Title

How does prior infection and vaccination relate to the risk of incident SARS-CoV-2 infection/reinfection? A prospective cohort study among nine clinical sites in the USA, February 2021 to January 2023

Journal

BMJ Open, 16(5)

ISSN

2044-6055

Authors

Judge, Ashley

Beidelman, Erika T

Allison, David B

et al.

Publication Date

2026-05-01

DOI

10.1136/bmjopen-2025-112893

Peer reviewed

BMJ Open

How does prior infection and vaccination relate to the risk of incident SARS-CoV-2 infection/reinfection? A prospective cohort study among nine clinical sites in the USA, February 2021 to January 2023

Ashley Judge ,[1] Erika T Beidelman,[1] David B Allison,[1] Stephanie Dickinson,[2] Kishore M Gadde,[3] Lilian Golzarri Arroyo,[2] John P Ioannidis ,[4,5] Anna LM Macagno,[2] Jonathan T Macy,[6] Kevin C Maki,[6,7] Stanley Perlman,[8] Michael J Ricciardi,[9,10] Molly Rosenberg,[1] Christina Ludema[1]

To cite: Judge A, Beidelman ET, Allison DB, *et al*. How does prior infection and vaccination relate to the risk of incident SARS-CoV-2 infection/reinfection? A prospective cohort study among nine clinical sites in the USA, February 2021 to January 2023. *BMJ Open* 2026;**16**:e112893. doi:10.1136/bmjopen-2025-112893

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-112893).

Received 24 October 2025
Accepted 17 April 2026

For numbered affiliations see end of article.

Correspondence to
Dr Christina Ludema;
cludema@iu.edu

ABSTRACT

Objective To estimate the relative effectiveness of vaccination (0, 1, 2, ≥3 doses) and prior infection, in combination, on risk of SARS-CoV-2 infection/reinfection.

Design Prospective cohort study.

Participants We recruited participants for the Aegis Study from nine clinics across five US states. Participants must have been 18 years or older, had a history of a positive PCR for SARS-CoV-2, SARS-CoV-2 antigen or antibody test for SARS-CoV-2 with documentation or had no suspected or documented prior SARS-CoV-2 infection, intended to remain in study area for the next 12 months, and had elevated risk of future SARS-CoV-2 exposure. Exclusion criteria included acute illness, contraindication to phlebotomy, use of immunosuppressants or receipt of systemic immunoglobulins.

Methods We used extended Cox regression with robust standard errors to estimate the association between time-varying number of vaccine doses and baseline prior infection on risk of infection/reinfection among a prospective cohort of US adults between February 2021 and January 2023, accounting for censoring using inverse probability of censoring weights. Additionally, to quantify possible exposure misclassification of prior infection by comparing prior infection operationalised as (1) documented/self-reported prior infection and (2) documented/self-reported prior infection plus nucleocapsid antibody indication of prior infection.

Results Of n=2178 who completed enrolment, n=1887 adults (63% female; 65% non-Latino White) contributed 366 905 days of observation. Participants contributed an average of 7.2 months of follow-up between February 2021 and January 2023. 28% (n=533) of individuals were infected or reinfected during the study period. Similar relative effectiveness was observed between the two different operationalisations of prior infection. After correction for prior infection status in the nearly 16% of those without study documentation of prior infection who had nucleocapsid antibody levels comparable to documented cases, relative to the unvaccinated with no prior infection, estimated effectiveness generally increased with increasing vaccine doses and prior infection (without prior infection: one (17%, 95% CI −31% to 47%), two (49%, 95% CI 31% to 63%), ≥three (71%, 95% CI 58% to 80%) vaccine doses; with prior infection: none (56%, 95% CI 30% to 72%), one (71%, 95% CI 42% to 86%), two (65%, 95% CI 49% to 76%), ≥three (80%, 95% CI 68% to 88%) vaccine doses). Pairwise comparisons at each vaccine dose (ref: no prior infection) revealed that prior infection provided additional protection, with stronger relationships for no and one dose (none: 56% (95% CI 30% to 72%), one: 66% (95% CI 28% to 84%), two: 31% (95% CI 7% to 49%), ≥three 31% (95% CI 0% to 53%)). There was a marked decrease in the protection offered by vaccination, prior infection, or both in the Omicron period versus pre-Omicron period.

Conclusion In our real-world observational sample, vaccination (with two and ≥three vaccine doses of any Food and Drug Administration Emergency Use Authorization approved vaccine) and prior infection

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This observational cohort study used a robust measure of prior infection via SARS-CoV-2 nucleocapsid IgG antibody measurements in combination with self-reported and documented prior infection.
⇒ This study examined the estimated effectiveness of vaccination against infection under real-world conditions, as compared with interventional vaccine trials which may more closely follow vaccination schedules.
⇒ The methodology employed accounted for time-varying number of vaccine doses and study attrition.
⇒ The study sample is a convenience, clinic-based sample, which may not be representative of all US adults.

conferred benefits for protection against infection/reinfection. Re-classification of prior infection status based on antibody levels had little effect on results.

INTRODUCTION

During the course of the COVID-19 pandemic, both vaccination and natural infection likely impacted the risk of reinfections, but their relative and combined merits have been debated. Several prior studies have shown waning of vaccination benefit on infection risk.[1 2] Other studies suggest that vaccination after infection may have increased time of protection against reinfection.[3 4] In fact, several studies have identified hybrid immunity as having greater protective effects than vaccination or prior infection alone on the risk of infection/reinfection.[3 5–10] Continued research examining the relative effectiveness of prior infection and vaccination(s) in combination is needed to strengthen the evidence base regarding determinants of immunological resistance against SARS-CoV-2.

Studies of vaccination and prior infection are limited by the fact that some individuals may have had undocumented and potentially unknown infections prior to entering the study. This would cause misclassification in prior infection status and potentially underestimate the association between prior infection or hybrid immunity on infection/reinfection. No study to our knowledge has assessed the additional use of antibody indicators of prior infection in conjunction with documentation of prior infection on estimates of vaccine effectiveness and protection conferred by prior infection against SARS-CoV-2 infection/reinfection. Detection of antibodies that bind to the SARS-CoV-2 nucleocapsid protein may help provide identification of unrecognised prior infection.[11] Anti-nucleocapsid (N) responses are not induced by typical vaccines approved in the USA, but are associated with prior infection, thus acting as a plausible indicator of natural infection.[11] Exploring the added benefit of using antibody indication of prior infection may provide a useful assessment of misclassification in settings where routine testing is not undertaken and history of SARS-CoV-2 infection is unverifiable, such as in survey research.

