The European Journal of Health Economics (2022) 23:969–978
https://doi.org/10.1007/s10198-021-01408-8

ORIGINAL PAPER

Optimising the impact of COVID-19 vaccination on mortality and hospitalisations using an individual additive risk measuring approach based on a risk adjustment scheme

Danny Wende[1,2] · Dagmar Hertle[1] · Claudia Schulte[1] · Pedro Ballesteros[1] · Uwe Repschläger[1]

Received: 26 May 2021 / Accepted: 3 November 2021 / Published online: 19 November 2021

Abstract
In this population-based cohort study, billing data from German statutory health insurance (BARMER, 10% of population) are used to develop a prioritisation model for COVID-19 vaccinations based on cumulative underlying conditions. Using a morbidity-based classification system, prevalence and risks for COVID-19-related hospitalisations, ventilations and deaths are estimated. Trisomies, behavioural and developmental disorders (relative risk: 2.09), dementia and organic psychoorganic syndromes (POS) (2.23) and (metastasised) malignant neoplasms (1.99) were identified as the most important conditions for escalations of COVID-19 infection. Moreover, optimal vaccination priority schedules for participants are established on the basis of individual cumulative escalation risk and are compared to the prioritisation scheme chosen by the German Government. We estimate how many people would have already received a vaccination prior to escalation. Vaccination schedules based on individual cumulative risk are shown to be 85% faster than random schedules in preventing deaths, and as much as 57% faster than the German approach, which was based primarily on age and specific diseases. In terms of hospitalisation avoidance, the individual cumulative risk approach was 51% and 28% faster. On this basis, it is concluded that using individual cumulative risk-based vaccination schedules, healthcare systems can be relieved and escalations more optimally avoided.

Keywords COVID-19 · Vaccination prioritisation · Immunization strategy · Severe outcomes · Risk adjustment scheme · Additive risk measuring

JEL Classification I18 · H84 · C41 · C63

Preface

Since 2020, the COVID-19-pandemic has negatively affected social life and economies worldwide. To control the pandemic, the world has been relying on vaccines. The first vaccines were approved between the end of 2020 and the beginning of 2021. At the beginning, vaccines were a rare resource, and their efficient allocation had a particularly high economic and social value. Meanwhile, many countries made good progress in vaccinating their population, but from a global point of view, and considering that new viruses or mutant variants may arise, it is of great importance to evaluate existing vaccination schedules and optimise methods of vaccine allocation. The process of vaccination itself takes time, and manufacturers repeatedly run into supply shortages (e.g. with AstraZeneca, see [1]). Ineffective vaccine allocation schedules potentially result in overburdened healthcare systems and preventable deaths occur. In addition to avoiding escalations connected to age or underlying conditions, the establishment of national vaccination strategies is aimed at ensuring early vaccination for groups particularly at risk due to occupation or living situation, or who work in front-line areas [2]. National vaccination strategies are built on international studies that have investigated specific diseases and age-differentiated risk cohorts with regard to COVID-19 (see Appendix for a selection of nations). Coronavirus Vaccination Regulations of all countries reviewed—including the USA, the largest European countries and a selection of

✉ Danny Wende
danny.wende@barmer.de

[1] Bifg Institute of BARMER, Wuppertal, Germany

[2] TU Dresden c/o Chair of Econometrics, Dresden, Germany

Asian countries—took into account only one characteristic at a time when considering which persons were to be prioritised for vaccination, i.e. age, morbidity, living situation or occupation. Given this fact, and the determination that each potential receiver of a vaccination is assigned to exactly one group, we refer to such a vaccination schedule, hereinafter, as the "cell approach". Definitions of cells are hardly comparable between countries. For example, there are no common specified diagnoses. The cell approach has significant limitations for vaccination strategy, namely the lack of comparability of disease definitions to primary studies, information on disease stage or therapy and consideration of interactions in terms of age and underlying conditions. Notably, lack of investigation of interactions between age and underlying conditions presents further limitations. Although age is by far the most significant risk factor [3], until now, it has not been possible to establish whether younger people with severe lung diseases, for example, have significantly higher mortality risk than relatively healthy 75-year-olds, and should, therefore, be assigned higher priority [4]. In addition, adopting the cell approach leads to inhomogeneous and oversized priority groups in the population (see Appendix).

Apart from reviewing national Vaccination Regulations, we conducted an orienting literature search with regard to international studies on COVID-19 vaccination strategies.[1] Literature related to COVID-19 vaccination strategies distinguishes between epidemiological simulation models [5], empirical probability models [2], empirical spatial dispersion models [6], cost-effectiveness studies [7] and observational medical studies [8] of individual diseases. The majority of previous studies consider the cell approach based on age groups or prioritisation of isolated diseases. Only few studies address individual cumulative risks based on a wide range of information that is compiled by means of machine learning techniques. Notable examples include Mellado et al. who studied the adoption of vaccination strategies at the individual level in Africa using artificial intelligence, and Russo et al. who used cumulative risks in Italy to prioritise groups in terms of Lasso techniques [9, 10]. However, these studies can only give a first exploratory impression of the effectiveness of cumulative approaches, as the relevant aspect of modelling COVID-19 exacerbating conditions is carried out on small samples and by means of machine intelligence instead of medical expertise, which carries the risk of missing relevant factors due to under-sampling and a lack of interpretability.

[1] We searched PubMed, medRxiv, bioRxiv and arXiv, for peer-reviewed articles, preprints, and research reports on COVID-19-associated vaccination strategies, using the search strategy „(covid [Titel] or corona [Titel]) and (vacc* [Titel]) and (strateg* [Titel] or prio* [Titel])" up to August 27, 2021.

