million reported deaths from Covid-19 worldwide in 2020, while the excess death metric indicated 3 million deaths more than expected based on previous years.

In this study, the relationship between excess death metrics and interventions was analyzed. The excess death metric indicates how the amount of deaths in the chosen period differs from several similar periods from before. The study found that the excess mortality metric was able to indicate the more effective interventions. Moreover, five interventions (school closing, workplace closing, public event limitation, gathering restriction, and staying at home) show a stronger relationship to the decrease in deaths, indicating the ability to reduce the death rate of the current pandemic.

Abstract

INTRODUCTION: The study focuses on finding a methodology for evaluating the effectiveness of the nonpharmaceutical intervention in the face of a new pathogen entering the population. Different interventions can have different effectiveness levels in different populations; thus, studying possible correlations and effectiveness among different groups is essential. With better knowledge of the topic, the outbreak management could be done more cost-effectively, reducing the need for antibiotics, vaccines, and possible reduction of infectious diseases caused burden in developing regions. Furthermore, the study aims to determine the ways of using excess mortality as an evaluation technique for nonpharmaceutical interventions used in the Covid-19 pandemic.

METHOD: The variables in time-series format were used to calculate a cross-correlation score alongside other correlation coefficient tests. With the cross-correlation, the lag will be established to estimate how the variables correlate in the timeline. In addition, the study will attempt to establish the connections between different nonpharmaceutical interventions and their strengths and different age groups.

RESULTS: The most frequent lag scores identified were 1 with 16 observations and 2 with 9 observations. The highest lag score was 4, which was observed once for the dataset of Hungary. The correlation between excess mortality and different harshness of NPI's was calculated. The correlation coefficient ranges from -0.3 to -0.39, indicating an overall low to medium correlation. The highest correlation was detected with stay at home requirements (-0.36), workplace closing (-0.37), and gathering restrictions (-0.39). In the final step, age-based correlations were established. The correlation ranged from 0.26 – 0.36, indicating an overall medium correlation. The lowest correlation can be seen in the youngest age group, 15-64 (correlation coefficient of 0.26), while the highest correlation of 0.36 can be seen in the 75-84 age group. Surprisingly the age group 85+ had a little lower correlation than the 75-84 age group.

DISCUSSION AND CONCLUSIONS: A stronger correlation between excess mortality and stringency index was detected in the countries with a higher death per capita. The two groups of intervention effectiveness were established: more effective (school closing, workplace closing, public event limitation, gathering restriction, and stay at home requirement) and less effective

(public transport limitation, restriction on internal movement, international travel control, public information campaigns, protection of elderly campaigns). This suggests that NPI effectiveness depends on population size. In the age-group-based analysis, the correlation became stronger with the age increase, indicating nonpharmaceutical intervention effectiveness against high mortality in older adults.

Acknowledgments

I would like to thank my supervisor, Raphaela Mayerhofer, for her support and insightful feedback through my thesis.

Secondly, I would like to express my gratitude to Patrik Dinnétz for the assistance while working on the methodology, data, and methods section.

I would like to thank all the teachers in the Infectious disease program and my classmates for the knowledge and insights throughout the year.

Finally, I would like to thank my family and partner for the support and understanding shown these past few months.

Table of content

1. Introduction

When an infectious disease causes an outbreak in the population, different types of interventions are employed to stop the disease spread. In this thesis, they were categorized into pharmaceutical and nonpharmaceutical interventions (NPI). The majority of the NPI is based on the public health theory, while pharmaceutical approaches rely on the availability and access to effective treatment (1). Pharmaceutical interventions are more costly (1) but provide a well-defined and more decisive outcome. The downside of pharmaceutical interventions is the availability aspect of such interventions. If the new pathogen was detected, there is no knowledge base for the creation of a specific treatment plan or medication and vaccination; moreover, there is no knowledge of the pathogens' genomic/protein structure, stopping the possibility of the treatment with readily available pharmaceuticals. Therefore, the situation requires the correct and straightforward implementation of NPI as a first-line defense against the outbreak's threat becoming a pandemic.

1.1 Background

Even though the cost-effectiveness of NPI is established (2), there is very little research into which interventions that have the best effects against pathogens (3). According to WHO, NPIs outside healthcare are:

- limits on the spread of the pathogen (travel screening and restrictions)
- local spread reduction (isolation, quarantine, social distancing, event cancellation)
- personal protection (hand hygiene, respiratory etiquette, face masks)
- communication to the public (4).

Policymakers are reluctant to implement the interventions due to their disruptive nature. In addition, lack of research or established evaluative effectiveness is another reason for the lack of NPI implementation.

The cluster-randomized trial conducted by Aiello et al. (2) analyzed the hand hygiene and face mask effects of the influenza-like disease (ILI) in the university and found a 48%-75% reduction in ILI cases in the hand hygiene and face masks intervention group. While the groups with both measures have shown a statistically significant reduction of the infection rate, the groups with just one intervention did not show a statistically significant difference (2). This outcome demonstrated the importance of analyzing how different interventions affect one another and what combinations work the best with specific pathogens.

A more extensive study into pediatric influenza by Torner et al. (3) explored how receiving information on influenza prevention, frequency of handwashing, use of alcohol-based sanitizers hand washing after touching contaminated surfaces affect influenza cases in different age groups. The paper concluded that some age groups (5-17 and 0-4) had no significant association with increased handwashing and reduced influenza cases (3), demonstrating that some interventions might be more or less effective for different age groups.

Different interventions can have different effectiveness levels in different populations, thus studying various interventions, possible correlation, and effectiveness among different groups could provide a knowledge base or encourage policymakers to implement the NPIs more often. With better knowledge of the effectivnes of NPIs, the outbreak management could be done more cost-effectively (4), reducing the need for antibiotics (5), vaccines (5), and reduction of infectious diseases caused burden in developing regions (6).

