BMJ Open

Effect of vaccination on the case fatality rate for COVID-19 infections 2020–2021: multivariate modelling of data from the US Department of Veterans Affairs

Glen H Murata ,[1] Allison E Murata,[2] Douglas J Perkins,[3] Heather M Campbell,[2] Jenny T Mao,[4] Brent Wagner,[1,4,5] Benjamin H McMahon,[6] Curt H Hagedorn[4]

To cite: Murata GH, Murata AE, Perkins DJ, *et al*. Effect of vaccination on the case fatality rate for COVID-19 infections 2020–2021: multivariate modelling of data from the US Department of Veterans Affairs. *BMJ Open* 2022;**12**:e064135. doi:10.1136/bmjopen-2022-064135

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-064135).

Received 27 April 2022
Accepted 04 November 2022

For numbered affiliations see end of article.

Correspondence to
Dr Brent Wagner;
brent.wagner@va.gov

ABSTRACT

Objectives To evaluate the benefits of vaccination on the case fatality rate (CFR) for COVID-19 infections.

Design, setting and participants The US Department of Veterans Affairs has 130 medical centres. We created multivariate models from these data—339 772 patients with COVID-19—as of 30 September 2021.

Outcome measures The primary outcome for all models was death within 60 days of the diagnosis. Logistic regression was used to derive adjusted ORs for vaccination and infection with Delta versus earlier variants. Models were adjusted for confounding factors, including demographics, comorbidity indices and novel parameters representing prior diagnoses, vital signs/baseline laboratory tests and outpatient treatments. Patients with a Delta infection were divided into eight cohorts based on the time from vaccination to diagnosis. A common model was used to estimate the odds of death associated with vaccination for each cohort relative to that of unvaccinated patients.

Results 9.1% of subjects were vaccinated. 21.5% had the Delta variant. 18 120 patients (5.33%) died within 60 days of their diagnoses. The adjusted OR for a Delta infection was 1.87±0.05, which corresponds to a relative risk (RR) of 1.78. The overall adjusted OR for prior vaccination was 0.280±0.011 corresponding to an RR of 0.291. Raw CFR rose steadily after 10–14 weeks. The OR for vaccination remained stable for 10–34 weeks.

Conclusions Our CFR model controls for the severity of confounding factors and priority of vaccination, rather than solely using the presence of comorbidities. Our results confirm that Delta was more lethal than earlier variants and that vaccination is an effective means of preventing death. After adjusting for major selection biases, we found no evidence that the benefits of vaccination on CFR declined over 34 weeks. We suggest that this model can be used to evaluate vaccines designed for emerging variants.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A large number of COVID-19 patients from many medical centres.
⇒ Robust control of the severity of potential confounders from pre-existing conditions, vital signs, laboratory tests and outpatient treatment.
⇒ Analysis stratified by time from vaccination to diagnosis to control for urgency (priority) of vaccination.
⇒ Results were limited to populations with characteristics similar to those of US veterans of military service.
⇒ Analysis requires extensive baseline data usually found in patients requiring long-term follow-up.

INTRODUCTION

Recent studies have shown an alarming decrease in the effectiveness of COVID-19 vaccines over time.[1–3] The metrics for vaccine effectiveness included infection and mortality rates.[4 5] It is imperative to have unbiased measures of vaccine effectiveness and an understanding of the mechanisms by which vaccines begin to fail. A robust framework for classifying vaccine effects facilitates such efforts. The most straightforward method is to use a probabilistic approach. Patients who succumb to a contagious disease must first be exposed, then develop an infection as a result of the exposure and then die as a consequence of the infection. The risk of death in a population observed for a given time is thus the joint probability of these three events or P(exposure, infection, death). From the chain rule of probability theory:

P(exposure, disease, death)=P(exposure)×P(infection | exposure)×P(death | exposure, disease).

Likewise, the risk of infection is the joint probability of the first 2:

P(exposure, disease)=P(exposure)×P(infection | exposure).

It is important to separate these metrics into their underlying risks because the latter represent separate targets for interventions. For example, COVID-19 precautions focus on the probability of exposure (P(exposure)), while antivirals target the probability of death (P(death | exposure, disease)). Vaccination

has favourable effects on the probability of infection (P(infection | exposure)) and the probability of death by promoting an immune response. Some projected that the probability of disease would decrease once herd immunity was achieved (depending on the duration of immunity). Thus, vaccine effectiveness might vary depending on which risk is targeted. Moreover, changes in one risk may be offset by changes in another—leading to erroneous conclusions about vaccine effectiveness. For example, the beneficial effects of the vaccine on the probability of infection may be diminished by abandoning COVID-19 precautions, which increases exposure.

One problem with some models for vaccine effectiveness is that they do not account for the risk of exposure. Doing so requires adjustments for patient behaviours and community-level effects. The former includes adherence to COVID-19 precautions such as masking, social distancing, handwashing, avoiding crowds, contact testing and telework. The latter include the prevalence of the virus, its infectivity, the extent to which the community embraces COVID-19 precautions and government mandates. As a result, increases in infection and mortality rates may be related to diminished vaccine effectiveness, changes in exposure or both. In this study, we focused on the third risk or probability of death. This term is analogous to the case fatality rate (CFR). We chose this outcome to assess vaccine effectiveness because, unlike infection and mortality rates, it is not affected by unmeasured patient behaviours and environmental factors. The purpose of this study was to analyse the magnitude and durability of a vaccine's effect on CFR. The hypothesis was that vaccinated patients with COVID-19 had a lower risk of death at 60 days than unvaccinated patients. A secondary hypothesis was that patients with remote vaccinations enjoyed a similar benefit. The analysis was done on existing medical records from the largest integrated healthcare system in the USA—the Department of Veterans Affairs (VA).

