

https://dspace.org/

SARSCoV2 correlates of protection from infection against variants of concern

K Sun; Jinal Bhiman; Stefano Tempia; Jacoba Kleynhans; Vimbai Madzorera ...

https://hdl.handle.net/10539/43373
WOS

Downloaded from a DSpace repository. https://dspace.org/

nature medicine

Article

https://doi.org/10.1038/s41591-024-03131-2

SARS-CoV-2 correlates of protection from infection against variants of concern

Received: 15 December 2023

Accepted: 11 June 2024

Published online: 26 July 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

Serum neutralizing antibodies (nAbs) induced by vaccination have been linked to protection against symptomatic and severe coronavirus disease 2019. However, much less is known about the efficacy of nAbs in preventing the acquisition of infection, especially in the context of natural immunity and against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) immune-escape variants. Here we conducted mediation analysis to assess serum nAbs induced by prior SARS-CoV-2 infections as potential correlates of protection against Delta and Omicron infections, in rural and urban household cohorts in South Africa. We find that, in the Delta wave, D614G nAbs mediate 37% (95% confidence interval: 34–40%) of the total protection against infection conferred by prior exposure to SARS-CoV-2, and that protection decreases with waning immunity. In contrast, Omicron BA.1 nAbs mediate 11% (95% confidence interval: 9–12%) of the total protection against Omicron BA.1 or BA.2 infections, due to Omicron's neutralization escape. These findings underscore that correlates of protection mediated through nAbs are variant specific, and that boosting of nAbs against circulating variants might restore or confer immune protection lost due to nAb waning and/or immune escape. However, the majority of immune protection against SARS-CoV-2 conferred by natural infection cannot be fully explained by serum nAbs alone. Measuring these and other immune markers including T cell responses, both in the serum and in other compartments such as the nasal mucosa, may be required to comprehensively understand and predict immune protection against SARS-CoV-2.

The acute phase of the coronavirus disease 2019 (COVID-19) pandemic has waned with the development of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) population immunity in most individuals through repeated episodes of vaccination, infection or both[1,2]. Owing to the unprecedented speed of SARS-CoV-2 vaccine development and distribution[3], considerable numbers of people were primed by vaccination, averting substantial morbidity and mortality[4]. However, due to immune-evasive variants, vaccine hesitancy and lack of global equity in vaccine access[5–7], a substantial proportion of the world's population acquired SARS-CoV-2 immunity through natural infections, especially in low- and middle-income countries[8,9]. Immune markers that reliably predict protection against infection or symptomatic disease are known as 'correlates of protection' (CoPs). The post-pandemic era is marked by rapid antigenic drift of Omicron subvariants leading to continued immune evasion[10–13]. Given this complex evolutionary landscape, it remains important to identify CoPs induced by natural infections and/or vaccinations against SARS-CoV-2 variants to monitor population susceptibility, anticipate future waves, optimize rollout of existing vaccines and facilitate design and approval of next-generation vaccines[14]. There has been substantial progress in defining serum neutralizing or binding antibodies to the spike protein as CoPs for COVID-19 vaccines, although most of these data are derived from early randomized controlled trials focused on peak immune responses shortly after vaccination and measured against

e-mail: kaiyuan.sun@nih.gov; cherylc@nicd.ac.za

https://doi.org/10.1038/s41591-024-03131-2

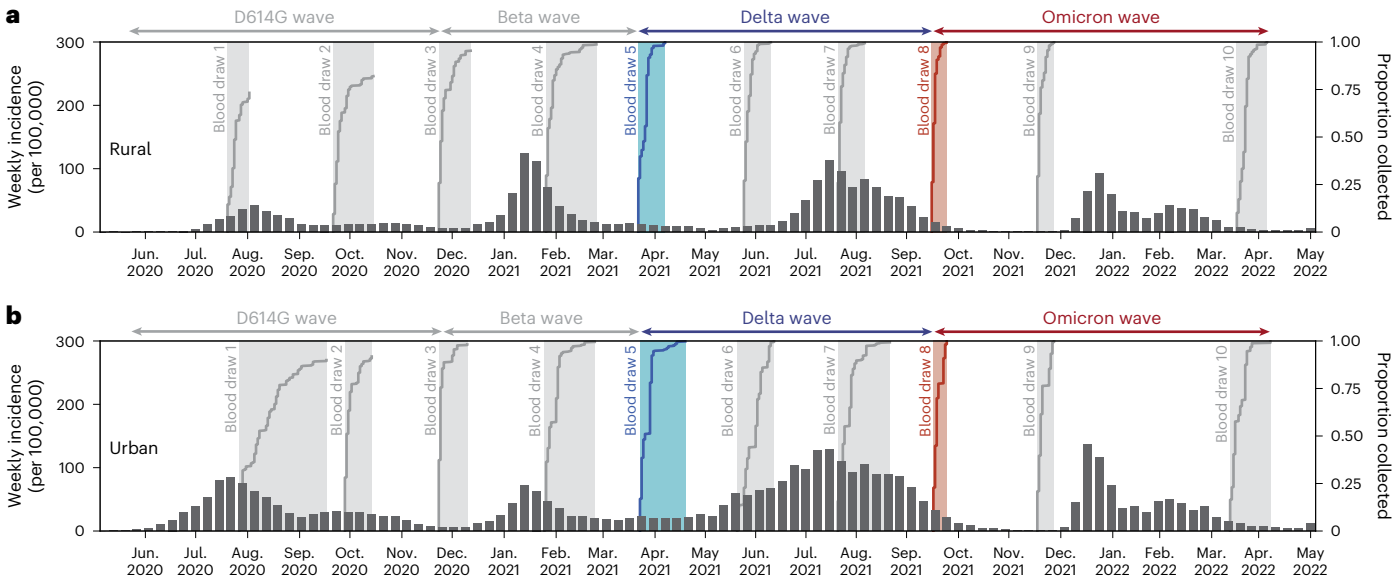


Fig. 1 | Timing of cohort sample collections with respect to SARS-CoV-2 variants' circulations in the two study sites. a, Timing of the blood draws with respect to the SARS-CoV-2 epidemic waves in the rural site (Agincourt) of the PHIRST-C cohort. The bar plot represents the weekly incidence (per 100,000 population) of SARS-CoV-2 cases from routine surveillance data collected from the Ehlanzeni district in the Mpumalanga province (where rural participants reside). The shaded areas represent the timing of the serum sample collections for the ten blood draws. Each curve within the shaded area indicates the cumulative proportion of participants' serum samples collected over time. The Delta wave subgroup analysis focuses on nAb titers among serum samples collected during BD5 (blue); the Omicron wave analysis focuses on nAb titers among serum samples collected during BD8 (red). **b**, Same as **a**, but for the urban site (Klerksdorp). The routine surveillance data (bar plot) were collected from the Dr. Kenneth Kaunda district in the North West province (where urban participants reside).

symptomatic disease caused by the ancestral strain, with updated data on variants[15–24]. In comparison, less is known about serum CoPs for infection-induced immunity[25], and protection against acquisition of subclinical infections.

CoPs may differ for immunity induced by infection versus vaccination: SARS-CoV-2 infections tend to induce more robust mucosal immunity despite lower serum antibody responses than intramuscularly delivered mRNA vaccines, as shown in a mouse model[26], and mucosal immunity may play a more important role in reducing risk of infection and transmission than systemic immunity[27,28]. Moreover, CoPs need to be interpreted in the context of viral evolution: in the pre-Omicron era, SARS-CoV-2 variants of concern emerged independently from one another, with the Alpha, Beta, Gamma, Delta and Omicron variants exhibiting distinct phenotypic characteristics. The Omicron variant stands out due to substantial genetic divergence from earlier strains and marked immune-evasion capabilities against antibody neutralization[29]. Equivalent antibody titers may not provide equivalent levels of protection against ancestral strains compared to more transmissible and immune-evasive variants like Omicron, and CoPs may, therefore, be variant dependent. Furthermore, serum antibody titers against SARS-CoV-2 wane with time.

The challenge of defining CoPs for infection-induced immunity partially stems from the difficulty of tracking immune exposures to SARS-CoV-2 infections, given that a substantial proportion of infections are asymptomatic or subclinical and cannot be fully captured by traditional symptom-based surveillance protocols. The SARS-CoV-2, Influenza and Respiratory Syncytial virus community burden, Transmission dynamics and viral interaction in South Africa (PHIRST-C) cohorts[30,31] overcame this challenge by implementing a rigorous sampling strategy, including collection of nasal swabs twice weekly during a period of intense follow-up, along with a total of ten sequential blood draws spanning the D614G (1st), Beta (2nd), Delta (3rd) and Omicron (4th) waves. Noting that the 1st wave was dominated by the SARS-CoV-2 ancestral strain with p.Asp614Gly substitution, we will refer to this variant as D614G in the rest of the paper for brevity. Additionally, the 4th wave was dominated by the Omicron BA.1 variant but also consisted of the Omicron BA.2 variant. We will refer to the 4th wave as the Omicron wave hereafter for brevity. This high-intensity sampling scheme allowed us to reconstruct the cohort participants' SARS-CoV-2 infection histories with high fidelity, and to monitor infection-induced antibody responses over time[31]. Blood samples collected immediately before Delta and Omicron waves offered a unique opportunity to investigate serum immune marker levels in close proximity to the next SARS-CoV-2 exposure. Furthermore, vaccine-derived immunity remained low at the onset of the Omicron wave, with less than 25% of the population fully immunized with Ad26.COV2.S (Janssen) and/or BNT162b2 (Pfizer BioNTech) vaccines[31,32]. In this study, we leveraged the PHIRST-C cohorts' unique serological and epidemiological data to perform mediation analysis and assess neutralizing antibody (nAb) titers induced by prior infection as CoPs against variants of concern. Specifically, we evaluated the role of D614G and Omicron BA.1 nAbs against the Delta and Omicron BA.1 and BA.2 infections.

Results

Cohort description and antibody titer measurements

We analyzed data from the multi-year PHIRST-C cohort study, covering the first four waves of SARS-CoV-2 infections including the Delta (3rd) and Omicron (4th) waves[30,31]. The study included a rural and an urban site in two provinces of South Africa. Households with more than two members and where at least 75% of members consented to participate were eligible. A total of 1,200 individuals from 222 randomly selected and eligible households among the two study sites were longitudinally followed from June 2020 through April 2022. The study was characterized by intense nasopharyngeal swab and serum sample collection from the peak of the SARS-CoV-2 D614G wave to after the peak of the Delta wave. After this initial follow-up period, nasopharyngeal swab sample collection stopped but serum samples continued with blood drawn immediately following the Omicron wave. The timing of the serum

https://doi.org/10.1038/s41591-024-03131-2

sample collection is visualized in Fig. 1. We previously reconstructed the detailed SARS-CoV-2 infection history of each individual in the cohort up to the Omicron wave and demonstrated that immunity conferred by prior infection reduced the risk of reinfection[31,33]. In this study, we extended this work to investigate how infection-induced nAb titers correlate with protection against SARS-CoV-2 reinfection with the Delta or Omicron BA.1 and BA.2 variants.

For the Delta wave, we focused on a subgroup of 797 participants from 196 households (Delta wave subgroup; Table 1 and Extended Data Fig. 1) who remained SARS-CoV-2 naive or had a single prior SARS-CoV-2 infection before the Delta wave (hence, excluding vaccinated and repeatedly infected individuals from the analysis; see Fig. 1 for the timing of the Delta wave). We define prior infection as positivity on the Roche Elecsys anti-nucleocapsid assay (an assay optimized to detect prior infection[34]), and/or real-time reverse-transcriptase polymerase-chain-reaction (rRT-PCR) positivity, at or before blood draw 5 (refer to BD5 hereafter). SARS-CoV-2 infections during the Delta wave were inferred based on the anti-nucleocapsid antibody level of two pre-Delta and one post-Delta wave serum samples, as previously described[31]. We focused on households with no more than six infected household members during this wave, due to computational constraints of the transmission model (Methods). Among the 797 subgroup participants, 34% (273/797) were infected during the Delta wave, with attack rates of 42% (229/544) and 17% (44/253) for naive and previously infected participants, respectively.

