THE IMPACT OF COVID-19 VACCINES ON ALL-CAUSE MORTALITY IN EU IN 2021: A MACHINE LEARNING PERSPECTIVE

Preprint · January 2022

DOI: 10.13140/RG.2.2.16851.71205

1 author:

Patrick E. Meyer
University of Liège
56 PUBLICATIONS **3,131** CITATIONS

SEE PROFILE

THE IMPACT OF COVID-19 VACCINES ON ALL-CAUSE MORTALITY IN EU IN 2021

A MACHINE LEARNING PERSPECTIVE BY PATRICK E. MEYER

Abstract. The question whether COVID-19 vaccines have no effect on all-cause mortality or perform as intended, that is mainly reduce excess mortality, has been debated recently in the scientific literature. By crossing the all-cause mortality data with the vaccine data from public European databases, we compare the impact on mortality of two variables of interest namely a vaccine-dose-rate and a covid-case-rate. Using classical machine learning strategies and graphical models, we are able to assess the conflicting hypothesis about the effect of vaccines on all-cause mortality, at least in Europe. Our conclusions differ for different age-categories investigated but, until a better predictive variable is found, our results clearly suggest that the benefit-risk balance for the 0-44 years old is not in favor of those vaccines.

1. INTRODUCTION

It can be sometimes difficult to navigate through the scientific literature when we face studies reaching opposite conclusions. This typically happens when data and subjects are new. Confronting hypothesis is a crucial part of the scientific process. Recently, two statistical studies based on data from the UK have reached the conclusion that COVID-19 vaccines may have no effect on the overall mortality. In other words, those vaccines may save people from COVID-19 in the same proportion than they may exacerbate other mortality causes [Crawford, 2021, Neil and Fenton, 2021]. Those studies are thus disagreeing with another study concerning US data and relayed on the CDC website [Xu et al., 2021]. Despite the fact that those studies use data from different countries, it could be expected that those reach similar conclusions rather than opposite ones. Assuming the same underlying effects are at play and in order to favor one of those conflicting hypothesis, we have attempted a machine learning approach to study the impact of those vaccines on the all-cause mortality in EU. Indeed, we have downloaded the EuroMOMO data giving the mortality z-scores of each age category across different EU countries [EuroMOMO, 2021]. We have crossed those with the ECDC vaccination data that also provide information by age category and countries. Finally, we extracted the ECDC 14-day-case-positivity-rate and death-rate in each of the targeted country [ECDC, 2021]. Next, combining graphical models with generalized linear models and random forests [Whittaker, 1990, Hastie et al., 2001, Breiman, 2001], we have evaluated the impact of a vaccine-dose-rate and a covid-case-rate variables on excess mortality of the current year. Our conclusions differ for different age-categories investigated but for the young cohorts, our analysis favor the studies showing no benefits from vaccination. We have discussed our results in the last section of this paper.

Date: 28th December 2021.

2. DATA

As stated in the introduction, we downloaded the mortality data from [EuroMOMO, 2021] and the vaccination data as well as the case-positivity-rate and deaths-poitivity-rate from [ECDC, 2021]. Unfortunately the list of countries and the list of age targets are not a perfect fit between our data sources. Hence, we have focused on the countries and age categories were an intersection was possible without creating strong distortions. There is a trade-off that could have manifested here, on the one hand removing too much data can lead us to a poor dataset in terms of number of samples, on the other hand, introducing too much distortions could also lead us to inaccurate conclusions. However, our intersecting data is already consequent. Indeed, in the end, data from 18 EU countries could be kept as is, namely Austria, Belgium, Cyprus, Denmark, Estonia, Finland, France, Hungary, Ireland, Israel, Italy, Luxembourg, Malta, Norway, Portugal, Slovenia, Spain and Sweden. Unfortunately data from Greece, Switzerland and UK were not present in both datasets. Also, the variables first dose and second dose from the ECDC dataset appeared incomplete for two countries: the Netherlands and Germany. Hence those five countries present in the EuroMOMO data have been removed from our resulting dataset.