The objective of this observational cohort study was to quantify the effects of combinations of prior infection and number of vaccine doses on subsequent risk of SARS-CoV-2 infection/reinfection in a geographically diverse, US prospective cohort during a time period that primarily spanned the Delta and Omicron variants. We compare risk of infection/reinfection among combinations of prior infection and number of vaccine doses in groups with prior infection operationalised as (1) history of self-reported or documented SARS-CoV-2 and (2) history of self-reported or documented SARS-CoV-2 and/or individuals without a history of prior SARS-CoV-2 but who had N-antibody indication of prior infection. We hypothesised that hybrid immunity would confer greater protection than prior infection or vaccination alone, especially as the number of vaccine doses increased, and that re-classification of prior infection based on N-antibodies would strengthen the estimated protection conferred by prior infection on reinfection.

METHODS

Protocol and reporting guideline

The protocol is posted in online supplemental material. The study is reported according to the STROBE reporting guidelines.

Study design and population

This study employed a prospective cohort design using data from the Aegis Study. To be included in the Aegis study sample, participants must have (1) been 18 years or older, (2) had a history of a positive PCR for SARS-CoV-2, SARS-CoV-2 antigen test, or antibody test for SARS-CoV-2 or had no suspected or documented prior SARS-CoV-2 infection, (3) intended to remain in study area for the next 12 months and (4) had elevated risk of future SARS-CoV-2 exposure (eg, being employed in a high-contact setting, living in a high-occupancy home, living and/or working in a location with high population density, regular use of congregate public transportation, being a member of a high-risk demographic group, regularly spending time with people outside of the household who are not wearing facial coverings, or anticipate attendance at events or in settings with 25 or more people at least twice in the next 12 months). Exclusion criteria included having been acutely ill or febrile within 72 hours prior to or at screening, having a bleeding disorder considered a contraindication to phlebotomy, being in an immunosuppressive or immunodeficient state, such as asplenia and/or recurrent severe infections, having received systemic immunosuppressants or immune-modifying drugs for >14 days in total within 6 months prior to screening, or having received systemic immunoglobulins within 3 months prior to the day of screening.

Nine study sites across the USA enrolled participants in the Aegis study between February 2021 and March 2022 (online supplemental figure 1). Sites were located in Illinois, Florida, Texas, Louisiana and Oklahoma. In order to detect significant effects between individuals with and without prior infection, we aimed to recruit 2100 participants, 1400 with and 700 without prior infection, assuming a 70% completion rate. A total of 2806 individuals were screened for participation with 2178 completing enrolment and at least the first study visit. Eight study sites completed data collection visits with participants by November 2022. One site continued data collection until January 2023 (online supplemental table 1). Participants were recruited from clinics and represent a convenience sample of US adults.

Patient and public involvement

Patients and the public were not involved in the design and conduct of this research.

Data collection

Participants attended up to seven visits, each 8 weeks apart. We collected self-reported surveys that included queries of vaccination documentation (Janssen, Moderna, Pfizer and Other), symptom reports, health status questionnaires, mental health questionnaires and COVID-19 related questions. Participant blood samples were also collected at each study visit for SARS-CoV-2 N IgG antibody titres and cardiometabolic testing. In addition to study visits, participants were provided at-home nasal swab/saliva SARS-CoV-2 tests and tabletop refrigerators to store viral samples. Participants were instructed to swab themselves every 2 weeks in between study visits. Between at-home swabbing and study visit swabbing, participants who completed all study procedures were tested for SARS-CoV-2 biweekly. Test swabs were collected at each study visit and tested for SARS-CoV-2 via PCR testing. PCR testing was undertaken by two labs, Sema4 and a lab at Indiana University (IU). 97% of tests were performed by Sema4. Sema4 used the Perkin Elmer New Coronavirus Nucleic Acid Detection Kit, which uses fluorescent detection of the *N* gene and *ORF1ab* gene and used an internal control (*MS2*). The IU lab targeted three viral genes (*S*, *N* and *ORF1ab*) and used an internal control (*MS2*).

Study oversight

This study was reviewed and all measures and protocols associated with this research were approved by MB Clinical Research (tracking #: 20203947, analysis protocol #: MB-2015), Indiana University's Human Subjects Office (IRB #: 22953) and Pennington Biomedical Research Center.

Definitions

Prior history of SARS-CoV-2 infection: Prior infection with SARS-CoV-2 was defined as a binary indicator variable at the date of first study visit. Prior infection was assigned if the participant provided documentation of a prior positive SARS-CoV-2 PCR test, self-reported SARS-CoV-2 infection, was classified by staff as having a prior SARS-CoV-2 infection at the first visit or self-reported a positive SARS-CoV-2 antibody test.

Vaccination: Number of vaccine doses was captured as a categorical variable (0, 1, 2 or ≥3 doses) based on self-report and verified with documentation in vaccine cards in 80% of cases among those who reported a date of vaccination.

SARS-CoV-2 infection/reinfection: The date of a participant's first positive SARS-CoV-2 PCR test during the observation period. This could be either the participant's first infection or a reinfection based on whether the participant reported a SARS-CoV-2 infection prior to the observation period. For individuals with dates of prior infection (91.3% of the analytic sample), an observation period infection was only included in analyses as a reinfection if the positive PCR test occurred 90 days or more after the date of the pre-study period infection. We assumed individuals without dates of prior infection had prior infections greater than or equal to 90 days prior to the reinfection observed during the study period. To ensure we captured all infections during follow-up, we also classified individuals who had a fourfold increase in N-binding IgG antibody levels between successive visits as infected.[12–16] We assigned these individuals a date of positivity based on the midpoint between the two visits.

Statistical analysis

We excluded 21 (1.0%) individuals who had irreconcilable dates (eg, documented prior infection after visit one (n=17), implausible vaccine dates (n=4)), did not have follow-up time recorded after 90 days from their baseline infection (n=25, 1.1%), received other vaccine types (n=37, 1.7%), were missing or only reported 'other' for variables used to create our final prior infection variable (n=26, 1.2%), or who self-reported COVID-19 vaccination but did not report dates of vaccination (n=180, 7.1%), totaling n=263 (12.1%) individuals. 133 individuals had start and end times on the same day. Most of these cases included individuals who only attended one visit. In order to avoid unnecessary exclusions, we added 0.1 to all individuals' end times.

To operationalise exposures, we categorised based on prior SARS-CoV-2 infection and number of vaccine doses (no doses with and without prior infection, one dose with and without prior infection, two doses with and without prior infection, ≥three doses with and without prior infection). Due to the dynamic nature of vaccination, number of vaccine doses was classified as a time-varying exposure. Participants who received additional vaccine doses during the study period reported the date of their additional doses at their next study visit. This was a relatively common experience due to the increasingly widespread availability of vaccines during the study period. Vaccination status was updated at the reported date of vaccination for each dose.

Our primary outcome was SARS-CoV-2 infection/reinfection. Because only a few individuals tested positive twice during the study, we censored individuals at their first positive test.

