Influenza Vaccination Rates Predict 30% of the Variance in Covid-19 Related Deaths in Europe – A Modeling Approach

Harald Walach[1,2,3] & Rainer J. Klement[4]

1 Poznan Medical University, Dept. Pediatric Gastroenterology, Poznan, Poland

2 University Witten-Herdecke, Dept. Psychology, Witten, Germany

3 Change Health Science Institute, Berlin, Germany

4 Department of Radiation Oncology, Leopoldina Hospital Schweinfurt, Germany

Contact:

Prof. Harald Walach

CHS-Institute

Schönwalder Str. 17

D – 13347 Berlin

harald.walach@uni-wh.de

Abstract

There is a high variability in Covid-19 related deaths whose origin is unclear. We used three variables, percent test-standardized number of SARS-CoV-2-cases in a country, influenza-vaccination coverage in the elderly in a country and number of non-pharmaceutical interventions, to predict the number of population standardized Covid-19 related deaths in European countries, using generalized linear models. With these variables we can clarify approximately 60% of the variation in Covid-19 related deaths, with flu-vaccination coverage in the elderly being the most important predictor, explaining nearly 30% of the variation. Thus, the higher the influenza vaccination coverage in the elderly in a country, the more Covid-19 related deaths we see. Also, the more non-pharmaceutical interventions, the more deaths are likely, explaining about 5% of the variation. Other variables, like life-expectancy, rapidity of a country's reaction to the epidemic and population size did not emerge as significant predictors. Thus, contrary to current opinion, flu-vaccination in the elderly might be an aggravating factor, when it comes to Covid-19 related deaths.

Influenza Vaccination Rates Predict 30% of the Variance in Covid-19 Related Deaths in Europe - A Modeling Approach

Introduction

Current public opinion seems to assume that a death following the infection with SARS-Corona-Virus 2 (CoV2) is largely due to this virus, because of its virulence. Untreated Covid-19 disease may lead to severe atypical pneumonia [1,2], a cytokine storm and other potentially lethal sequelae [3-5]. Other potential factors, such as host factors or population factors, are not much considered. We know that initially mostly elderly patients with a mean age above 70 years have been severely affected [6-8]. But in due course also younger patients became severely ill. However, there is a wide variation across countries and regions. This variation is partially shrouded by the fact that most agencies and their dashboards propagate unstandardized figures of cases and deaths. A recent publication that estimated excess death rates in the US during the time of the CoV2 pandemic as compared with the same months of previous years reveals a wide variation from -71,9 deaths per 100.000 inhabitants in North Dakota to 299,1 deaths per 100.000 inhabitants in New York City [9], with seven states actually exhibiting less excess mortality than in the previous comparison years, and 12 US states presenting with excess mortality figures below 10 per 100.000 inhabitants. The same is true for Europe: Miles and colleagues [10], citing various sources, list excess deaths of 21% for Spain, 20% for the UK, 18% for Italy down to 6% for Sweden, 3% for Portugal, -1% for Germany, -3% for Denmark and -4% for Norway. Inspecting the European Mortality Database (https://www.euromomo.eu/graphs-and-maps) one can see a large, but very sharp peak in excess mortality of 88.598 at the peak across all European countries that exceeds every other peak by a z-score of 56,68 and thus is without doubt a clear sign of excess mortality during the weeks 8 to 22 of 2020, but looking at single countries one can see again that this peak is mainly due to excess deaths in Belgium, France, Italy, Spain, Netherlands and the UK, and to some degree in Sweden and Ireland, while in other countries there is rather a negative excess of mortality. This is so despite the fact that some of the severely affected countries, like Spain, Italy, France have imposed severe restrictions on their populations, banned gatherings and issued stay-at-home orders, while others that have comparatively mild figures of excess deaths like Sweden or negative excess deaths like Norway have ordered less or no restrictions.

Thus, there is clearly a need to identify other drivers of mortality than the infection itself, or rather, to understand what might mediate the course from infection to death. Are there population variables, public health variables, or host factors that can be identified that make this variation understandable? This was the guiding question of this modeling study. In a companion paper [11] we used data from the European Center for Disease Prevention as of 15th of May. We found that the main predictors of CoV2-Cases are life-expectancy, i.e. age, number of tests, and the fact whether borders were closed or not, and that the main predictors of Covid-19 related deaths are again life-expectancy, the time the virus had been in the country and school closures, which was a positive, i.e. accelerating predictor. This study used a diverse set of countries, and a point can be raised that not all of them had reliable data at the time, and also that the 15th of May was somewhat early. Meanwhile an interesting paper showed that the influenza vaccination rate in the elderly was significantly correlated at $r = .68$ with Covid-19 related deaths in Europe [12]. This, and further developments prompted us to

repeat our modeling exercise with a different set of countries and data. We used exclusively European data as published on the site "Worldometers" (https://www.worldometers.info/coronavirus/#countries) as of 30th of August, restricted the analysis to standardized deaths (cases per 1 million inhabitants as presented by the database and checked against population size), and used as predictors test-standardized case numbers, population and life-expectancy data, as well as influenza vaccination rates, the number of non-pharmaceutical measures and the rapidity of action in a country as predictors. We found that the most important predictor is the percentage of flu-vaccination rate in the elderly (explaining about 30% of the variance), followed by the test-standardized number of cases (explaining another 26% of the variance) with number of non-pharmaceutical interventions as a positive, i.e. accelerating predictor, explaining 7% of the variance.