The ECDC writes with regard to the vaccination strategy: "The individual risk of hospitalisation and death increases with the number of preconditions. […] The causality and magnitude of risk from each (of these) underlying conditions should be monitored and periodically reviewed" [11].

Our study endeavours to close this gap by developing a prioritisation model for COVID-19 vaccinations based on patient-individual cumulative underlying conditions to demonstrate that this approach can be used to vaccinate vulnerable individuals more rapidly. In the following, we refer to this approach as the "individual approach". For the selection of persons based on age and underlying conditions, an individual approach determines the personal risk for each individual. In addition to simultaneous consideration of age and underlying conditions, the determination of cumulative risks, as they arise when several underlying conditions are present, is also considered. Prioritisation based on occupation or living situation is not considered in this study. To the best of our knowledge, this individual approach is the first to provide a classification of COVID-19 relevant underlying conditions across the full ICD catalogue. Using German health insurance data, prevalence and risks for COVID-19-related hospitalisations, ventilations and deaths are estimated on a sample of circa 9 million insured persons. Moreover, optimal vaccination priority schedules for participants are established on the basis of individual escalation risk and in comparison with the procedure chosen by the German Government. We estimate how many people would have already received a vaccination prior to severe escalation.

In the following section, Germany's vaccination strategy is described, and it is outlined why we consider Germany to be a suitable basis for evaluating our approach. Nevertheless, what has been stated can be applied to all countries that have the necessary database. The second section presents the method used to build the COVID-19 conditional risk model. The model itself is available in the Appendix. The third section presents the empirical goodness of fit of the model and the fourth section demonstrates impact simulations of the model against Germany's vaccination strategy. The paper concludes with discussion, limitations and conclusion.

The German COVID-19 vaccination strategy

In the following, the Federal German Coronavirus Vaccination Regulations of 14 December 2020 (CoronaImpfV) is considered [12].

In the CoronaImpfV, the Federal Government of Germany defined four groups with descending priority for the vaccination schedule [4, 13]. As in most western countries, this prioritisation is based on recommendations of an expert council (see Appendix)—in Germany, this is

Table 1 Relevant underlying conditions of priority Group 3 (§ 4 CoronaImpfV)

People at increased risk of severe escalation or death following infection with SARS-CoV-2 coronavirus
People with obesity (persons with body mass index over 30)
People with chronic kidney disease
People with chronic liver disease
People with immunodeficiency or HIV infection
People with diabetes mellitus
People with heart failure, arrhythmia, atrial fibrillation, coronary artery disease or arterial hypertension
People with cerebrovascular disease or apoplexy
People with cancer
People with COPD or bronchial asthma
People with autoimmune diseases or rheumatic diseases

the Standing Committee on Vaccination (STIKO). Allocations of persons to each group can overlap; therefore, group sizes cannot be conclusively determined. The age- and underlying condition-based elements were compiled exclusively on the basis of literature and without independent raw data analysis. Moreover, only underlying conditions that were investigated in systematic reviews could be considered in the recommendation. Rare underlying conditions or conditions not investigated so far may be absent.

Priority Group 1 of the Federal Coronavirus Vaccination Regulations is restricted to all people 80 or older, nursing home residents and healthcare workers, but no persons were selected on the basis of underlying conditions. Priority Group 2 is restricted to persons over 70 and persons with dementia, Down's syndrome, or organ transplant recipients. Priority Group 3 includes more than 10 million persons between 60 and 69 years, and at least five million persons selected on the basis of occupation, as well as all persons with relevant underlying conditions (Table 1). Since the underlying conditions were only vaguely defined, a further 10 to 15 million persons could be added who suffer from at least one named underlying condition. Vaccinating every willing person in Group 3 took several months due to size. Vaccination within this group occurred at a quasi-random schedule.

In Germany, about 90% of the population (73.36 million) is insured in the statutory health insurance scheme. Similar to the NHS, among other systems, the German health insurance system has the ability to provide comprehensive data for vaccination prioritisation. Therefore, the German vaccination strategy lends itself to comparison between a cell approach and an individual risk approach.

Methodology

Systematic identification of underlying conditions that exacerbate COVID-19

Routine data-based classification of underlying conditions that exacerbate COVID-19 is needed for a personal-risk-based vaccination strategy. Conditions are coded internationally according to ICD-10. To our knowledge, there is no consolidated classification of ICD diagnoses for underlying conditions that exacerbate COVID-19.

Accordingly, current classification of risk adjustment in health insurance exchange (RA) is used to detect underlying conditions that exacerbate COVID-19. The RA distributes statutory health insurance contributions to statutory health insurance funds in a way that corresponds to expected medical resource consumption of insured persons. In this process, anticipated resource consumption is determined individually and primarily on the basis of morbidity groups whose risk weightings are determined by prospective multiple regression, i.e. resource consumption originates from year t and group membership from $t-1$. Multimorbid insured persons may belong to several or many morbidity groups whose risk weightings then accumulate. Comparable RAs exist in countries such as Germany, the Netherlands, Switzerland, Austria and the USA [14]. It is widely accepted in the literature that risk structure adjustment models serve as the basis for resource allocation models [15].