Observation times varied across participants. To avoid immortal time bias, we assigned origin time based on timing of prior infection in relation to visit one. For participants with no prior infection, an infection more than 90 days prior to visit one, or without dates of prior infection, the observation time (time origin) began at the visit one date. For participants with a prior SARS-CoV-2 infection less than 90 days prior to visit one, the time origin was set at 90 days after the documented date of prior infection. The observation time ended at date of last study visit, first SARS-CoV-2 infection or date of censoring. We divided participants' overall observation time into the time under observation corresponding to the number of vaccine doses they had received. Observation time for each individual for each vaccine dose (0–3) was calculated as beginning at either visit one, 90 days after their prior infection or 14 days after date of vaccination and ending at 13 days after

subsequent date of vaccination, infection or last visit. We administratively censored follow-up after 300 days due to the limited number of observations past 300 days.

We used an extended Cox proportional hazards model with robust standard errors to estimate the effectiveness of prior infection and vaccine dose combinations against SARS-CoV-2 infection. We used robust standard errors to account for individuals possibly contributing multiple observations (ie, if they changed exposure groups over time). Extended Cox proportional hazards models have the advantage of addressing non-proportionality by allowing covariates to vary over time. The extended Cox model also allowed us to address our expectation that hazard rates would vary across exposure groups.

The time periods used in the extended Cox models were defined by the receipt of a new vaccine dose, beginning at 0 doses up to 3 doses. This allowed us to estimate the hazard rate for each level of vaccine exposure, structure risk sets unique to each vaccine dose, and allow key covariates to vary over time. A participant moved to the next exposure category each time they received a new vaccine dose. As a result, individuals could contribute to multiple exposure categories as they received new vaccinations over the follow-up period. To control for confounding, we adjusted for age in years (using a squared term), sex (male or female), self-reported race/ethnicity (Hispanic/Latino, non-Latino Asian, non-Latino Black or African American, non-Latino White and non-Latino Other or Unknown), educational attainment (high school graduate (including General Educational Development (GED)) or less, some college or technical school, or college graduate or higher), income (<US$25 000, US$25 000–49 999, US$50 000–74 999, US$75 000–99 999 and ≥US$100 000),[17] essential worker status (essential worker, not an essential worker or not applicable/unemployed), site, and the calendar date of the start of follow-up for each exposure period. We examined estimated effectiveness of combinations of vaccine doses and prior infection relative to having no prior infection and no vaccinations. Estimated effectiveness was expressed as (1−hazard ratio)*100.

We mitigated the potential bias of selective attrition over time by applying inverse probability weights to all models. We handled missing data through multiple imputation. We created 20 datasets over 10 iterations. We imputed baseline depression status, measured by the Center for Epidemiological Studies Depression (CES-D) scale (where each item was scored from 1 to 5) defined as a score greater than or equal to 36, income, educational attainment, financial difficulty, self-reported baseline health status, essential worker status, type of residence and number of individuals residing in the home.[18]

To create inverse probability of censoring weights, in each multiply imputed dataset we estimated time-varying Cox proportional hazards models based on the following predictors of study dropout: sex, age, race/ethnicity, essential work status, income, educational attainment, site, baseline self-reported health status, type of residence, number of individuals residing in the home, financial difficulty (ie,'how difficult is it to live on your household income right now' measured via a Likert scale from 1=not at all difficult to 5=extremely difficult), CES-D indication of depression, number of vaccine doses (time-varying) and prior infection status. We pooled our results across multiply imputed datasets according to Rubin's rules.[19]

To assess whether our estimates from our first sensitivity analysis differed in the pre-Omicron vs Omicron time periods, we stratified follow-up time from our first sensitivity analysis to pre-Omicron and Omicron time periods defined as before or on/after 11 December 2021, respectively. For the pre-Omicron (ie, Alpha/Delta) period, we used the same start times as in previous analyses and ended all individuals who had observation times greater than 10 December 2021 at this date, and only classified individuals as infected if they were infected by this date. For the Omicron period, we started observation times for all individuals without incident infections by 11 December 2021 on this date and used the same end date as in previous analyses.

All statistical analyses were performed in RStudio V.2024.12.0.467.[20] Extended Cox proportional hazards models were estimated using the *survival* package,[21 22] multiple imputation and pooling were implemented with the *mice* package,[23] and inverse probability of censoring weights were calculated using the *ipw* package.[24] Statistical significance was set with a two-tailed alpha of 0.05.

Sensitivity analyses

We conducted seven sensitivity analyses. To address measurement error, we conducted sensitivity analyses for our exposure and outcome. To assess misclassification of prior infection, our first sensitivity analysis examined whether including individuals with indication of prior infection via a data-driven threshold for the presence of N-binding IgG antibodies altered results. Detailed methods have been described elsewhere (Beidelman 2025 accepted Am J Epidem).[25] We calculated the area under the curve (AUC) of the serial dilutions of participant serum samples using the trapezoid method, which represents the magnitude of the antibody response through an estimation of the absorbance of the samples. Because values were skewed, we used the natural log plus one of the anti-nucleocapsid AUC values in this analysis. We created a binary indicator of antibody-indicated prior infection, with values of the ln(AUC+1) over 6.46 indicating prior infection and values less than or equal to that indicating no prior infection. We chose this value by examining the distribution of values among individuals with N-antibody data from the first study visit who had a study documented prior infection and chose the first quartile among the documented cases as our cut-off value (online supplemental table 2 and online supplemental figures 2 and 3). This cut-off value also closely corresponded to the third quartile among the undocumented. Based on this cut-off value, we recategorised only individuals without

documentation of prior infection who had values over the cut-off into the prior infected group.

To assess the robustness of our first sensitivity analysis in light of missing data, we compared results from our main analysis and first sensitivity analysis, where 27 individuals with missing N-antibody results were classified for the first sensitivity analysis into the prior infection category that they had in the main analysis, to results if these individuals had been excluded from both the main analysis and first sensitivity analysis.

To assess the robustness of our first sensitivity analysis to the data-driven threshold for reclassification, we compared results from our first sensitivity analysis to those if we chose a cut-off to reclassify individuals into the prior infected group of the median ln(AUC+1) value (ie, 5.545) among individuals who did not report prior infection and did not have an incident infection under observation across all study visits.

To check the robustness of our results to our outcome assessment based on increases in N-binding IgG antibodies, we used a twofold increase in N-binding IgG antibody threshold, rather than the fourfold threshold, to identify individuals without positive PCR tests who likely had an infection during follow-up. Research suggests that the twofold increase has higher sensitivity, in exchange for lower specificity,[12 14 26 27] and may be more appropriate for population surveillance rather than individual diagnostics.