Results

The results of the generalized linear models assuming a Gamma distribution for the outcome variable are presented in Table 1, using the full model with all predictors, the two simplest models with only one predictor (influenza vaccination coverage and test-standardized number of cases, respectively), and the best model with three significant predictors. It is obvious that the full model using all predictors had a slightly worse fit than the best model (Akaike Information Criterion - AIC =446.9 versus 444.9). The most parsimonious models that would assume that Covid-19 related deaths are only driven by flu vaccination coverage or the number of infections, respectively, are obviously too simple and had comparatively bad fits (AIC = 467.1 for flu vaccination coverage and AIC=478.1 for test-standardized number of cases). The three parameters that are useful predictors are the flu-vaccination rate in the elderly, the number of test-standardized cases and the number of NPIs (model 4 in Table 1). Note that all three parameters are positive, i.e. we see more deaths in countries with more test-standardized cases, which is expected, and in countries with a *higher* flu-vaccination coverage and *a greater number of* political interventions, which is counter-intuitive. Thereby, the flu vaccination rate was able to explain most of the variance with a Kullback-Leibler-based R^2 value of 0.323 (adjusted 0.304)[13], which was further increased by 0.26 by adding the test-standardized cases as a predictor (Table 1).

For a more natural interpretation on the amount of variance that could be explained by the various predictors, we also calculated a standard multiple linear regression model on the ln-transformed outcome. The result for the best model built via feature forward selection is given in Table 2. The stepwise procedure used again flu vaccination rate as the strongest predictor first, which explains 28% of the variance, followed by test-standardized cases, which explains another 25% of the variance, and finally the number of NPIs, which explains an additional 7% of the variance. Note that the results replicate the Gamma-generalized linear model (Table 1) in strength and sequence of the predictors. In addition, this last model gives us a more widely known estimation of the variance explained. The full model explains 57% of the variance ($R^2_{adj.} = 0.57$), very similar to the corresponding Gamma-generalized linear model (Kulback-Leibler-based $R^2_{adj.} = 0.60$).

Residual diagnostics (inspection of normal-probability plots, residual distribution, distributions of predicted versus observed values) confirmed that the models were adequate. After removal of the outlier Albania the linearity assumptions were met.

Table 1 - Result of generalized linear models (intercept not reported)

Model 1 – Full Model	**Parameters**			**Goodness-of-fit**				
Variables	Coefficient Estimate (SE)	Standardized regression coefficient estimate (SE)	p-value	KL-R^2	KL-$R^2_{adj.}$	AIC	Log likelihood	Deviance/DF
Flu vaccination coverage (%)	0.0260 (0.0062)	0.615 (0.147)	2.22×10^{-4}	0.673	0.610	446.9	-215.46	1.81
Test-standardized cases (%)	0.239 (0.042)	0.581 (0.103)	3.58×10^{-6}					
Number of NPIs	0.158 (0.078)	0.240 (0.118)	0.051					
Life expectancy (years)	0.069 (0.038)	0.231 (0.129)	0.083					
Rapidity of reaction (days)	0.0086 (0.012)	0.128 (0.176)	0.473					
Population size (10^6)	5.8×10^{-4} (5.7×10^{-3})	0.018 (0.172)	0.919					
Model 2 – Best univariable Model	**Parameters**			**Goodness-of-fit**				
Variable	Coefficient Estimate (SE)	Standardized regression coefficient	p-value	KL-R^2	KL-$R^2_{adj.}$	AIC	Log likelihood	Deviance/DF

		estimate (SE)						
Flu vaccination coverage (%)	0.0266 (0.006)	0.629 (0.143)	9.18×10^{-5}	0.323	0.304	467.1	-230.53	10.00
Model 3 – 2nd-best univariable Model	**Parameters**			**Goodness-of-fit**				
Variable	Coefficient Estimate (SE)	Standardized regression coefficient estimate (SE)	p-value	KL-R^2	KL-$R^2_{adj.}$	AIC	Log likelihood	Deviance/DF
Test-standardized cases (%)	0.197 (0.074)	0.478 (0.179)	0.012	0.127	0.103	478.1	-236.04	12.89
Model 4 – Best Model	**Parameters**			**Goodness-of-fit**				
Variables	Coefficient Estimate (SE)	Standardized regression coefficient estimate (SE)	p-value	KL-R^2	KL-$R^2_{adj.}$	AIC	Log likelihood	Deviance/DF
Flu vaccination coverage (%)	0.035 (0.004)	0.827 (0.106)	4.18×10^{-9}	0.323	0.304	444.9	-217.43	3.20
Test-standardized cases (%)	0.238 (0.041)	0.579 (0.099)	1.35×10^{-6}	0.583	0.559			
Number of NPIs	0.189 (0.069)	0.288 (0.105)	0.0096	0.639	0.608			

AIC: Akaike Information Criterion; KL-R^2: Kullback-Leibler-based R^2; KL-R^2_{adj}: adjusted Kullback-Leibler-based R2; SE: Standard error

Table 2 - Result of multiple linear regression of the three predictors (intercept not reported) from best model on ln-transformed outcome: Model $R^2_{adj.} = 0.574$, $F(3/34) = 17.64$, $p = 4.4 \times 10^{-7}$

Variable	Regression coefficient (SE)	Standardized regression coefficient estimate (SE)	t (34)	p-value	R^2	Change in R^2
Flu vaccination coverage (%)	0.0344 (0,006)	0.813 (0.134)	6.07	7.04×10^{-7}	0,28	0,28
Test standardized cases (%)	0.24 (0,05)	0.588 (0.125)	4.68	4.40×10^{-5}	0,54	0,25
Number of NPIs	0.22 (0,08)	0.332 (0.133)	2.50	0.0176	0,61	0,07

SE: Standard error

Discussion

This modeling exercise, predicting Covid-19 related death rates in European countries, unravels some interesting findings:

a) Unsurprisingly, test-standardized CoV2-cases predict the number of deaths. This variable explains about 25% of the variance.
b) Surprisingly, more important is the flu-vaccination coverage in the elderly: the higher this vaccination rate is, the more Covid-19 related deaths we see in a country. This is even more important than the number of cases and explains about 30% of the variance.