The current classification of German RA is used. The classification of German RA is based on the respective valid version of the ICD-10-GM (German Modification) and classifies all valid diagnoses according to the respective reporting procedure into hierarchised groups. Thus, the spectrum of conditions coded in the ICD is fully represented. Its hierarchical structure at the individual level avoids the double recording of morbidities of the same aetiology but different severity. Coding errors are minimised by a sophisticated system of technical verification and medical validation procedures. For example, in addition to the coded ICD diagnoses, drug use, hospitalisation and dialysis identification are among the criteria required for morbidity assignment. The current classification of RA is, therefore, recommended for estimating the impact in Germany of underlying conditions that exacerbate COVID-19. However, the aim in this case is to predict not the use of resources, but the probability of hospitalisation, ventilation or death on the basis of known underlying conditions.

Yet, in a cohort-based study such as this, it is not legitimate to assume causality and thus determine a number of avoided hospitalisations, ventilations or deaths. Instead, it

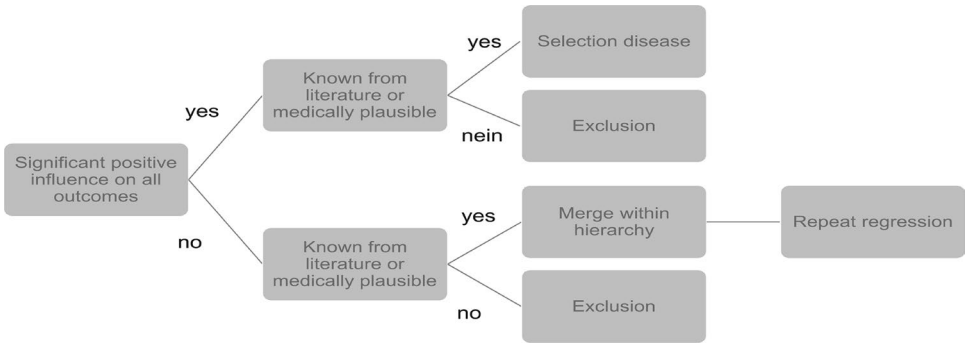

Fig. 1 Identification strategy to qualify relevant underlying conditions

is appropriate to determine, in retrospect, whether alternative vaccination schedules would have prioritised people so that they could have been vaccinated before they became infected and their COVID-19 disease escalated. The key outcome is, therefore, the proportion of people who could have been vaccinated in advance of suffering their escalation.

The COVID-19 model

For the COVID-19 model, outcomes "hospitalisation", "ventilation" and "death" (each following laboratory-confirmed COVID-19 infection) are treated as severe escalations.

The classification of RA, which was developed for prospective use, is examined to see how well it predicts these outcomes. For this purpose, in a multi-stage procedure, RA classification was restricted to underlying conditions for which an association with severe COVID-19 escalations could be identified (Fig. 1):

The selection of underlying conditions that exacerbate COVID-19 was based on a regression model and medical evaluation of results. For this purpose, all insured persons of BARMER were grouped according to the classification model of German RA (495 morbidity groups and 40 age and sex groups). Logistic regressions on the outcomes were performed using the group affiliations as independent variables. Underlying conditions whose regression coefficients did not significantly indicate a more severe escalation were questioned. In contrast, a significant impact on the temporal sequence with infection, hospitalisation, ventilation and death was considered a significant COVID-19-related medical history. Nevertheless, as no causal statements can be made, current literature and medical plausibility were used to substantiate the relevant underlying conditions.

The entire decision-making process was iteratively repeated until all underlying conditions were clearly identified as included or excluded.

The analysis was limited to the German prioritisation Groups 2–4, as the vaccination of prioritisation Group 1 (people in residential care facilities, people ≥ 80 years and caregivers) has already begun.

Regression model and vaccination strategy

A vaccination strategy, $rang : \mathbb{R} \to \mathbb{N}, z \mapsto r$, assigns each member of a society $i \in \{1, \ldots, N\}$ an individual position within a vaccination schedule $r \in \{1, \ldots, N\}$, accoding to an expectation z. The observation of an escalation is y_{ik}, with $k \in \{\text{hospitalisation, ventilation, death}\}$. Let us assume that individuals are only differentiated according to their underlying conditions, $x_i^{'} = \left(x_{ij}\right)_i$.

In the cell approach, individuals are assigned to a position based on the most relevant underlying condition:

$$r_{ik}^{cell} = rang\left(z_{ik}\right) \quad \text{with} \quad z_{ik} = \max_j \left(x_{ij}\alpha_k\right)$$

The government determines the importance of underlying conditions, α_k, in terms of prioritisation groups. In the individual approach, individuals receive their positions according to their individual risk of experiencing an escalation, $P\left[y_{ik} = 1\right] = P[y_{ik} = 1 | X = x_i^{'}]$:

$$r_{ik}^{ind} = rang\left(P\left[y_{ik} = 1\right]\right) \quad \text{with} \quad P\left[y_{ik} = 1\right] = \frac{\exp\left(x_i^{'}\beta_k\right)}{1 + \exp\left(x_i^{'}\beta_k\right)}$$

Assuming an independent and identically Gumbel distributed error term, the individual risk can be determined by a logistic regression, with coefficients $\beta_k = \left(\beta_{1,k}, \ldots, \beta_{j,k}\right)$ [16].