To identify CoPs against the Delta variant, for the 113 and 140 participants who had been infected by D614G and Beta variants prior to the Delta wave, we measured their D614G nAb titers (measured as the inhibitory dilution at which 50% neutralization is attained, referred to as ID_{50} hereafter), using the blood draw immediately preceding the Delta wave (BD5). To evaluate the potential impact of antibody waning, we also measured the peak nAb level for each participant (defined as the highest D614G nAb titer among the first five blood draws). We then calculated the degree to which nAbs had waned from peak level to that at BD5 by calculating the difference between peak nAb titer and nAb titer at BD5 (denoted as ΔnAb^{W} hereafter). If the peak response was already below the nAb assay detection threshold (which is set at 20), then ΔnAb^{W} was also assigned to be below the threshold, as further titer drop was not detectable. Notably, 28% (32/113) and 58% (81/140) of individuals previously infected with D614G and Beta exhibited D614G nAb titers below the detection threshold at BD5, respectively (Extended Data Table 1). The proportion below the detection threshold was higher for individuals previously infected with the Beta variant than the D614G variant, given the Beta variant has eight amino acid differences in the spike protein, resulting in an antigenically distinct receptor-binding domain compared to the D614G variant used in the neutralization assay. However, more than 90% of individuals remained positive on the Roche Elecsys anti-nucleocapsid assay for both prior D614G and Beta infected individuals[34], despite low nAb titer level (Extended Data Table 1).

Figure 2a shows the Delta wave participants' D614G nAb titers at peak and at BD5. The ID_{50} geometric mean titer (GMT) was 125 (95% confidence interval (CI): 97–161) at peak and waned to 85 (95% CI: 69–104) at BD5, representing an average 1.47-fold (95% CI: 1.32–1.67) reduction due to waning. The D614G nAb titers (in log scale) at peak and at BD5 were highly correlated (Pearson correlation coefficient 0.89, $P < 0.0001$). Comparing the nAb titers between individuals who were infected during the Delta wave versus those who were not infected, we found that the GMTs of infected individuals were significantly lower than those of uninfected individuals for both D614G nAbs at peak level and at BD5 (Fig. 2b,c). In contrast, we did not find a significant difference in the degree of antibody loss due to waning (ΔnAb^{W}) between infected and uninfected individuals (Fig. 2d).

Similarly, for the Omicron wave, we focused on a subgroup of 535 participants from 184 households who had only one prior SARS-CoV-2 infection (vaccinated and repeatedly infected individuals were excluded from the analysis) or remained naive just before the Omicron wave (see Table 1 and Extended Data Fig. 2 for a description of participants and Fig. 1 for the timing of the Omicron wave). Prior SARS-CoV-2 infection was ascertained in a similar fashion as for the Delta wave (that is, positivity by anti-nucleocapsid assay and/or rRT–PCR for the time period spanning the first eight blood draws). Infections during the Omicron wave were inferred based on the anti-nucleocapsid antibody level of two pre-Omicron and one post-Omicron wave serum samples, as previously described[31]. Two-thirds, or 67% (359/535), of participants included in the Omicron wave analysis were infected by these variants, with attack rates of 77% (149/193) and 61% (210/342) for naive and previously infected individuals, respectively.

To evaluate nAbs as CoPs in the context of Omicron's extensive immune escape, we measured both the D614G nAb titers and Omicron BA.1 nAb titers for serum samples collected at blood draw 8 (the blood draw taken shortly before the onset of the Omicron wave, referred to as

Table 1 | Characteristics of the PHIRST-C cohort's Delta and Omicron wave subgroup populations

	Delta wave subgroup 196 households	Omicron wave subgroup 184 households
Characteristics	Number of individuals (%)	Number of individuals (%)
All	797 (100)	535 (100)
Study site		
Rural	427 (54)	300 (56)
Urban	370 (46)	235 (44)
Age group, in years		
0–4	90 (11)	77 (14)
5–12	270 (34)	231 (43)
13–18	111 (14)	80 (15)
19–34	126 (16)	84 (16)
35–59	126 (16)	43 (8)
60+	74 (9)	20 (4)
Sex		
Male	324 (41)	229 (43)
Female	473 (59)	306 (57)
Household size		
3–5	372 (47)	254 (48)
6–8	264 (33)	197 (37)
9–12	124 (15)	72 (13)
13+	37 (5)	12 (2)
HIV status		
Negative	673 (85)	496 (93)
PLWH	97 (12)	31 (6)
Unknown	27 (3)	8 (1)
Prior immunity		
Naive	544 (68)	193 (36)
Prior D614G infection	113 (14)	61 (11)
Prior Beta infection	140 (18)	120 (22)
Prior Delta infection	–	161 (31)
Infected[a]		
Yes	273 (34)	359 (67)
No	524 (66)	176 (33)

[a]Indicates if a participant of the Delta/Omicron wave subgroup was infected (either primary or repeat infection) during the Delta/Omicron wave. PLWH, people living with HIV.

https://doi.org/10.1038/s41591-024-03131-2

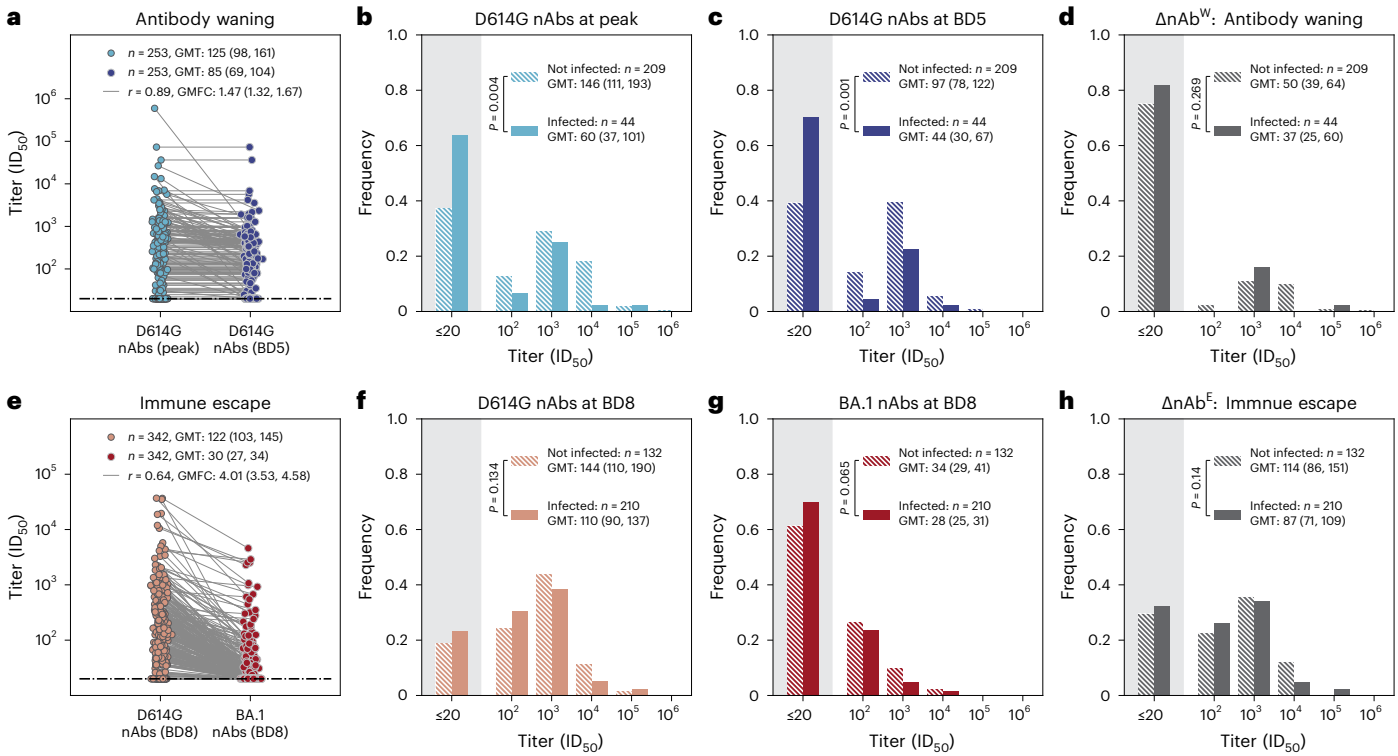


Fig. 2 | D614G and BA.1 nAb titers for the Delta wave and the Omicron wave analysis. a, For the Delta wave subgroup, the distribution of the peak D614G nAb titer up to BD5 (light blue dots) and the D614G nAb titer at BD5 (dark blue dots), among individuals who had one prior SARS-CoV-2 infection before BD5. Each dot represents one individual, with two measurements of the same individual connected through a gray line. GMFC, geometric mean fold change from peak D614G titer to that at BD5; *r*, Pearson correlation coefficient. **b**, For the Delta wave subgroup, the distribution of the peak D614G nAb titer up to BD5, stratified by individuals who were infected during the Delta wave (solid bar) versus those who were not infected (dashed bar). The independent-samples *t*-test (two-sided) was used to determine the statistical significance (*P* value reported in the legend) of the difference between the GMTs of the two groups. **c**, Same as **b** but for D614G nAb titers at BD5. **d**, Same as **b** but for ΔnAb^W (defined as the difference between D614G titers at peak and at BD5). **e**, For the Omicron wave subgroup, the distribution of D614G nAb titers (light red dots) and BA.1 titers at BD8 (dark red dots), among individuals who had one prior SARS-CoV-2 infection before BD8. Each dot represents one individual, with two measurements of the same individual connected through a gray line. **f**, For the Omicron wave subgroup, the distribution of the D614G nAb titer at BD8, stratified by individuals who were infected during the Omicron wave (solid bar) versus those who were not infected (dashed bar). The independent-samples *t*-test (two-sided) was used to determine the statistical significance (*P* value reported in the legend) of the difference between the GMTs of the two groups. **g**, Same as **f** but for BA.1 nAb titers at BD8. **h**, Same as **f** but for ΔnAb^E (defined as the difference between BA.1 and D614G titers at BD8).

BD8 hereafter). Given that none of the participants had been infected by Omicron before BD8, the Omicron BA.1 neutralizing activity at this time point originated from cross-reactive antibodies elicited by prior variant infections. Thus, the difference between D614G and BA.1 nAb titers at BD8 represents the quantity of D614G nAbs that failed to recognize mutated epitopes on Omicron BA.1, resulting in a lack of neutralizing function against Omicron BA.1. For the remainder of the paper, we will use ΔnAb^E to represent the quantity of antibodies able to neutralize D614G but not Omicron BA.1 due to mutations in the Omicron spike. Similarly to the Delta wave subgroup, a substantial proportion of previously infected individuals in the Omicron wave subgroup exhibited D614G and Omicron nAb titers below the detection threshold at BD8 (Extended Data Table 1). The absence of detectable nAbs was also more pronounced when the variant of prior infection and the variant's spike used in the neutralization assay were mismatched (Extended Data Table 1). Roche Elecsys anti-nucleocapsid assay remained robust in detecting prior infection[34], despite low nAb titer level (Extended Data Table 1).

Figure 2e shows the D614G and the BA.1 nAb titers at BD8 for participants included in the Omicron wave analysis. The nAb GMT against D614G was 122 (95% CI: 103–145) and 30 (95% CI: 27–34) for antibodies that could neutralize BA.1, representing an average 4.01-fold (95% CI: 3.53–4.58) reduction attributed to the immune-evasive properties of Omicron. The D614G and BA.1 nAb titers (in log scale) at BD8 were modestly correlated (Pearson correlation coefficient 0.64, $P < 0.0001$). Comparing the nAb titers between individuals who were infected during the Omicron wave versus those who were not infected, we did not find significant differences in GMT levels for D614G nAbs, BA.1 nAbs or ΔnAb^E (Fig. 2f–h). However, it is worth noting that the point estimates of GMTs were higher for uninfected individuals compared to infected individuals across all three measurements.

Pre-exposure nAb titer as a CoP against variant infection

We conducted mediation analyses in a household transmission modeling framework to investigate how nAb titers against SARS-CoV-2 variants at the onset of a SARS-CoV-2 wave mediate the risk of infection during the corresponding epidemic wave[35,36]. Specifically, following the causal inference framework proposed by Halloran and Struchiner[37], we introduced SARS-CoV-2 transmission probabilities as causal parameters, representing either the risk of acquiring infection from the general community or the per-contact transmission risk within the household. Transmission probabilities were dependent on an individual's prior infection history, the level of preexisting nAb titers (mediators) and other confounding factors (age, sex and comorbidities). We fitted a chain-binomial household transmission model, parameterized by the transmission probabilities, to the infection outcomes of the Delta and Omicron waves among all subgroup participants. To evaluate how the level of nAb titers mediated SARS-CoV-2 transmission probability, we use the fitted transmission model to project potential infection outcomes among previously infected individuals in counterfactual

scenarios absent prior infection exposure or nAbs. The details of the mediation analysis are described in Methods.