These findings are strengthened by the fact that two different models reach the same conclusions: a generalized linear model predicting a gamma-distributed outcome variable with log-linked predictors and a standard multiple linear regression model with identity link functions of predictors on a log-transformed outcome variable. The best model also includes the number of non-pharmaceutical interventions as a positive predictor, i.e. there are *more* deaths in countries with more NPIs implemented. This predictor explains roughly another 5% of the variation. It is reassuring that both models reach the same conclusions in terms of importance and sequence of the predictor variables.

How might these findings be explained? It is easy to understand that more CoV2 cases translate into more Covid-19 related deaths. What is unexpected is the fact that the importance of this predictor is comparatively minor. That such a model is too simple can be seen when comparing the three models (full model 1, simple model 3 and adequate model 4) in Table 1. The univariable model 3 using only number of test-standardized cases is not adequate and produces the worst fit of all models. Thus, obviously, there remains variance to be explained. Far from assuming that we have captured all variables, we have captured at least some and thus are able to explain about 60% of the total variance with these variables. Most surprising and most counterintuitive are the two findings that there are more Covid-19 related deaths in countries with higher flu vaccination coverage in the elderly, and, in addition, that the number of non-pharmaceutical interventions is a positive predictor of Covid-19 related deaths.

How can this negative impact be explained? A meta-analysis of 60 influenza outbreaks in care homes for the elderly found that the attack rate varied widely from 1.3% to 65%, mostly dependent on non-pharmaceutical and some pharmaceutical interventions [14]. But the influenza vaccination rate was unrelated to attack rate. Although the influenza vaccination rate increased over the years, the attack rate did not decrease. A careful randomized trial of flu vaccination in children showed that children who were vaccinated against influenza were better protected against influenza but suffered a fourfold risk of other respiratory virus dependent diseases [15]. This might have to do with unknown mechanisms that disturb the ecology of pathogens, known as the virus interference phenomenon. A study conducted during the 2017/2018 influenza season revealed that flu vaccination was associated with a 36% increased odds of contracting respiratory coronavirus diseases (odds ratio 95% confidence interval 1.14-1.63, $p<0.01$), while affording specific protection against influenza and parainfluenza viruses[16]. A recent study showed that influenza vaccination in people over 65 was ineffective in preventing mortality or hospitalizations due to influenza [17].

Thus, the negative impact of flu vaccination might have to do with several mechanisms: First, the virus interference phenomenon as shown for non-CoV2 coronaviruses[16]; second, the fact that the immunological load on an organism that has to deal with a flu vaccine binds resources that cannot be mustered against a new and dangerous pathogen like CoV2. Third, it might also be the case that immune-enhancers in vaccines, such as aluminum derivates, which are potentially toxic, burden the organism and hamper natural immunity. It has been shown that aluminum toxicity is a widely underreported and unrecognized issue [18-20]. Thus, widely propagated flu vaccination programs, especially for the elderly, might be problematic when at the same time another pathogen, such as CoV2, is arriving that needs the full prowess of the immunity of a population. Furthermore, it has been argued that influenza vaccines are produced in eggs and other cell-systems that are not routinely tested against corona-viruses. Hence, corona-virus proteins from other corona-viruses might be present in these vaccines and induce allergic reactions against the novel CoV2. [21] Our finding is in contrast to data from the US[22,23]. However, the correlation between influenza vaccination and COVID-19 death rate in the US is much lower than in Europe[12], probably because there is little variation in influenza vaccine coverage in the US. Our results are derived from population level data in Europe in the elderly, which might be a specifically susceptible fraction of the population.

Non-pharmaceutical interventions are widely hailed in modeling studies as having prevented higher incidence figures of cases and deaths e.g. [24,25,26]. While this might be true for some countries and some single interventions, we [27-29], and others [30-33], are skeptical. Careful modeling studies for Germany, for instance, show that, although Germany was comparatively early to react – first measures were introduced on March 8 and shortly after this a full country lockdown was enacted – the peak of the infection and of the reproduction numbers was reached in nearly all 420 German districts on or around March 8 and thus none of the NPIs could have been causally related to the reduction of cases, and hence deaths [34,35]. A careful study of the cases of the first outbreak in Wuhan, China,[36] shows this as well [37,38]. The cases can be seen to peak on the 21st of January (there is a later peak a few days later, but this is due to a reporting outlier as the authors explain). This is the same day as the Chinese spring festival ends that had brought guests and travel to Wuhan and from Wuhan. After that the number of cases drops. The cordon sanitaire and lockdown of the airport and the city started on the 23rd of January. Considering that the reporting delay has to be also reckoned with, the true peak of the infections was likely already earlier. This means: even in Wuhan was the peak of the infection reached before the public health service had time to react. The ensuing reduction of cases is a misattribution: it is not due to the lockdown, but obviously to the fact that the virus follows its own dynamic which needs to be better understood. Thus, there is evidence that non-pharmaceutical interventions are less effective than often thought. Our analysis suggests: they seem to be rather an expression of fear and an escalating situation in a country. This would explain the positive association with Covid-19 related deaths. The association is not strong, but highly significant. Countries that saw a steep rise in cases and associated hospitalizations were of course more inclined to react by implementing more NPIs than others, and hence the number of NPIs might have been an expression of a potentially dangerous development. In our analysis a larger number of NPIs did not act as a preventive, else we would see a negative sign of the predictor. This does not mean that some NPIs, such as wearing face masks in high-risk settings or preventing mass gatherings might not have been important. This question cannot be answered by our analysis. But what we can clearly say is that when it comes to NPI the equation the more the better is simply not true.