METHODS

Measuring effectiveness in observational studies requires a robust approach to confounding factors because treatment is not randomly allocated across a population. Variables that impact a measured outcome may confound retrospective analyses. Regarding COVID-19, these include the likelihood of vaccination, complete vaccination, anergy to inoculation, pre-existing conditions and comorbidity severities. In most cases, treatment benefits are offset by their preferential use in patients with a poorer prognosis. A multivariate analysis separates the independent effects of treatment and associated comorbidities (eg, obesity[6]), removing bias and revealing beneficial factors. The challenge is that hundreds of conditions may serve as confounders. Comorbidity scores may not be suitable for this purpose because they do not represent all conditions that pose a risk. Critical findings in vital signs and laboratory tests may also serve as confounders. The most robust solution is to do a systematic survey of all high-risk conditions from several domains in the medical record and adjust the effect of vaccination by some aggregate measure of their impact. In this study, we developed and applied such procedures for individual International Classification of Diseases (ICD)-10 codes, vital signs, commonly used laboratory tests and outpatient medications.

Finally, patients can have certain traits that majorly affect prognosis but are not easily measured or well represented by their underlying diagnoses. For example, nursing home patients have a poor prognosis but are not easily identified if such care is delivered through a nursing home contract or private arrangement. Nevertheless, it is essential to control for these confounders to get an unbiased estimate of vaccine effect. We felt that the timing of vaccination might be used as a proxy for these traits because VA prioritised its delivery of the vaccine. The COVID-19 Vaccination Plan for the Veterans Health Administration (VHA) acknowledged that elderly patients, certain ethnic groups, and those with major comorbidities were at high risk of death or complications. It also authorised a population-based risk stratification plan for vaccine administration and its implementation when supplies were limited. For this reason, we examined vaccine effect for cohorts defined by the time from vaccination to diagnosis. Members of each cohort had the same priority for vaccination—removing the criteria as potential confounders. Multivariate analysis within each cohort was then used to adjust the effect of vaccination for several patient covariates. This dual approach to confounders also reduces the bias resulting from patient self-selection—that is, patients seeking early vaccination if they believed their health was poor or deferring vaccination if they believed their health was good. Because the estimates were unbiased, it was possible to compare vaccine effectiveness across the cohorts to determine whether it declined over time. Our choice of endpoints, a more robust approach to measured confounders that includes their severity, and a stratified analysis to handle the urgency (priority) of vaccination provided new insights into the benefits of vaccination on CFR.

Cases were identified through the VA's COVID-19 Shared Data Resource (CSDR). The case definition conforms to the US Centers for Disease Control, which requires nucleic acid amplification or antigen testing. CSDR contains cases reported by 130 medical centres and may include non-veterans referred to VA by other agencies. Clinical data were retrieved from VA's Corporate Data Warehouse (CDW) through the VA Informatics and Computing Infrastructure. CDW has been the central repository for all patient data entered into VA's electronic medical records since 2006.

Subjects were included in this study if their index infections occurred before October 2021. Because of the large number of covariates, we wanted to identify the largest sample size possible, so we included all those in the CSDR. Based on surveillance data from the US Centers

for Disease Control, the Delta variant was considered the infecting agent for those presenting in July, August or September of 2021. Although sporadic cases of the Delta variant were reported in late May, Delta was the predominant variant over the time we selected. The primary outcome was death within 60 days of the diagnosis. The outcome was retrieved from the CSDR, which assigns a 1 to those who died and 0 otherwise. CSDR uses several strategies and data sources to ascertain major outcomes. The cohort was followed through November 2021 so that each subject reached a definitive endpoint.

VA maintains two databases containing information on COVID-19 vaccination. CSDR has a robust and highly vetted registry of patients who have been vaccinated within and outside of the agency. The CDW immunisation domain contains similar information but is less structured and possesses duplicates. The CDW data were scrubbed and reorganised to match the CSDR format. Cases identified in CDW, but not in CSDR, were added to the latter to create a pooled vaccine registry. Patients were considered vaccinated if they had received one dose of the Johnson & Johnson product or two doses of any other formulation at least 14 days prior to the diagnosis of COVID-19.

Our study used three novel parameters representing potential confounders from major domains of the medical record. PDeathDx refers to the predicted probability of death based on 153 ICD10 category diagnoses.[7] Pre-existing conditions were identified by reviewing all diagnoses entered into the electronic medical record during outpatient visits, as updates to the patient problem list, or at the time of hospital discharge. 'Pre-existing' refers to entries made up to 14 days before the COVID-19 diagnosis. ICD9 codes were converted to ICD10 using a crosswalk provided by the Centers for Medicare/Medicaid Services. A 'category diagnosis' was defined as all characters preceding the decimal point for ICD10 codes or the ICD9 equivalent. Each patient was deemed to have (or not have) each category diagnosis before COVID-19. A proprietary computer programme was used to identify all patients with a given condition who died or survived, as well as all patients without that condition who died or survived. The software used these cell frequencies to derive the relative risk (RR) of death associated with the condition along with the CI. CIs were adjusted for multiple comparisons by the Bonferroni method. A category diagnosis was considered to have a significant effect on the outcome if the lower limit for the CI was ≥1.5 or the upper limit for the CI was ≤0.80. The procedure was thus used to identify high-risk or protective. Stepwise logistic regression identified those diagnoses that were independent predictors of death. The model was then used to generate a predicted probability of death (PDeathDx) for each subject.

PDeathLabs refers to the predicted probability of death based on 49 parameters derived from complete value sets for four vital signs (systolic blood pressure, diastolic blood pressure, O2 saturation and body mass index) and 7 routine laboratory tests (estimated glomerular filtration rate, alanine aminotransaminase (ALT), haematocrit, serum albumin, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol and haemoglobin A1c). Entries for these 11 clinical measurements were retrieved if their recorded dates were ≥14 days before the diagnosis of COVID-19. A total of 13 parameters were derived for each type of measurement to reflect criteria used by practitioners to assess metabolic control (total=13×11=143). Logistic modelling showed that 49 of these parameters were independently predictive of death.[8] The model assigned a predicted probability of death (PDeathLabs) based on these clinical measurements to each subject.