1.2 Problem

Public health or nonpharmaceutical interventions have been proven to be beneficial and financially effective in preventing the development of chronic disease or stopping the spread of infectious diseases (1). However, different pathogen characteristics and distribution patterns make the choices for interventions against infections hard. Therefore constant evaluation and estimation of effectiveness are essential. Case or death counts are used as metrics to indicate the effect of the intervention (7).

There is not much research on what other health indicators could be used to get a more precise evaluation. In this study, the possibility of using excess death (ED) or excess mortality indicators as an evaluation metric for NPI's in infectious disease control will be explored. The need for new metrics rises from the insights gained in the Covid-19 pandemic. The different definitions of death by Covid-19 caused severe over or under-reporting of deaths associated with Covid-19 around the world (8). Because of these differences, some policymakers in different regions did not have information on how severe the pandemic was, which may have caused the pandemic to last longer (7). The case reporting faced similar issues; because of the zoonotic nature of the pathogen, the disease definition for the Covid-19 pandemic was not clear (8). The rushed disease definition was not exact and did not address all the characteristics of the disease, causing under or over-reporting (9). The suggested ED indicator does not rely on the knowledge base for the pathogen or the disease. ED are counted based on all deaths accruing in the period, not the cases or deaths indicated by surveillance systems as Covind-19 related. Finally, if analyzed by age groups, the measure might help indicate which age groups are affected more severely by the pathogen.

Different interventions were implemented throughout the Covid-19 pandemic, in most cases using trial and error or analyzing the effects in other countries. This process indicated the need for a new metric that would rely on the information that is available when the new pathogen is introduced. The measure should be able to evaluate the interventions and their effects on the population effectively.

1.3 Aim and Objectives

The study aims to determine different possibilities of using excess deaths metric as an evaluation metric for NPIs used in the Covid-19 pandemic. The objectives for the thesis were to conduct a statistical analysis of excess death measure and stringency scores correlation throughout the pandemic and determine if there is a correlation between the most common interventions and excess death measure.

1.4 Research questions

Can excess death (ED) be used as an evaluation metric for the interventions used in the Covid-19 pandemic?

What is the relationship between EDM and stringency index in different countries?

Which of the nonpharmaceutical interventions show the strongest negative correlation with EDM?

Is there an age effect acting on the relationship between EDM and NPI?

2. Methodology, data, and methods

2.1 Methodology

In this study, the positivistic approach was used, where research focuses on a statistical analysis of the includes attributes. The experimental design was chosen as a study methodology. It can be defined as a type of design with at least one of the independent variables manipulated, one of the variables is controlled, and one or more variables are observed (11). In this study, the correlation coefficients were observed; correlation tests were manipulated, and NPI types, strengths and ED were considered controlled variables.

2.2 Data

Three datasets and two measure types (stringency scores and excess deaths expressed in P scores) for all the countries worldwide that track their excess deaths were used in the project. The stringency indexes indicate a total score of interventions and their strengths awarded to each country at the given time; the intervention list and points awarded for each severity level implemented can be seen in Table 1. Excess deaths were expressed in the P score, which is the percentile excess mortality between the current time frame (2020) and the average over past years (2015-2019) (8). The overall stringency index dataset was downloaded from ourworlddata.org (12); the data contains all countries with daily stringency indexes. The excess death metric dataset was downloaded from ourworldata.org (13) too. For the relationship between different interventions and excess death measures, the dataset from Blavatnik school of government, University of Oxford (14), was used. The research group responsible for the dataset has a GitHub library (15) that includes a frequently updated dataset and explanation for all the variables. The selected interventions and their score meaning are explained in Table 1;

In the project, from the first dataset, overall stringency scores for each country were used. From the second dataset, the overall P scores and P scores for different age groups were used and the third dataset provided data for different interventions and their strengths implemented in different countries. Due to time limitations, just the countries that belong to the EU (Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden.) were chosen for this analysis.

Table 1 Intervention list. *Interventions selected for the analysis and their severity explanations. ()*

Intervention	Rank meaning
School closing	1- no measure 2- recommendations to close 3- the requirement to close just some levels 4- the requirement to close all
Workplace closing	1- No measure 2- Recommended closing or less one hours 3- Recommended closing or some sectors 4- Required closing to all nonessential businesses
Public event limitation	1- No measure 2- Recommended canceling 3- Required canceling
Public transport limitation	1- No measure 2- Recommended closing or significant reduction 3- Required closing
Stay at home requirements	1- No measure 2- Recommended not to leave 3- Required not to leave besides essentials 4- Required not to leave with minimal exceptions
Restrictions on internal movement	1- No measure 2- Recommended not to travel between cities 3- No internal movement

Public information campaigns	1- no campaigns 2- Public official urging caution Coordinated public health information campaigns
Protection of elderly people	1- No measure 2- Recommended isolation 3- Recommended isolation with some limitations to visitors 4- Extensive limitations for visiting required isolation

The analysis was conducted using python v3.8 programming language. For data preparation and engineering, pandas (16) and numpy (17) packages were used. Both packages provide data manipulation functionality and are used as a pre-processing step for the data analysis later.

Statistical analysis will be conducted using *scikit* (18) package. Scikit is a popular package with an extensive application for statistics and machine learning. It includes classification, regression, clustering, dimension reduction, model selection, and pre-processing methods (18).

2.3 Methods

2.3.1 Data preparation

The data preparation began with missing value identification. Missing values or 0 were investigated. If the country has missing values, it was removed from the study. If 0 in the dataset refers to the score or value was 0, no actions were taken. Irland was removed because of the missing values.