In terms of age category, the EuroMOMO dataset has the 0-14, 15-44, 45-64, 65-74, 75-84 and 85+ age categories while the ECDC data have as target group : 0-4, 5-9, 10-14, 15-17, 18-24, 25-49, 50-59, 60-69, 70-79, 80+. As a result, grouping the three first categories from the ECDC would match perfectly the 0-14 years category of the EuroMOMO data. The three next age categories could also be grouped to obtain a 15-49 age group matching closely the 15-44 category from the first dataset. The next grouping 50-59 and 60-69 had an acceptable 5 years shift from the first dataset. However, the other age categories were not attempted to be matched in order to avoid too strong distortions from our age-matching strategy. Hopefully, our analysis based on the population targeting the 0-64 years from 18 different EU countries could deliver us with sufficient evidence to favor one of the contradicting hypothesis debated above. Each of the downloaded dataset provide us with a weekly monitoring of their respective variables. We opted for grouping by 4 weeks periods. It seemed logical to opt for a multiple of 14 days because the case-positivity-rate and deaths-rate used by the ECDC dataset is precisely based on 14-day period. Another reason for this particular choice is that the 52 weeks of a year can be easily divided by 4. Finally, having a time-interval big enough to obtain both a smoothing effect and a higher likelihood for capturing time-delayed effects has been intended. Our data starts at week 46 of 2021 and goes down to week 47 of 2020 by groups of 4 weeks, that is for each variable. Other choices of starting week could be chosen but 46 is the last week where we obtained data from all three sources (1 dataset from all-cause mortality and 2 datasets from ECDC). Finally we defined a new variable called *DoseRate* which is simply the sum of the number of first doses and second doses of all the COVID-19 vaccines that were administered to the targeted age-group during the 4-weeks-period, divided by the total population of the targeted country. The age-targeted-population could have been a better choice for our rate. However, since each age-category is treated separately, the analysis should not suffer from this choice. Since age-categories are not a perfect match in between the different data sources, we applied a crude corrective term on the variable *DoseRate* for our third age-category (equivalent to remove 10% of the

total doses) to account for the fact that the 50-69 years old have likely received more doses during each period than the 45-64 of interest. We also applied a corrective term to our second age-group to account for the excess 5 years over the 34 years involved in the 14-49 category (this corrective term thus amount to 5/34 of the total doses received). Although those corrective terms are crude, the removal of those appear to have no consequence on the subsequent statistical analysis. In the end, we produced 3 datasets, one for each age category of the EuroMOMO data treated (i.e. 0-14, 15-44, 45-64). Each dataset has 13 periods multiplied by 18 countries, hence 234 samples. There are 9 variables in total, namely *ZscoresCurrent*, *ZscoresPast1Y* (i.e. mainly 2020), *ZscoresPast2Y* (i.e. mainly 2019), *Where* (i.e. country), *When* (i.e. which group of 4 weeks), *Target* (i.e. age-category), *DoseRate* (i.e. reflecting the administered doses during the period), *CaseRate* (i.e. average ECDC 14-day-positivity-rate during the period) and *CovDeathRate* (i.e. average ECDC 14-day-deaths-rate during the period). Our R script to extract data and our resulting datasets are now freely available [Meyer, 2021].

3. ANALYSIS AND METHODS

It must be emphasized that it is near impossible to perform a statistical analysis without making any assumption or without any bias. For example, the use of the variable "all-cause mortality" introduces in itself many biases. What if a higher rate of suicides become visible in the all-cause mortality not because of the pandemic itself but because of the various political measures limiting freedoms? What if cancer treatments have been delayed by the same political measures but results in a higher death rate in the following years? What if crimes and accidents increase because of some recovered freedom of movements from the feeling of safety due to the vaccination campaign? Despite those weaknesses, not using the all-cause mortality also suffer from many biases [Classen, 2021, Neil and Fenton, 2021]. For example, what if the spike protein present in each COVID-19 vaccine has an inherent toxicity that do increase mortality through cardiac and/or cancer related unknown mechanisms? What if crimes, accidents and suicides are in fact the consequences of some neurological impact of a vaccination more than the results of the various political measures taken? In the end, we might have to be humbly conscious of the trade-offs that are connected to each choice of mortality measure and state clearly the implicit assumptions behind those choices. For example, we should pinpoint that [Xu et al., 2021] may also have made some choices that could importantly bias their conclusion. For example, « *To ensure comparable healthcare–seeking behavior among persons who received a COVID-19 vaccine and those who did not (unvaccinated persons), eligible unvaccinated persons were selected from among those who received ≥1 dose of influenza vaccine in the last 2 years.* ». If this choice indeed removes one bias w.r.t. healthcare seeking behavior, it could well introduce other biases. It could preferentially select weaker population for the younger cohort. Indeed we could argue that, at least in EU, young populations do not make a wide use of the influenza vaccine unless they suffer from health issues [Mereckiene, 2018]. Also [Xu et al., 2021] exclude all COVID19 deaths. This could be a very delicate operation that the authors recognize themselves : « *...although deaths associated with COVID-19 were excluded, causes of death were not assessed. It is possible that the algorithm used might have misclassified some deaths associated with COVID-19 because of lack of testing or because individual mortality reviews*