Current treatment was identified by reviewing all outpatient medicines active on the 14 day before the COVID-19 diagnosis. A patient was considered on treatment if (s) he still had a supply of medications from their most recent 'fill' on the cut-off date. The VA system assigns each formulation to one or more drug classes. A process identical to the one above was used to assign an RR and CI to each of the 343 VA drug classes. AggRiskRx refers to the protective effect of eight VA drug classes with an upper boundary for CI≤0.80. This definition presumes that a protective effect goes beyond neutralising the underlying condition and is therefore likely independent of its initial indication. An aggregate effect for all eight classes was derived by log transforming the RR for each and adding the transformed values. This approach assumed independent effects, and an aggregate impact was the product of the individual RR. We did not examine high-risk drugs because the RR of pre-existing conditions reflects the underlying disease and the drugs used to treat the condition.

Age at diagnosis, gender, self-reported race and ethnicity, veteran status, smoking history, and use of supplemental oxygen were retrieved from the CSDR. The CSDR was interrogated for Charlson Comorbidity scores (both the 2-year and the lifetime) and Elixhauser Comorbidity scores (2-year, Elix2Yrs, and lifetime, ElixEver).

Statistical methods

Univariate analysis was used to compare the attributes of patients who died and survived. Group differences in nominal variables were tested by χ^2 analysis. Group differences in continuous variables were examined by the student's t-test or Mann-Whitney U test.

Main model

Stepwise logistic regression was used to construct a multivariate model for COVID-19 death in the entire sample. The dependent variable was death within 60 days of the diagnosis. The predictor of interest was prior vaccination for COVID-19. Covariates included age, gender, race, ethnicity, veteran status, current smoking, use of supplemental oxygen, probability of death/diagnosis (PDeathDx), laboratory-derived death probability (PDeathLabs), drug class protective effect (AggRiskRx), Charlson comorbidities (2 years/lifetime), Elixhauser Comorbidity Score (2 years/lifetime) and infection with

the Delta versus earlier variants. Variables were entered in a stepwise fashion with a P-to-enter of 0.01 and to remove of 0.05. The model was used to derive each patient's overall predicted probability of death (PDeath). The ability of the predicted probability of death (PDeath) to discriminate between the two groups was assessed by the area under the receiver operator characteristic (ROC) curve. An adjusted OR and its 95% CI were derived for the vaccination term. A standard on-line calculator was used to convert the adjusted OR to an equivalent RR. The identical procedure was used to evaluate the Delta term.

Early variants versus Delta

Separate models were developed for early variants (pre July) and for Delta (July–September) using the methods described above. The objective was to determine if predictors of death had changed significantly and if the effectiveness of prior vaccination differed for the two groups.

Vaccine cohorts

Eight patient cohorts infected with Delta were assembled based on the time from vaccination to the date of diagnosis (VxToDx). Each cohort was comprised of patients whose VxToDx fell within a 4-week interval. Cohort 1 was vaccinated≥2 and <6 weeks prior to diagnosis, cohort 2 (≥6 and <10 weeks), cohort 3 (≥10 and <14 weeks), cohort 4 (≥14 and <18 weeks), cohort 5 (≥18 and <22 weeks), cohort 6 (≥22 and <26 weeks), cohort 7 (≥26 and <30 weeks) and cohort 8 (≥30 and <34 weeks). Patients vaccinated≥34 weeks prior to diagnosis were excluded because the cohorts were small. CFR was calculated for each cohort and plotted against cohort number (figure 1).

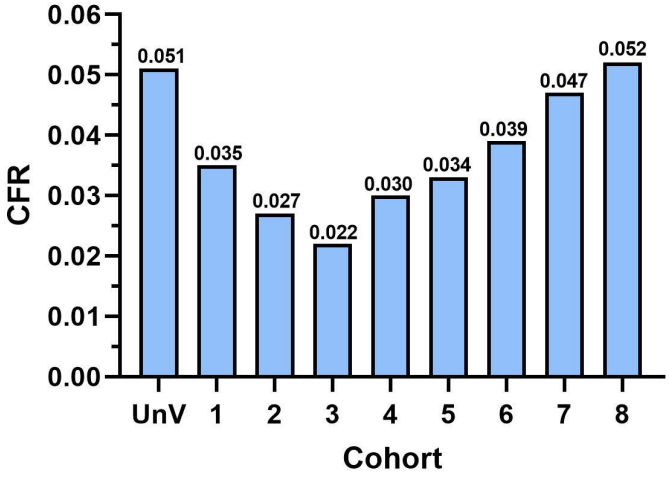

Figure 1 Unadjusted case fatality rates (CFR) for Delta infections by time from vaccination to diagnosis (in 4-week blocks). The CFR for unvaccinated (UnV) patients was 5.06%. Cohort 1 was vaccinated≥2 and <6 weeks prior to diagnosis, cohort 2 (≥6 and <10 weeks), cohort 3 (≥10 and <14 weeks), cohort 4 (≥14 and <18 weeks), cohort 5 (≥18 and <22 weeks), cohort 6 (≥22 and <26 weeks), cohort 7 (≥26 and <30 weeks) and cohort 8 (≥30 and <34 weeks). The CFR is shown above each column for the unvaccinated individuals and the eight cohorts. Note that CFR reached a nadir for cohort 3 and rose monotonically across cohorts 3–8.