2.3.2 The correlation between the measures

The problem was phrased as a time series, and the goal was to calculate a cross-correlation score. Time series-based analysis implies that the data is indexed in time order, and data points cannot be moved between indexes. Cross-correlation of time-series data can be defined as the correlation between two variables, where one variable is displaced (20). Before the cross-correlation can be calculated, the cross-covariance function needs to be defined (20):

Equation 1 Cross-variance g in x-y time series with time-shift k

$$g_k^{xy} = \frac{1}{n}\sum_{t=1}^{n-k} \quad (y_t - \underline{y})(x_{t+k} - x)$$

Then, cross-correlation can be expressed as (20):

Equation 2 Cross-correlation r between x and y time series, with the time shift k. Where SD - standard deviation, g_k^{xy} *- cross variance.*

$$r_k^{xy} = \frac{g_k^{xy}}{\sqrt{SD_x SD_y}}$$

These calculations were used to identify how well the excess death measure correlates to the stringency index in all selected counties, and lag between two variables was used to align them. Afterward, other correlation methods were used, based on the assumption the variables fulfill.

Pearson correlation was used for the two continuous variable correlations. The test assumes that variables are independent, have a linear relationship, and are homoscedastic (21).

Equation 3 Pearsons correlation formula where cov - covariance,σ standard deviation

$$\rho_{x.y} = \frac{cov\ (X, Y)}{\sigma_x \sigma_y}$$

Spearman correlation is a non-parametric correlation coefficient that was used for the data where at least one of the variables is ordinal or does not fulfill the assumptions for Person correlation.

Equation 4 Spearman rank correlation, where d_i^ - difference between the two ranks of each pair, n - no. of observations

$$\rho = 1 - \frac{6\sum \quad d_i^2}{n(n^2 - 1)}$$

Kendall tau is a non–parametric test used with continuous variables alongside cross-correlation and spearman correlation to determine if the lag alignment was necessary for these variables.

Equation 5 Kandel tau formula,where C - no.of concordant pairs,D- no.of discordant pairs

$$\tau = \frac{C - D}{C + D}$$

The chosen correlation will be calculated on a country basis, so high-low mortality rates in different counties do not affect the integrity of the correlation coefficient.

2.3.3 Age groups-based correlation

In this step, the excess death metric was separated into the age groups from the dataset (15-64, 65-74, 75-84, 85+) and correlated to the stringency index. For the correlation, the data with applied lag will be used, based on the assumption fulfilled either Pearson (Equation 3) or Spearman (Equation 4) correlation is going to be used. This step will show how different age groups respond to the NPI intervention severity in different countries.

3. Results

3.1 Cross-correlation and lag

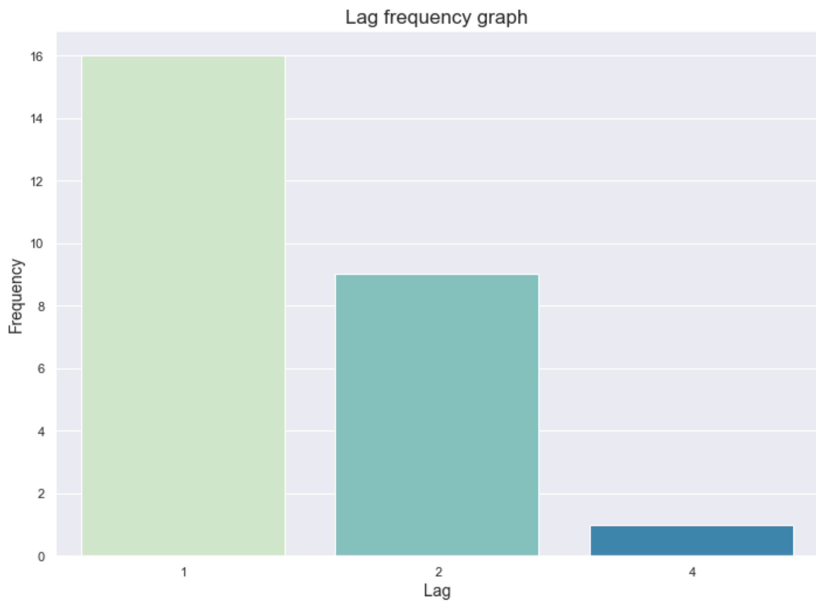

Figure 1 Lag frequency between stringency index and excess deaths, measured in weeks*. The lag was calculated using the cross-correlation method.*

From the cross-correlation test, best-fit lag scores were calculated for each country. Lag for this analysis indicates how many weeks after the change of stringency index, the excess deaths start to decrease. The test was conducted from lag scores in a range of [0; 8], meaning the time period of 8 weeks after the stringency index change was tested, to find the best fitting results. The best fit scores can be seen in **Figure 1**. The most frequent lag score identified was 1 with 16 observations and 2 with 9 observations. The highest score was of 4 was observed once for the data from Hungary. The lag scores suggested that Covid-19 intervention effectivity can be assessed by using excess mortality scores from 1-2 weeks after the implementation. The case of Hungary being outside the 1-2 weeks range might indicate a case reporting issue.

In Figure 2, the comparison of correlation coefficients for three correlation tests for each of the countries can be seen. The tests conducted were: cross-correlation with lag score appropriate of the country, Spearman correlation, and Kendall tau. All correlation coefficients are highly dependent on the country. Results from the correlation follow a similar trend, except Bulgaria, where the Spearman correlation is 0.013, Kandel tau – -0.04, but cross-correlation is -0.15, indicating that cross-correlation was able to show a much stronger correlation than other tests. The strongest coefficients were identified in France (cross-correlation – -0.64, spearman – -0.64, Kendall tau – -0.5) and Italy (cross-correlation – -0.73, spearman – -0.756, Kendall tau – -0.55). The weakest coefficients were observed in Bulgaria (Spearman correlation- 0.013, Kandel tau – -0.04, cross-correlation is -0.15), Finland (cross-correlation – -0.14, spearman – -0.14, Kendall tau – -0.09), Greece (spearman – -0.15, Kendall tau – -0.11)). This wide range of correlations between countries indicates that the relationship between index and excess deaths might be more complex.