were not conducted. ». Hence, the authors remove all deaths happening within 30 days of a COVID-19 diagnosis. If this choice could make sense to assess the efficacy of any vaccine, it does not when it comes to assess the security of it. For example, what if COVID-19 vaccines are the cause of a temporary drop in immunity that would increase the probability of catching COVID-19?

In our analysis, we use the z-scores of excess mortality. Due to the limited amount of variables investigated, we cannot eliminate all the biases connected to the use of all-cause mortality measure. However, it is worth noting that the 18 EU countries in our data have used different restriction measures, at different point in time of the pandemic, and all have different healthcare providing services and capacities. This could results into an averaging out of some biases. Indeed, it would be quite astonishing (though not impossible) that suicides and crimes happen exactly at the same moment with respect to either the viral waves or the vaccination campaign in each country and also with a similar intensity. It should be clear also from the data that we do not hold into account or adjust for socio-economic status, health conditions and other confounders. We do not use either the standard mortality rate (SMR) because we perform the same analysis for each age category separately. In fact, we deliberately intended to alter minimally all the variables. Finally, it should also be stressed that inadequate assumptions or biases can always lead to correct conclusions, that is why the scientific approach usually evaluates a hypothesis or a model, not so much through the lens of biases, but rather by using its quality of predictions on new data [Pearl, 2000]. It is also for those reasons that a machine learning based perspective is defended here. This is of course our own research bias [Meyer, 2008]. Although epidemiologists would have rather used more classical tools of their field, we would like to emphasize that we do not attempt to provide any hypothesis in this paper, we are merely trying to favor one of the conflicting hypothesis stated above with as much neutrality as possible. We also deem that both, the data and the analysis provided here are valuable for epidemiologists to pursue more advanced modeling strategies should they enquire it.

Now that our disclaimer has been clearly stated, let us start our analysis with a few classical machine-learning definitions of variable relevance. Three degrees of relevance are defined in [Kohavi and John, 1997]:

Definition 1. Let Y be a target variable to predict, X be a set of input variables, and X_{-j} be the same set without the jth variable: $X \setminus \{X_j\}$:

A variable X_j is said "strongly relevant" in X iff there exists some x_j, y and x_{-j} for which $p(x) > 0$, such that

$$p(y|x) \neq p(y|x_{-j})$$

A variable X_j is said "weakly relevant" iff it is not strongly relevant, but there exists a subset X_S of variables of X_{-j} for which there exists some x_j, y and x_S with $p(x_j, x_S) > 0$ such that

$$p(y|x_j, x_S) \neq p(y|x_S)$$

A variable is said irrelevant iff it is not relevant (weakly or strongly).

In other words, an input variable X_j is strongly relevant if the removal of X_j alone will result in a change of the conditional probability distribution of Y. An input variable is weakly relevant if it is not strongly relevant, but in some context X_S it may change the conditional probability distribution of Y.