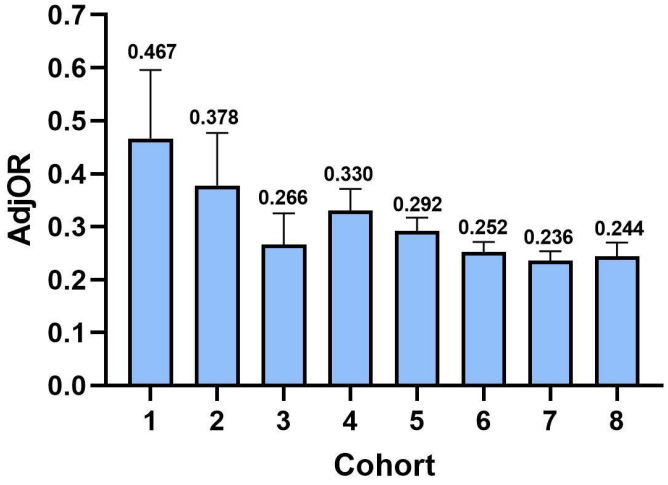

Figure 2 Adjusted ORs for vaccination (AdjOR) by time from vaccination to diagnosis (in 4-week blocks). Cohorts are defined in the caption for figure 1. A cohort-specific AdjOR is the odds of death for that cohort relative to that of all unvaccinated patients. The AdjOR is shown above the SE of the OR for the eight cohorts. Note that the benefit of vaccination for preventing COVID-19 death remained relatively stable across cohorts 3–8 (10–34 weeks).

Logistic modelling was used to derive cohort-specific adjusted ORs for vaccination. For cohort 1, patients in vaccine cohorts 2–8 were excluded from the data set of all Delta patients. A logistic model was then fitted to the remaining cases. This model was comprised of age at diagnosis, male gender, use of supplemental oxygen, current smoking, prior vaccination, PDeathDx, PDeathLabs and Charlson comorbidities (2 years/lifetime). This model was chosen because preliminary regressions showed that all variables were significant predictors of death for every cohort. The subset models also had similar power to discriminate between non-survivors and survivors (ROC areas from 0.810 to 0.816). The vaccine effect was therefore adjusted for eight other demographic and clinical variables and expressed as the odds of death for the cohort relative to that of all unvaccinated patients. This process was repeated for the remaining cohorts. The adjusted ORs for vaccination were plotted against cohort number (figure 2).

Statistics software

Statistical analysis was done using Stata MP V.17.

Patient and public involvement

Not applicable as a secondary data analysis. Patients were not involved with recruitment because it was based on existing data. They were also not involved in the design or conduct of this study because of the highly technical nature of the protocol. However, patients not employed by the medical centre participate in meetings of the institutional review board and regularly provide feedback to principal investigators. Results will be disseminated to patients through conventional patient

Table 1 Characteristics of the study population*

	Overall, n (%), (N=339 772)
Age in years	58.6±16.7
Sex	
Female	53 945 (15.9)
Male	285 827 (84.1)
Race/ethnicity	
American Indian/Native Alaskan	2899 (0.9)
Asian	3169 (0.9)
Black/African American	68 784 (20.2)
Native Hawaiian/Pacific Islander	3086 (0.9)
Unknown	45 795 (13.5)
White	216 039 (63.6)
Hispanic or Latino	30 472 (9.0)
Veterans	325 428 (95.8)
Supplemental oxygen	2421 (0.7)
Full vaccinated≥14 days from COVID-19 diagnosis	30 817 (9.1)
Median interval between vaccination and diagnosis (days)	154
Injections after 1 July 2021	73 117 (21.5)
Death within 60 days of diagnosis	18 120 (5.3)

*A total of 339 772 individuals out of 347 220 had pre-existing conditions.

education programmes, informing their providers of the results and conventional publications.

RESULTS

On 30 September 2021, there were 347 220 patients in VA's CSDR (online supplemental figure 1) (https://digitalrepository.unm.edu/kinm/4/#attach_additional_files). A total of 339 772 (or 97.9%) had at least one pre-existing condition and forming this report's basis (table 1). The mean age at the time of diagnosis was 58.6±16.7 years, 84.1% were men, 22.9% were members of a racial minority, 9.0% were Hispanic, 95.8% were veterans, 0.7% were on supplemental oxygen and 11.8% were current smokers. Overall, 9.1% had been fully vaccinated at least 14 days prior to the COVID-19 diagnosis. The median interval between vaccination and diagnosis was 154 days (IQR 111–185). Overall, 21.5% acquired their infections after 1 July 2021 and were presumed to have the Delta variant. Overall, 18 120 patients (5.33%) died within 60 days of their diagnosis.

Table 2 shows the results of univariate analysis comparing non-survivors and survivors. Non-survivors were older and more likely to be men, white and on supplemental oxygen but less likely to be Hispanic or current smokers. Vaccinated patients were less likely to die than the unvaccinated (3.95% vs 5.47%, respectively; p<0.001). The CFR was lower for those acquiring Delta than earlier variants (4.64% vs 5.52%; p<0.001). This finding persisted even

Table 2 Characteristics of non-survivors and survivors

Attribute	Non-survivors	Survivors	P value
Age at diagnosis (years)	76.1±11.2	57.6±16.4	<0.001
Male	97.5%	83.4%	<0.001
White	72.5%	63.1%	<0.001
Hispanic	7.0%	9.1%	<0.001
O_2 supplementation	1.8%	0.7%	<0.001
Current smoker	8.9%	12.0%	<0.001
Charl2Yrs	3.23±2.65	1.41±2.00	<0.001*
CharlEver	4.99±3.31	2.28±2.77	<0.001*
Elix2Yrs	11.87±12.75	4.77±8.45	<0.001*
ElixEver	21.28±16.62	9.47±12.76	<0.001*
PDeathDx	0.139±0.128	0.049±0.064	<0.001*
PDeathLabs	0.136±0.106	0.058±0.065	<0.001*
AggRiskRx	−0.0854±0.2720	−0.2025±0.5424	<0.001*

*Mann-Whitney U test; differences in ages were tested by student's t-test.
AggRiskRx, drug class protective effect; CharlEver, Charlson comorbidity score, lifetime; Charl2Yrs, Charlson comorbidity score, 2-years; ElixEver, Elixhauser comorbidity score, lifetime; Elix2Yrs, Elixhauser comorbidity score, 2-years; PDeathDx, probability of death/diagnosis; PDeathLabs, laboratory-derived death probability.