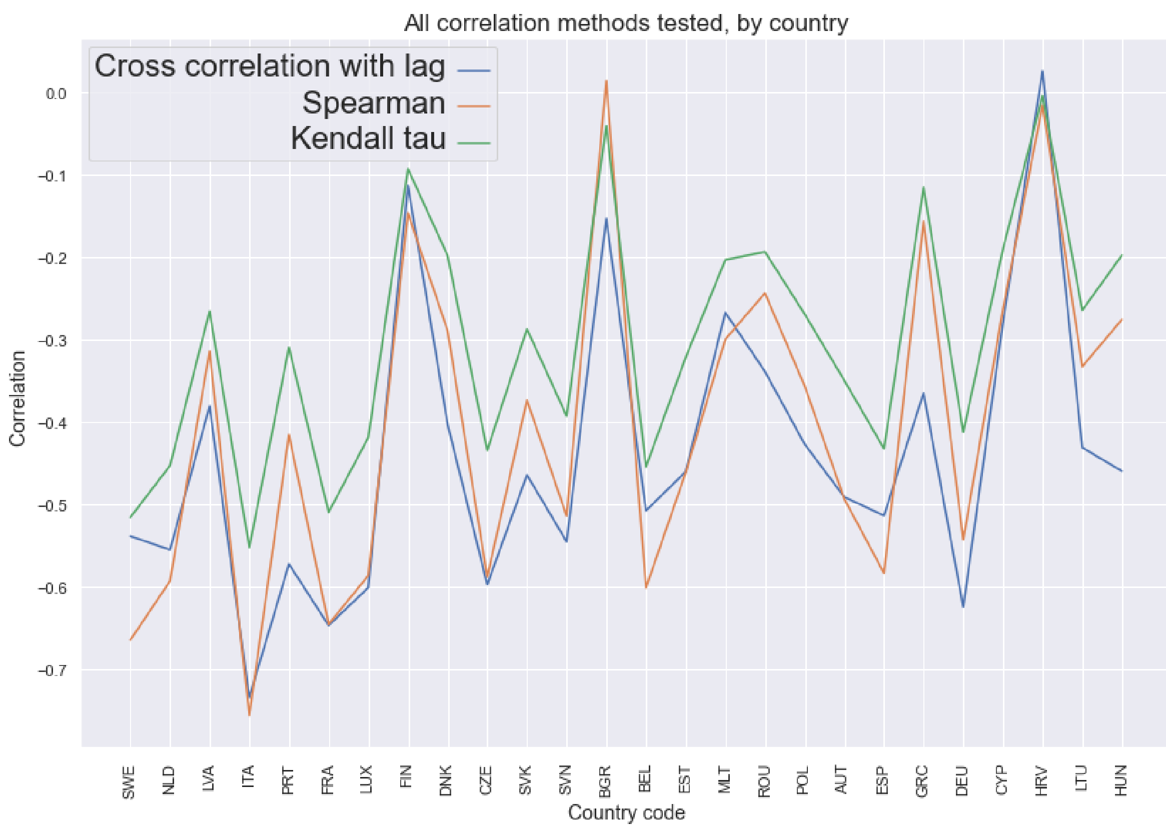

Figure 2 ***All correlations between stringency index and excess death coefficients, categorized by country.*** *A negative correlation indicates that a higher stringency index (more strict interventions) correlates with lower P-scores.*

At the same time, only the spearman correlation test for data from Finland had a p-value higher than the 0.05 threshold (0.2). The most significant difference between the methods used can be seen between Greece (cross-correlation – -0.36, spearman – -0.15, Kendall tau – -0.11), where the cross-correlation method was the best, and Italy (cross-correlation – -0.73, spearman – -0.75, Kendall tau – -0.552), where Kandel tau method shown a significantly weaker correlation. Overall, cross-correlation with lag had stronger correlation coefficients on average.

3.2 Correlation between different NPI harshness and excess mortality

The correlation coefficient between excess mortality and NPI harshness ranges from 0.46 to -0.77, indicating an overall low to medium correlation. The strongest correlation was detected with gathering restrictions (average -0.42), workplace closing (average -0.42). The weakest correlation was observed with public transport limitations (average -0.24), international travel restrictions (average -0.24), and restrictions on internal movement(average -0.25).