Strong relevance can be associated to the notion of causality because under the causal sufficiency assumption (i.e. all the causes of an effect-variable are also present in that dataset [Neapolitan, 2003]) then strong relevance implies direct causality. This results from the fact that being unable to cancel the dependency between two variables in a dataset (containing causal variables) can only be explained by having one of them be the most direct cause of the other [Neapolitan, 2003, Pearl, 2000]. Weak relevance is more difficult to interpret because it means that in some context the variable improves prediction but not in others. This typically happen either with redundant variables or with a variable that is not the most direct causal explanation of the other. Irrelevance appears to be the easiest definition to interpret but unfortunately it also requires the causal sufficiency assumption to be assuredly meaningful. Indeed it can be shown that missing a strongly relevant variable can turn another strongly relevant variable into an irrelevant one. In fact, that is the underlying principle behind cryptography: the coded message (strongly relevant variable) is irrelevant to the decoded message (target variable) unless you have the private key (the other strongly relevant variable). As a result, we face two major problems when modeling a) we do not have the real probabilities underlying our model but only some data that allow us to estimate those and b) the causal sufficiency assumption is a very strong hypothesis. For the former issue, machine-learning algorithms have a long track record of estimating quite well underlying probabilities without having recourse to too strong hypothesis [Mitchell, 1997, Hastie et al., 2001]. The latter issue, i.e. the causal sufficiency assumption, explains why Science always attempts to provide the explanation that lead to the most accurate predictions and admit it as true until a better explanation can replace the previous one [Pearl, 2000]. Indeed one can never be completely sure to have all the causal variables in the dataset.

Let us now look at the "a priori" causal network of our data in order to grasp the extent of our causal (in)sufficiency (see Fig. 3.1). Simply stated, our variable COVID-19-death-rate depends on the number of COVID-19-positive-cases in the population (i.e. represented by our variable *CaseRate*). However, the *CaseRate* variable is likely altered by the vaccination in two positive ways (represented here by green arrows) 1) vaccination should protect from deaths related to the virus 2) vaccination may reduce the transmission. On the other side COVID-19 vaccines may increase, via some spike-protein toxicity for example, a hidden variable called here vaccines-death-rate. The excess mortality of the current year can thus be predicted with four impacting variables: the COVID-19, the related vaccines, the government measures and all the usual/cyclic causes of deaths. The latter is a hidden variable (i.e. yellow in the figure) that is observable indirectly through the all-cause excess deaths of the previous years. It is worth noting that the excess mortality of the year before (i.e. 2020) is a bit more complex to handle because that year already dealt with COVID-19 while also being subject to political measures meant to reduce viral transmission, such as lockdowns, but yet without vaccines (at least not before the week 46 used in our dataset).

3.1. **Correlation Analysis.** Some methods infer graphical models using only correlations [Whittaker, 1990, Meyer et al., 2007]. Although there are well-known dangers to connect pairwise correlations with causality, correlations can at least offer us with two valuable information: 1) a ranking of variables by pairwise relevance and 2) the directionality of the pairwise relevance (correlated vs anti-correlated). For those reasons, we report correlations for all the connecting paths of our graphical