Table 3 Main multivariate model

Variable	OR	SE	Z score	P value	Lower CI	Higher CI
AgeAtDx	1.0656	0.0010	65.8100	0.0000	1.0636	1.0676
PDeathDx	5.6789	0.5630	17.5200	0.0000	4.6760	6.8970
PDeathLabs	15.5174	1.7617	24.1500	0.0000	12.4217	19.3845
VaccFlag	0.2802	0.0110	−32.5100	0.0000	0.2595	0.3026
Delta	1.8656	0.0494	23.5400	0.0000	1.7712	1.9651
CharlEver	1.0216	0.0060	3.6200	0.0000	1.0098	1.0334
Male	1.8282	0.1037	10.6400	0.0000	1.6359	2.0431
AggRiskRx	1.2671	0.0397	7.5700	0.0000	1.1918	1.3473
White	0.8599	0.0176	−7.3800	0.0000	0.8261	0.8951
ElixEver	1.0065	0.0009	6.9700	0.0000	1.0047	1.0083
Hispanic	1.2367	0.0433	6.0700	0.0000	1.1548	1.3245
Charl2Yrs	1.0323	0.0061	5.3500	0.0000	1.0203	1.0443
Veteran	2.9615	1.5124	2.1300	0.0340	1.0884	8.0579
Constant	0.0001	0.0000	−18.2200	0.0000	0.0000	0.0002

n=239 393; receiver operator characteristic=0.824, VaccFlag is the vaccination status of the patient.

AgeAtDx, Age at diagnosis; AggRiskRx, drug class protective effect; CharlEver, Charlson comorbidity score, lifetime; Charl2Yrs, Charlson comorbidity score, 2-years; ElixEver, Elixhauser comorbidity score, lifetime; PDeathDx, probability of death/diagnosis; PDeathLabs, laboratory-derived death probability; VaccFlag, vaccine flag.

when vaccinated patients were removed from the analysis (5.06% vs 5.55%; p<0.001).

The primary multivariate model is shown in table 3. A total of 239 393 patients (70.5%) had complete data sets available for multivariate modelling. Overall, 13 variables were identified as statistically significant and independent determinants of death at 60 days. A poorer prognosis was observed for the elderly, men and Hispanics while being white was protective. PDeathDx, PDeathLabs, AggRiskRx and three of four comorbidity measures were all significant predictors of death. The adjusted OR for Delta infection was 1.87±0.05, which corresponds to an RR of 1.78. The adjusted OR for prior vaccination was 0.280±0.011, which corresponds to an RR of 0.291. This observation suggests that the Delta variant is substantially more lethal than earlier variants—an effect that is largely offset by prior vaccination.

Tables 4 and 5 show the multivariate models for early COVID-19 variants and Delta, respectively. Of 11 variables identified as predictors before 1 July 2021, 8 were still significant after the emergence of Delta. The adjusted OR for vaccination prior to July 2021 was 0.404±0.033, while the OR thereafter was 0.259±0.012. This observation suggests that prior vaccination was more effective in reducing the CFR for Delta than earlier variants. However, only 4649 (or 15.1%) of 18 120 breakthrough infections occurred before July 1. The earlier ORs were therefore based on a relatively small number of deaths in the vaccinated group.

Overall, 73 117 patients were presumed to have been infected with the Delta variant based on their date of infection. A total of 26 168 (35.8%) patients had previously been vaccinated. Overall, 25 818 vaccines were assigned to 8 cohorts defined by the time from completed vaccination to diagnosis (VxToDx) (in 4-week blocks, online supplemental figure 2) (https://digitalrepository.unm.edu/kinm/4/#attach_additional_files). Cohort 1 was comprised of the most recent vaccines, while cohort eight had the most remote vaccinations. The cohorts varied in size from 457 to 6896 subjects. Figure 1 shows that CFR fell across the lowest cohorts and reached a nadir of 2.19% for cohort 3. It then increased monotonically across cohorts 4–8. One possibility for the latter trend is that vaccine effectiveness declined after 10–14 weeks. However, patients in the later cohorts (4–8) also received their vaccinations the earliest because their needs were the most urgent. For this reason, we used a dual approach to control for confounding. A total of 52 613 patients (72.0% of all Delta cases) were available for this analysis. For each cohort of interest, patients in the other cohorts were excluded, and a common model fitted to the remaining cases. Thus, members of each cohort had the same priority for vaccination, and the effect of vaccination for each cohort was adjusted for eight other patient attributes. Figure 2 shows that the adjusted OR for vaccination declined across the lowest cohorts and remained low for the remaining ones. Patients with the most remote vaccinations still had an adjusted OR for