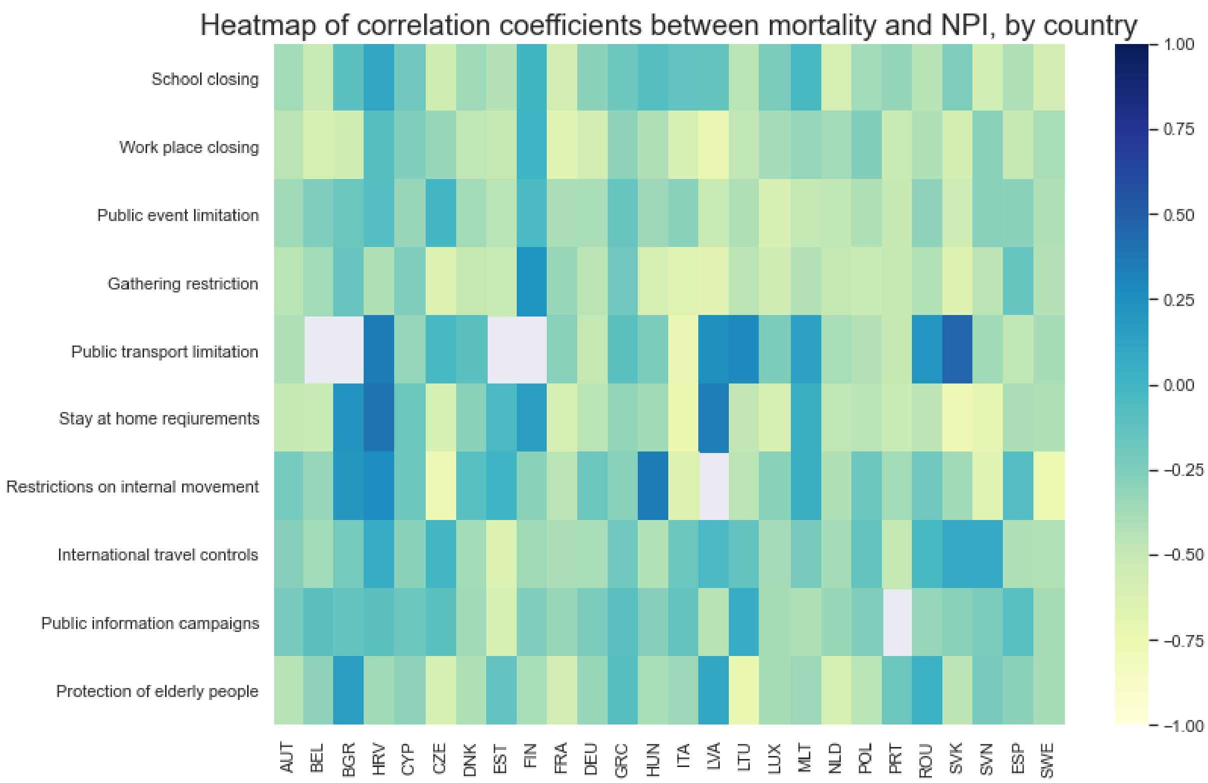

Figure 3 Heat map of correlation coefficients between excess mortality and nonpharmaceutical interventions*. The Pearson correlation method was used with data from all selected EU countries. Grey squares indicate a missing stringency index for the interventions in the country. The missing information might indicate an error in reporting, or there was no intervention implemented*

Strong above 0 correlations were identified in Croatia, with public transport limitations (0.34) and stay at home requirement (0.39), in Hungary with internal movement restrictions (0.34), in Latvia with a stay at home requirement (0.33), and in Slovakia with public transport limitations (0.46).

Overall public transport limitation had the widest range of correlations (0.46- -71), indicating possible adherence to the limitations issue. On the other hand, this variable is highly dependant on the country's infrastructure development level, so such a wide range of coefficients might be an indicator of that. Public even limitation was the only intervention with only the negative correlation coefficients (ranging from -0.01- -0.6), indicating the high independence from the population or economic differences. The correlation coefficients were lower than the statistical significance p-values threshold of 0.05, indicating their statistical significance.

3.3 Correlation based on age groups

In the **Figure 4** the Spearman correlation coefficients between stringency indexes and excess deaths in different age groups is shown. All of the correlation coefficients were statistically significant, with p-values < 0.05. The correlation ranges from 0.15 – -0.71, indicating an overall medium correlation.

The lowest correlation can be seen in the youngest age group, 15-64 (average -0.28), while the highest correlation of -0.39 average can be seen in the 75-84 age group. Surprisingly age group 85+ had a little lower correlation than the 75-84 age group. This might be caused by a smaller sample size for the 85+ age group, or be indicative of a smaller social circle, thus concluding to fewer contacts. The 85+ age group in most countries has a lower correlation than the younger group, so any cultural or country-specific implications can not be considered.

When separated by the countries, the strongest correlation can be seen in Czechia (respectively -0.6, -0.71, -0.63, -0.49) and Italy (-0.53,-0.59,-0.59,-0.57). The weakest correlation was seen in Croatia (0.045, 0.071, 0.033,0.047), being the only country with all positive correlation coefficients. The general trend of stronger correlation with the aging population indicates that interventions are more effective at protecting the older generation.