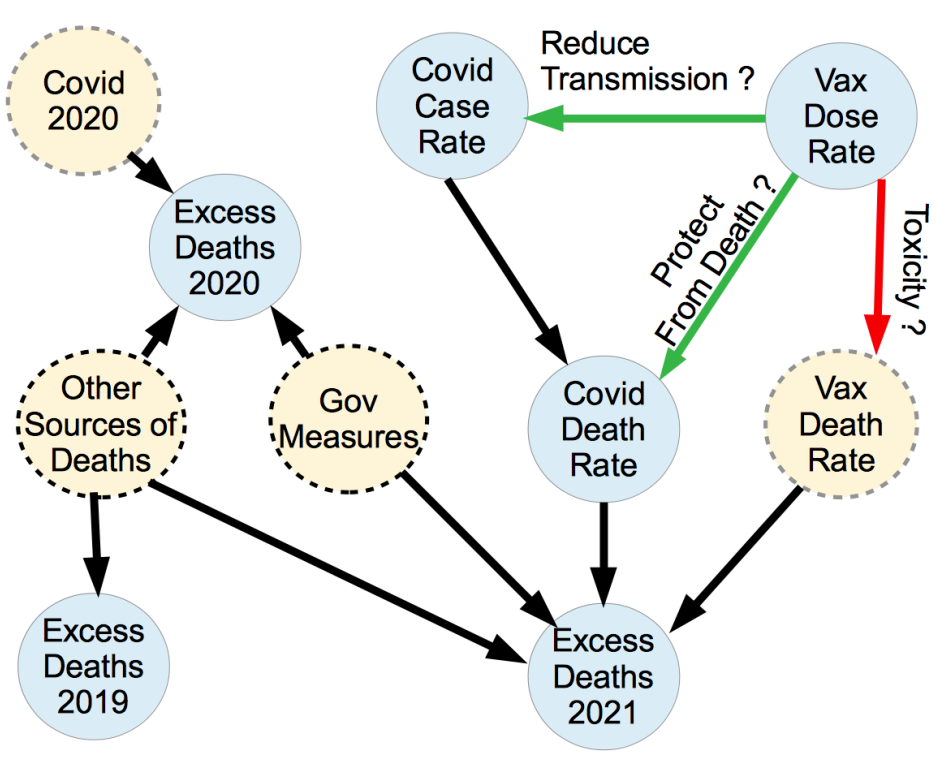

FIGURE 3.1. A priori causal network underlying our extracted dataset. The main variables of our dataset are in blue. Hidden variables are in yellow. Arrows of interest are in green and red.

Correlation	Zscores Excess mortality 2021 with					CovDeathR with	
	2019	2020	DoseR	CaseR	CovDthR	DoseR	CaseR
0-14	0.132	0.407	0.159	0.011	-0.092	-0.160	0.546
15-44	0.296	0.292	0.051	0.214	0.424	-0.320	0.546
45-64	0.360	0.234	-0.011	0.402	0.720	-0.179	0.546

TABLE 1. Pairwise correlations computed along all the paths of our graphical model. Those correlations are computed on 234 samples (18 countries times 13 4-weeks-periods).

model in the Table 1. It is worth noting that using Spearman's correlation instead of Pearson's do not change neither the ranking of variables nor the directionality of our pairwise dependencies.

At first glance, several values appear interesting. First we observe that 2020 has a better correlation than 2019 for the young probably because the impact of lockdowns has been stronger for them than for the other age categories. Second the *CaseRate* and the *CovDeathRate* (columns 4 and 5) which are constant for all age-groups reflect the fatality rate of the disease for each age-group, that is no

impact on the 0 to 14, a small impact on the 15 to 44 and a stronger one on the 45-64. The correlation between *CaseRate* and *CovDeathRate* is constant and strong as expected. The *DoseRate* has a negative correlation with the *CovDeathRate* which is to be expected since vaccines are meant to protect from COVID-19-death (at least on the span of our 4-weeks period). Finally, the variable *DoseRate* has a strong correlation with the current excess mortality for the younger ones while almost no correlation for the two other age groups. Although it would be tempting to conclude that COVID-19 vaccines have no beneficial effects on the older groups and a deleterious one on the younger one, several elements can explain those correlations. First, there are many very low-values both in the *DoseRat*e and in the excess mortality for the younger group, hence a spurious value can appear there. Second, the excess mortality is a variable that is affected by many other variables as shown in our graphical model. Hence, those relations should be investigated more deeply as we are doing in the next section.

3.2. **Strong Relevance Analysis.** In order to check for strongly relevant variables (i.e. causal relationships) we make use of two different models (with their default parameters): the generalized linear model and the random forest. The first one makes an assumption of linearity of relationships between variables and the other one is known to capture a vast set of non-linear dependencies [Hastie et al., 2001]. The loss function used for the random forest is the out-of-bag mean-squared-error (MSE) whereas for the glm the Akaike criterion (AIC) is used [Sakamoto and Kitagawa, 1987, Devroye et al., 1996, Breiman, 2001], both are computed internally by their respective R functions. Indeed the statistical language R, has been used both in the extraction of the data and in the statistical analysis [Gentleman and Ihaka, 1996].

To assess which variables is strongly relevant, we make use of a unilateral paired Wilcoxon Rank statistical test [Dietrich, 1998]. In other words, each variable that, when removed from our full initial model, increases statistically significantly the prediction error across the 2 models times 3 age-groups, is considered as strongly relevant. The other variables are not impacting our models enough when removed. As a result, those are considered weakly relevant. The p-values are not corrected here simply because each variable is evaluated independently of the others. In other words, those are only in competition with the full model but not with the other variables. It differs from a strategy that aims to select the best among competitive ones where an adjustment would be advised.