The higher the quality of a vaccination strategy, the better it succeeds in offering vaccination to as many vulnerable people as possible at an early stage. To determine this quality, a thought experiment is used. Let us assume that a vaccine had already been available at the beginning of the COVID-19 pandemic and that vaccination had begun according to one of the presented strategies. Assuming further that all escalations would have been prevented by prior vaccination, one would ask what proportion of the population must be vaccinated until a significant share of escalations would have been prevented. Given a very small COVID-19 escalation rate in the population, this thought experiment is equivalent to the information given by the receiver operating characteristic (ROC). Therefore, the

primary measure of the vaccination strategy is the area under the ROC curve (AUROC). In addition, the coefficient of determination according to Nagelkerke is specified [17]. It can be argued that the effectiveness of vaccination is not 100% and not every severe escalation can be prevented. Even though the effectiveness of vaccination might not be equal among the entire population, general certainty about the presence of similar effectiveness is sufficient for measuring the relative quality of vaccination schedules by AUROC. A precise consideration is, however, negligible because COVID-19 diseases that escalate despite vaccination are not preventable by any vaccination schedule.

The thought experiment is limited by the fact that the probability model is trained a posteriori on COVID-19 escalations. In reality, however, future COVID-19 escalations are unknown. To address this issue, the performance measures are additionally determined via ten-fold cross-validation. Performance measures determined this way are referred to as "Out of sample" (Oos).

In the following, the individual and cell approaches according to the Federal Coronavirus Vaccination Regulations are compared.

Results

Sample description

The analysis for vaccination prioritisation was based on observations of approximately nine million insured persons of BARMER. The focus was placed on inpatient cases between 1 January 2020 and 30 November 2020 with a clinico-epidemiologically confirmed COVID-19 infection (U07.1). The classification of underlying conditions was based on 2019 data. A classification whose data may already include circumstances that were first caused by COVID-19 would lead to distorted results. Table 2 shows key sample figures.

Observed escalations are differentiated according to which ventilation took place and escalations in which the patient died, during, or up to 30 days after hospitalisation. Death was more common than ventilation (20.5% of hospitalised escalations, 11.4%). The average age of persons in the present sample was 48 years; 70 years at hospitalisation, and 82 years at death (both with COVID-19). These figures correspond to comparable study results and statistics from the Robert Koch-Institute [4]. Men are overrepresented among severe COVID-19 escalations, especially in cases of ventilation. The temporal progression of inpatient COVID-19 escalations and deaths corresponds to known developments within the entire population (cf. Fig. 2).

Table 2 Key sample figures. Source: BARMER data, without restriction to priority Groups 2 to 4 CoronaImpfV

Key figure	Baseline data	Inpatient with COVID-19	Received ventilation	Deceased
Persons	9,154,806	13,464	1539	2753
Average age	48.19	70.42	71.23	81.67
Proportion of men	43.31%	43.73%	60.03%	48.18%
Number of underlying conditions*	4.08	8.70	9.55	10.66
Admission days	1.56	20.86	41.53	25.63
Ventilation hours	0.59	34.32	300.25	80.48

*Underlying conditions according to RA classification model 2021 with morbidity 2019

Fit of the COVID-19 model

The RA includes 535 cells (40 age and sex groups). However, due to the restriction on priority Groups 2 to 4, eight age groups denoting an age of 80 years or more are omitted. The final model adapted to COVID-19 has 98 cells with 32 age groups.

When the observed escalations are explained using one of two models, the performance measures shown in Table 3 are obtained. The coefficients of determination range between 7 and 21%. Taking into account the extreme nature of escalations, the values can be interpreted as moderate to good. The coefficient of determination increases with more severe escalations. The RA model shows a slightly better fit than the final COVID-19 model.

The AUROC ranges from 76 to 94%, i.e. models succeed in prioritising a random patient with escalation over a random patient without escalation in up to 94% of cases. Such classification rates are interpreted as good to very good.

The consideration of the out-of-sample performance measures shows that the RA model, with the exception of hospitalisation, has a significantly deteriorating performance. The COVID-19 model remains stable. The difference is most obvious in the criterion OoS-AUROC. For example, while the prioritisation rate for ventilated patients decreases by 7% using the RA model, the rate falls by only 1% in the COVID-19 model. It can be concluded that the RA model suffers from overfitting and that optimisation of the model with regard to the COVID-19 underlying condition is purposeful.

Estimations

Identified underlying conditions, together, account for 38.4% of explained variation in hospitalisation risk, 51.5%

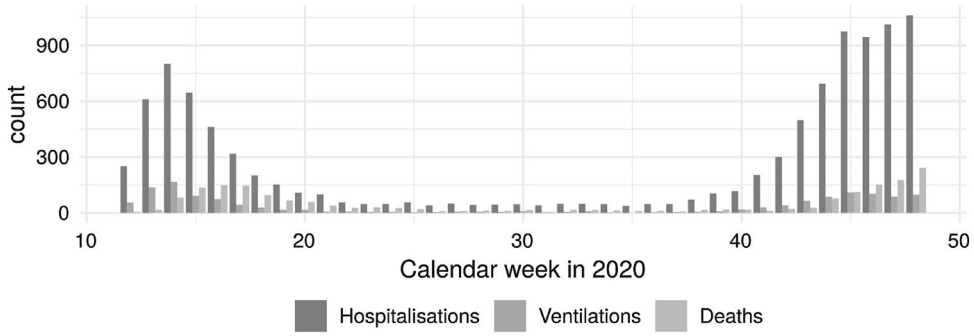

Fig. 2 Time course of COVID-19 cases

Table 3 Performance measures of classification models

	Hospitalisation		Ventilation		Death	
	RA (%)	COVID-19 model (%)	RA (%)	COVID-19 model (%)	RA (%)	COVID-19 model (%)
R^2	6.9	6.6	13.4	11.3	21.2	18.3
AUROC	76.9	75.8	88.6	87.2	93.9	92.9
Oos-R^2	5.5	6.0	1.8	9.2	7.6	17.0
OoS-AUROC	75.6	75.6	81.7	86.0	87.7	92.4