For the Delta wave mediation analysis, we considered D614G nAb titer at BD5 as the candidate mediator of protection and the quantity of antibodies that had waned from peak (ΔnAb^W) as the putative negative control (that is, we hypothesized that antibodies lost due to waning could not conceivably contribute to protection). For the Omicron wave, we considered BA.1 nAb titer at BD8 as the candidate mediator of protection and the quantity of nAbs that escape Omicron neutralization (ΔnAb^E) at BD8 as the putative negative control. We used the term 'direct effect' from the causal inference framework to refer to the effect of exposure (prior infection) on the outcome (repeat infection during the Delta or Omicron wave) in the absence of the mediators (nAb titers). Conversely, the term 'indirect effect' represents the effect of exposure (prior infection) on the outcome (repeat infection) that operates through the mediators (nAb titers). We estimated both the direct effect of prior infection and effects mediated through specific nAb titers against serologically confirmed SARS-CoV-2 infections. We report the estimates of the mediation analysis for both Delta and Omicron wave in Table 2. For the ease of interpretation, we then translate the estimated odds ratios into risk reductions (1 − odds ratio), along with other estimates in causal diagrams depicted in Fig. 3.

Our findings indicate that immunity derived from prior infection, overall, reduced the risk of contracting a Delta wave infection by 61% (95% CI: 59–63%; Fig. 3a). Notably, nAbs represented an important mediator of this overall protection: for every 10-fold increase in the D614G nAb titers at BD5, the risk of infection decreased by 40% (95% CI: 19–56%). In contrast, the decline in nAbs from peak levels to BD5 (ΔnAb^W) showed no contribution to the overall protection, with a risk reduction per 10-fold increase of −1% (95% CI: −21–16%). This result indicated that the waning of nAbs leads to waning of protection, in agreement with our hypothesis. Furthermore, we estimated that the protection mediated through D614G nAbs at BD5 accounted for 37% (95% CI: 34–40%) of the overall protection derived from prior infection, suggesting that over half of the protection against Delta was not mediated by serum nAbs against D614G. Lastly, our analysis indicated that individuals reinfected with the Delta variant were 78% (95% CI: 24–94%) less likely to transmit the infection to other household members compared to those who experienced primary infections (Fig. 3a). This finding suggested that even in cases where prior immunity is not sufficient to block reinfection with the Delta variant, infection-induced immunity still offered sizable mitigation against onward transmission.

The causal diagram depicting the mediation analysis for the Omicron wave is illustrated in Fig. 3b. Our findings indicate that, overall, prior infection-derived immunity resulted in a 37% (95% CI: 35–38%) reduction in the risk of contracting an Omicron wave infection, a notably lower effect compared to that of the Delta wave. We observed that Omicron BA.1 nAbs at BD8 significantly mediated protection against the Omicron BA.1 and BA.2 variants: for every 10-fold increase in Omicron BA.1 nAb titers, the risk of Omicron BA.1 or BA.2 infection decreased by 28% (95% CI: 6–44%). Conversely, antibodies unable to neutralize Omicron due to immune escape (ΔnAb^E) did not mediate protection against Omicron BA.1 or BA.2 infection, with a risk reduction of −1% (95% CI: −21–16%) per 10-fold titer increase. Furthermore, we estimated that the protection mediated through Omicron BA.1 nAbs at BD8 accounted for only 11% (95% CI: 9–12%) of the total protection conferred by prior exposure. This, coupled with the observation that Omicron BA.1 caused an average 4.01-fold drop in nAb titers (Fig. 2e), underscores the ability of Omicron BA.1 and BA.2 to evade host protective immunity mediated through nAbs. Additionally, in contrast to the Delta wave, individuals reinfected with the Omicron BA.1 or BA.2 variant were as likely to transmit the infection to other household members compared to those who experienced primary infections (risk reduction of −17%, 95% CI: −110–35%). These observations suggest that Omicron not only evades prior immunity's protection against acquisition of infection but also escapes protection against onward transmission.

Table 2 | Mediation analysis for nAbs as CoPs against Delta and Omicron wave infections, with a waning model for direct effect

Wave			Delta	Omicron
Protection against reinfection	Direct effect (protection absent of nAbs)	Effect size (odds ratio, absent of waning)	0.34 (0.17, 0.68)	0.29 (0.17, 0.50)
		Waning half-life (days)	121 (72, 242)	
	Mediator effect (protection from nAbs)	D614G nAb (odds ratio, per 10-fold increase)	0.60 (0.44, 0.81)	–
		ΔnAb^W (odds ratio, per 10-fold increase)	1.01 (0.74, 1.37)	–
		Omicron BA.1 nAb (odds ratio, per 10-fold increase)	–	0.72 (0.56, 0.94)
		ΔnAb^E (odds ratio, per 10-fold increase)	–	1.01 (0.84, 1.21)
	Total protection (relative risk compared to naive individuals)		0.39 (0.37, 0.41)	0.63 (0.62, 0.65)
	Proportion of protection mediated by nAbs		37% (34%, 40%)	11% (9%, 12%)
	Protection against onward transmission (odds ratio compared to naive individuals)		0.22 (0.06, 0.76)	1.17 (0.65, 2.10)

Averages and 95% CIs are provided for each of the model parameters. ΔnAb^W: the quantity of D614G nAbs waned from peak level to that at BD5. ΔnAb^E: the quantity of antibodies that can neutralize D614G but fail to neutralize Omicron BA.1 at BD8 due to Omicron's immune escape.

Although neutralizing titers measured at BD5 and BD8 offered a temporally proximate evaluation of protective immunity preceding the onset of the Delta and Omicron waves, we could not identify the immune mediators responsible for the direct effects of prior immunity (that is, the fraction of protection that was not mediated by nAbs) due to lack of additional serum biomarkers. We could, however, estimate the potential for these direct effects to wane over time. To do so, we modeled an exponential decline for the direct effect based on the time elapsed since prior infection and jointly estimated the duration of protection for both the Delta and Omicron waves' analysis. We found that protection not mediated by nAbs decreased with time, with a waning half-life of 121 (95% CI: 72–242) days (Fig. 3 and Table 2). After adjusting for waning, the effect sizes of protection from direct effects were similar for both variants, with odds ratios of acquiring infection in the absence of waning (compared to naive individuals) of 0.34 (95% CI: 0.17, 0.68) and 0.29 (95% CI: 0.17, 0.50) for the Delta and Omicron wave, respectively (Table 2). These results suggest that, while Omicron escaped preexisting nAbs, protection from other immune effectors was preserved against this variant. The waning half-life of protection not mediated by nAbs was estimated at approximately 4 months in our study, comparable to the reported waning timescale of T cell immunity[38,39]. Several sensitivity analyses demonstrating the robustness of the findings of the mediation analysis are reported in Methods.

Discussion

In this cohort of unvaccinated individuals, we found that nAb titers immediately before the onset of the Delta wave (that is, D614G nAb level at BD5) correlated with protection against Delta wave infections. Moreover, we demonstrated that nAb titers lost over time due to waning (that is, ΔnAb^W) were not associated with protection, aligning with the

https://doi.org/10.1038/s41591-024-03131-2

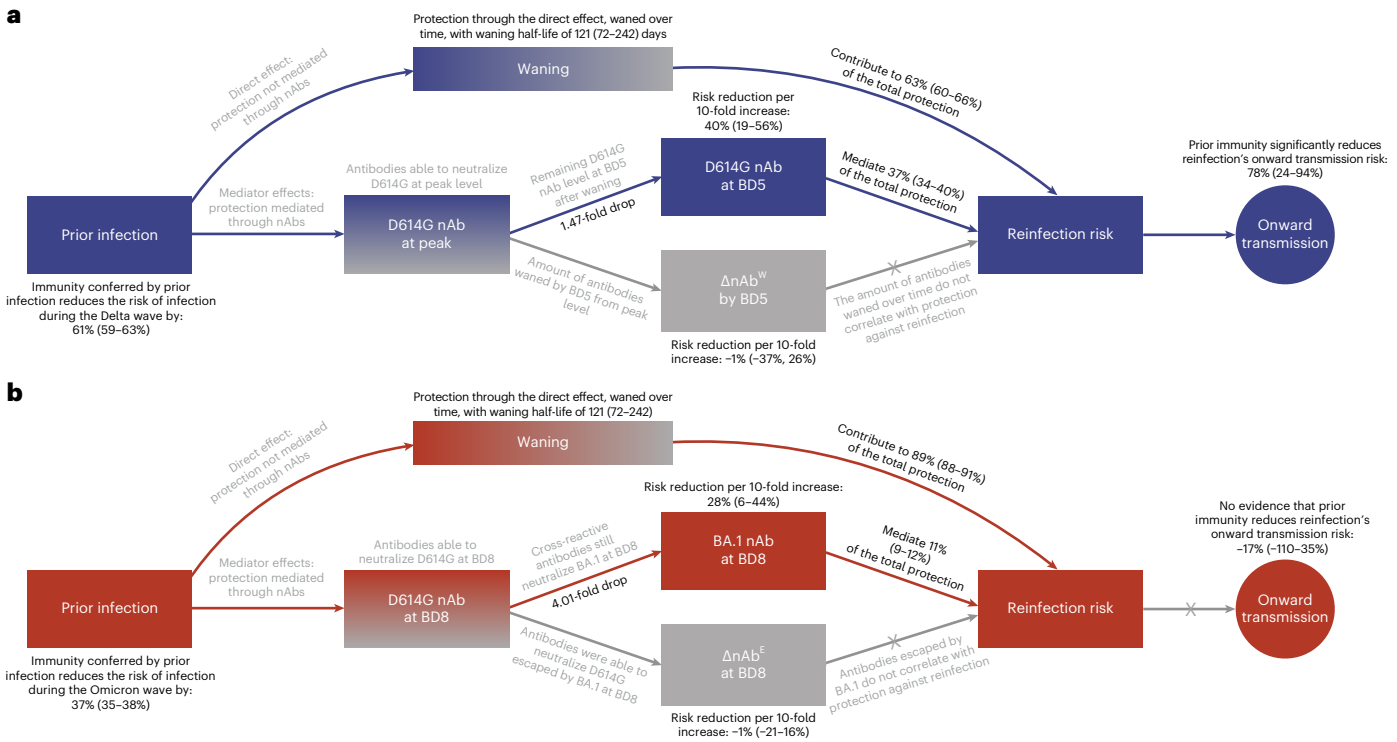


Fig. 3 | Causal diagrams for the mediation analyses. a, Causal diagram of the Delta wave mediation analysis showing the hypothesized relationship between prior immunity (induced by prior SARS-CoV-2 infection) and SARS-CoV-2 infection (outcome of interest) during the Delta wave. The mediators of interest are D614G nAbs at BD5 and ΔnAb^{W} (the quantity of D614G nAbs waned from peak level to that at BD5). The direct effect represents protection operating through immune mechanisms other than the mediators of interest. We hypothesized that the direct effect could wane over time since the initial immune exposure. For the prospective cohort data, both mediator–outcome confounding and exposure–outcome confounding factors need to be adjusted for the mediation analysis, as the immune exposure (prior SARS-CoV-2 infection) was not randomly assigned (unlike SARS-CoV-2 randomized control vaccine trials where vaccination was randomly assigned to the participants). Furthermore, cohort participants may experience heterogeneous levels of SARS-CoV-2 exposure due to different intensity SARS-CoV-2 transmission in their household settings. We adjusted this by embedding the mediation analysis in a mechanistic household transmission model (Methods). We also look at the impact of prior immunity on the reduction of onward transmission, conditional on the failure of preventing reinfection. The estimates of the Delta wave mediation analysis are presented in Table 2. **b**, Same as **a** but for the Omicron wave analysis. The mediators of interest are BA.1 nAbs at BD8 and ΔnAb^{E} (the quantity of antibodies that can neutralize D614G but fail to neutralize Omicron BA.1 at BD8 due to Omicron's immune escape). The estimates of the Omicron wave mediation analysis are presented in Table 2.

expectation that waning of nAbs in serum corresponds to waning of clinical protection. For the Omicron wave subgroup, we further investigated the impact of immune escape against protection mediated through nAbs. We found that only Omicron BA.1 nAbs correlated with protection against infection during the Omicron wave, whereas D614G nAbs that were unable to neutralize Omicron BA.1 in vitro due to spike escape mutations did not show clinical protection. The identification of variant-neutralizing antibodies derived from infection-induced immunity as CoPs against infections for both Delta and Omicron variants aligns with findings from studies on variant-specific correlates for vaccine-induced or hybrid immunity[21–24]. Considering that antibody-mediated protection against acquisition of infection likely operates at the mucosal site rather than in serum, it is interesting that serum antibody levels can anticipate protection[27]. In a recent analysis of the data from the COVE trial, Zhang et. al. further demonstrated that boosting of nAb titers against Omicron by a third dose of mRNA-1273 vaccine, afforded additional protection against Omicron compared to individuals who only received two doses of the mRNA-1273 vaccine[22]. Collectively, these empirical data lend support for using nAbs against circulating variants as immunobridging markers for periodic vaccine updates.