We find it quite remarkable that only three variables help to explain roughly 60% of the variation in Covid-19 related deaths. It might be important to study other potential variables pertaining to the host[39], such as vitamin D status[40], which we were unable to study. In our other modeling study [11] we did not find evidence that health related variables, such as obesity, physical activity, or diabetes had a relevant influence. We think that the exposure to heavy metals, perhaps mercury, that are immunologically relevant as immune-suppressors, or widely used pharmacological substances such as anti-inflammatories might be relevant, but those data are difficult to glean. Another potentially important variable that has recently emerged, blood groups, we did not consider either [41,42].

One might criticize our study on various accounts and the limitations of such an approach need to be kept in mind:

First, there might have been collinearity between the three predictors found to be significant in the best model. However, both scatterplots (Supplementary Figures 3 and 4) and variance inflation factors (all <1.162) showed that there was no collinearity between these three variables.

Second, we were unable to find vaccination data for all countries. Although it might have been a strategy to interpolate those cases with neighboring countries, for instance Andorra by Spain, or San Marino by Italy, or Monaco by France, or using multiple imputation by chained equations, we did not do that, because we have not enough knowledge about the situation in those countries. This reduced the number of cases, variability and also power. But statistical power was certainly not the problem of this study, as the clear significances show. We avoided overfitting as much as possible, and the goodness-of-fit statistics of the generalized linear models support this.

Third, one potential problem we cannot remedy is the notorious unreliability of data or differences in the definition of cases, of deaths, and in reporting standards. This can be seen in the fact that Belgium is a clear outlier in all analyses that decreases the fit of the model. It is well known that the definition of Covid-19 related deaths in Belgium is more lenient than in other countries. Also, reporting systems might be less reliable in some countries compared with others. These are the limits of our data and our analyses. But considering the fact that the whole world, politicians and public health officials use exactly the same data for their decisions should allow us to use them for analysis. The data base was most shaky when it comes to non-pharmaceutical interventions. The reporting standards on the dedicated pages in Wikipedia are not uniform, and thus for all countries that were not already mentioned in [26] we had to decide what measures were implemented. If in doubt we attempted to err on the conservative side and assumed a stricter regime of NPIs, for instance if a stay-at-home order was issued or a country lockdown it was logical to assume that also schools, non-essential businesses were closed and mass gatherings forbidden. It is exactly because test strategies vary a lot across countries, with some countries having tested nearly half the populations and others testing only in symptomatic cases or special groups, that we have adopted the strategy to use case numbers standardized on tests conducted in a country. This allows us to better compare the variable across countries. The fact that the relationship between Covid-19 related deaths and test-standardized cases is weaker than one would expect is exactly due to this situation and to the fact that being a case, when considering the number of tests in a country, has only a weak relationship with becoming a fatality. It has been shown that the case fatality rate is much less than previously assumed and estimated to be 0.25% [43]. In Germany the case-fatality rate has recently been calculated from well documented cohorts to be 0.12 to

0.16% [44]. The still widely circulating higher case fatality rates are due to the fact that they are largely calculated using raw, absolute figures without knowledge of the real prevalence.[45] But also standardized figures might be unreliable. Often the same person is tested multiple times. Thus, we likely overestimate the number of cases by some margin. This would mean: the true link between being a case and becoming a fatality is probably even weaker.

Considering all these weaknesses our paper also has some strengths: restricting the analysis to Europe means that we have a comparatively homogeneous sample which nevertheless has enough variability. While all countries issued warnings the way it was implemented differed widely, from suggestions and recommendations in Sweden to very strict stay-at-home orders that were policed in Spain, from nearly no regard in Belarus to very strict political measures in France and Germany. Thus, we likely see a representative laboratory for the world, except that we do not cover any variance in ethnicity.

In conclusion we see that Covid-19 related deaths are most importantly dependent on the flu-vaccination rate among the elderly in a country: the higher the vaccination rate, the higher the Covid-19 death toll, explaining about 30% of the total variation. The number of cases is the second, unsurprising driver, but its relationship is weaker than one would assume. Non-pharmaceutical interventions seem to be more an indicator of the seriousness of the situation in a country than a prevention of deaths. These three variables explain about 60% of the variability in Covid-19 related deaths. This might encourage others to look for other, perhaps even more important host factors that can explain why we see such a wide variability in cases and deaths.

Methods

We followed in essence the same procedures as outlined in the previous protocol deposited ahead of analysis (https://osf.io/x93np/). The only difference was that we used another database and different predictors. But otherwise the procedure was similar. We tried to balance the need for flexibility in an exploratory framework with the need for parsimony to avoid overfitting and sham significances as outlined in standard texts [46-50].

Data Sources

We used data on standardized COVID-19 cases and deaths (per 1.000.000 inhabitants), cumulative numbers of PCR-tests for CoV2 and population data, as well as life expectancy data as provided by https://www.worldometers.info/coronavirus/#countries (accessed 30th of August) for 47 European countries (Albania, Andorra, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Channel Islands, Croatia, Czechia, Denmark, Estonia, Faroe Islands, Finland, France, Germany, Gibraltar, Greece, Hungary, Iceland, Ireland, Isle of Man, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, UK, Ukraine; the "Europe" tab of Worldometers' Covid-19 site does not contain data for Cyprus). Since numbers of deaths were zero for Faroe Islands and Gibraltar, these two countries were excluded from the database, because the appropriate model, a generalized linear model to predict a gamma-distributed variable, does not allow for zeros in the predictor variable.