The three strongly relevant variables identified here already deliver a minimal model with quite accurate prediction. Indeed, the second column named "20+CDR+CR" standing for the three strongly relevant variables, namely *Zscores-Past1Y, CovDeathRate, CaseRate* in the Table 3, shows that removing the two weakly relevant variables together do not impact significantly our prediction errors (in fact column 2 of Table 3 has similar errors than column 2 of Table 2). The fact that *DoseRate* is a strongly relevant variable can only be explained in our graphical model through the red arrow. In order to better quantify this negative impact of vaccines we can further our analysis by replacing *DoseRate* by *CaseRate* in that minimal model. Indeed, this should results in an increase in errors because the *CaseRate* information should already be captured by the *CovDeathRate* variable. Since machine learning algorithms can sometimes be highly sensitive to the number of variables [Meyer, 2008], replacing one rate (i.e. *DoseRate*) by another of similar structure (i.e. *CaseRate*) allows us to eliminate a potential bias. The columns 2

Error with	All vars	-2019	-2020	-CovDthR	-CaseR	-DoseR
AIC-0-14	408.883	407.974	**452.099**	**413.976**	408.304	**411.159**
MSE-0-14	0.331	0.337	**0.383**	0.328	0.341	0.330
AIC-15-44	564.756	**570.051**	**575.865**	**595.423**	562.756	**576.098**
MSE-15-44	0.697	0.708	**0.766**	**0.772**	0.700	**0.737**
AIC-45-64	715.590	**719.995**	**731.761**	**824.469**	714.575	**723.814**
MSE-45-64	1.436	**1.567**	**1.585**	**1.860**	**1.546**	**1.546**
p-values	ref	0.109	**0.016**	**0.031**	0.844	**0.031**
Relevance	-	Weak	**Strong**	**Strong**	Weak	**Strong**

TABLE 2. Strong-relevance evaluated with an unilateral paired Wilcox Rank test on the Akaike criterion of generalized linear models (AIC) and the out-of-bag mean-square-error of a random forests (MSE). In bold are the values bigger than their corresponding reference of the first column.

Error with	All vars	20+CDR+DR	20+CDR+CR	20+DR	20+CR
AIC-0-14	408.88	407.44	410.32	410.86	417.98
MSE-0-14	0.33	0.35	0.35	0.35	0.35
AIC-15-44	564.76	568.05	580.48	617.78	610.56
MSE-15-44	0.70	0.71	0.75	0.90	0.81
AIC-45-64	715.59	718.15	726.48	900.61	861.69
MSE-45-64	1.44	1.52	1.66	2.75	2.38
Conclusion	-	close from ref	CR is worse	bad model	CR is better

TABLE 3. Errors showing the inversion of relevance between *CaseRate* (CR) and *DoseRate* (DR) in function of *CovDeathRate* (CDR), i.e. present (col. 2 and 3) and absent (col. 4 and 5)

and 3 of Table 3 show the results of this strategy. Another way to check if our network makes sense, is to recompute the previous columns (i.e. *CaseRate* instead of *DoseRate*) but this time with the variable *CovDeathRate* removed. This strategy should allow us to observe if the *CaseRate* becomes then more relevant than the *DoseRate*. Indeed, since the *CovDeathRate* is preventing the flow of information from the *CaseRate* variable toward the 2021 excess mortality variable, it is expected that once removed, the effect of the disease will become more relevant than the effect of the vaccines (in order to predict excess mortality). The next two columns (i.e. 5 and 6) of the Table 3 show precisely the inversion of relevance of those two variables, thereby further reinforcing the validity of our graphical model.

However, we can already note that the first age-category (i.e. 0-14) seems quite unaffected by the removal of the *CovDeathRate* variable and it is quite explainable due to the very low amount of deaths in that age-category. On the other side, the age-category 45-64 is strongly impacted by the removal of the *CovDeathRate* variable. Indeed, in that age-category the relevance inversion is much stronger. To make that effect more visible, we can report the ratio of error measures when we replace *CaseRate* by *DoseRate* and reciprocally. In the Table 4, we report the ratio between the model using the most impactful variable (either *CaseRate* or