While a comprehensive understanding of the role of nAbs in SARS-CoV-2 protection is important, a key finding of our study is that serum nAb titers did not fully mediate protection conferred by prior infection. In the case of the Delta wave subgroup, we estimate that D614G nAbs mediate about one-third (37%) of protection. In comparison, nAbs mediate about two-thirds of the mRNA-1273 vaccine efficacy[15]. For the Omicron wave subgroup, Omicron BA.1 nAbs are estimated to mediate only 11% of protection, which was substantially lower than that observed for the Delta wave and vaccine-induced immunity. This low percentage of protection mediated by nAbs for the Omicron wave could be attributed to the highly immune-evasive nature of Omicron against neutralizing activity. Omicron effectively rendered a substantial proportion of serum D614G nAbs nonfunctional against Omicron. The large proportion of overall protection that was not mediated by nAbs may be explained by a variety of immune mechanisms, including the Fc-effector function of binding antibodies, and T cell functions, both of which are resilient against mutations in variants of concern[14,40]. Additionally, SARS-CoV-2 initially infects and predominantly transmits through the upper respiratory tract. Mucosal immunity in the upper respiratory tract, therefore, likely plays a key role in preventing SARS-CoV-2 infection, and may not be fully represented by immune markers in serum[41]. Our study validates the use of serum nAbs as a CoP against reinfection but also suggests potential important roles for other candidate immune markers that could act as 'co-correlates' of protection[42]. This is particularly important because these mechanisms may be more broadly cross-protective against future variants than nAbs. Future CoP analyses incorporating measurements of T cell immunity and non-nAb functions, ideally at the mucosal site, could potentially disentangle these important protective mechanisms and inform the design of next-generation vaccines[27,43–45].

Our study has several limitations. Firstly, the vaccination rate in the PHIRST-C cohort was low at the time of the analysis; with <20% of participants fully vaccinated before the Omicron wave (thus excluded from our analysis). Consequently, we lacked sufficient statistical power to assess CoPs for vaccine-induced (or hybrid) immunity and compare with our findings for infection-induced immunity in the same cohort. Secondly, we focused on SARS-CoV-2 infections that were ascertained by seroconversion or amnestic boosting of the anti-nucleocapsid antibodies. However, not all PCR-positive SARS-CoV-2 infections led to systemic antibody response[31,46,47]. Thus, our CoP analysis does not account for protection against abortive or transient infections that lack systemic antibody responses. We also could not evaluate CoPs against symptomatic cases, as there was no systemic monitoring of SARS-CoV-2 symptoms for the cohort population during the Omicron waves. Further, severe outcomes (hospitalizations and deaths) due to SARS-CoV-2 infections were rare during the PHIRST-C study, and evaluation of protection against those outcomes is underpowered. Identifying correlates of protection against severe outcomes is important from both clinical and public health perspectives, thus warranting further studies. Thirdly, the strains of antigens used in the neutralization assay were not perfectly matched to the circulating variants in the CoP analysis. For the Delta wave analysis, we evaluated D614G antibody titers (rather than nAb titers against Delta). Although Delta is not as immune evasive as Omicron with respect to D614G, there are substitutions on the spike of Delta (that is, L452R and T478K) that are linked to moderate antigenic escape[48,49]. In addition, although infections were predominantly caused by the Delta variant during the Delta wave epidemic, other variants also circulated at low levels during the same time period, including Alpha and C.1.2 (ref. 31). Similarly, genomic surveillance revealed that while Omicron BA.1 accounted for the majority of infections during the Omicron wave, Omicron BA.2 also co-circulated, with potential antigenic spike substitutions (for example, T376A, D405N, R408S) that were not present in BA.1 (refs. 31,49,50). Thus, using a BA.1-specific neutralizing assay may introduce bias in our CoP analysis, particularly against Omicron BA.2. Lastly, we only measured serum antibodies, but did not have any information on antibody response at the mucosal site or on cell-mediated immunity. While serum IgG nAbs may transudate into the nasal mucosa and thereby play a role in protection, the contribution of locally produced nasal IgA nAb remains to be investigated.

Moving forward, future works focusing on understanding how protective immunity accumulates through repeated infections, vaccinations and hybrid immunity, and identifying a suite of predictive markers of protection reflecting different arms of immune responses[51], are key to anticipating long-term SARS-CoV-2 burden, optimizing vaccine boosters and designing next-generation SARS-CoV-2 vaccines.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-03131-2.

References

1. World Health Organization. Statement on the Fifteenth Meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic (2023).
2. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/ (accessed 20 June 2024).
3. World Health Organization. WHO COVID19 Vaccine Tracker. https://covid19.trackvaccines.org/agency/who/ (accessed 20 June 2024).
4. Watson, O. J. et al. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect. Dis.* **22**, 1293–1302 (2022).
5. Gozzi, N. et al. Estimating the impact of COVID-19 vaccine inequities: a modeling study. *Nat. Commun.* **14**, 3272 (2023).
6. Wang, Q. et al. Mapping global acceptance and uptake of COVID-19 vaccination: a systematic review and meta-analysis. *Commun. Med.* **2**, 113 (2022).
7. Lazarus, J. V. et al. A survey of COVID-19 vaccine acceptance across 23 countries in 2022. *Nat. Med.* **29**, 366–375 (2023).
8. Bergeri, I. et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: a systematic review and meta-analysis of standardized population-based studies. *PLoS Med.* **19**, e1004107 (2022).
9. Lewis, H. C. et al. SARS-CoV-2 infection in Africa: a systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *BMJ Glob. Health* **7**, e008793 (2022).
10. Wang, Q. et al. Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4 and BA.5. *Nature* **608**, 603–608 (2022).
11. Ito, J. et al. Convergent evolution of SARS-CoV-2 Omicron subvariants leading to the emergence of BQ.1.1 variant. *Nat. Commun.* **14**, 2671 (2023).
12. Cao, Y. et al. BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature* **608**, 593–602 (2022).
13. Cao, Y. et al. Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature* **614**, 521–529 (2023).
14. Goldblatt, D., Alter, G., Crotty, S. & Plotkin, S. A. Correlates of protection against SARS-CoV-2 infection and COVID-19 disease. *Immunol. Rev.* **310**, 6–26 (2022).
15. Gilbert, P. B. et al. Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science* **375**, 43–50 (2022).
16. Fong, Y. et al. Immune correlates analysis of the PREVENT-19 COVID-19 vaccine efficacy clinical trial. *Nat. Commun.* **14**, 331 (2023).
17. Fong, Y. et al. Immune correlates analysis of the ENSEMBLE single Ad26.COV2.S dose vaccine efficacy clinical trial. *Nat. Microbiol* **7**, 1996–2010 (2022).
18. Feng, S. et al. Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 2032–2040 (2021).
19. Khoury, D. S. et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021).
20. Earle, K. A. et al. Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* **39**, 4423–4428 (2021).
21. Atti, A. et al. Antibody correlates of protection against Delta infection after vaccination: a nested case–control within the UK-based SIREN study. *J. Infect.* **87**, 420–427 (2023).
22. Zhang, B. et al. Omicron COVID-19 immune correlates analysis of a third dose of mRNA-1273 in the COVE trial. Preprint at *bioRxiv* https://doi.org/10.1101/2023.10.15.23295628 (2023).
23. Hertz, T. et al. Correlates of protection for booster doses of the SARS-CoV-2 vaccine BNT162b2. *Nat. Commun.* **14**, 4575 (2023).
24. Gilboa, M. et al. Factors associated with protection from SARS-CoV-2 Omicron variant infection and disease among vaccinated health care workers in Israel. *JAMA Netw. Open* **6**, e2314757 (2023).
25. Maier, H. E. et al. An immune correlate of SARS-CoV-2 infection and severity of reinfections. Preprint at medRxiv https://doi.org/10.1101/2021.11.23.21266767 (2021).
26. Tang, J. et al. Respiratory mucosal immunity against SARS-CoV-2 after mRNA vaccination. *Sci. Immunol.* **7**, eadd4853 (2022).
27. Knisely, J. M. et al. Mucosal vaccines for SARS-CoV-2: scientific gaps and opportunities-workshop report. *NPJ Vaccines* **8**, 53 (2023).
28. Miyamoto, S. et al. Infectious virus shedding duration reflects secretory IgA antibody response latency after SARS-CoV-2 infection. *Proc. Natl Acad. Sci. USA* **120**, e2314808120 (2023).

29. Markov, P. V. et al. The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **21**, 361–379 (2023).
30. Cohen, C. et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020-21. *Lancet Infect. Dis.* **22**, 821–834 (2022).
31. Sun, K. et al. Rapidly shifting immunologic landscape and severity of SARS-CoV-2 in the Omicron era in South Africa. *Nat. Commun.* **14**, 246 (2023).
32. Pulliam, J. R. C. et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa. *Science* **376**, eabn4947 (2022).
33. Sun, K. et al. SARS-CoV-2 transmission, persistence of immunity, and estimates of Omicron's impact in South African population cohorts. *Sci. Transl. Med.* **14**, eabo7081 (2022).
34. Roche Diagnostics. Elecsys Anti-SARS-CoV-2. https://diagnostics.roche.com/us/en/products/params/elecsys-anti-sars-cov-2.html (accessed June 20, 2024).
35. Cowling, B. J. et al. Influenza hemagglutination-inhibition antibody titer as a mediator of vaccine-induced protection for influenza B. *Clin. Infect. Dis.* **68**, 1713–1717 (2019).
36. Lim, W. W., Shuo, F., Wong, S.-S., Sullivan, S. G. & Cowling, B. J. Hemagglutination inhibition antibody titers mediate influenza vaccine efficacy against symptomatic influenza A(H1N1), A(H3N2), and B/Victoria infections. *J. Infect. Dis.* **7**, jiae122 (2024).
37. Halloran, M. E. & Struchiner, C. J. Causal inference in infectious diseases. *Epidemiology* **6**, 142–151 (1995).
38. Cohen, K. W. et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021).
39. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
40. Eser, T. M. et al. Nucleocapsid-specific T cell responses associate with control of SARS-CoV-2 in the upper airways before seroconversion. *Nat. Commun.* **14**, 2952 (2023).
41. Havervall, S. et al. Anti-spike mucosal IgA protection against SARS-CoV-2 Omicron infection. *N. Engl. J. Med.* **387**, 1333–1336 (2022).
42. Plotkin, S. A. Vaccines: correlates of vaccine-induced immunity. *Clin. Infect. Dis.* **47**, 401–409 (2008).
43. Morens, D. M., Taubenberger, J. K. & Fauci, A. S. Rethinking next-generation vaccines for coronaviruses, influenzaviruses, and other respiratory viruses. *Cell Host Microbe* **31**, 146–157 (2023).
44. Topol, E. J. & Iwasaki, A. Operation Nasal Vaccine-Lightning speed to counter COVID-19. *Sci. Immunol.* **7**, eadd9947 (2022).
45. US Department of Health and Human Services. Fact Sheet: HHS details $5 billion 'Project NextGen' initiative to stay ahead of COVID-19. https://aspr.hhs.gov/newsroom/Pages/ProjectNextGen-May2023.aspx (accessed 20 June 2024).
46. Liu, W. et al. Predictors of nonseroconversion after SARS-CoV-2 infection. *Emerg. Infect. Dis.* **27**, 2454–2458 (2021).
47. Lindeboom, R. G. H. et al. Human SARS-CoV-2 challenge uncovers local and systemic response dynamics. *Nature* https://doi.org/10.1038/s41586-024-07575-x (2024).
48. McCallum, M. et al. Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. *Science* **374**, 1621–1626 (2021).
49. Wilks, S. H. et al. Mapping SARS-CoV-2 antigenic relationships and serological responses. *Science* **382**, eadj0070 (2023).
50. Tegally, H. et al. Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat. Med.* https://doi.org/10.1038/s41591-022-01911-2 (2022).
51. Fourati, S. et al. Pan-vaccine analysis reveals innate immune endotypes predictive of antibody responses to vaccination. *Nat. Immun.* **23**, 1777–1787 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024