R^2 Nagelkerke coefficient of determination, *AUROC* area under the receiver operating characteristic, *Oos*-out-of-sample estimator via tenfold cross-validation

Table 4 Top 10 underlying conditions with risk association to COVID-19 mortality

Underlying conditions	Insured person (statutory scheme)	Hospitalisation	Ventilation	Death
Trisomies, behavioural and developmental disorders	439,231	2.09*** (1.59–2.75)	2.18* (1.03–4.61)	5.73*** (3.55–9.27)
Dementia and organic psychoorganic syndromes	362,189	2.23*** (1.98–2.51)	2.11*** (1.60–2.78)	5.54*** (4.57–6.72)
(Metastasised) malignant neoplasms	211,613	1.99*** (1.66–2.38)	1.60* (1.00–2.54)	3.85*** (2.80–5.29)
Haematological neoplasms	243,932	1.68*** (1.42–2.00)	2.40*** (1.69–3.41)	2.95*** (2.15–4.04)
Mental illnesses	1,576,220	1.70*** (1.51–1.92)	2.02*** (1.45 – 2.82)	2.86*** (2.16–3.79)
Severe kidney disease, dialysis	335,354	2.15*** (1.89–2.45)	2.37*** (1.80–3.14)	2.83*** (2.18–3.67)
HIV Tuberculosis Systemic mycoses	67,885	1.47 (0.96–2.26)	3.35*** (1.65–6.79)	2.62** (1.15–5.98)
Cirrhosis of the liver Liver failure	184,728	1.23 (0.99–1.54)	1.36 (0.84–2.20)	2.32*** (1.62–3.33)
Infections with multi-resistant germs/opp. Pathogens	439,231	1.51*** (1.26–1.81)	1.79** (1.22–2.62)	2.26*** (1.65–3.10)
Severe neurological diseases	362,189	1.72*** (1.46–2.01)	1.45 (0.95–2.20)	2.16*** (1.58–2.94)

For the underlying condition definition used here, please refer to the allocation table published. The complete list of underlying conditions is available in the results table, which has also been published

*pr(>|z|) < 0.05; **pr(>|z|) < 0.025; ***pr(>|z|) < 0.001; alpha-95% confidence interval in parentheses

in ventilation risk and 51.4% in mortality risk. On average, the value of the information of a single underlying condition equals the value of information of three years of age for hospitalisation, two years of age for ventilation and half a year of age for death. The high combined determination of underlying conditions results from the high multimorbidity of patients (cf. Table 3) and the multiplicative effect of risk ratios in a logistic regression model.

Table 4 shows estimated results for the ten most important underlying conditions with risk association to COVID-19 mortality.