Table 4 Multivariate model for early variants

Variable	OR	SE	Z score	P value	Lower CI	Higher CI
AgeAtDx	1.0681	0.0012	60.0600	0.0000	1.0658	1.0704
PDeathDx	5.7727	0.6256	16.1800	0.0000	4.6680	7.1388
PDeathLabs	14.6853	1.8320	21.5400	0.0000	11.5000	18.7530
CharlEver	1.0255	0.0067	3.8400	0.0000	1.0124	1.0388
VaccFlag	0.4039	0.0332	−11.0400	0.0000	0.3438	0.4744
Male	1.8551	0.1233	9.2900	0.0000	1.6285	2.1133
White	0.8241	0.0187	−8.5500	0.0000	0.7883	0.8614
AggRiskRx	1.3161	0.0485	7.4500	0.0000	1.2244	1.4147
Hispanic	1.2593	0.0488	5.9500	0.0000	1.1672	1.3585
ElixEver	1.0059	0.0010	5.7100	0.0000	1.0039	1.0080
Charl2Yrs	1.0250	0.0067	3.7600	0.0000	1.0119	1.0383
Constant	0.0002	0.0000	−87.3900	0.0000	0.0002	0.0003

n=186 505; receiver operator characteristic=0.826.
AgeAtDx, Age at diagnosis; AggRiskRx, drug class protective effect; CharlEver, Charlson comorbidity score, lifetime; Charl2Yrs, Charlson comorbidity score, 2-years; ElixEver, Elixhauser comorbidity score, lifetime; PDeathDx, probability of death/diagnosis; PDeathLabs, laboratory-derived death probability; VaccFlag, vaccine flag.

vaccination of 0.244±0.012. Thus, there was no evidence that the vaccine effect on CFR declined over the observation period in this study.

DISCUSSION

In this study, we stress the importance of a robust system for classifying vaccine effects, including the severity of confounding factors and the priority of vaccination. The reason is that the usual methods for evaluating effectiveness are composite measures reflecting three underlying risks (exposure, infection, death) and may not precisely define the mechanisms by which vaccines have failed. CFR was chosen as the outcome because it is not affected by exposure rates and is a more direct measure of the vaccine's biological properties. In this study, we determined the extent to which vaccination provides protection against mortality using data from the largest integrated healthcare system in the USA.

In the absence of potent countermeasures (ie, durable vaccinations and acute therapies), predictive modelling was an attempt to stem the viral tsunami that emanated from Wuhan, China, in late 2019.[9] VHA data are an

Table 5 Multivariate model for Delta variant

Variable	OR	SE	Z score	P value	Lower CI	Higher CI
AgeAtDx	1.0570	0.0021	27.5400	0.0000	1.0529	1.0612
VaccFlag	0.2593	0.0119	−29.4700	0.0000	0.2370	0.2836
Charl2Yrs	1.0704	0.0119	6.1100	0.0000	1.0473	1.0941
PDeathLabs	20.0041	5.5087	10.8800	0.0000	11.6604	34.3181
PDeathDx	4.5774	1.1316	6.1500	0.0000	2.8196	7.4311
Male	1.9115	0.2051	6.0400	0.0000	1.5490	2.3589
ElixEver	1.0094	0.0018	5.2800	0.0000	1.0059	1.0129
O_2 Supp	1.8408	0.2729	4.1200	0.0000	1.3766	2.4616
AggRiskRx	1.1373	0.0664	2.2000	0.0270	1.0144	1.2751
Constant	0.0007	0.0001	−46.2300	0.0000	0.0005	0.0010

n=52 888; receiver operator characteristic=0.818.

AgeAtDx, Age at diagnosis; AggRiskRx, drug class protective effect; Charl2Yrs, Charlson comorbidity score, 2-years; ElixEver, Elixhauser comorbidity score, lifetime; PDeathDx, probability of death/diagnosis; PDeathLabs, laboratory-derived death probability; VaccFlag, vaccine flag.

excellent substrate for modelling emerging infections and the impact of countermeasures. Prior to licensed vaccines and when therapeutics were being tested, some mathematical models suggested focus on disease transmission coefficient (β, an estimate of the probability of contracting the disease from an infectious individual) and clinical outbreak rates.[10] Early in the pandemic, interpretations of mathematical models naïvely hinted that mankind could triumph over nature.[11] Awareness campaigns[12] have proven futile as over 57% of the world's population has been infected or reinfected despite the availability of vaccinations. The spread and evolution of the SARS-CoV-2 cannot be undone.

Observational studies of vaccine effectiveness are heavily biased. Patients at the highest risk of death are more likely to receive the vaccine for several reasons, including personal choice, the concern of their physicians and national policies driven by vaccine shortages and stressed delivery systems. This prioritisation confounds the relationship between the intervention and outcome because the benefits of vaccination are offset by their preferential use in patients with the poorest prognosis. In 2021, (before the deluge of infections and reinfections in fully vaccinated and boosted individuals), there was optimism that vaccination campaigns prevented new cases while decreasing hospitalisations and fatalities.[13] Humoral and cellular immunity may have at least an 8-month duration.[14] We found that the adjusted OR for vaccination was 0.280. This value corresponds to a 71% reduction in the risk of death. This benefit was observed at a median of 5.1 months after vaccination. Substantial benefits of vaccination were observed before and after the emergence of Delta, although the former was significantly less.

Our cohort studies showed that the CFR for vaccines declined with time to a nadir 10–14 weeks after vaccination and then rose thereafter. Since CFR is not affected by patient behaviours or environmental factors, this pattern is consistent with the acquisition and subsequent loss of a physiological factor that promoted recovery from an established infection. The other possibility is selection bias. Patients with the longest time from vaccination to diagnosis received their vaccines the earliest. For example, for Delta infections acquired in August 2021, cohort 8 contained the very first vaccines, while cohort 1 contained patients who deferred their vaccinations until July. Because early vaccinations were directed at those with the highest risks of death, CFR would be higher for early vaccines regardless of vaccine effect. For this reason, we used a dual approach to confounders to control for the urgency of vaccination and other patient attributes. Our results showed that the rising CFR with time was due to selection bias and not loss of vaccine effect. The benefits of vaccination remained large even at 30–34 weeks—the longest observation period in this study. This finding contrasts sharply with prior studies (1–3) showing a loss of vaccine effect over time. The difference may be due to the use of CFR instead of infection or mortality rates, more robust handling of confounders and avoidance of selection biases introduced by the way that VA rolled out its vaccination programme. Unlike infection or mortality rates, CFR is not affected by unmeasured personal behaviours or environmental factors that affect the probability of exposure. The ORs for vaccination in our models were also adjusted for a much larger number of pre-existing conditions, vital sign abnormalities, laboratory results and medications than previously reported. Finally, because our cohorts were assembled over a short time frame, our approach was not affected by VA's changing priorities for vaccination.