To compare to prior research that reported on completion of primary series rather than number of shots, we assessed whether grouping vaccination by completion of primary series (ie, one dose of Janssen as equivalent to two doses of Pfizer or Moderna vaccines) rather than number of vaccine doses altered results as compared with our first sensitivity analysis. We excluded nine individuals for this analysis who were missing vaccine type.

To assess waning vaccine effectiveness on infection/reinfection, compared with our first sensitivity analysis, we truncated the length of time of each exposure period that included any number of vaccinations to 90 days, which can be considered the period after vaccination which offers the most protection.

There were 29 infections during the 14-day window from vaccination to when we operationalised individuals to be adequately protected from vaccination. We conducted a sensitivity analysis operationalising the day after vaccination as the start period of an individual's time under a new vaccine category. To account for the non-binary nature of vaccination, these analyses assess the impact of classification timing at the upper and lower bounds.

RESULTS

Sample characteristics

We had a final sample size of 1887 individuals. Participants had a median age of 54 (IQR 38–64) years old, 1196 (63%) were female, 467 (26%) had an annual household income of US$100 000 or more, 1012 (54%) had a college degree or higher, 1219 (65%) were non-Latino White and 663 (35%) were essential workers (table 1).

Participants in the study sample attended a median of 5 visits (IQR: 4–6) and contributed 366 905 days (12 063 months) of follow-up (median: 117 days, range: 0.1–300 days). 396 individuals (21%) had a prior SARS-CoV-2 infection at baseline (individuals could be included in multiple categories; categorised hierarchically in the following order: n=115 documented, n=235 self-reported positive PCR test, n=25 classified by staff, n=21 self-reported positive antibody test). Among individuals with a prior infection who had documented or self-reported dates of prior infection (n=344, 91% of all prior infections), approximately two-thirds had been infected during a period of primary circulation of the wildtype variant (wildtype: 241 individuals (68%), Alpha: 31 individuals (9%), Delta: 80 individuals (23%)).

At baseline, 23.1% of participants were unvaccinated (n=435), 11% had one vaccine dose (n=213), 62% had two vaccine doses (n=1165) and 4% had ≥three vaccine doses (n=74). During follow-up, 44% of participants were vaccinated with additional dose(s) (n=838).

Over a quarter of participants (28%, n=533) tested positive for SARS-CoV-2 during the study period (n=483 via positive PCR, n=50 additional based on x4 N-binding antibody between two successive visits), 88 of whom were classified as reinfections (22% of baseline prior infections). Among these participants, the majority were infected during a period of primary circulation of the Omicron variant (Omicron: 456 individuals (86%), Alpha: 3 individuals (0.6%), Delta: 74 individuals (14%)). Infections and reinfections peaked in early January 2022 and declined by June 2022 (online supplemental figure 4). Three individuals (two with re-infection) self-reported hospitalisation due to COVID-19 infection during follow-up. No deaths were reported.

Online supplemental table 3 provides information on the number of tests per 100 days under observation by prior infection and number of vaccine doses. Individuals in each group contributed approximately eight tests per 100 days.

Estimated effectiveness of vaccination and prior infection (study documented)

In our main analysis, we used commonly reported indicators of prior infection (eg, study documented: self-reported with or without formal documentation) to operationalise prior infection. Compared with being unvaccinated and having no prior infection, the estimated effectiveness of having a prior infection without vaccination was 48% (95% CI 15% to 68%) (figure 1A). The estimated effectiveness of having one vaccine dose with and without prior infection was 66% (95% CI 29% to 84%) and 18% (95% CI −27% to 47%), respectively. The estimated effectiveness of having two vaccine doses with and without prior infection was 63% (95% CI 43% to 76%) and 45% (95% CI 25% to 59%), respectively. The

Table 1 Demographic characteristics by baseline prior infection status

	Overall	No prior infection	Prior infection
	(n=1887)	(n=1491)	(n=396)
Age			
Mean (SD)	51.4 (16.3)	52.3 (16.5)	47.9 (15.4)
Median (min, max)	54.0 (18.0, 86.0)	56.0 (18.0, 86.0)	49.0 (18.0, 84.0)
Sex			
Female	1196 (63.4%)	944 (63.3%)	252 (63.6%)
Male	691 (36.6%)	547 (36.7%)	144 (36.4%)
Annual household income			
Less than US$24 999	425 (23.6%)	360 (25.4%)	65 (17.1%)
US$25 000 to US$49 999	386 (21.5%)	305 (21.5%)	81 (21.3%)
US$50 000 to US$74 999	292 (16.2%)	220 (15.5%)	72 (18.9%)
US$75 000 to US$99 999	228 (12.7%)	177 (12.5%)	51 (13.4%)
US$100 000 or higher	467 (26.0%)	356 (25.1%)	111 (29.2%)
Missing[+]	89 (4.7%)	73 (4.9%)	16 (4.0%)
Educational attainment			
High school graduate or less (including GED)	268 (14.4%)	225 (15.3%)	43 (10.9%)
College 1 year to 3 years (some college or technical school)	586 (31.4%)	444 (30.2%)	142 (35.9%)
College 4 years or more (college graduate or higher)	1012 (54.2%)	802 (54.5%)	210 (53.2%)
Missing[+]	21 (1.1%)	20 (1.3%)	1 (0.3%)
Race/ethnicity			
Latino	127 (6.7%)	86 (5.8%)	41 (10.4%)
Non-Latino—American Indian or Alaska Native	62 (3.3%)	45 (3.0%)	17 (4.3%)
Non-Latino—Asian	141 (7.5%)	118 (7.9%)	23 (5.8%)
Non-Latino—Black or African-American	202 (10.7%)	173 (11.6%)	29 (7.3%)
Non-Latino—White	1219 (64.6%)	961 (64.5%)	258 (65.2%)
Non-Latino—other or unknown	136 (7.2%)	108 (7.2%)	28 (7.1%)
Essential worker status			
Essential worker	663 (35.2%)	484 (32.5%)	179 (45.2%)
Not an essential worker	514 (27.3%)	404 (27.1%)	110 (27.8%)
NA/unemployed	708 (37.6%)	601 (40.4%)	107 (27.0%)
Missing*	2 (0.1%)	2 (0.1%)	0 (0%)
Vaccination status at baseline			
0 doses	435 (23.1%)	298 (20.0%)	137 (34.6%)
1 dose	213 (11.3%)	164 (11.0%)	49 (12.4%)
2 doses	1165 (61.7%)	968 (64.9%)	197 (49.7%)
≥3 doses	74 (3.9%)	61 (4.1%)	13 (3.3%)

*Missing values were imputed for the final analysis.
GED, General Educational Development; NA, Not Applicable.

estimated effectiveness of having ≥three vaccine doses with and without prior infection was 79% (95% CI 64% to 88%) and 68% (95% CI 54% to 77%), respectively. For the time under observation and number of infections in each stratum, see online supplemental table 4. For survival curves, see online supplemental figure 5. The median interval of time from vaccination to infection, end date of study, or censoring by number of vaccines and prior infection ranged from 28 to 133 days (online supplemental figure 6). The median interval of time from prior infection to infection, end date of study or censoring by number of vaccine doses ranged from 296 to 483 days, increasing with increasing number of vaccine doses (online supplemental figure 7).