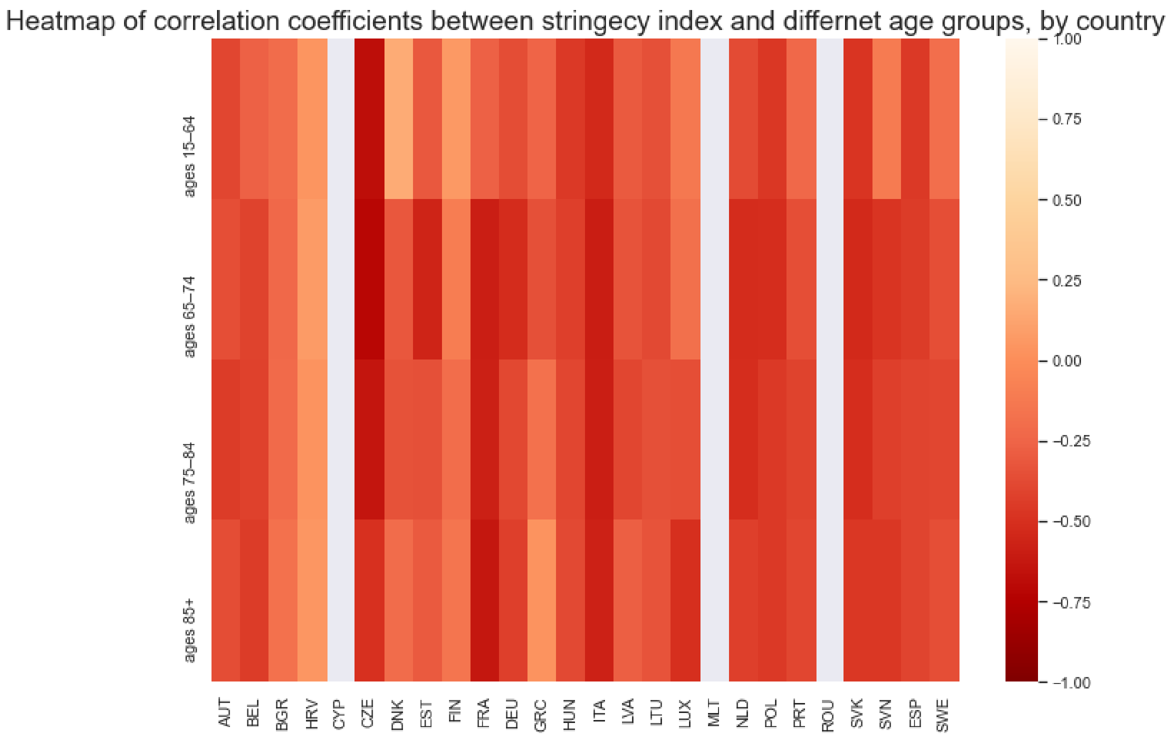

***Figure 4 Heatmap of correlation coefficients between stringency index and excess deaths in different age groups**. Spearman's correlation was used to determine the correlation between the stringency index and excess deaths in different age groups. Grey countries had missing age-separated excess death data.*

Next, the correlation coefficients were calculated for each age group to every intervention method to determine how stricter interventions affect excess mortality in different age groups, the calculation was done for each country separately. In figure 5, the averages of all counties are represented. Again, all of the coefficients had a p-value lower than 0.05. The overall correlation coefficient ranged from -0.04 to -0.34, indicating low to medium correlation. The weakest correlated intervention was public transport limitation, ranging from -0.04- -0.13, international travel control (-0.093- -0-18), and elderly protection campaigns (-0.083 - -0.24) increasing with age groups. The strongest correlation overall age groups was with workplace closing ranging -0.26- -0.34. Similarly, school closing and public event limitation had a relatively strong (-0.29- -0.31) correlation in 65+ age groups. Surprisingly, public information campaigns had a relatively weak (-0.11- -0.17) correlation within all age groups. As in the analysis before, the 85+ group had weaker or the same correlation as 75-84 age group.

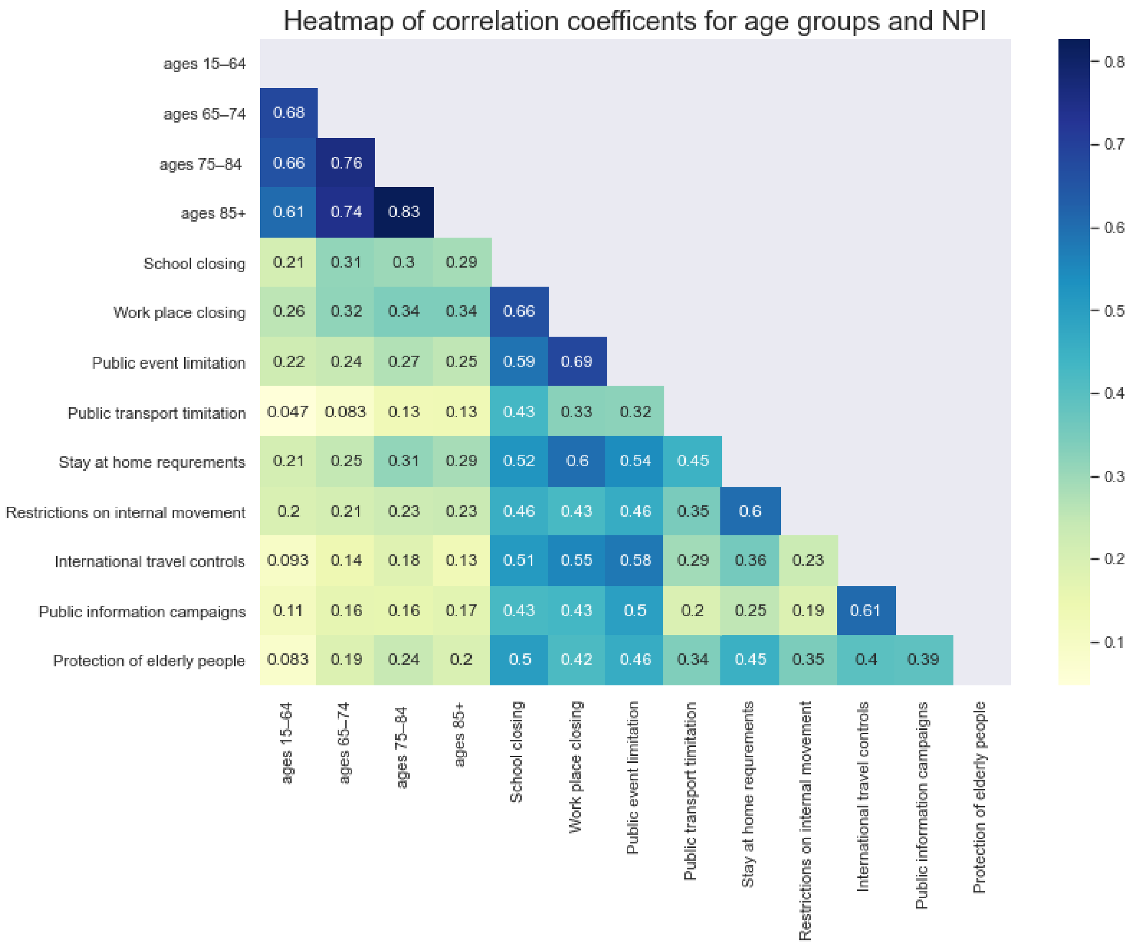

Figure 5 Correlation coefficient heatmap between different age groups and the NPI strictness.*Spearmans correlation was used to determine how effective different interventions and their strictness is in different age groups during the pandemic.*

4. Discussion and conclusion

The goal of the suggested methodology was to evaluate how NPIs affected the excess mortality rates in the 2020 Covid-19 pandemic. At the same time, the excess death mortality measure was analyzed as an evaluatory metric for the estimation of NPI effectiveness. This was achieved by conducting three different tests for the correlation of excess death mortality and stringency index, indicating the lag scores for each country, and applying them to the data. Afterward, analysis of the correlation of different NPIs and their influence on the excess death measure and an age-based analysis of NPIs were conducted.