Missing life-expectancy data were retrieved from World-Bank registers or Wikipedia in some single cases. Start dates of the epidemic were retrieved from Wikipedia (https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Europe) with their associated country sites. Number of non-pharmaceutical interventions (NPIs: restrictions of mass gatherings, business closures, educational facilities closed, non-essential services closed, stay at home order issued, compulsory wearing of face masks) was retrieved from Hunter et al. [26] and checked against information on Wikipedia, which was also used as a source for those countries not covered by Hunter et al. [26]. We added the number of NPIs to a simple numerical index that could range from 0 to 6 but in fact ranged from 1 to 6. We determined the date of the first political intervention in a country using the data provided on Wikipedia and calculated the time it took a government to react to the pandemic as the number of days from the date when the first case was known in a country to the date when the first political intervention was made public. We used influenza vaccination rates in the elderly (usually in persons aged 65 and older) as presented in [12], which retrieved the data from EUROSTAT and the European Center for Disease Prevention and Control. Data for countries that were missing in this publication were retrieved from https://gateway.euro.who.int/en/indicators/infl_8-influenza-vaccination-coverage-elderly/visualizations/#id=31628 (accessed August 31st 2020). Vaccination data for Channel Islands and Isle of Man were interpolated with UK data. Data for Albania were taken from [51]. We were not able to locate influenza vaccination data for Andorra, Bosnia, Liechtenstein, Moldovia, Monaco and San Marino. These six countries were therefore excluded from the analysis.

Statistics

The outcome of interest for this modeling study was the number of deaths per 1.000.000 inhabitants. The following variables were used as putative predictors of this dependent variable: (i) the test-standardized number of cases (in %), calculated as the number of cases in a country divided by the number of tests in that country × 100; (ii) the population size; (iii) the influenza vaccination rate in the elderly; (iv) life expectancy (in years); (v) rapidity of a country's government reaction (days from the first case to the first political intervention); (vi) the number of NPIs.

In order to explore whether non-linear functions might be necessary, a generalized additive model was used [52,53]. This indicated in fact that the variable "Percent test-standardized cases" was nonlinearly related to the outcome, while all others did not exhibit non-linearity. Closer inspection showed that this non-linearity was solely due to one outlier, Albania. We therefore decided to remove Albania from the database for the subsequent analysis.

Because the distribution of the outcome variable followed a gamma distribution well (Supplementary Figure 1), we calculated generalized linear models on a gamma-distributed variable with a log-link function. To build the best predictive model, we applied feature forward selection by successively reducing model uncertainty which was measured by a Kullback-Leibler divergence-based R^2 measure [13]. To avoid overfitting, the maximum number of variables was constrained to three. The Akaike Information Criterion (AIC), log-likelihood and deviance divided by degrees of freedom were also used as model goodness-of-fit indicators.

We also calculated a standard multiple linear regression model on a log-transformed dependent variable. The log-transformation produced an outcome variable with an

approximately normal distribution (Shapiro-Wilk normality test p=0.530, S-Figure 2). The best model was again built using feature forward selection based on R^2 as the measure of uncertainty.

All analyses were calculated independently both with Statistica Version 13.3, and R version 4.0.2 and converged on the same results.

Data Availability

Data and analysis scripts are available at https://osf.io/852dc/ once the paper is published and until then for reviewers.

References

1 Wu, C. *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Internal Medicine*, doi:10.1001/jamainternmed.2020.0994 (2020).

2 Adhikari, S. P. *et al.* Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infectious Diseases of Poverty* **9**, 29, doi:10.1186/s40249-020-00646-x (2020).

3 Qin, C. *et al.* Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clinical Infectious Diseases*, doi:10.1093/cid/ciaa248 (2020).

4 Shi, S. *et al.* Association of Cardiac Injury With Mortality in Hospitalized Patients With COVID-19 in Wuhan, China. *JAMA Cardiology*, doi:10.1001/jamacardio.2020.0950 (2020).

5 Richardson, S. *et al.* Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* **323**, 2052-2059, doi:10.1001/jama.2020.6775 (2020).

6 Onder, G., Rezza, G. & Brusaferro, S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*, doi:10.1001/jama.2020.4683 (2020).

7 Grasselli, G. *et al.* Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. *JAMA* **323**, 1574-1581, doi:10.1001/jama.2020.5394 (2020).

8 Modi, C., Boehm, V., Ferraro, S., Stein, G. & Seljak, U. Total COVID-19 Mortality in Italy: Excess Mortality and Age Dependence through Time-Series Analysis. *medRxiv*, 2020.2004.2015.20067074, doi:10.1101/2020.04.15.20067074 (2020).

9 Weinberger, D. M. *et al.* Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Internal Medicine*, doi:10.1001/jamainternmed.2020.3391 (2020).

10 Miles, D., Stedman, M. & Heald, A. Living with covid-19: Balancing costs against benefits in the face of the virus. *National Institute Economic Review* **253**, R60-R76, doi:10.1017/nie.2020.30 (2020).

11 Walach, H. & Hockertz, S. What association do political interventions, environmental and health variables have with the number of Covid-19 cases and deaths? A linear modeling approach. *medRxiv*, 2020.2006.2018.20135012, doi:10.1101/2020.06.18.20135012 (2020).