Kaiyuan Sun[1,13] ✉, Jinal N. Bhiman[2,3,13], Stefano Tempia[4,5,6], Jackie Kleynhans[4,5], Vimbai Sharon Madzorera[2,3], Qiniso Mkhize[2,3], Haajira Kaldine[2,3], Meredith L. McMorrow[6,7], Nicole Wolter[4,8], Jocelyn Moyes[4,5], Maimuna Carrim[4,8], Neil A. Martinson[9,10], Kathleen Kahn[11], Limakatso Lebina[9], Jacques D. du Toit[11], Thulisa Mkhencele[4], Anne von Gottberg[4,8], Cécile Viboud[1,14], Penny L. Moore[2,3,12,14], Cheryl Cohen[4,5,14] ✉ & PHIRST-C group*

[1]Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA. [2]SAMRC Antibody Immunity Research Unit, University of the Witwatersrand, Johannesburg, South Africa. [3]Centre for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. [4]Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa. [5]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [6]Influenza Division, Centers for Disease Control and Prevention, Atlanta, GA, USA. [7]COVID-19 Response, Centers for Disease Control and Prevention, Atlanta, GA, USA. [8]School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [9]Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa. [10]Johns Hopkins University Center for TB Research, Baltimore, MD, USA. [11]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [12]Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa. [13]These authors contributed equally: Kaiyuan Sun, Jinal N. Bhiman. [14]These authors jointly supervised this work: Cécile Viboud, Penny L. Moore, Cheryl Cohen. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: kaiyuan.sun@nih.gov; cherylc@nicd.ac.za

PHIRST-C group

Jinal N. Bhiman[2,3,13], Amelia Buys[4], Maimuna Carrim[4,8], Cheryl Cohen[4,5,14], Linda de Gouveia[4], Mignon du Plessis[4,8], Jacques du Toit[11], Francesc Xavier Gómez-Olivé[11], Kathleen Kahn[11], Kgaugelo Patricia Kgasago[9], Jackie Kleynhans[4,5], Retshidisitswe Kotane[4], Limakatso Lebina[9], Neil A. Martinson[9,10], Meredith L. McMorrow[6,7], Tumelo Moloantoa[4], Jocelyn Moyes[4,5], Stefano Tempia[4,5,6], Stephen Tollman[11], Anne von Gottberg[4,8], Floidy Wafawanaka[11] & Nicole Wolter[4,8]

Methods

Inferring Delta and Omicron wave infections based on longitudinal serum samples

We have previously described the serologic inference method for SARS-CoV-2 infections among the PHIRST-C cohort participants during the Delta wave (3rd SARS-CoV-2 wave) and the Omicron wave (4th SARS-CoV-2 wave)[31]. To briefly summarize, ascertainment of Delta wave infections was based on the serial serologic readout of blood draws 5 and 6 (both before the Delta wave; Fig. 1a,b) and blood draw 8 (after the Delta wave), measured by the Roche Elecsys Anti-SARS-CoV-2 nucleocapsid assay[34]. The participants' serologic trajectories were then grouped into 13 categories of distinct serum antibody patterns, reflecting the rise, waning and/or amnestic boosting of anti-nucleocapsid antibody levels. Because the Delta wave was also covered by intense virologic sampling with twice-weekly nasopharyngeal swab collection, we grouped the 13 serologic categories into indicators of either presence or absence of SARS-CoV-2 infection to achieve the highest concordance with rRT–PCR-confirmed Delta infections. The Omicron wave was not covered by the intense rRT-PCR testing; however, the timing of blood draws 8, 9 and 10 with respect to the Omicron wave was similar to that of blood draws 5, 6 and 8 with respect to the Delta wave (Fig. 1a,b). We thus applied the same classification method of serial serologic trajectories defined by blood draws 8, 9 and 10 to infer SARS-CoV-2 infections during the Omicron waves.

Laboratory methods

Serum nAb titers against SARS-CoV-2 D614G and BA.1 variants (lentiviral pseudovirus production and neutralization assay). Virus production and pseudovirus neutralization assays were done as previously described[52]. Briefly, 293T/ACE2.MF cells modified to overexpress human ACE2 (kindly provided by M. Farzan, Scripps Research) were cultured in DMEM (Gibco BRL Life Technologies) containing 10% heat-inactivated serum (FBS) and 3 μg ml^{-1} puromycin at 37 °C, 5% CO_2. Cell monolayers were disrupted at confluency by treatment with 0.25% trypsin in 1 mM EDTA (Gibco BRL Life Technologies). The SARS-CoV-2, Wuhan-1 spike, cloned into pCDNA3.1 was mutated using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) and NEBuilder HiFi DNA Assembly Master Mix (NEB) to include D614G (wild-type) or lineage defining mutations for Delta (T19R, 156–157del, R158G, L452R, T478K, D614G, P681R and D950N), Omicron BA.1 (A67V, 69–70del, T95I, G142D, 143–145del, 211del, L212I, 214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F) and Omicron BA.2 (T19I, L24S, 25–27del, G142D, V213G, G339D, S371F, S373P, S375F, T376A, D405N, R408S, K417N, N440K, S477N, T478K, E484A, Q493R, Q498R, N501Y, Y505H, D614G, H655Y, N679K, P681H, N764K, D796Y, Q954H, N969K). Pseudoviruses were produced by co-transfection in 293T/17 cells with a lentiviral backbone (HIV-1 pNL4. luc encoding the firefly luciferase gene) and either of the full-length SARS-CoV-2 spike plasmids with PEIMAX (Polysciences). Culture supernatants were clarified of cells by a 0.45-μm filter and stored at −70 °C. Plasma samples were heat inactivated and clarified by centrifugation. Pseudovirus and serially diluted plasma/sera were incubated for 1 h at 37 °C, 5% CO_2. Cells were added at 1×10^4 cells per well after 72 h of incubation at 37 °C, 5% CO_2. Luminescence was measured using a PerkinElmer Life Sciences Model Victor X luminometer. Neutralization was measured as described by a reduction in luciferase gene expression after single-round infection of 293 T/ACE2. MF cells with spike-pseudotyped viruses. Titers were calculated as the reciprocal plasma dilution (ID_{50}) causing a 50% reduction of relative light units.

We measured neutralization titer using a lentiviral-backboned pseudovirus neutralization assay. A systematic review of Omicron neutralization data showed that pseudovirus neutralization assays tend to report higher neutralizing titers compared to live-virus assays. The titer drops from wild type to Omicron also tend to be less pronounced for pseudovirus platforms, suggesting the pseudovirus assay may underestimate Omicron's capability to escape neutralization[53].

SARS-CoV-2 spike ELISA. For ELISA, Hexapro SARS-CoV-2 full spike protein with the D614G substitution was expressed in human embryonic kidney 293F suspension cells by transfecting the cells with the respective expression plasmid. After incubating for 6 days at 37 °C, proteins were first purified using a nickel resin followed by size exclusion chromatography. Relevant fractions were collected and frozen at −80 °C until use. In total, 2 μg ml^{-1} of D614G spike protein was used to coat 96-well, high-binding plates (Corning) and incubated overnight at 4 °C. The plates were incubated in a blocking buffer consisting of 1× PBS, 5% skimmed milk powder and 0.05% Tween 20. Plasma samples were diluted to a 1:100 starting dilution in a blocking buffer and added to the plates. IgG secondary antibody (Merck) was diluted to 1:3,000 in blocking buffer and added to the plates followed by TMB substrate (Thermo Fisher Scientific). Upon stopping the reaction with 1 M $H2SO_4$, optical density (OD) was measured at 450 nm. The monoclonal antibodies CR3022 and palivizumab were used as the positive and negative controls respectively.

Statistical analysis

Mediation analyses and household transmission model fitted to observed infections in the cohort. Here we blend concepts from causal inference and infectious disease transmission models. The no-interference assumption in causal inference stipulates that the outcome of an individual does not depend on the outcome of others, which is often violated in infectious disease dynamics[37,54,55]. This is because the spread of infectious diseases requires pathogens to be transmitted from one host to another. In other words, the infection outcome of one individual inherently depends on the infection outcome of others, and this is particularly pronounced in a household setting[37]. The 'dependent happening' nature of infectious disease dynamics violates the no-interference assumption. As a result, the traditional regression approach for causal inference analysis cannot be applied to infectious disease outcomes among individuals who can in theory transmit the disease from one to another. To overcome this, Halloran and Struchiner[37] introduced the probability of infection conditional on exposure to already infected individuals (transmission probability), as the causal parameter. Using this proposed framework, we can investigate how the presence/absence of preexisting immunity along with the immunologic marker of interest could modulate probability of infection, after adjusting for levels of exposure to the infectious source(s). The corresponding causal inference framework requires modeling the transmission process explicitly. Under this framework, we conduct mediation analyses to investigate how nAb titers against variants at the start of a SARS-CoV-2 wave correlate with SARS-CoV-2 transmission risk, using the Delta and Omicron waves as examples[35,36]. We focus on the Delta and Omicron subgroup participants who have had a single or no prior infection, and fit a chain-binomial model to their infection outcomes during the corresponding Delta/Omicron wave[56]. Specifically, we introduce the causal parameters:

- p_{ij}^k: the per-contact SARS-CoV-2 household transmission probability from infected individual i to individual j in household k.
- q_j^k: the overall probability of acquiring SARS-CoV-2 infection from outside the household by individual j of household k (probability of infection from the community).

We use e_j to indicate individual j's prior SARS-CoV-2 infection history, with $e_j = 0$ representing no prior infection reported before

the start of Delta/Omicron wave and $e_j = 1$ representing one prior infection by the start of Delta/Omicron wave. A prior SARS-CoV-2 infection ($e_j = 1$) would induce immunologic responses, measured by a set of immune markers (that is, candidate mediators) $\{m_j|e_j = 1\}$ (for example, nAb titers level). Then, the household transmission probability $p_{ij}^k = p_{ij}^k(e_j, \{m_j|e_j = 1\}, \{c_i, c_j, c_k\})$ can be expressed as a function of prior infection status e_j, immunologic mediators of SARS-CoV-2 transmission probability $\{m_j|e_j = 1\}$ and additional adjustment terms $\{c_i, c_j, c_k\}$, representing a set of potential confounding factors of individual i, individual j, and household k (for example, age of the donor and/or recipient, comorbidities and household size). Similarly, the community infection probability $q_j^k = q_j^k(e_j, \{m_j|e_j = 1\}, \{c_j\})$ can be expressed as a function of individual j's prior exposure history e_j, immunological markers $\{m_j|e_j = 1\}$, and additional adjustment terms $\{c_j\}$, representing a set of potential confounding factors of individual j (for example, age or comorbidities).

The causal diagram of the mediation analysis framework is shown in Fig. 3. We fit a household transmission model to the imputed household transmission chains based on an expectation–maximization (EM) algorithm (detailed in 'Transmission chains imputation and parameters estimation based on an EM algorithm'). Specifically, for the Delta/Omicron wave, if we look into a specific household k of size N, there are a total of n individuals infected belonging to L distinct chains of transmission due to L independent introductions of SARS-CoV-2 into the household. The uninfected individuals are $N - n$. We denote P_j^k the likelihood of any individual j of household k having the observed infection status over the Delta/Omicron wave (that is, either infected or not) in a particular realization of the model. There are a few scenarios to write down P_j^k:

- Within a given transmission chain $l \in L$, the initial generation $g_j^l = 0$ always has an individual j acquiring infection from the general community (outside the household k). Thus, the probability of individual j being infected is $P_j^k = q_j^k$ if j is the first individual to be infected in the chain.
- For infected individual j in the first generation of transmission chain l, that is, $g_j^l = 1$, this individual would have to escape infection risk from the general community but get infected by the infected household member of $g_i^l = 0$. Thus, the probability of individual j being infected can be written as $P_j^k = \left(1 - q_j^k\right) p_{ij}^k$.
- For infected individual j in transmission chain l with generation greater than 1, that is, $g_l > 1$, this individual has escaped infection risk from the general community as well as infected individuals i two generations away ($g_i^l \le g_j^l - 2$) but got infected by an infector i' of j's previous generation on the same transmission chain l. Thus, the probability of individual j being infected can be written as $P_j^k = \left(1 - q_j^k\right) \times \prod_{i \in \{g_i^l \le g_j^l - 2\}} \left(1 - p_{ij}^k\right) \times p_{i'j}^k$.
- For uninfected individual j within household k, this individual has escaped infection risk from the general community as well as all the n infected individuals within the same household. Thus, the probability of individual j remaining uninfected can be written as $P_j^k = \left(1 - q_j^k\right) \times \prod_{i \in \{n\}} \left(1 - p_{ij}^k\right)$.