Moreover, from a mechanistic perspective, emerging reports indicate that current COVID-19 vaccines elicit robust T-cell responses that may last 8–15 months.[14–16] The duration of long-lasting T-cell responses is in keeping with and may account for our findings demonstrating that the benefit of vaccination on CFR is sustained for at least 30–34 weeks. It is conceivable that while specific components of the immune responses may wane over time, thereby increasing the risk of breakthrough infection, the population of memory T cells recognising the virus in vaccinated individuals remains relatively stable for at least 8 months. As such, when vaccinated individuals are exposed to the SARS-CoV-2 virus, specific memory T cells quickly reactivate, expand in numbers and rapidly elicit host defence mechanisms capable of mitigating the risk of death from COVID-19 infection.

Prior vaccination has been included in other prediction models for COVID-19 mortality. For example, Hippisley-Cox *et al*[17] published a multivariate model for COVID-19 death in a large, vaccinated cohort in England. The investigators found that the adjusted HR for full versus partial vaccination was 0.17 (0.13–0.22), suggesting a major effect. Differences in vaccine effectiveness between our study and theirs may be explained by the populations studied. Individuals cannot have a debilitating congenital abnormality to be eligible for military service (and thus inclusion in the present study). It is also possible that veterans may have diminished immune responses to vaccination that may not be fully explained by covariates even in a highly specified model. The effect of vaccination on the CFR may depend on the population studied.

Although our findings show the effectiveness of vaccination, there are still many other variables that potentially impact survival. For example, while we controlled for as many clinical variables as possible, there is no way of measuring all relevant patient attributes. Vaccination could be a marker for many traits that affect recovery from a serious illness, such as physical fitness, nutrition, medication compliance, preventive care and so forth. This possibility is suggested by the observation that vaccinated patients have a lower risk of death than unvaccinated persons, even when COVID-19 deaths are excluded.[18]

Our conclusions are limited to patients with characteristics of the veteran population and the Delta variant.[19] Because our approach requires an extensive amount of baseline data, the results are also biased towards patients

with chronic conditions that require periodic evaluation. Data were most often missing for routine laboratory tests. This event happens for patients who use the VA only for specific services or benefits either because they do not need routine care or receive such care in the private sector. Since patients with and without complete data are quite different, it is unclear if values imputed from the former should substitute for missing data in the latter. The most conservative approach is to conclude that the benefits of vaccination for those missing such data are unknown.

It is important to determine if our results are sensitive to changes in the definition of CFR (ie, the proportion of patients dying by 60 days). It is possible that patients can succumb to 'long COVID-19' even if their initial symptoms were mild. Whether vaccination prevents these deaths is unknown. We have not done these studies because no vaccinated patient with the Delta variant has been followed for more than 8 months. Our analysis is the first to report characteristics potentially relevant to mortality in immunised US Veteran patients during the Delta portion of the 2020 pandemic. The multivariate models described herein are unique in employing US Veteran patient data (age, a combination of vital signs/laboratory values, Charlson Comorbidity indices, pharmaceuticals and Elixhauser Comorbidity scores).

In conclusion, vaccination has significant and sustained beneficial effects on the CFR for patients with established COVID-19 infection. Our methods represent a new approach to evaluating the effectiveness of interventions in observational studies. If validated by others for COVID-19 and other diseases, the approaches presented here represent an alternative and perhaps more robust method for reconciling confounders.

Author affiliations
[1]Research Service, New Mexico VA Health Care System, Albuquerque, New Mexico, USA
[2]Clinical Research Pharmacy Coordinating Center, VHA Cooperative Studies Program, Albuquerque, New Mexico, USA
[3]Center for Global Health, University of New Mexico Health Sciences Center, Albuquerque, New Mexico, USA
[4]Medicine Service, New Mexico VA Health Care System, Albuquerque, New Mexico, USA
[5]Kidney Institute of New Mexico, University of New Mexico Health Sciences Center, Albuquerque, New Mexico, USA
[6]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Acknowledgements This work was supported by the US Department of Veterans Affairs, Veterans Health Administration through the use of data, resources and/or facilities of the New Mexico VA Health Care System, VA Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, the VA COVID-19 Shared Data Resource and the VA Informatics and Computing Infrastructure, VA HSR RES 13-457. BW is supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences of the National Institutes of Health through Grant Number UL1TR001449 (CTSC/DCI Kidney Pilot Project CTSC004-12, CTSC/Environmental Health Signature Program and Superfund Research Center Pilot Project). BW is funded by a Veterans Administration Merit Award (I01 BX001958), a National Institutes of Health R01 grant (DK-102085), Dialysis Clinic and partially supported by the University of New Mexico Brain and Behavioral Health Institute (BBHI 2018-1008, 2020-21-002) and in part by the University of New Mexico's Signature Program in Cardiovascular and Metabolic Disease. BW is further supported by the University of New Mexico School of Medicine Research Allocation Committee (C-2459-RAC, New Mexico Medical Trust). BW is an Associate Member of the University of New Mexico Health Sciences Center Autophagy, Inflammation and Metabolism Center of Biomedical Research Excellence supported by NIH grant P20GM121176. BW has a user agreement with the Center for Integrated Nanotechnologies (CINT, Los Alamos National Laboratory & Sandia National Laboratories, 2019AU0120, 2021BC0021).