12 EBMPHET Consortium. COVID-19 Severity in Europe and the USA: Could the Seasonal Influenza Vaccination Play a Role? *SSRN*, doi:10.2139/ssrn.3621446

13 Cameron, C. A. & Windmeijer, F. A. G. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* **77**, 329-342, doi:https://doi.org/10.1016/S0304-4076(96)01818-0 (1997).

14 Rainwater-Lovett, K., Chun, K. & Lessler, J. Influenza outbreak control practices and the effectiveness of interventions in long-term care facilities: a systematic review. *Influenza Other Respir Viruses* **8**, 74-82, doi:10.1111/irv.12203 (2014).

15 Cowling, B. J. *et al.* Increased risk of noninfluenza respiratory virus infections associated with receipt of inactivated influenza vaccine. *Clin Infect Dis* **54**, 1778-1783, doi:10.1093/cid/cis307 (2012).

16 Wolff, G. G. Influenza vaccination and respiratory virus interference among Department of Defense personnel during the 2017–2018 influenza season. *Vaccine* **38**, 350-354, doi:https://doi.org/10.1016/j.vaccine.2019.10.005 (2020).

17 Anderson, M. L., Dobkin, C. & Gorry, D. The Effect of Influenza Vaccination for the Elderly on Hospitalization and Mortality. *Annals of Internal Medicine* **172**, 445-452, doi:10.7326/M19-3075 (2020).

18 Kim, H., Lim, K. Y., Kang, J., Park, J. W. & Park, S.-H. Macrophagic myofasciitis and subcutaneous pseudolymphoma caused by aluminium adjuvants. *Scientific Reports* **10**, 11834, doi:10.1038/s41598-020-68849-8 (2020).

19 Guimarães, L. E., Baker, B., Perricone, C. & Shoenfeld, Y. Vaccines, adjuvants and autoimmunity. *Pharmacological Research* **100**, 190-209, doi:http://dx.doi.org/10.1016/j.phrs.2015.08.003 (2015).

20 Gherardi, R. K. & Authier, F. J. Macrophagic myofascitis: characterization and pathophysiology. *Lupus* **21**, 184-189 (2012).

21 Arumugham, V. Proteins that contaminate influenza vaccines have high homology to SARS-CoV-2 proteins thus increasing risk of severe COVID-19 disease and mortality. *Zenodo*, doi:http://doi.org/10.5281/zenodo.3997694 (2020).

22 Zanettini, C. *et al.* Influenza Vaccination and COVID19 Mortality in the USA. *medRxiv*, 2020.2006.2024.20129817, doi:10.1101/2020.06.24.20129817 (2020).

23 Zein, J. G., Whelan, G. & Erzurum, S. C. Safety of Influenza Vaccine during COVID-19. *Journal of Clinical and Translational Science* **in print**, 1-6, doi:10.1017/cts.2020.543 (2020).

24 Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, doi:10.1038/s41586-020-2405-7 (2020).

25 Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*, eabb9789, doi:10.1126/science.abb9789 (2020).

26 Hunter, P. R., Colon-Gonzalez, F., Brainard, J. S. & Rushton, S. Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental study. *medRxiv*, 2020.2005.2001.20088260, doi:10.1101/2020.05.01.20088260 (2020).

27 Kuhbandner, C., Homburg, S., Walach, H. & Hockertz, S. Comment on Dehning et al (Science, 15 May 2020, eabb9789: Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions). *advance Social Sciences and Humanities Preprint* **Preprint**, doi:https://doi.org/10.31124/advance.12362645.v1 (2020).

28 Kuhbandner, C., Homburg, S., Walach, H. & Hockertz, S. Was Germany's Corona Lockdown Necessary? *Advande - Sage Preprint*, doi:10.31124/advance.12362645.v3 (2020).

29 Klement, R. J. The SARS-CoV-2 Crisis: Has Medicine Finally Entered a Reductionist Era? *Complementary Medicine Research* **27**, 207-208, doi:10.1159/000510453 (2020).

30 Daunizeau, J., Moran, R. J., Mattout, J. & Friston, K. On the reliability of model-based predictions in the context of the current COVID epidemic event: impact of outbreak peak phase and data paucity. *medRxiv*, 2020.2004.2024.20078485, doi:10.1101/2020.04.24.20078485 (2020).

31 Moran, R. J. *et al.* Estimating required lockdown cycles before immunity to SARS-CoV-2: Model-based analyses of susceptible population sizes, S0, in seven European countries including the UK and Ireland. *medRxiv*, 2020.2004.2010.20060426, doi:10.1101/2020.04.10.20060426 (2020).

32 Friston, K. J. *et al.* Tracking and tracing in the UK: a dynamic causal modelling study. *arXiv* **2005.07994** (2020).

33 Friston, K. J. *et al.* Second waves, social distancing, and the spread of COVID-19 across America. *arxiv* **1104.3344v1** (2020).

34 Wieland, T. Flatten the Curve! Modeling SARS-CoV-2/COVID-19 Growth in Germany on the County Level. *medRxiv*, 2020.2005.2014.20101667, doi:10.1101/2020.05.14.20101667 (2020).

35 Wieland, T. A phenomenological approach to assessing the effectiveness of COVID-19 related nonpharmaceutical interventions in Germany. *Safety Science* **131**, 104924, doi:https://doi.org/10.1016/j.ssci.2020.104924 (2020).

36 Pan, A. *et al.* Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China. *JAMA* **online first**, doi:10.1001/jama.2020.6130 (2020).

37 Walach, H. & Hockertz, S. Wuhan Covid19 data – more questions than answers. *Toxicology* **440**, 152486, doi:https://doi.org/10.1016/j.tox.2020.152486 (2020).