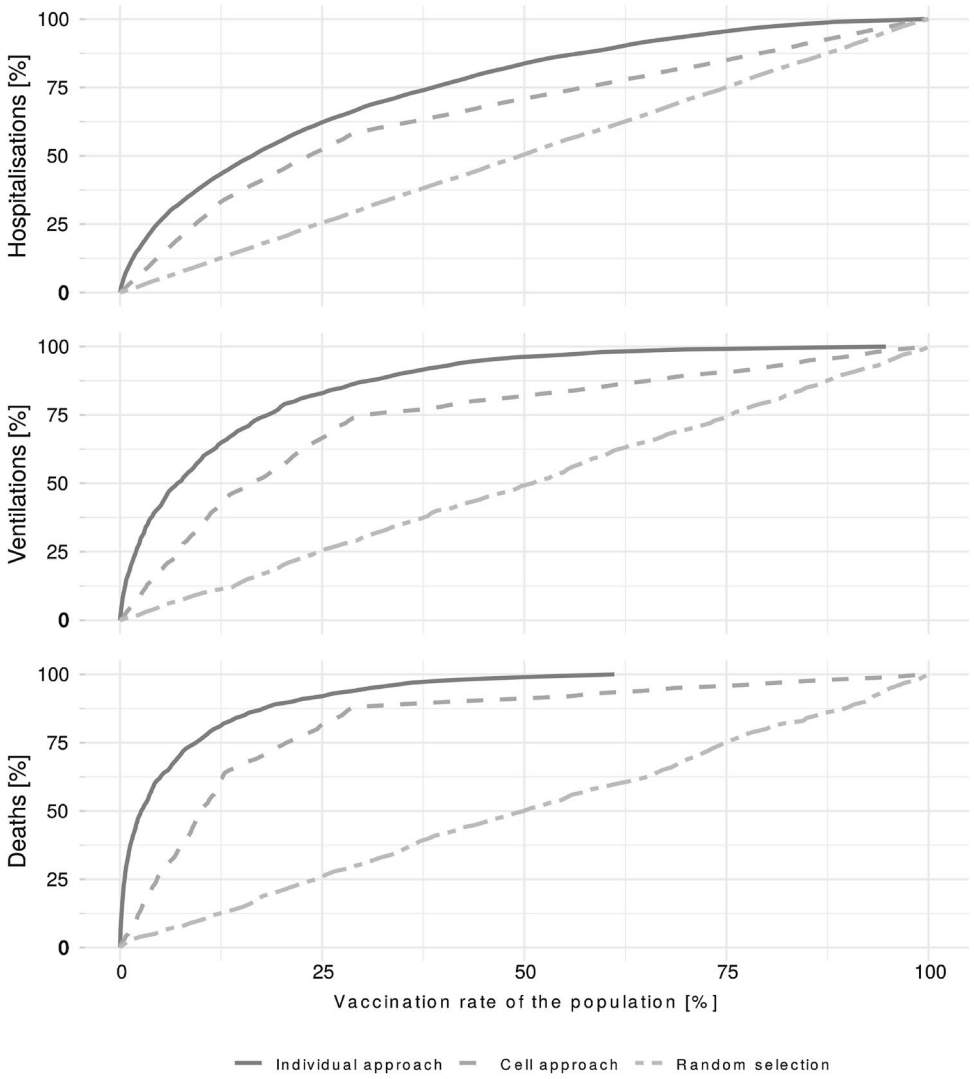

Fig. 3 Proportion of vaccinated hospitalised, ventilated, and deceased patients in 2020, for different vaccination schedules, at a given population vaccination rate

The values in each row represent the risk ratio for escalation of COVID-19 patients with a corresponding underlying condition compared to reference patients with COVID-19 of the same age and sex but without an underlying condition. The risk ratios from comorbidities can be determined line by line by multiplying the individual risk ratios. If a person with Down's syndrome also suffers from psychosis, for example, the risk for hospitalisation is multiplied by 3.6, for ventilation by 4.4 and for death by 16.4 compared to a healthy patient of the same age. The estimation results of all underlying conditions is available in the Appendix.

Impact of the individual vaccination strategy on hospitalisations, ventilations, and deaths

In the following, three graphs are shown for hospitalised, ventilated and deceased patients with COVID-19 (Fig. 3). The graphs are differentiated according to the vaccination strategy of an individual approach, cell approach and random vaccination. Each graph shows what proportion of the respective patient group (in 2020) would have been vaccinated at a certain level of vaccination coverage in the population. The higher this proportion, the sooner the most vulnerable patients would have been vaccinated, and the greater the number of hospitalisations, ventilations and deaths possibly avoided.

If 1% of the population had been vaccinated, 10% of hospitalised patients, 17% of ventilated patients and 34% of deceased patients would have already been vaccinated using the individual approach. Using the cell approach, on the other hand, there would have been 3% hospitalised, 4% ventilated and 6% deceased. With random vaccination, the figures would have been 1% in each case. Expressed in vaccine doses, if prioritisation is based on the COVID-19 model, the same prospective reduction in deaths is achieved at the start of the second priority group's immunisation with approximately 500,000 vaccine doses (0,7% vaccination coverage) as would be the case

with the currently planned prioritisation with four million vaccine doses (5.2% vaccination coverage).

With the individual approach, the required vaccinated population proportion, in which 50% of patients with a severe escalation would have already been vaccinated, is 16% for hospitalisation, 7% for ventilation, and 3% for deceased. For the cell approach, on the other hand, it would be 23%, 16%, and 10%, respectively. Random vaccination requires a 50% vaccination rate in the population.

Regardless of herd immunity, 100% of the population would have needed to be vaccinated in all three approaches to have substantially minimised all hospitalisations and, thus, also all ventilations. With regard to deceased patients, a vaccination rate of 62% would have been sufficient in the individual approach—nevertheless, as the pandemic progresses in subsequent years, a vaccination rate of 100% is necessary to reduce mortality to a minimum.

Vaccinating 1% of the population at a constant rate represents one time unit. Accordingly, in the case of random vaccination, the first percentage of the population wait one time unit, the second percentage two time units, and the last percentage 100 time units. In total, a waiting time of 5050 time units exists for the random approach, independent of severe COVID-19 courses. For the individual approach, this waiting time would be only 2474 cumulated time units for hospitalised people, which is 51% faster (1-2,474/5,050). Applying similar reasoning, a ventilated person would be vaccinated 74% faster and a deceased person 85% faster than with a random vaccination strategy. Compared to the cell approach, a hospitalised person would be vaccinated 29% faster, a ventilated person 49% faster and a deceased person 57% faster.

Discussion

International comparison of vaccination strategies and morbidity groups

In all western nations considered, prioritisations are determined according to age, underlying conditions and exposed population groups. With regard to the priority criterion "underlying conditions", the differentiation of information in various countries varies considerably. The spectrum ranges from unspecified "risk criteria" (Spain) to indications of severity based on laboratory values (e.g. HbA1c > 8% for diabetics in Switzerland). In all countries, only one risk factor is taken into account for grouping the vaccination schedule, i.e. the highest individual risk in each case. Underlying conditions mentioned most frequently, but with varying degrees of differentiation, are trisomy 21, dementia, cancer, chronic kidney and liver disease, obesity, immunodeficiencies of various causes, cardiovascular disorders and chronic lung conditions (cf. the Appendix).

Prioritisation by morbidity differs between countries. Only in Germany, do mental illnesses appear in the STIKO list of underlying conditions to be considered—but a connection to more severe COVID-19 escalations has been described in international literature [18, 19]. The present analysis confirms this and shows, as does the study by Wang et al. [20], that mental illness is associated with an increased risk of a severe COVID-19 escalation independent of age and other comorbidities.

The reverse is true for obesity, which is listed in almost all countries (albeit in different BMI figures). Obesity also occurs as a sole factor in the present analysis of the 66 relevant underlying conditions, but in the lower third. This example also shows the advantage of a cumulative view, because obesity does not cause escalation independently, but rather in connection with other underlying conditions, such as diabetes and hypertension. Both examples show the importance of assessing possible interdependencies of different comorbidities on the risk of severe COVID-19 escalations. There are international studies that also consider age and comorbidities cumulatively with regard to COVID-19 escalation [2, 21, 22]. Results and statements of these studies are essentially congruent with those presented. However, to our knowledge, no systematic analysis of comorbidities on the full ICD spectrum, nor the impact on possible vaccination strategies, has been reviewed to date. This study promises to make progress in this regard. The presented approach can be implemented in most western healthcare systems.