Then, within household k of size N, we can express the likelihood of transmission chain l as $\prod_{j \in l} P_j^k$; the likelihood of observing all infections within k can be expressed as $\prod_{l \in L} \prod_{j \in l} P_j^k$; the likelihood of observing $N - n$ uninfected individuals can be expressed as $\prod_{j \in N-n} P_j^k$. Putting these together, the likelihood of observing one realization of the imputed (details of the EM imputation method described in the next section) households' transmission trees for Delta/Omicron wave can be expressed according to equation (1):

$$L^{\text{Delta/Omicron}} = \prod_k L_k^{\text{Delta/Omicron}} \tag{1}$$

Where the likelihood of a given household's transmission chain configuration $L_k^{\text{Delta/Omicron}}$ can be expressed according to equation (2):

$$L_k^{\text{Delta/Omicron}} = \prod_{l \in L} \prod_{j \in l} P_j^k\left(p_{ij}^k, q_j^k\right) \times \prod_{j \in N-n} P_j^k\left(p_{ij}^k, q_j^k\right) \tag{2}$$

In the remainder of the section, we will consider a few versions of the transmission model with slightly different implementations for p_{ij}^k and q_j^k.

Model 1: waning model for prior exposure with serologically ascertained Delta and Omicron wave infections. This is the transmission model presented in the main analysis of the paper (results of the model shown in Table 2. In this model, we consider that protection from prior infection unexplained by nAb titers wanes over time but is not dependent on the variant responsible for prior infection (that is, prior D614G or Beta infections for the Delta wave analysis, and prior D614G, Beta or Delta infections for the Omicron wave analysis). Additionally, in this model, both the Delta and Omicron wave infections were ascertained by serology based on the approach described in a prior session in Methods.

More specifically, for the Delta wave, p_{ij}^k and q_j^k can be expressed according to equations (3) and (4):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon\left(\frac{1}{2}\right)^{\frac{\Delta t}{\tau}} e_j + \left(\delta_{\text{nAb}}^{\text{D614G}} m_j^{\text{D614G}} + o_{\text{ΔnAb}}^{\text{waning}} m_j^{\text{waning}}\right)^{e_j} + \lambda e_i + \\ \sum_{c_i \in \{c_i\}} \gamma_{c_i} c_i + \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \alpha_s\end{array}\right) \tag{3}$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon\left(\frac{1}{2}\right)^{\frac{\Delta t}{\tau}} e_j + \left(\delta_{\text{nAb}}^{\text{D614G}} m_j^{D614G} + o_{\text{ΔnAb}}^{\text{waning}} m_j^{\text{waning}}\right)^{e_j} + \\ \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \beta_s\end{array}\right) \tag{4}$$

As described before, e_j indicates individual j's prior SARS-CoV-2 infection history, with $e_j = 0$ representing uninfected individuals at the start of the Delta wave, $e_j = 1$ representing one prior infection, and ϵ representing the effect size of the immune protection by prior infection not mediated through D614G nAbs (direct effect; Table 2). Δt is the elapsed time between prior infection and BD5 (the blood draw taken before the Delta wave, which we use in this model) and τ is the waning half-life of ϵ (direct effect; Table 2). m_j^{D614G} represents the D614G nAb titer at BD5 and $\delta_{\text{nAb}}^{\text{D614G}}$ represents the effect size of m_j^{D614G} in mediating infection probability p_{ij}^k against the Delta wave infection (mediator effect; Table 2) at BD5. While m_j^{waning} represents the quantity of D614G nAbs waned from peak level (measured as the highest D614G nAb titer level among the first five blood draws) to that at BD5 and $o_{\text{ΔnAb}}^{\text{waning}}$ represents the effect size of m_j^{waning} in mediating transmission probability p_{ji}^k against the Delta wave infection (mediator effect; Table 2) at BD5. Note that the term $\delta_{\text{nAb}}^{\text{D614G}} m_j^{\text{D614G}} + o_{\Delta\text{nAb}}^{waning} m_j^{\text{waning}}$ only exists when $e_j = 1$.

We further evaluate whether breakthrough infections have reduced infectiousness compared to primary infections and may in turn affect p_{ij}^k. We use e_i to indicate individual's (the donor) prior SARS-CoV-2 infection history ($e_i = 0$ means no infection, and $e_i = 1$ represents one prior infection at the start of Delta wave). Further, λ represents the effect size of prior infection (in i) in reducing the infectiousness of reinfections.

We also consider confounding factors for donor i and recipient j, where c_i and γ_{c_i} represent infector i's confounding factor (i's age and sex) and effect size, respectively; c_j and γ_{c_j} represent j's confounding factor (j's age-/sex-specific susceptibility (biology), age-/sex- and site-specific susceptibility (behavioral), HIV infection status) and effect size, respectively; c_k and γ_{c_k} represent household k's confounding factor (household size) and effect size, respectively. Lastly, α_s and β_s are logits of the baseline risks for household and community exposures. All parameters' effect sizes are measured in the log of odds ratios.

Similarly, for the Omicron wave, p_{ij}^k and q_j^k can be expressed according to equations (5) and (6):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon\left(\frac{1}{2}\right)^{\frac{\Delta t}{\tau}} e_j + \left(o_{\text{nAb}}^{\text{BA.1}} m_j^{\text{BA.1}} + o_{\Delta\text{nAb}}^{\text{escape}} m_j^{\text{escape}}\right)^{e_j} + \lambda e_i \\ + \sum_{c_i \in \{c_i\}} \gamma_{c_i} c_i + \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \alpha_s\end{array}\right) \quad (5)$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon\left(\frac{1}{2}\right)^{\frac{\Delta t}{\tau}} e_j + \left(o_{\text{nAb}}^{\text{BA.1}} m_j^{\text{BA.1}} + o_{\Delta\text{nAb}}^{\text{escape}} m_j^{\text{escape}}\right)^{e_j} \\ + \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \beta_s\end{array}\right) \quad (6)$$

As described before, e_j indicates individual j's prior SARS-CoV-2 infection history, with $e_j = 0$ representing individual j who remained naive to SARS-CoV-2 at the start of Omicron wave, while $e_j = 1$ represents individual j who had one prior infection at the start of Omicron wave, and ϵ represents the effect size of the immune protection by prior infection not mediated through D614G nAbs (direct effect; Table 2). Δt is the elapsed time between prior infection and BD8 (the blood draw taken before the Omicron wave), and τ is the waning half-life of ϵ (direct effect; Table 2). Here we consider that parameter τ is shared between the Delta and Omicron wave and will be jointly estimated (described in the next session). $m_j^{\text{BA.1}}$ represents the BA.1 nAb titer at BD8 and $o_{\text{nAb}}^{\text{BA.1}}$ represents the effect size of $m_j^{\text{BA.1}}$ in mediating transmission probability p_{ji}^k against the Omicron wave infection (mediator effect; Table 2) at BD8. m_j^{escape} represents the difference in titer from D614G nAb to BA.1 nAb at BD8, and $o_{\Delta\text{nAb}}^{\text{escape}}$ represents the effect size of m_j^{escape} in mediating transmission probability p_{ji}^k against the Omicron wave infection (mediator effect; Table 2) at BD8. Note that the term $o_{\text{nAb}}^{\text{BA.1}} m_j^{\text{BA.1}} + o_{\Delta\text{nAb}}^{\text{escape}} m_j^{\text{escape}}$ only exists when $e_j = 1$. All other parameters have the same definition of the Delta wave.

$\alpha_s, \beta_s, \epsilon, \tau, o_{\Delta\text{nAb}}^{\text{escape}}, o_{\text{nAb}}^{\text{BA.1}}, \{\gamma_{c_i}\}, \{\gamma_{c_j}\}, \{\gamma_{c_k}\}$ are estimated through maximizing the likelihood function L for each of the 100 bootstrapped realizations, and bootstrap mean and CIs are calculated for each of the parameters.

Sensitivity analysis

Model 2: sensitivity analysis considering variant-specific prior exposure for the direct effects. A potential confounding factor in understanding the waning of protection through direct effects is the diversity of prior SARS-CoV-2 exposures, with the dominance of the D614G variant in the first wave, Beta variant in the second wave and Delta variant in the third wave (Fig. 1). The effectiveness of protection may vary depending on the specific variant of prior exposure that induced the immune response at play. We conducted a sensitivity analysis (model 2) using a variant-specific model for the direct effects, which accounted for distinct types of SARS-CoV-2 variants conferring prior immunity, instead of considering a generic waning model. Specifically, in model 2, we considered a more complex version of model 1, where protection from prior infection depends on the type of infecting variant (that is prior D614G or Beta infections for the Delta wave analysis, and prior D614G, Beta or Delta infections for the Omicron wave analysis). We consider waning in neutralizing titers as in model 1, but we eliminate waning in the effect of prior infection that is not captured by neutralizing titers. More specifically, for the Delta wave, p_{ij}^k and q_j^k can be expressed according to equations (7) and (8):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon^{\text{D614G}} e_j^{\text{D614G}} + \epsilon^{\text{Beta}} e_j^{\text{Beta}} + \left(\delta_{\text{nAb}}^{\text{D614G}} m_j^{\text{D614G}} + o_{\Delta\text{nAb}}^{\text{waning}} m_j^{\text{waning}}\right)^{e_j} + \lambda e_i \\ + \sum_{c_i \in \{c_i\}} \gamma_{c_i} c_i + \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \alpha_s\end{array}\right) \quad (7)$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon^{\text{D614G}} e_j^{\text{D614G}} + \epsilon^{\text{Beta}} e_j^{\text{Beta}} + \left(\delta_{\text{nAb}}^{\text{D614G}} m_j^{\text{D614G}} + o_{\Delta\text{nAb}}^{\text{waning}} m_j^{\text{waning}}\right)^{e_j} + \\ \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \beta_s\end{array}\right) \quad (8)$$

Here, $e_j^{\text{D614G(Beta)}} = 1$ indicates individual j, before the Delta wave, was infected with D614G (Beta) variant. If $e_j^{\text{D614G}} = e_j^{\text{Beta}} = 0$, individual j was naive at the beginning of the Delta wave. ϵ^{D614G} and ϵ^{Beta} represent the effect size of immune protection by prior D614G and Beta infection not mediated through D614G nAbs, respectively.

For the Omicron wave, p_{ij}^k and q_j^k can be expressed per equations (9) and (10):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon^{\text{D614G}} e_j^{\text{D614G}} + \epsilon^{\text{Beta}} e_j^{\text{Beta}} + \epsilon^{\text{Delta}} e_j^{\text{Delta}} \\ + \left(o_{\text{nAb}}^{\text{BA.1}} m_j^{\text{BA.1}} + o_{\Delta\text{nAb}}^{\text{escape}} m_j^{\text{escape}}\right)^{e_j} + \lambda e_i + \\ \sum_{c_i \in \{c_i\}} \gamma_{c_i} c_i + \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \alpha_s\end{array}\right) \quad (9)$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon^{\text{D614G}} e_j^{\text{D614G}} + \epsilon^{\text{Beta}} e_j^{\text{Beta}} + \epsilon^{\text{Delta}} e_j^{\text{Delta}} \\ + \left(o_{\text{nAb}}^{\text{BA.1}} m_j^{\text{BA.1}} + o_{\Delta\text{nAb}}^{\text{escape}} m_j^{\text{escape}}\right)^{e_j} + \\ \sum_{c_j \in \{c_j\}} \gamma_{c_j} c_j + \sum_{c_k \in \{c_k\}} \gamma_{c_k} c_k + \beta_s\end{array}\right) \quad (10)$$

Here, $e_j^{\text{D614G(Beta,Delta)}} = 1$ indicates individual j, before the Omicron wave, was infected with the D614G (Beta, Delta) variant. If $e_j^{\text{D614G}} = e_j^{\text{Beta}} = e_j^{\text{Delta}} = 0$, individual j was naive at the beginning of the Omicron wave. ϵ^{D614G}, ϵ^{Beta} and ϵ^{Delta} represent the effect size of the immune protection by prior D614G, Beta and Delta infection not mediated through D614G nAbs, respectively.

Additionally, similarly to model 1, both the Delta and Omicron wave infections were ascertained by serology for model 2. All other settings of model 2 were kept the same as model 1. The results of model 2 are presented in Extended Data Table 2.

Our analysis revealed that for both the Delta and Omicron waves, more recent variants conferred stronger protection than earlier variants, albeit with overlapping CIs (Extended Data Table 2). This temporal trend aligns with the expectations of the waning model. Both waning and variant-specific immunity may modulate the direct effects of prior immunity; however, our study lacked sufficient statistical power to jointly estimate the relative contributions of these two factors. Full estimates of this sensitivity analyses are presented in Extended Data Table 2.