Contributors GHM was responsible for conceptualisation, methodology, software, validation, formal analysis, original draft and revision. AEM was responsible for conceptualisation, software, validation, data curation, review and editing. DJP was responsible for conceptualisation, methodology, visualisation, review and editing. HMC was responsible for conceptualisation, methodology, review and editing. JTM was responsible for conceptualisation, methodology, review and editing. BW was responsible for conceptualisation, methodology, visualisation, review, editing and revision. BM was responsible for conceptualisation, methodology, review, editing and revision. CHH was responsible for conceptualisation, methodology, project administration, review, editing and revision. GHM is the guarantor of the study, data, and the decision to publish.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. BW is supported by Dialysis Clinic and has projects supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences of the National Institutes of Health through Grant Number UL1TR001449 (CTSC/DCI Kidney Pilot Project CTSC004-12 and CTSC/Environmental Health Signature Program Pilot Project CTSC003-13).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants. This study was approved by the New Mexico VA Health Care System Institutional Review Board (study number 20-H319). Because the study was minimal risk, the Board waived the requirements for patient consent and the HIPAA amendment (Health Insurance Portability and Accountability Act of 1996) in accordance with the policies of the Department of Veterans Affairs. This study was approved by the New Mexico VA Health Care System Institutional Review Board (study number 20-H319). Because the study was minimal risk, the Board waived the requirements for patient consent and the HIPAA amendment (Health Insurance Portability and Accountability Act of 1996) in accordance with the policies of the Department of Veterans Affairs.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The data that support the findings of this study are available from the Veterans Health Administration. Restrictions apply to the availability of these data, which were used under an IRB-approved protocol for this study. Data are available from GHM with the permission of the Veterans Health Administration. Data can be shared with outside investigators subject to the policies and procedures of the Department of Veterans Affairs and approval by the New Mexico VA Health Care System Institutional Review Board. Inquiries should be directed to the corresponding author.

ORCID iD
Glen H Murata http://orcid.org/0000-0001-8067-8281

REFERENCES

1 Cohn BA, Cirillo PM, Murphy CC, *et al*. SARS-CoV-2 vaccine protection and deaths among US veterans during 2021. *Science* 2022;375:331–6.
2 Tartof SY, Slezak JM, Fischer H, *et al*. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. *Lancet* 2021;398:1407–16.
3 Goldberg Y, Mandel M, Bar-On YM, *et al*. Waning immunity after the BNT162b2 vaccine in Israel. *N Engl J Med* 2021;385:e85.
4 Moghadas SM, Vilches TN, Zhang K, *et al*. The impact of vaccination on coronavirus disease 2019 (COVID-19) outbreaks in the United States. *Clin Infect Dis* 2021;73:2257–64.
5 Arbel R, Hammerman A, Sergienko R, *et al*. BNT162b2 vaccine booster and mortality due to Covid-19. *N Engl J Med* 2021;385:2413–20.
6 Osayomi T, Adeleke R, Yaya S, *et al*. Do pre-existing medical conditions affect COVID-19 incidence and fatality in Nigeria? A geographical perspective. *Open Health* 2022;3:50–9.
7 Campbell HM, Murata AE, Mao JT, *et al*. A novel method for handling pre-existing conditions in multivariate prediction model development for COVID-19 death in the Department of Veterans Affairs. *Biol Methods Protoc* 2022;7:p.bpac017.
8 Murata GH*et al*. Baseline metabolic profiling and risk of death from COVID-19. *Medrxiv* 2022.
9 Samui P, Mondal J, Khajanchi S. A mathematical model for COVID-19 transmission dynamics with a case study of India. *Chaos Solitons Fractals* 2020;140: :110173.
10 Khajanchi S, Sarkar K, Mondal J, *et al*. Mathematical modeling of the COVID-19 pandemic with intervention strategies. *Results Phys* 2021;25: :104285.
11 Sarkar K, Khajanchi S, Nieto JJ. Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals* 2020;139: :110049.
12 Rai RK, Khajanchi S, Tiwari PK, *et al*. Impact of social media advertisements on the transmission dynamics of COVID-19 pandemic in India. *J Appl Math Comput* 2022;68:19–44.
13 Somekh I, KhudaBukhsh WR, Root ED, *et al*. *Quantifying the Population-Level Effect of the COVID-19* Mass Vaccination Campaign in Israel: A Modeling Study. *Open Forum Infect Dis* 2022;9:p. ofac087.
14 Barouch DH, Stephenson KE, Sadoff J, *et al*. Durable humoral and cellular immune responses 8 months after Ad26.COV2.S vaccination. *N Engl J Med* 2021;385:951–3.
15 Wragg KM, Lee WS, Koutsakos M, *et al*. Establishment and recall of SARS-CoV-2 spike epitope-specific $CD4^+$ T cell memory. *Nat Immunol* 2022;23:768–80.
16 Moss P. The T cell immune response against SARS-CoV-2. *Nat Immunol* 2022;23:186–93.
17 Hippisley-Cox J, Coupland CA, Mehta N, *et al*. Risk prediction of covid-19 related death and hospital admission in adults after covid-19 vaccination: national prospective cohort study. *BMJ* 2021;374:n2244.
18 Xu S, Huang R, Sy LS, *et al*. *COVID-19 Vaccination and Non-COVID-19 Mortality Risk - Seven Integrated Health Care Organizations,* United States, December 14, 2020-July 31, 2021. *MMWR Morb Mortal Wkly Rep* 2021;70:1520–4.
19 Zhao S, Lou J, Cao L, *et al*. Differences in the case fatality risks associated with SARS-CoV-2 Delta and non-Delta variants in relation to vaccine coverage: an early ecological study in the United Kingdom. *Infect Genet Evol* 2022;97:105162.