38 Walach, H. & Hockertz, S. A Reply to Dr. Pan's and Dr. Wu's Response: "Wuhan COVID-19 data – an example to show the importance of public health interventions to fight against the pandemic". *Toxicology* **441**, 152524, doi:10.1016/j.tox.2020.152524 (2020).

39 Klement, R. Systems thinking about SARS-CoV-2. *Frontiers in Public Health* **in print**, doi:10.3389/fpubh.2020.585229 (2020).

40 Grant, W. B. *et al.* Evidence that vitamin D supplementatoin could reduce rsik of influenza and COVID-19 infections and deaths. *Nutrients* **12**, 988, doi:10.3390/nu12040988 (2020).

41 Breiman, A., Ruvën-Clouet, N. & Le Pendu, J. Harnessing the natural anti-glycan immune response to limit the transmission of enveloped viruses such as SARS-CoV-2. *PLOS Pathogens* **16**, e1008556, doi:10.1371/journal.ppat.1008556 (2020).

42 Miotto, M., Di Rienzo, L., Gosti, G., Milanetti, E. & Ruocco, G. Does blood type affect the COVID-19 infection pattern? *arXiv:2007.06296* (2020).

43 Ioannidis, J. The infection fatality rate of COVID-19 inferred from seroprevalence data. *medRxiv*, 2020.2005.2013.20101253, doi:10.1101/2020.05.13.20101253 (2020).

44 Schrappe, M. *et al.* Thesenpapier 4.0: Die Pandemie durch SARS-CoV-2/Covid-19 - der Übergang zur chronischen Phase - Verbesserung der Outcomes in Sicht - Stabile Kontrolle: Würde und Humanität wahren - Diskursverengung vermeiden: Corona nicht politisieren. (www.matthias.schrappe.com, Köln, 2020).

45 Klement, R. J. & Bandyopadhyay, P. S. The Epistemology of a Positive SARS-CoV-2 Test. *Acta Biotheoretica*, doi:10.1007/s10441-020-09393-w (2020).
46 Breen, R., Karlson, K. B. & Holm, A. Interpreting and understanding logits, probits, and other nonlinear probability models. *Annual Review of Sociology* **44**, 4.1-1.16, doi:10.1146/annurev-soc-073117-04142 (2018).
47 Burnham, K. P. & Anderson, D. R. *Model Selection and Inference: A Practical Information-Theoretic Approach*. (1998).
48 McQuarrie, A. D. R. & Tsai, C.-L. *Regression and time-series model selection*. (World Scientific Publishers, 1998).
49 Liao, T. F. *Interpreting Probability Models: Logit, Probit, and Other Generalized Models*. (Sage, 1994).
50 Cohen, J. & Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. (Erlbaum, 1983).
51 Preza, I., Nelaj, E. & Bino, s. Influenza vaccination in Albania (Abstract). *International Journal of Infectious Diseases* **79**, 102-103 (2019).
52 James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning with an Introduction to R*. (Springer, 2017, orig. 2013).
53 Hastie, T. & Tibshirani, R. *Generalised Additive Models*. (Chapman & Hall, 1990).

Acknowledgement

We thank the EBMPHET-Consortium for help with vaccination data and advice on Generalised Additive Models.

Author Contributions

HW initiated this study, collated the data, calculated the first analyses using Statistica and wrote the first draft of the MS. RK checked the data, calculated the GLM using R and contributed to writing and discussion of the results.

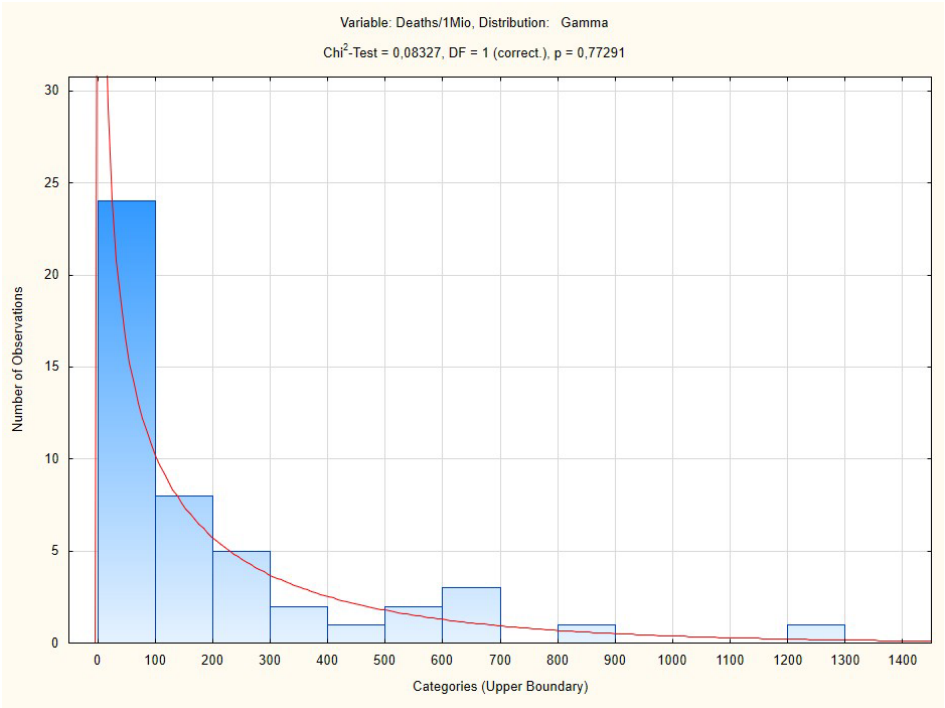
Conflict of Interests

None of the authors has a conflict of interest.