Model 3: sensitivity analysis with Delta wave infections ascertained by PCR and/or serology. For model 1, both the Delta and Omicron wave infection outcomes were inferred using the kinetics of anti-nucleocapsid antibodies from longitudinal serologic sampling, as detailed in previously published studies of the PHIRST-C cohort[31,33]. This approach for inferring infections based on serology was calibrated against virological evidence of infection during the Delta wave, established through twice-weekly rRT–PCR tests regardless of symptom presentation. However, it should be noted that this calibration did not achieve perfect concordance; the serology approach demonstrated 93% sensitivity and 89% specificity when compared to infections identified by rRT–PCR tests[31]. To address the uncertainties arising from the imperfect concordance between the two approaches for ascertaining infections, we conducted a sensitivity analysis (model 3) for the Delta wave, where we considered infections based on rRT–PCR positivity and/or anti-nucleocapsid antibody serology. We identified an additional 17 infections during the Delta wave through this more sensitive infection ascertainment approach, bringing the total number of Delta wave infections to 290. All other

settings of model 3 were kept the same as model 1. The results of the Model 3 are presented in Extended Data Table 3.

Notably, estimates of the direct and indirect effects of the mediation analysis were comparable between this sensitivity analysis and the main analysis (compare Extended Data Table 3 to Table 1). These findings provide support for the utilization of anti-nucleocapsid serology to ascertain Omicron wave infections in the studied cohorts, in a period where twice-weekly rRT–PCR testing was not available and confirms the robustness of our CoP analyses.

Model 4: D614G spike binding antibodies as mediators of protection. We conducted sensitivity analysis (model 4) to explore the role of D614G spike binding antibodies (referred to as bAbs hereafter), as potential CoPs for both Delta and Omicron infections. Using an in-house ELISA, we quantified the level of D614G spike bAbs by measuring absorbance at 450 nm at an OD at peak levels and BD5 (DB8) for the Delta (Omicron) wave analysis (Extended Data Fig. 3). The reduction in binding antibody levels from peak (ΔbAb^W) was determined as the difference between OD values at peak and BD5 (BD8) for the Delta (Omicron) wave (Extended Data Fig. 3).

Model 4 builds on model 2 but replaces nAb titers with D614G spiking binding ELISA readouts as mediators of protection, in order to compare the protection afforded by neutralizing versus binding antibodies. More specifically, for the Delta wave, p_{ij}^k and q_j^k can be expressed according to equations (11) and (12):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon^{D614G}e_j^{D614G} + \epsilon^{Beta}e_j^{Beta} + \left(\delta_{bAb}^{D614G}m_j^{D614G} + o_{\Delta bAb}^{waning}m_j^{waning}\right)^{e_j} + \lambda e_i + \\ \sum_{c_i\in\{c_i\}}\gamma_{c_i}c_i + \sum_{c_j\in\{c_j\}}\gamma_{c_j}c_j + \sum_{c_k\in\{c_k\}}\gamma_{c_k}c_k + \alpha_s\end{array}\right) \quad (11)$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon^{D614G}e_j^{D614G} + \epsilon^{Beta}e_j^{Beta} + \left(\delta_{bAb}^{D614G}m_j^{D614G} + o_{\Delta bAb}^{waning}m_j^{waning}\right)^{e_j} + \\ \sum_{c_j\in\{c_j\}}\gamma_{c_j}c_j + \sum_{c_k\in\{c_k\}}\gamma_{c_k}c_k + \beta_s\end{array}\right) \quad (12)$$

Here, m_j^{D614G} represents the D614G bAbs ELISA readout at BD5 and δ_{bAb}^{D614G} represents the effect size of m_j^{D614G} in mediating transmission probability p_{ij}^k against the Delta wave infection at BD5. Further, m_j^{waning} represents the drop from peak D614G bAbs readout before BD5 (measured as the highest D614G bAb titer level among the first five blood draws) to that at BD5 and $o_{\Delta bAb}^{waning}$ represents the effect size of m_j^{waning} in mediating transmission probability p_{ji}^k against the Delta wave infection at BD5.

For the Omicron wave, p_{ij}^k and q_j^k can be expressed per equations (13) and (14):

$$p_{ij}^k = \text{expit}\left(\begin{array}{c}\epsilon^{D614G}e_j^{D614G} + \epsilon^{Beta}e_j^{Beta} + \epsilon^{Delta}e_j^{Delta} + \\ \left(\delta_{bAb}^{D614G}m_j^{D614G} + o_{\Delta bAb}^{waning}m_j^{waning}\right)^{e_j} + \lambda e_i + \\ \sum_{c_i\in\{c_i\}}\gamma_{c_i}c_i + \sum_{c_j\in\{c_j\}}\gamma_{c_j}c_j + \sum_{c_k\in\{c_k\}}\gamma_{c_k}c_k + \alpha_s\end{array}\right) \quad (13)$$

$$q_j^k = \text{expit}\left(\begin{array}{c}\epsilon^{D614G}e_j^{D614G} + \epsilon^{Beta}e_j^{Beta} + \epsilon^{Delta}e_j^{Delta} + \\ \left(\delta_{bAb}^{D614G}m_j^{D614G} + o_{\Delta bAb}^{waning}m_j^{waning}\right)^{e_j} + \\ \sum_{c_j\in\{c_j\}}\gamma_{c_j}c_j + \sum_{c_k\in\{c_k\}}\gamma_{c_k}c_k + \beta_s\end{array}\right) \quad (14)$$

Here, m_j^{D614G} represents the D614G bAbs ELISA readout at BD8 and δ_{bAb}^{D614G} represents the effect size of m_j^{D614G} in mediating transmission probability p_{ij}^k against the Omicron wave infection at BD8. m_j^{waning} represents the drop from peak D614G bAbs readout before BD8 (measured as the highest D614G bAb titer level among the first eight blood draws) to that at BD8 and $o_{\Delta bAb}^{waning}$ represents the effect size of m_j^{waning} in mediating transmission probability p_{ji}^k against the Omicron wave infection at BD8. All other settings of model 4 were kept the same as model 2. The results of the model 4 are presented in Extended Data Table 4.

We found that binding antibody levels at BD5 (BD8) correlate with protection against Delta (Omicron) wave infections: the risk of infection decreased by 74% (95% CI: 41–88%) and 40% (95% CI: 33–54%) per unit increase in OD value for the Delta and Omicron wave analyses, respectively. Conversely, the decline in bAbs from peak levels to BD5/BD8 (ΔbAb^W) demonstrated no contribution to the overall protection, with a risk reduction per 10-fold increase of −2% (95% CI: −91–55%) for Delta wave infections and −2% (95% CI: −87–55%) for Omicron wave infections. These findings underscore the correspondence between waning of binding antibodies and a waning of protection. Furthermore, our estimations indicate that the proportion of protection conferred through D614G spike bAbs at BD5 is 35% (95% CI: 32–38%) against Delta wave infections, a figure comparable to the estimation based on D614G nAbs (37%, 95% CI: 34–40%; Extended Data Table 4). Notably, D614G spike bAbs at BD8 accounted for 27% (95% CI: 25–29%) of protection against the Omicron wave infection, representing a larger proportion compared to BA.1 nAbs (11%, 95% CI: 9–12%; Extended Data Table 4).

Transmission chain imputation and parameters estimation based on an EM algorithm. Here we describe the process to fit the models described in 'Statistical analysis' (model 1) and 'Sensitivity analysis' (models 2–4) to the household infection data. The serologic data available for the Delta and Omicron only provide information on the total number of infections within the household between two blood draws collected before and after the SARS-CoV-2 wave. The data do not provide the details of the transmission chains within the household, the order of infections among infected individuals, nor the infection dates. To account for the uncertainties of the transmission tree structure within households given only the total number of infections, we enumerate and reconstruct all possible transmission chains among the infected individuals, where each infected individual may have been infected by members of their own household or the general community. Supplementary Fig. 1 illustrates all 16 possible configurations of transmission chains for a household with 3 infected individuals. We limited our analysis to households with no more than 6 infected individuals, as the possible configurations of transmission chains among 6 infected individuals already reaches 16,807. Enumeration of all possible transmission chain configurations would be computationally intractable for households with more than 6 infected individuals. Additionally, the probability of each possible transmission chain depends on the parameter estimates of the transmission model described in the previous section. To address the statistical uncertainties due to unresolved transmission chains (which would affect the statistical confidence of mediation analysis detailed in the prior section), we jointly fit the household transmission model and impute the topological structure of the transmission trees. We use an EM algorithm, as described below[57].

To resolve who infected whom within the household in a probabilistic manner, we considered an EM algorithm that iteratively estimates the transmission model parameters α_s, β_s, ϵ, τ, $\delta_{nAb}^{D614G/BA.1}$, $o_{\Delta nAb}^{waning/escape}$, $\{\gamma_{c_i}\}$, $\{\gamma_{c_j}\}$, $\{\gamma_{c_k}\}$ through maximizing the likelihood function L as described in equation (1) in the previous section and then updates the imputed probability of each transmission tree configuration within each household based on the fitted transmission model. The process is as follows:

(1) Initial imputation of the household transmission trees with equal sampling probability for all configurations: For each

household, we randomly sample one transmission tree with equal probability among all transmission tree configurations that are compatible with the number of infections. We iterate through all households so that each household has a simulated transmission tree. We then repeat the imputation 1,000 times to obtain 1,000 realizations of each household's transmission tree.

(2) Maximization step: We consider the waning parameter τ a hyper-parameter (nonlinear term in equations (3–6), cannot be estimated by logistic regression). For a fixed value of τ, for each of the 1,000 realizations of the simulated household transmission chains, we estimate transmission model parameters $\alpha_s, \beta_s, \epsilon, \delta_{\text{nAb}}^{\text{D614G/BA.1}}, o_{\Delta\text{nAb}}^{\text{waning/escape}}, \{\gamma_{c_i}\}, \{\gamma_{c_j}\}, \{\gamma_{c_k}\}$ through maximizing the likelihood function L described in equation (1). The maximization of the likelihood function is achieved through fitting a logistic regression of the infection/exposure outcomes for all participants using R package 'brglm' (version 0.7.2). We then pool the estimates from the 1,000 realizations using the 'pool' function in the R package 'mice' (version 3.16.0). The full likelihood of the combined Delta and Omicron waves fitting in this EM step m can be expressed as $L_m(\tau) = L_m^{\text{Delta}}(\tau) \times L_m^{\text{Omicron}}(\tau)$

(3) Expectation step: for a fixed value of hyper-parameter τ, based on the pooled estimates of the transmission model parameters $\alpha_s, \beta_s, \epsilon, \delta_{\text{nAb}}^{\text{D614G/BA.1}}, o_{\Delta\text{nAb}}^{\text{waning/escape}}, \{\gamma_{c_i}\}, \{\gamma_{c_j}\}, \{\gamma_{c_k}\}$, we calculate the likelihood all configurations of transmission chains within each household based on equation (2). We use these configuration-specific likelihoods to resample transmission chains: For each household, we randomly sample one transmission tree among all transmission tree configurations with probability proportional to the transmission tree likelihood described in equation (2), given the parameters estimated by the most recent maximization step. We iterate through all households so that each household is assigned one simulated transmission tree. We repeat the process 1,000 times to obtain 1,000 realizations of the household transmission trees.

(4) For each fixed value of hyper-parameter τ over a plausible range (30–500 days), we iterate over the EM steps (2) and (3) until $L_m(\tau)$ converge to the maximum value of the EM algorithm. We scan through the values of τ from 30 to 500 days at 10-day steps. The EM algorithm convergence curve is shown in Supplementary Fig. 2 for each of the τ values. The EM algorithm converges at step 50, irrespective of the value of τ. The marginal likelihood of the model at τ, $L(\tau)$ is estimated by taking the average of $L_m(\tau)$ for EM steps 50 through 100. Supplementary Fig. 3 shows the log of the likelihood $L(\tau)$ as a function of τ, based on a spline interpolation. The point estimate of τ is taken from the maximum of $\log(L(\tau))$, while the 95% CI is estimated by finding τ values with log-likelihood value at the maximum minus 1.92 (Supplementary Fig. 3).

(5) We then take the best estimate of hyper-parameter τ and repeat the EM algorithm until convergence to estimate transmission model parameters $\alpha_s, \beta_s, \epsilon, \delta_{\text{nAb}}^{\text{D614G/BA.1}}, o_{\Delta\text{nAb}}^{\text{waning/escape}}, \{\gamma_{c_i}\}, \{\gamma_{c_j}\}, \{\gamma_{c_k}\}$ as shown in Table 2. The same EM algorithm was applied to models 2–4 for the sensitivity analysis as well.