We found that all point estimates were higher in the pre-Omicron period than the Omicron period, with

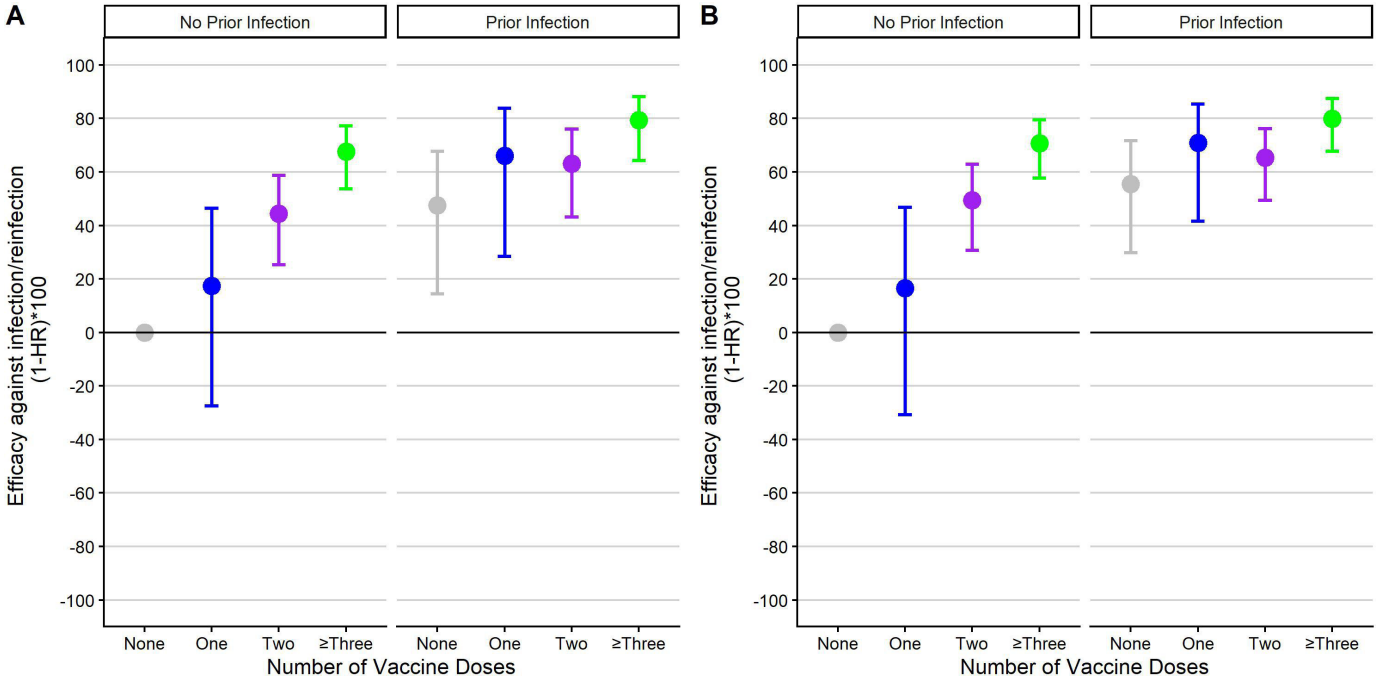

Figure 1 (A) Main analysis: estimated effectiveness of prior infection (study documented) and number of vaccine doses on incident infection/reinfection; (B) Sensitivity analysis: estimated effectiveness of prior infection (study documented and reclassified based on N-antibodies) and number of vaccine doses on incident infection/reinfection.

non-overlapping CIs for two vaccine doses with or without prior infection and ≥three doses without prior infection, with the caveat that we did not have enough precision to interpret estimates for one dose and three doses with prior infection (online supplemental figures 8 and 9, online supplemental table 6). Protection against infection/reinfection with two vaccine doses without prior infection attenuated to the null in the Omicron period (estimated effectiveness pre-Omicron: 72% (48%, 85%); Omicron: 21% (-15%, 46%)). In contrast, ≥three vaccine doses without prior infection remained statistically significantly protective against infection/reinfection (estimated effectiveness pre-Omicron: 89% (70%, 96%); Omicron: 53% (30%, 68%)). Point estimates of the estimated effectiveness of prior infection with or without vaccination were reduced in the Omicron compared with pre-Omicron time period and were similar to the estimated efficacy of ≥three vaccine doses without prior infection in the Omicron period (estimated effectiveness of prior infection with or without vaccination: pre-Omicron 77–89%, Omicron 49–67%; Omicron ≥three vaccine doses without prior infection: 53%).

Sensitivity analyses

In our first sensitivity analysis, we compared our results using prior infection defined based on documentation *and* elevated anti-nucleocapsid levels. Using N-antibody levels to reclassify individuals without prior infection into the prior infection group, 234 individuals who did not have documentation of prior infection (16% of participants without documentation of prior infection) were subsequently added to the prior infection group. Comparing relative associations from the main analysis, point estimates did not substantially change and CIs overlapped estimates from the main analysis (table 2, figures 1B and 2). Using the no prior infection group for each vaccine dose as the referent group, pairwise comparisons revealed that prior infection provided additional protection, with stronger relationships for no and one dose (none: 56% (95% CI 30% to 72%), one: 66% (95% CI 28% to 84%), two: 31% (95% CI 7% to 49%), ≥three 31% (95% CI 0% to 53%)) (online supplemental table 5).

We compared our results from the main analysis and first sensitivity analysis to those if 23 individuals without prior infection who were missing N-antibody results were excluded from both the main analysis and first sensitivity analysis. Point estimates were effectively unchanged, and CIs overlapped from the original analyses (online supplemental figure 10, online supplemental table 6).

We compared our results from our first sensitivity analysis to those using a lower N-antibody threshold to determine probable prior infections. This resulted in 593 individuals who did not have documentation of prior infection (39.8% of participants without documentation of prior infection) to be added to the prior infection group. Point estimates for prior infection with or without vaccination were slightly attenuated, while CIs overlapped estimates from the first sensitivity analysis (online supplemental figure 11, online supplemental table 6).