Comparison of individual approach and cell approach using the example of Germany (COVID-19 model vs CoronaImpfV)

Firstly, the present study demonstrates that individual underlying conditions or morbidity groups acquire a different vaccination prioritisation when age, sex and several comorbidities are considered in combination, since a purely literature-based prioritisation does not consistently consider further comorbidities. Secondly, it is demonstrated that each individual should be offered vaccination on the basis of individual risks.

The strength of the individual approach (COVID-19 model) compared to the cell approach (Federal Coronavirus Vaccination Regulations) is clearly demonstrated by the more precise specification of underlying conditions and multiple risk consideration in the example of persons insured with BARMER. This more precise vaccination schedule made possible by our individual approach can prevent COVID-19-related deaths, and can relieve hospitals of COVID-19-related admissions in the medium term (Fig. 3). Based on the present study's findings, calculations

can be made as to how different vaccination strategies affect the optimal use of available vaccine doses. Considering vaccine availability, this appears to be of great relevance.

Limitations

In Fig. 3, it was assumed that all people within a priority level participate equally in vaccination, regardless of their individual risk. Already in priority Group 1, vaccination registration usually has to be done actively via hotlines or the internet. In the later stages, the need to actively obtain a medical certificate is added for people with underlying conditions. It is possible that self-selection occurs, so that people with higher risk independently obtain higher prioritisation.

The investigation includes inpatient cases between 1 January 2020 and 30 November 2020. As the treatment and the success of the treatment have changed in this period, our results may include inaccuracies.

The results enable a risk assessment for morbid and multimorbid individuals. Other goals of a vaccination strategy, such as to prevent the spread of the virus, or to protect important front-line worker groups, are closed to analysis.

Although none of the vaccines has been licensed in Germany for children under 12, minors were included in the analyses because it is unknown when such a license will be granted. The exclusion of children would have only a minor impact on the model itself due to the lower number of severe escalations and relevant underlying conditions in children.

No valid statements can be made for orphan diseases, or those where only very few BARMER-insured persons were to be affected, due to the limited number of cases.

The linear modelling proposed by this study limits the ways in which comorbidity is covered. A number of other studies use supervised learning algorithms to circumvent this limitation. However, complex algorithms such as gradient boosting are a black box, at least for the general public. Simple rule-based algorithms, such as classification trees, have a different problem; in the case of a low frequency of outcomes, the chance of ecological fallacies rises as random events become dominant even among large comorbidity groups. To demonstrate our point, we additionally model the hospitalisation risk using a classification tree and a boosted tree. Both algorithms outperform the linear model in terms of in-sample AUC (0.814, 0.772). However, the out-of-sample AUC for the classification tree falls below the linear model (0.700). The boosted tree performs well (0.763), but generates an incommunicable large set of rules. Nevertheless, in terms of validity and applicability, it is essential for vaccination strategies to incorporate medical knowledge, which limits the methodology.

Concluding remarks

The most efficient vaccination strategy is a crucial factor in how quickly countries' healthcare systems can be relieved and how many escalations can be avoided. The vaccination strategies of western countries are built on a literature-based grouping of the population into priority cells based on age, occupation and selected underlying conditions with varying definitions. The presented study shows a methodological approach that complements the literature-based prioritisation in a data-driven setting. The approach places the diverse underlying condition definitions on a uniform basis and extends the consideration to the entire spectrum of the International Statistical Classification of Diseases and Related Health Problems. Such a population classification allows the assessment of cumulative risks, and, consequently, a vaccination strategy oriented towards the individual risk of escalation. Our data demonstrate that such a vaccination strategy is advantageous for achieving essential vaccination targets faster, especially when vaccine availability is initially limited. The chosen approach is also applicable in other healthcare systems.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10198-021-01408-8.

References

1. European Commission: Press statement by Commissioner Kyriakides on vaccine deliveries and on the vaccine export transparency scheme (2021)
2. Williams, J., Degeling, C., McVernon, J., Dawson, A.: How should we conduct pandemic vaccination? Vaccine (2021). https://doi.org/10.1016/j.vaccine.2020.12.059
3. Bubar, K.M., Reinholt, K., Kissler, S.M., Lipsitch, M., Cobey, S., Grad, Y.H., Larremore, D.B.: Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. Science (2021). https://doi.org/10.1126/science.abe6959
4. Vygen-Bonnet, S., Bogdan, C., Heininger, U., Littmann, M., Meyer, H., Schmid-Küpke, N., Terhardt, M., Überla, K., van der Sande, M., Wichmann, O., Wicker, S., Wiedermann, U., Wild, V., von Kries, R.: Beschluss der STIKO zur 2. Aktualisierung der COVID-19-Impfempfehlung und die dazugehörige wissenschaftliche Begründung. Robert Koch Institute. Robert-Koch-Institut. https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2021/Ausgaben/02_21.pdf?__blob=publicationFile (2021)
5. Libotte, G.B., Lobato, F.S., Platt, G.M., Silva Neto, A.J.: Determination of an optimal control strategy for vaccine administration in COVID-19 pandemic treatment. Comput. Methods Programs Biomed. (2020). https://doi.org/10.1016/j.cmpb.2020.105664
6. Grauer, J., Löwen, H., Liebchen, B.: Strategic spatiotemporal vaccine distribution increases the survival rate in an infectious disease like COVID-19. Sci. Rep. (2020). https://doi.org/10.1038/s41598-020-78447-3
7. Kohli, M., Maschio, M., Becker, D., Weinstein, M.C.: The potential public health and economic value of a hypothetical COVID-19 vaccine in the United States: Use of cost-effectiveness modeling