The excess mortality measure has shown a low to moderate correlation with different stringency indexes. However, the specific interventions with a higher correlation shown in **Figure 3** correspond with the research with a similar goal (23), implying that the lag indicated by cross-correlation and excess mortality provided similar results to a more advanced methodology. Even though the cross-correlation between independent variables has its criticisms(20,21), overleaping results with different approaches support the possibility to use the method

The correlation seems to be higher in countries with higher death rates seen in Figure 2. Simultaneously, through the results in **Figure 4** and comparison to similar research, it can be suggested that excess death metric reflects in the stringency index. However, data manipulation through the lag application might be necessary. This might indicate that the excess mortality does not directly correlate to the interventions, but the relationship is delayed, which would support the general understanding of NPI effectiveness and timeliness when assessed with death-related metrics.

Moreover, the best correlation was detected with school closing, workplace closing, public event limitation, gathering restriction, and stay-at-home requirements. All of these interventions affect the lives of large parts of the population. This leads to the assumption that interventions that apply to a bigger part of the population have better chances of reducing the deaths caused by the outbreak.

Finally, the analysis based on age showed significant differences between interventions based on the age groups. This conclusion suggests that assessment of the intervention effectiveness based on age could assist in outbreak management and should be investigated further, with different outbreaks or smaller age ranges. Smaller age groups could provide an insight into the effects of the interventions for working-age (18-64), school-age (7-18) populations since they have the most extensive social circles and the highest amount of contact with other people.

4.1 Results interpretation

The overall correlation was calculated for each country. A stronger correlation was detected in the countries with higher deaths per capita. For example, Italy, France, Luxembourg, the Czech Republic, and Belgium had 280- 133.4 deaths per 100 000 pop., while the lowest death rate in Europe was 8.6 per 100 000 pop. in Iceland (23). Even though the counties mentioned were not in the top 5 (highest Hungary 306 per 100 000 pop.), all of them were ranked in the upper third highest deaths rated in Europe (23). While this does not provide a direct conclusion, it can be assumed that all correlation methods have better applicability on the data with more considerable variance.

The analysis between different NPI strengths and their correlation with excess mortality was evaluated. In this step, the two groups of effectiveness can be established (Figure 8), more correlated (school closing, workplace closing, public event limitation, gathering restriction, and stay art home requirement) and lower correlation (public transport limitation, restriction on internal movement, international travel control, public information campaigns, protection of elderly campaigns). From these groups, it can be gathered that NPI effectiveness depends on population size. In contrast, school and workplace closing will affect the significant size of the population, and public transport limitation is highly dependant on the infrastructure and travel culture of the country. Similarly, internal or international travels will affect a minor part of the population that woulöd choose to travel in a pandemic.

Age group-based analysis shown the correlation became stronger with the age increase. Two of the NPIs had a more significant gap going from 15-64 to 65-74 gap. These interventions were school closing, workplace closing, which is an exciting and unexpected outcome, suggesting that reduced working population movement slows down the spread among the elderly population.

Other research analyzing reproductive rate changes indicated that school closing, business closing, gathering limitation, (24) traveling restrictions, and workplace closing (25) had a higher effect.

4.2 Further research

With advancements in computer and data sciences, advanced techniques are becoming more accessible and explainable. Epidemiology produces high amounts of data, which would provide enough test/training information for machine learning-based methods. Outside the scope of the study, the multi-layer perceptron (MLP) based time series forecasting algorithm was tested. The model results showed that not enough data was provided for the training process of the model. The model should be tested again with a longer timeline.

By conducting more comprehensive analyses of nonpharmaceutical interventions, a knowledge database could be developed, indicating which interventions and strengths are best suited for different types of outbreaks. This goal would require much research, collaboration, and investment, but the possible result could be groundbreaking and applicable worldwide.

4.3 Limitations

The proposed idea of excess mortality metric had its strengths and downfalls. Even though the over or under-reporting of cases is not an issue when using excess mortality, the information provided by the analysis is limited. The metric indicated changes in death, but just the most severe cases in an outbreak end in death, so it is hard to estimate how widespread the outbreak is. It is also hard to formulate a conclusion that would indicate the severity of an outbreak.

The method chosen for this thesis can be improved by a more detailed analysis of the statistical tests available and applicable for the chosen problem. Similarly, the more advance, AI-powered techniques could be tested. From the analysis, it can be assumed that the relationship between variables might be more complex than the statistical test can detect; therefore, more advanced techniques like artificial neural network models could have been used.

4.5 Conclusion

The objective of this thesis was to evaluate NPI and determine to what extent excess mortality can be used as an evaluator. The results show promise because they do reflect similar research that used different indicators. Different countries proved to have different relationships between the chosen metrics, indicating that population, culture, and economics of the environment need to be considered when evaluating the interventions. Similarly, the connection between stronger negative correlations and how much of the population interventions are aimed at was established. The increasing effect of the interventions was observed with the increase of age.