Error ratios	20+CDR +CR/+DR	20 +DR/+CR	19+CDR +CR/+DR	19 +DR/+CR	19+20+CDR +CR/+DR	19+20 +DR/+CR
AIC-0-14	1.01	0.98	1.01	0.99	1.01	0.98
MSE-0-14	1.00	**1.01**	0.97	**1.04**	0.98	**1.06**
AIC-15-44	1.02	1.01	1.02	1.01	1.02	1.01
MSE-15-44	1.06	**1.11**	1.12	1.06	1.05	1.02
AIC-45-64	1.01	**1.05**	1.01	**1.05**	1.01	**1.05**
MSE-45-64	1.09	**1.16**	1.14	**1.20**	0.98	**1.19**

TABLE 4. Error ratios of models that measure the impact of *DoseRate* (DR) versus the impact of *CaseRate* (CR). In bold the results supporting that the variable *CaseRate* is more impactful than the variable *DoseRate*. Results favor *CaseRate* only in the older category.

DoseRate depending on the presence of *CovDeathRate*) over the model using the other variable. We also report the same values when our models uses the excess mortality of 2019 rather than of 2020 and also with both years jointly.

We observe that *DoseRate* and *CaseRate* variables appear equally impactful in the 0-14 category independently of the presence of *CovDeathRate*. Indeed, not only all the ratio are close to one, thereby showing that both variables have similar impact on predictions but also the best model is not always using the same variable, i.e. *DoseRate* is slightly more impactful when using linear models and *CaseRate* when using random forests. In the 15-44 years, we observe also close-to-one ratios at least with linear models. There is a stronger unbalance with random forests. However, the unbalance is again not always in favor of the same variable. In fact, the negative impact of *DoseRate* seem equal or even slightly stronger than the negative impact of *CaseRate*. It is only in the last age-category, i.e. 45-64, that a clear message is conveyed. In the latter case, the *CaseRate* variable (in absence of *CovDeathRate*) is more impactful than the *DoseRate* (in presence of *CovDeathRate*). It should be emphasized that our models do not, as is, evaluate those impacts in absolute number of deaths because we are predicting z-scores on excess all-cause mortality. This is also the reason why the whole strategy defended in this paper has been focused on comparing impacts on predictions rather than quantifying them in absolute terms. We deem this approach as more crude but also more reliable since it does not require any transformation of the downloaded variables. As a consequence, it is quite worrying that two age categories have excess mortality similarly impacted by the negative effects of vaccines than they are impacted by the disease. However, those age categories were known to have a low fatality rate initially [Semenzato et al., 2021].

4. CONCLUSION

Our goal in this study has been to favor one of the conflicting hypothesis stated in the introduction: either the COVID-19 vaccines increase the all-cause mortality in the same proportion than they protect, in each age-category [Neil and Fenton, 2021, Crawford, 2021] or it does not increase the non-COVID-19 mortality at all [Xu et al., 2021]. We provided a graphical model and studied the relevance of several key variables

in order to achieve our goal. Interestingly our results, based on EU data, agree with [Neil and Fenton, 2021, Crawford, 2021] for the 0-44 years, that is vaccines have clearly no net benefits on excess mortality. The fact that the vaccines have been delivered to an important proportion of the population means that even a small toxicity could be responsible for as many deaths than the disease itself. Indeed not everyone contract the virus and additionally in those that contract the virus, very few dies in the youngest categories [Semenzato et al., 2021]. However, our third category shows a different signal. It would be tempting to conclude that the benefit-risk balance for the oldest category is favorable, but the fact that the mortality is better explained by the variable *CaseRate* than the variable *DoseRate* does not really allow to assess vaccines efficacies. In other words, we have not tried to compare benefits of vaccines versus costs of vaccines as implicitly done in [Neil and Fenton, 2021, Crawford, 2021], we have rather compared costs of vaccines versus costs of the disease. As a consequence, our results do not necessarily oppose those results even for the third age-group. However, our results disagree with [Xu et al., 2021] for the 0-44 years old. We believe that the flu-vaccine induced bias mentioned above could explain the different conclusions reached. Beside, for all studies mentioned we cannot put aside all the possible multiple confounding variables impacting at least partially the statistics like different underlying populations, healths and healthcare systems, type of vaccines used, delays applied in between doses,... Nonetheless, the excess mortality of 2021 in EU is well above the excess mortality of 2020 that is itself well above 2019. It appears that the variable COVID-19-Death-Rate is not sufficient to explain the surge in 2021. Our variable *DoseRate,* related to COVID-19 vaccines, apparently explains a major part of the signal observed in the 0-44 years category. As a consequence, until a better predictive variable is found, our results clearly suggest that the benefit-risk balance for the 0-44 years old is not in favor of those COVID-19 vaccines. This could change in the future, for example with the emergence of less favorable variants or equivalently with more favorable COVID-19 vaccines.