Comparing results from the first sensitivity analysis to those using vaccination operationalised as completion of primary series, estimates for one partially vaccinated with

Table 2 Estimates effectiveness of prior infection (study documented and reclassification based on N-antibodies) and vaccination against incident infection/reinfection

Vaccine dose—prior infection	Person-days	Unadjusted estimated effectiveness*	Adjusted† estimated effectiveness* (%) (95% CI)	Probability of remaining uninfected‡ (%) (95% CI)					
		Overall	Overall	50 days	100 days	150 days	200 days	250 days	300 days
None—no	25579.6	Reference	Reference	88.3 (82.7 to 93.9)	66.5 (49.7 to 83.3)	53.1 (27.4 to 78.8)	35.2 (–7 to 77.4)	24.5 (–32.5 to 81.5)	12.6 (–71.7 to 96.9)
None—yes	23916	48.8 (20.3 to 67.2)	55.5 (29.8 to 71.8)	94.7 (91.9 to 97.6)	83.7 (75 to 92.4)	75.9 (62.5 to 89.3)	63.5 (41.5 to 85.6)	54.2 (24.4 to 84.1)	40.7 (–3.6 to 85)
One—no	16474.2	29.3 (–10.5 to 54.8)	16.6 (–30.7 to 46.8)	90.2 (84.6 to 95.7)	71.2 (54.2 to 88.2)	59 (32.8 to 85.2)	41.9 (–1.3 to 85.1)	31 (–27.5 to 89.4)	17.8 (–68.7 to 100.0)
One—yes	9287.8	59.9 (21.1 to 79.6)	70.9 (41.7 to 85.5)	96.5 (93.9 to 99.2)	89.1 (80.8 to 97.3)	83.6 (70.8 to 96.3)	74.4 (53.4 to 95.4)	67.1 (38.8 to 95.5)	55.7 (13.9 to 97.4)
Two—no	131916.8	51.1 (34.3 to 63.6)	49.4 (30.8 to 62.9)	93.9 (91.3 to 96.5)	81.4 (73.8 to 88.9)	72.6 (61.2 to 84.1)	59.1 (40.3 to 77.9)	49.2 (23.7 to 74.6)	35.2 (–2.4 to 72.8)
Two—yes	56903.2	58.9 (41.5 to 71.2)	65.3 (49.4 to 76.2)	95.8 (93.8 to 97.8)	86.9 (81 to 92.8)	80.4 (71.4 to 89.4)	69.8 (55.1 to 84.5)	61.6 (41.7 to 81.4)	49 (19.6 to 78.5)
Three—no	73887	51.3 (34.3 to 63.6)	70.7 (57.8 to 79.6)	96.5 (95 to 98.1)	89 (84.5 to 93.4)	83.4 (76.8 to 90.1)	74.2 (63.6 to 84.8)	66.9 (52.9 to 80.9)	55.3 (35.2 to 75.5)
Three—yes	28940	64.0 (45.4 to 76.2)	79.9 (67.7 to 87.5)	97.6 (96.2 to 98.9)	92.2 (88.2 to 96.2)	88.2 (82.1 to 94.3)	81.3 (71.4 to 91.1)	75.6 (62.4 to 88.8)	66.3 (47.1 to 85.5)

*Expressed as (1–hazard ratio)×100, where 0 indicates no effectiveness.
†Adjusted for age, sex, race/ethnicity, income, education, essential worker status, date of entry into vaccine-prior infection category, and site.
‡Covariates held at the most common value: race/ethnicity=non-Latino-White, mean age, sex=female, income=US$50 000–US$74 999, education=college graduate or higher, essential worker status=NA/unemployed, entry into vaccination category=06 October 2021, site=4, and mean weight=1.
NA, Not Applicable.

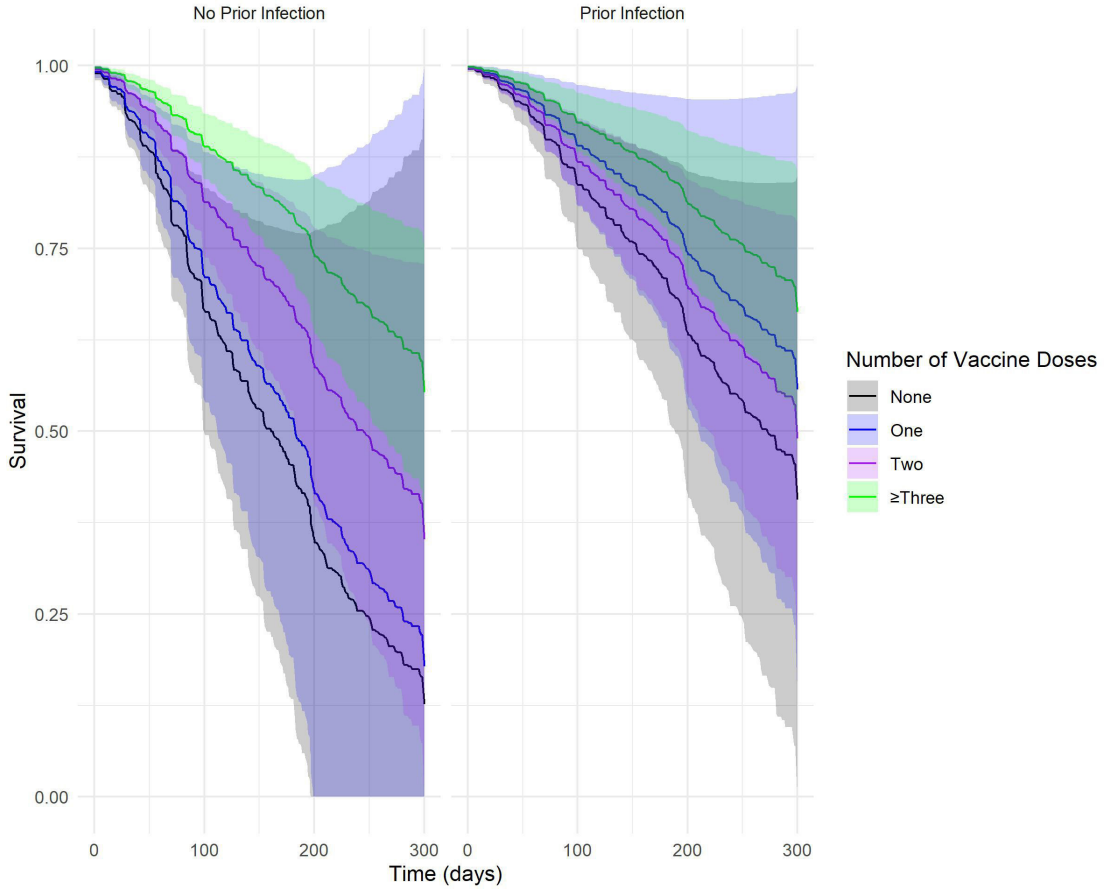

Figure 2 Survival curves by prior infection status (study documented and reclassified based on N-antibodies) and number of vaccine doses.* *Modelled holding covariates at their mode/median values (non-Latino-White, female, US$50 000–74 999 household income, college graduate, unemployed, vaccine dose start date of 6 October 2021, site 4). Inverse probability of censoring weights were set to 1.

or without infection differed from the first sensitivity analysis mainly due to small sample size. Other estimates did not substantially change, and CIs overlapped estimates from the main analysis (online supplemental figure 12, online supplemental table 4).