Even though the correlation found was not strong, it is reasonable to assume that the death metric will not correlate directly with the interventions that do not target extreme cases of infection. In this case, a lower correlation can be expected because the interventions are targeted at a healthy population. At the same time, the excess mortality metric explains how the pandemic affected the number of deaths. Thus, the thesis provides an insight into how these effects can be interpreted. On the other hand, cofounding can be affecting the correlation too. Excess mortality does not refer to any specific reason for the death. In this case, Covid-19 might be causing the increase, but it can not be assumed that it is the only reason.

5 Reference list

1. Campbell ME, Gardner CE, Dwyer JJ, Isaacs SM, Krueger PD, Ying JY. Effectiveness of public health interventions in food safety: A systematic review. Can J Public Heal. 1998;89(3):197–202.
2. Masters R, Anwar E, Collins B, Cookson R, Capewell S. Return on investment of public health interventions: A systematic review. J Epidemiol Community Health. 2017;71(8):827–34.
3. Aledort JE, Lurie N, Wasserman J, Bozzette SA. Non-pharmaceutical public health interventions for pandemic influenza: An evaluation of the evidence base. BMC Public Health [Internet]. 2007 Dec 15 [cited 2021 Apr 8];7(1):208. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-7-208
4. WHO. Nonpharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. A checklist for pandemic influenza risk and impact management: building capacity for pandemic response [Internet]. 2019. Available from: https://www.who.int/publications/i/item/non-pharmaceutical-public-health-measuresfor-mitigating-the-risk-and-impact-of-epidemic-and-pandemic-influenza
5. Aiello AE, Perez V, Coulborn RM, Davis BM, Uddin M, Monto AS. Facemasks, hand hygiene, and influenza among young adults: A randomized intervention trial. PLoS One [Internet]. 2012 Jan 25 [cited 2021 Apr 8];7(1). Available from: https://pubmed.ncbi.nlm.nih.gov/22295066/
6. Torner N, Soldevila N, Garcia JJ, Launes C, Godoy P, Castilla J, et al. effectiveness of nonpharmaceutical measures in preventing pediatric influenza: A case-control study. BMC Public Health [Internet]. 2015;15(1):1–8. Available from: http://dx.doi.org/10.1186/s12889-015-1890-3
7. Staadegaard L, Taylor RJ, Spreeuwenberg P, Caini S, Simonsen L, Paget J. Monitoring the mortality impact of COVID-19 in Europe: What can be learned from 2009 influenza H1N1p mortality studies? Int J Infect Dis [Internet]. 2021;102:115–7. Available from: https://doi.org/10.1016/j.ijid.2020.10.037
8. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. N Engl J Med [Internet]. 2020 Mar 26 [cited 2021 Apr 8];382(13):1194–6. Available from: http://www.nejm.org/doi/10.1056/NEJMp2002125
9. Beaney T, Clarke JM, Jain V, Golestaneh AK, Lyons G, Salman D, et al. Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide? J R Soc Med. 2020;113(9):329–34.
10. Positivism - Research Methodology [Internet]. [cited 2021 May 2]. Available from: https://research-methodology.net/research-philosophy/positivism/
11. Kothari CR. Reaserch methodology: methods and techniques. 2nd ed. New age international publisher; 2004. 414 p.
12. COVID-19: Stringency Index [Internet]. [cited 2021 May 2]. Available from:

https://ourworldindata.org/grapher/covid-stringency-index?tab=chart&time=earliest..latest

13. Excess mortality during the Coronavirus pandemic (COVID-19) - Statistics and Research - Our World in Data [Internet]. [cited 2021 May 2]. Available from: https://ourworldindata.org/excess-mortality-covid

14. COVID-19 Government Response Tracker | Blavatnik School of Government [Internet]. [cited 2021 May 2]. Available from: https://www.bsg.ox.ac.uk/research/research-projects/Covid-19-government-response-tracker

15. Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav. 2021 Apr 1;

16. pandas - Python Data Analysis Library [Internet]. [cited 2021 May 2]. Available from: https://pandas.pydata.org/

17. NumPy [Internet]. [cited 2021 May 2]. Available from: https://numpy.org/

18. scikit-learn: machine learning in Python — scikit-learn 0.24.2 documentation [Internet]. [cited 2021 May 2]. Available from: https://scikit-learn.org/stable/

19. What is a z-score? — EUROMOMO [Internet]. [cited 2021 May 2]. Available from: https://www.euromomo.eu/how-it-works/what-is-a-z-score/

20. Holmes H.H, Scheuerell M. D., Ward E. J. 4 Basic timeseries functions in R. In: Applied Time Series Analysis for Fisheries and Environmental Sciences [Internet]. 2021 [cited 2021 May 2]. Available from: https://nwfsc-timeseries.github.io/atsa-labs/

21 Dean, R.T., Dunsmuir, W.T.M. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. Behav Res 48, 783–802 (2016).

22. Pearson’s Correlation Coefficient - Statistics Solutions [Internet]. [cited 2021 May 19]. Available from: https://www.statisticssolutions.com/pearsons-correlation-coefficient/

23. Mortality Analyses - Johns Hopkins Coronavirus Resource Center [Internet]. [cited 2021 Jun 15]. Available from: https://coronavirus.jhu.edu/data/mortality

24. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the Effectiveness of government interventions against COVID-19. Science (80-) [Internet]. 2021 Feb 19 [cited 2021 May 20];371(6531). Available from: https://doi.org/10.1126/science.abd9338

25. Liu Y, Morgenstern C, Kelly J, Lowe R, Munday J, Villabona-Arenas CJ, et al. The impact of nonpharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. BMC Med [Internet]. 2021 Dec 1 [cited 2021 May 20];19(1):1–12. Available from: https://doi.org/10.1186/s12916-020-01872-8