to inform vaccination prioritization. Vaccine (2021). https://doi.org/10.1016/j.vaccine.2020.12.078
8. Ribas, A., Sengupta, R., Locke, T., Zaidi, S.K., Campbell, K.M., Carethers, J.M., Jaffee, E.M., Wherry, E.J., Soria, J.-C., D'Souza, G.: Priority COVID-19 vaccination for patients with cancer while vaccine supply is limited. Cancer Discov. (2020). https://doi.org/10.1158/2159-8290.CD-20-1817
9. Mellado, B., Wu, J., Kong, J.D., Bragazzi, N.L., Asgary, A., Kawonga, M., Choma, N., Hayasi, K., Lieberman, B., Mathaha, T., Mbada, M., Ruan, X., Stevenson, F., Orbinski, J.: Leveraging artificial intelligence and big data to optimize COVID-19 clinical public health and vaccination roll-out strategies in Africa. Int. J. Environ. Res. Public Health (2021). https://doi.org/10.3390/ijerph18157890
10. Russo, A.G., Decarli, A., Valsecchi, M.G.: Strategy to identify priority groups for COVID-19 vaccination: A population based cohort study. Vaccine (2021). https://doi.org/10.1016/j.vaccine.2021.03.076
11. European Centre for Disease Prevention and Control (ECDC): COVID-19 vaccination and prioritisation strategies in the EU/EEA, Stockholm (2020)
12. Bundesministerium für Gesundheit: Verordnung zum Anspruch auf Schutzimpfung gegen das Coronavirus SARS-CoV-2 (Coronavirus-Impfverordnung). ImpfVO (2020)
13. Vygen-Bonnet, S., Bogdan, C., Heininger, U., Littmann, M., Meyer, H., Schmid-Küpke, N., Terhardt, M., Überla, K., van der Sande, M., Wichmann, O., Wicker, S., Wiedermann, U., Wild, V., von Kries, R.: Beschluss der STIKO für die Empfehlung der COVID-19-Impfung und die dazugehörige wissenschaftliche Begründung. STIKO-Empfehlung zur COVID-19-Impfung. Robert-Koch-Institut. https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2021/Ausgaben/02_21.pdf?__blob=publicationFile (2021)
14. Ellis, R.P., Martins, B., Rose, S.: Risk Adjustment for Health Plan Payment. In: Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets, pp. 55–104. Elsevier (2018)
15. Dixon, J., Smith, P., Gravelle, H., Martin, S., Bardsley, M., Rice, N., Georghiou, T., Dusheiko, M., Billings, J., Lorenzo, M.D., Sanderson, C.: A person based formula for allocating commissioning funds to general practices in England: development of a statistical model. BMJ (Clin. Res. Ed.) (2011). https://doi.org/10.1136/bmj.d6608
16. Cameron, A.C., Trivedi, P.K.: Microeconometrics. Methods and applications, 7th edn. Cambridge Univ. Press, Cambridge (2008)
17. Nagelkerke, N.J.D.: A note on a general definition of the coefficient of determination. Biometrika (1991). https://doi.org/10.1093/biomet/78.3.691
18. de Hert, M., Mazereel, V., Detraux, J., van Assche, K.: Prioritizing COVID-19 vaccination for people with severe mental illness. World Psychiatry (2021). https://doi.org/10.1002/wps.20826
19. Li, L., Li, F., Fortunati, F., Krystal, J.H.: Association of a prior psychiatric diagnosis with mortality among hospitalized patients with coronavirus disease 2019 (COVID-19) infection. JAMA Netw. Open (2020). https://doi.org/10.1001/jamanetworkopen.2020.23282
20. Wang, Q., Xu, R., Volkow, N.D.: Increased risk of COVID-19 infection and mortality in people with mental disorders: analysis from electronic health records in the United States. World Psychiatry (2021). https://doi.org/10.1002/wps.20806
21. Clift, A.K., Coupland, C.A.C., Keogh, R.H., Diaz-Ordaz, K., Williamson, E., Harrison, E.M., Hayward, A., Hemingway, H., Horby, P., Mehta, N., Benger, J., Khunti, K., Spiegelhalter, D., Sheikh, A., Valabhji, J., Lyons, R.A., Robson, J., Semple, M.G., Kee, F., Johnson, P., Jebb, S., Williams, T., Hippisley-Cox, J.: Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. BMJ (Clin. Res. Ed.) (2020). https://doi.org/10.1136/bmj.m3731
22. McKeigue, P.M., Weir, A., Bishop, J., McGurnaghan, S.J., Kennedy, S., McAllister, D., Robertson, C., Wood, R., Lone, N., Murray, J., Caparrotta, T.M., Smith-Palmer, A., Goldberg, D., McMenamin, J., Ramsay, C., Hutchinson, S., Colhoun, H.M.: Rapid epidemiological analysis of comorbidities and treatments as risk factors for COVID-19 in Scotland (REACT-SCOT): A population-based case-control study. PLoS Med. (2020). https://doi.org/10.1371/journal.pmed.1003374