Comparing results from our first sensitivity analysis to those limiting the time under observation of vaccine exposure groups to 90 days, where protection from vaccination should be strongest, we observed increased point estimates, but CIs overlapped with those from the first sensitivity analysis (online supplemental figure 13, online supplemental tables 4 and 6).

Using a lower threshold to identify individuals with probable incident infections based on N-antibodies across subsequent visits, an additional 14 individuals were re-classified as having an incident infection. Estimated effectiveness did not substantially change from the first sensitivity analysis (online supplemental figure 14, online supplemental table 6).

Beginning an individual's contribution to time under a vaccine category at the day after vaccination rather than 14 days after vaccination produced similar results, although slight reductions in the point estimate for the prior infection-no vaccination category (online supplemental figure 15, online supplemental table 6).

DISCUSSION

We found that vaccination (in particular with ≥three vaccine doses) and prior infection both provided substantial protection against infection/reinfection among a cohort of US adults during the study follow-up, with substantial decline of protection during the Omicron period. Additional information of prior infection provided by nucleocapsid antibodies reclassified nearly 16% of individuals without documentation of prior infection into the prior infected category but did not substantially alter estimates of protection against infection/reinfection.

Our findings align with those from prior studies assessing the effectiveness of combinations of vaccination and prior infection on SARS-CoV-2 infection/reinfection. During the Omicron variant timeframe, investigators observed a similar magnitude of protective effects conferred by three vaccine doses without prior infection (40–48% increased efficacy compared with primary series[28]). Our estimates of protection conferred by ≥three vaccine doses plus prior infection are similar to those previously reported for reinfection (reinfection: 80% vs 83%[10] and 70% increased efficacy compared with primary series[6]). Moreover, the protective effects of vaccination against infection were reduced during the Omicron

variant compared with previous variants.[2 5 6 28] A systematic review of nine studies found that the protection conferred by primary series completion against SARS-CoV-2 infection during the Omicron variant was reduced by nearly 20% from one to 4 months after vaccination,[29] 2 months sooner than that reported among previous variants.[30] Our results showing reduced estimated effectiveness for two and ≥three vaccine doses with or without prior infection align with this literature.

This is the first study, to our knowledge, to demonstrate that correction for misclassification of prior infection using an anti-nucleocapsid antibody test does not substantially alter estimates of protection conferred by combinations of vaccination and prior infection. We found nearly 16% of individuals without documentation of prior infection had antibody levels similar to those with documentation of prior infection. Research during the initial stages of the pandemic demonstrated that antibody measurements are a useful tool to identify prior infection and showed that individuals with antibody indication of prior SARS-CoV-2 infection had greater protection against SARS-CoV-2 reinfection relative to those without.[31 32] While antibodies wane over time, other investigators have shown the persistence of nucleocapsid antibody detection in approximately 70% of individuals with a prior history of infection at 18 months after infection,[11] with lower values among individuals who were vaccinated prior to infection.[33] Based on this research, we do not expect waning to have substantially affected our results, as the median time from prior infection to antibody measurement was 8.4 months (min: 17 days, max: 22 months). It is also worth noting that we did not find material waning of N-antibodies over time among the uninfected with prior infection. Nevertheless, it is possible that we did not capture all individuals who may have had an undocumented prior infection due to waning of nucleocapsid antibodies over time. Further research is needed to confirm our results.

The strengths of this study are in its prospective design, real-world conditions, repeated SARS-CoV-2 testing of participants, geographically diverse sample of participants, use of immunologic measures to complement documentation of prior infection and analytic method to account for changing vaccination status over time. The limitations of our study include a relatively small sample. Additionally, the many exposure categories likely reduced the power of the study. Our participants represent a convenience sample of US adults, which may differ from the general US adult population in health status, limiting the generalisability of results.[34] This study did not involve an ideal randomised trial of vaccination or prior infection. We may not have captured all the confounding factors that influence vaccination behaviour and likelihood of incident infection. Therefore, though we may use language that implies causation, in all cases that should be seen as shorthand for associations.[35]

CONCLUSION

Among a geographically diverse cohort of US adults, we found that ≥three vaccine doses of a Pfizer, Moderna or Janssen vaccine without prior infection provided substantial protection against SARS-CoV-2 infection/reinfection, as did prior infection without vaccination. Vaccination and prior infection had protective effects, but effectiveness decreased in the Omicron period. Additionally, correction for misclassification of prior infection using antibodies did not alter estimates of effectiveness of prior infection and vaccination in combination.

Author affiliations
[1]Department of Epidemiology and Biostatistics, Indiana University Bloomington School of Public Health, Bloomington, Indiana, USA
[2]Biostatistics Consulting Center, Indiana University Bloomington School of Public Health, Bloomington, Indiana, USA
[3]Department of Surgery, University of California Irvine School of Medicine, Irvine, California, USA
[4]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA
[5]Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, California, USA
[6]Department of Applied Health Science, Indiana University Bloomington School of Public Health, Bloomington, Indiana, USA
[7]Midwest Biomedical Research Addison, Addison, Illinois, USA
[8]Department of Microbiology and Immunology, University of Iowa, Iowa City, Iowa, USA
[9]Mabloc LLC, Washington, District of Columbia, USA
[10]Department of Pathology, The George Washington University School of Medicine and Health Sciences, Washington, District of Columbia, USA

Acknowledgements We sincerely thank Danielle Ompad and John Lowe for their comments on prior versions of this manuscript. We also would like to extend our thanks to site PIs, sub-Is and staff for their work enrolling participants and gathering data, in particular Rupal Trivedi, Dawn Beckman, Indika Edirisinghe, Britt Burton-Freeman, Allison Blomer, Ronald Z Surowitz, Nikhil Dhurandhar, Kristina Petersen, Brandon Lenox, Tom Looke, John Lowe and Melessa Kelley. Finally, we thank the participants of this study for their time and participation.

Contributors AJ and ETB conducted data analysis, interpreted results and prepared the manuscript; AJ is the guarantor of the analysis and manuscript; DBA and KCM conceived of the study, oversaw data collection and reviewed for critical intellectual content; MJR and CL oversaw data collection, conceived the analytic methodology, interpreted results and reviewed for critical intellectual content; JTM oversaw data collection and reviewed for critical intellectual content; SD, LGA and ALMM managed data acquisition, performed data cleaning, assisted with data analysis, reviewed for reproducibility and critical intellectual content; KMG critiqued the study methods that led to protocol amendments, collected data, reviewed for critical intellectual content; JPI, MR and SP reviewed for critical intellectual content.