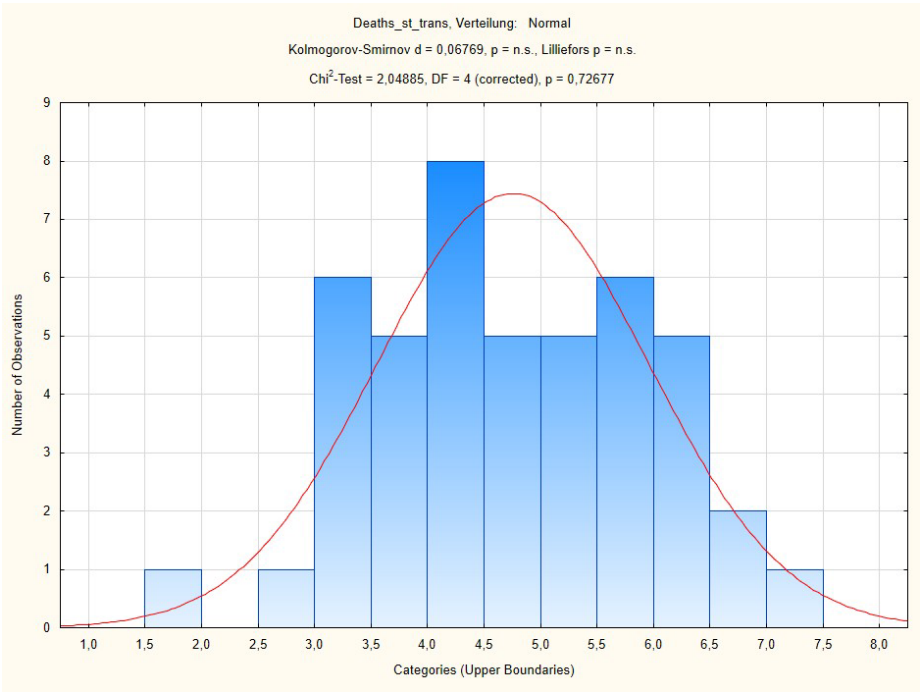
Supplementary Material

Figures

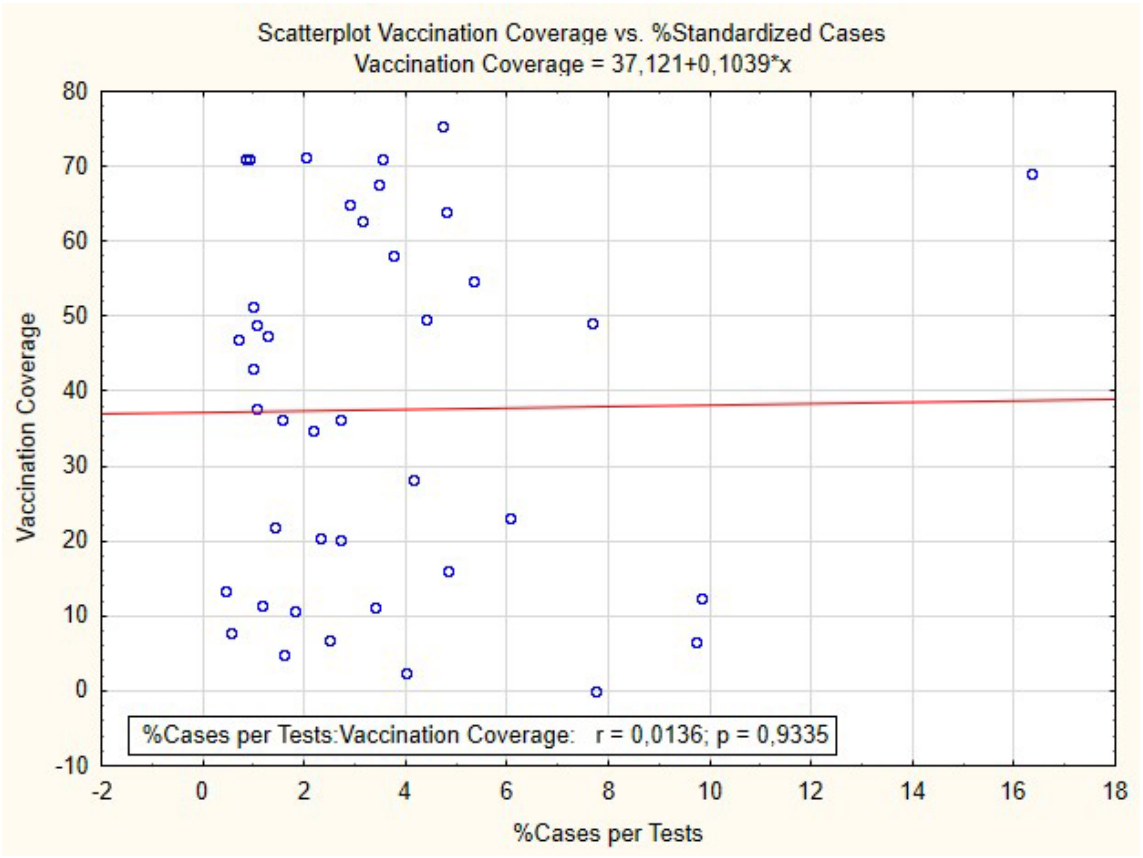
S-Figure 1 – Distribution of dependent variable (deaths/1.000.000 Inhabitants)



S-Figure 2 – Distribution of ln-transformed dependent variable (Number of deaths/1.000.000 inhabitants)



S-Figure 3 – Scatterplot of the variables %test standardized cases and vaccination coverage



S-Figure 4 – Scatterplot of the variables %test standardized cases and number of NPIs