Acknowledgments: I would like to thank several colleagues (who will recognize themselves) for their helpful comments on this work. I hope that those who have expressed concern with the possible political use of our conclusion will understand that we are dealing here with life and death matters (literally). In those situation, we should all agree that scientists must rapidly provide as many tools as possible to analyze a perilous situation.

References

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine Learning*, 45.

[Classen, 2021] Classen, J. B. (2021). Us covid-19 vaccines proven to cause more harm than good based on pivotal clinical trial data analyzed using the proper scientific endpoint, "all cause severe morbidity". *Trends in Internal Medicine*.

[Crawford, 2021] Crawford, M. (2021). Uk data shows no all-cause mortality benefit for covid-19 vaccines. https://roundingtheearth.substack.com/p/uk-data-shows-no-all-cause-mortality?

[Devroye et al., 1996] Devroye, L., Györfi, L., and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer-Verlag.

[Diettrich, 1998] Diettrich, T. G. (1998). Approximate statistical tests for comparing supervised learning algorithms. *Neural Computation*, 10.

[ECDC, 2021] ECDC (2021). Europe's journal on infectious disease surveillance, epidemiology, prevention and control. - https://www.ecdc.europa.eu/en/publications-data.

[EuroMOMO, 2021] EuroMOMO (2021). Euromomo bulletin, week 47, 2021 - https://www.euromomo.eu/graphs-and-maps/.

[Gentleman and Ihaka, 1996] Gentleman, R. and Ihaka, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5.

[Hastie et al., 2001] Hastie, T., Tibshirani, R., and Friedman, J. H. (2001). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. Springer Series in Statistics.

[Kohavi and John, 1997] Kohavi, R. and John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324.

[Mereckiene, 2018] Mereckiene, J. (2018). Seasonal influenza vaccination and antiviral use in eu/eea member states. ECDC Technical Report.

[Meyer, 2008] Meyer, P. E. (2008). *Information-theoretic variable selection and network inference from microarray data*. PhD thesis, Université Libre de Bruxelles.

[Meyer, 2021] Meyer, P. E. (2021). Shared datasets - http://www.bioinfo.uliege.be/meyer/covid.html.

[Meyer et al., 2007] Meyer, P. E., Kontos, K., Lafitte, F., and Bontempi, G. (2007). Information-theoretic inference of large transcriptional regulatory networks. *EURASIP Journal on Bioinformatics and Systems Biology*, Special Issue on Information-Theoretic Methods for Bioinformatics.

[Mitchell, 1997] Mitchell, T. (1997). *Machine Learning*. McGraw Hill.

[Neapolitan, 2003] Neapolitan, R. E. (2003). *Learning Bayesian Networks*. Prentice Hall.

[Neil and Fenton, 2021] Neil, M. and Fenton, N. (2021). Latest statistics on england mortality data suggest systematic mis-categorisation of vaccine status and uncertain effectiveness of covid-19 vaccination. Preprint.

[Pearl, 2000] Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

[Sakamoto and Kitagawa, 1987] Sakamoto, Y. and Kitagawa, G. (1987). *Akaike information criterion statistics*. Kluwer Academic Publishers.

[Semenzato et al., 2021] Semenzato, L., Botton, J., Drouin, J., Cuenot, F., Dray-Spira, R., Weill, A., and Zureik, M. (2021). Maladies chroniques, états de santé et risque d'hospitalisation et de décès hospitalier pour covid-19 : analyse comparative de données des deux vagues épidémiques de 2020 en france à partir d'une cohorte de 67 millions de personnes. Rapport EPIPHARE - Groupement d'intérêt scientifique (GIS) ANSM-CNAM.

[Whittaker, 1990] Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley.

[Xu et al., 2021] Xu, S., Huang, R., Sy, L. S., Glenn, S. C., and et al. (2021). Covid-19 vaccination and non–covid-19 mortality risk. MMWR Early Release.

Bioinformatics and Systems Biology Lab, Univeristé de Liège, Belgium.