Funding This work was supported by multiple anonymous donations through the IU Foundation and by Eli Lilly and Company.

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 and was reviewed and all measures and protocols associated with this research were approved by MB Clinical Research (tracking #: 20203947, analysis protocol #: MB-2015), Indiana University's Human Subjects Office (IRB #: 22953) and Pennington Biomedical Research Center. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Ashley Judge https://orcid.org/0000-0003-4501-9744
John P Ioannidis https://orcid.org/0000-0003-3118-6859

REFERENCES

1 Hu W-H, Cai H-L, Yan H-C, *et al*. Protective effectiveness of previous infection against subsequent SARS-Cov-2 infection: systematic review and meta-analysis. *Front Public Health* 2024;12:1353415.
2 Stein C, Nassereldine H, Sorensen RJD, *et al*. Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis. *Lancet* 2023;401:833–42.
3 Sánchez-de Prada L, Martínez-García AM, González-Fernández B, *et al*. Impact on the time elapsed since SARS-CoV-2 infection, vaccination history, and number of doses, on protection against reinfection. *Sci Rep* 2024;14:353.
4 Bobrovitz N, Ware H, Ma X, *et al*. Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression. *Lancet Infect Dis* 2023;23:556–67.
5 Altarawneh HN, Chemaitelly H, Ayoub HH, *et al*. Effects of previous infection, vaccination, and hybrid immunity against symptomatic Alpha, Beta, and Delta SARS-CoV-2 infections: an observational study. *EBioMedicine* 2023;95:104734.
6 Altarawneh HN, Chemaitelly H, Ayoub HH, *et al*. Effects of Previous Infection and Vaccination on Symptomatic Omicron Infections. *N Engl J Med* 2022;387:21–34.
7 Hall V, Foulkes S, Insalata F, *et al*. Protection against SARS-CoV-2 after Covid-19 Vaccination and Previous Infection. *N Engl J Med* 2022;386:1207–20.
8 Vicentini M, Venturelli F, Mancuso P, *et al*. Risk of SARS-CoV-2 reinfection by vaccination status, predominant variant and time from prior infection: a cohort study, Reggio Emilia province, Italy, February 2020 to February 2022. *Euro Surveill* 2023;28:2200494.
9 Chin ET, Leidner D, Lamson L, *et al*. Protection against Omicron from Vaccination and Previous Infection in a Prison System. *N Engl J Med* 2022;387:1770–82.
10 Carazo S, Skowronski DM, Brisson M, *et al*. Estimated Protection of Prior SARS-CoV-2 Infection Against Reinfection With the Omicron Variant Among Messenger RNA-Vaccinated and Nonvaccinated Individuals in Quebec, Canada. *JAMA Netw Open* 2022;5:e2236670.
11 Jarlhelt I, Pérez-Alós L, Bayarri-Olmos R, *et al*. Distinguishing SARS-CoV-2 infection and vaccine responses up to 18 months post-infection using nucleocapsid protein and receptor-binding domain antibodies. *Microbiol Spectr* 2023;11:e01796-23.
12 Meyer B, Torriani G, Yerly S, *et al*. Validation of a commercially available SARS-CoV-2 serological immunoassay. *Clin Microbiol Infect* 2020;26:1386–94.
13 Long Q-X, Liu B-Z, Deng H-J, *et al*. Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* 2020;26:845–8.
14 Cauchemez S, Horby P, Fox A, *et al*. Influenza infection rates, measurement errors and the interpretation of paired serology. *PLoS Pathog* 2012;8:e1003061.
15 Miyamoto S, Numakura K, Kinoshita R, *et al*. Serum anti-nucleocapsid antibody correlates of protection from SARS-CoV-2 re-infection regardless of symptoms or immune history. *Commun Med (Lond)* 2025;5:172.
16 Lindsey KM, Farrell Z, Tutino R, *et al*. From first infection to reinfection: Comparing Nucleocapsid antibody kinetics in vaccinated and unvaccinated adults. *Vaccine (Auckl)* 2025;62:127593.
17 Yasmin F, Najeeb H, Moeed A, *et al*. COVID-19 Vaccine Hesitancy in the United States: A Systematic Review. *Front Public Health* 2021;9:770985.
18 Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1:385–401.
19 Rubin DB. *Multiple imputation, in flexible imputation of missing data*. Chapman and Hall/CRC, 2018: 29–62.
20 Posit Team. RStudio: integrated development environment for R. Posit Software; 2024. Available: http://www.posit.co/
21 Therneau T. A package for survival analysis in R. 2024.
22 Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. Springer, 2000.
23 van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
24 van der Wal WM, Geskus RB. ipw: An R package for inverse probability weighting. *J Stat Softw* 2011;43:1–23.
25 Beidelman ET, Judge A, Allison DB, *et al*. The relationship between psychological stress and SARS-CoV-2 nucleocapsid-antibody response following natural infection: Longitudinal findings from the Aegis Study. *Am J Epidemiol* 2026.:kwag048.
26 Hoeve CE, Neppelenbroek N, Vos ERA, *et al*. Using SARS-CoV-2 nucleoprotein antibodies to detect (re)infection. *Epidemiol Infect* 2025;153:e38.
27 Rasmussen M, Neumann A, Moghaddassi M, *et al*. Anti-SARS-CoV-2 nucleocapsid antibodies to detect exposure to SARS-CoV-2: results from a prospective cohort study on COVID-19 vaccination. *Infect Dis (Lond)* 2025;57:782–92.
28 Andeweg SP, de Gier B, Eggink D, *et al*. Protection of COVID-19 vaccination and previous infection against Omicron BA.1, BA.2 and Delta SARS-CoV-2 infections. *Nat Commun* 2022;13:4738.
29 Higdon MM, Baidya A, Walter KK, *et al*. Duration of effectiveness of vaccination against COVID-19 caused by the omicron variant. *Lancet Infect Dis* 2022;22:1114–6.
30 Feikin DR, Higdon MM, Abu-Raddad LJ, *et al*. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *The Lancet* 2022;399:924–44.
31 Hall VJ, Foulkes S, Charlett A, *et al*. SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN). *Lancet* 2021;397:1459–69.
32 Harvey RA, Rassen JA, Kabelac CA, *et al*. Association of SARS-CoV-2 Seropositive Antibody Test With Risk of Future Infection. *JAMA Intern Med* 2021;181:672–9.
33 Dhakal S, Yu T, Yin A, *et al*. Reconsideration of Antinucleocapsid IgG Antibody as a Marker of SARS-CoV-2 Infection Postvaccination for Mild COVID-19 Patients. *Open Forum Infect Dis* 2023;10:ofac677.
34 Lindsted KD, Fraser GE, Steinkohl M, *et al*. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist Health Study. *J Clin Epidemiol* 1996;49:783–90.
35 Hernán MA. The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. *Am J Public Health* 2018;108:616–9.