The 'treatment effect' by prior infection is estimated by simulating from the best-fit model. We first sample 1,000 realizations of the imputed household transmission trees, with imputation probability proportional to the best estimates of the transmission model using the EM algorithm and hyper-parameter τ. For each of the 1,000 realizations, we focus on the subset of individuals who had one prior SARS-CoV-2 infection, denoted as $S_j = \{j | e_j = 1\}$). We use the fitted transmission model to predict the probability of infection (that is, $P_j = p_{ij}^k$ or q_j^k, with) of these non-naive subsets under three scenarios:

a. Scenario 1: the probability of infection estimated with predictors as reported in the data, denoted as P_j^{obs}.

b. Scenario 2: a counterfactual scenario of potential outcome where the probability of infection is estimated with predictor $e_j = 0$ (that is, a counterfactual naive individual), whereas all other covariates (confounders) are the same as observed, removing both direct and mediator effects. We denote the infection probability in this counterfactual scenario as $P_j^{\text{counterfactual}}(e_j = 0)$.

c. Scenario 3: a counterfactual scenario of potential outcome where the probability of infection is estimated with predictor $e_j = 1$, but setting $m_{\text{nAb}}^{\text{BA.1}} = 0$ (or $m_{\text{nAb}}^{\text{D614G}} = 0$), effectively removing the mediator effect of nAb on preventing transmission, but keeping the direct effect. We denote the infection probability in this counterfactual scenario as $P_j^{\text{counterfactual}}(e_j = 1; m_{\text{nAb}} = 0)$.

We then calculate the total protection conferred by prior infection as the population average of $P_j^{\text{counterfactual}}(e_j = 0)/P_j^{\text{obs}}$, based on bootstrap resampling with replacement (maintaining the same number of observations) of each of the 1,000 realizations of the household transmission chains. Point estimates and 95% CIs are based on the median and 95% quantiles of 1,000 realizations' estimates.

Similarly, we calculate the proportion of protection mediated by nAbs as the population average of $1 - \frac{P_j^{\text{counterfactual}}(e_j=1; m_{\text{nAb}}=0)/P_j^{\text{obs}}}{P_j^{\text{counterfactual}}(e_j=0)/P_j^{\text{obs}}}$. We use the same bootstrapping approach as that used for total protection.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Aggregate data to reproduce the figures are available at Zenodo via https://doi.org/10.5281/zenodo.11375487 (ref. 58). Individual-level data cannot be publicly shared because of ethical restrictions and the potential for identifying included individuals. Accessing individual participant data and a data dictionary defining each field in the dataset would require provision of protocol and ethics approval for the proposed use. To request individual participant data access, please submit a proposal to C.C. who will respond within 1 month of request. Upon approval, data can be made available through a data sharing agreement.

Code availability

Code to reproduce the figures, using Python version 3.8.11 and SciPy version 1.7.1, is available at Zenodo via https://doi.org/10.5281/zenodo.11375487 (ref. 58).

References

52. Wibmer, C. K. et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* https://doi.org/10.1038/s41591-021-01285-x (2021).
53. Netzl, A. et al. Analysis of SARS-CoV-2 Omicron neutralization data up to 2022-01-28. Preprint at *bioRxiv* https://doi.org/10.1101/2021.12.31.474032 (2023).
54. Cox, D. R. *Planning of Experiments* (Wiley, 1958).
55. Imbens, G. W. & Rubin, D. B. *Causal Inference in Statistics, Social, and Biomedical Sciences* (Cambridge University Press, 2015).
56. Longini, I. M. Jr & Koopman, J. S. Household and community transmission parameters from final distributions of infections in households. *Biometrics* **38**, 115–126 (1982).
57. Yang, Y., Longini, I. M. Jr, Halloran, M. E. & Obenchain, V. A hybrid EM and Monte Carlo EM algorithm and its application to analysis of transmission of infectious diseases. *Biometrics* **68**, 1238–1249 (2012).
58. Kaiyuan, S. Data and code for SARS-CoV-2 correlates of protection from infection against variants of concern. Zenodo https://doi.org/10.5281/zenodo.11375487 (2024).

Acknowledgements
We thank all the participants who kindly agreed to take part in the study, as well as the PHIRST-C group. We are grateful to B. J. Cowling and M. E. Halloran for their insightful feedback on the paper. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institutes of Health or the US Centers for Disease Control and Prevention. This work was supported by the National Institute for Communicable Diseases of the National Health Laboratory Service and the US Centers for Disease Control and Prevention (cooperative agreement no. 6 U01IP001048) and the Wellcome Trust (grant no. 221003/Z/20/Z) in collaboration with the Foreign, Commonwealth and Development Office, United Kingdom. P.L.M. and J.N.B. are supported by the Bill and Melinda Gates Foundation through the Global Immunology and Immune Sequencing for Epidemic Response (GIISER) program (INV-030570) and receive funding from the Wellcome Trust (226137/Z/22/Z). P.L.M. is supported by the South African Research Chairs Initiative of the Department of Science and Innovation and National Research Foundation of South Africa and the SA Medical Research Council SHIP program.

Author contributions
K.S., J.N.B., S.T., J.K., A.v.G., M.L.M., N.W., J.M., N.A.M., K.K., L.L., C.V., P.L.M. and C.C. designed the experiments. J.N.B., C.C., J.K., P.L.M. and S.T. accessed and verified the underlying data. J.N.B., S.T., J.K., V.S.M., Q.M., H.K., A.v.G., M.L.M., N.W., J.M., M.C., N.A.M., K.K., L.L., J.d.T., T.M., P.L.M. and C.C. collected the data and performed laboratory experiments. K.S., J.N.B., S.T., J.K., A.v.G., M.L.M., N.W., J.M., M.C., N.A.M., K.K., L.L., J.d.T., T.M., C.V. and C.C. analyzed the data and interpreted the results. K.S., J.N.B., C.V., P.L.M. and C.C. drafted the paper. All authors critically reviewed the paper. All authors had access to all the data reported in the study.

Competing interests
C.C. has received grant support from Sanofi Pasteur, the US Centers for Disease Control and Prevention, the Bill & Melinda Gates Foundation, the Taskforce for Global Health, the Wellcome Trust and the South African Medical Research Council. A.v.G. has received grant support from Sanofi Pasteur, Pfizer related to pneumococcal vaccine, the US Centers for Disease Control and Prevention and the Bill & Melinda Gates Foundation. N.W. reports grants from Sanofi Pasteur and the Bill & Melinda Gates Foundation. N.A.M. has received an institutional grant from Pfizer to conduct research in patients with pneumonia and from Roche to collect specimens to assess a novel tuberculosis assay. J.M. has received grant support from Sanofi Pasteur. The other authors declare no competing interests.

Ethics
The PHIRST-C protocol was approved by the University of Witwatersrand Human Research Ethics Committee (ref. 150808), and the US Centers for Disease Control and Prevention's Institutional Review Board relied on the local review (no. 6840). The protocol was registered on ClinicalTrials.gov on 6 August 2015 and updated on 30 December 2020 (NCT02519803). Participants receive grocery store vouchers of ZAR50 (USD 3) per visit to compensate for time required for specimen collection and interview. All participants provided written informed consent for study participation. For minors, consent was obtained from the parent or guardian.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41591-024-03131-2.

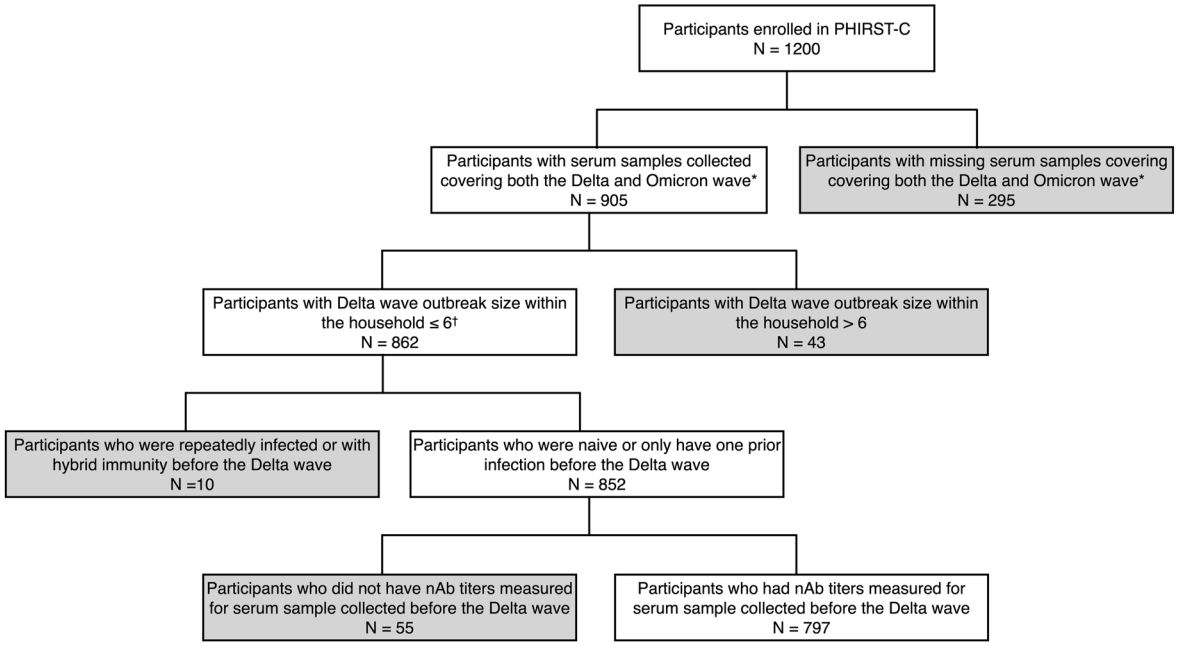
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-024-03131-2.

Correspondence and requests for materials should be addressed to Kaiyuan Sun or Cheryl Cohen.

Peer review information *Nature Medicine* thanks Sophie Valkenburg, Bo Zhang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Alison Farrell, in collaboration with the *Nature Medicine* team.

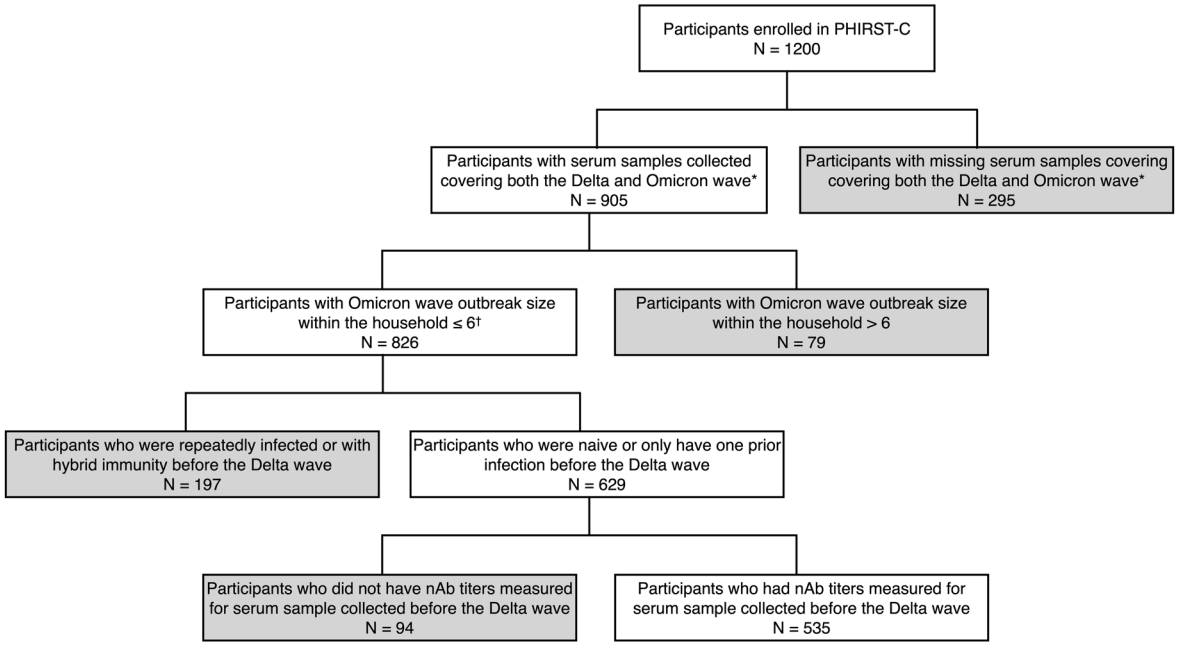
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41591-024-03131-2



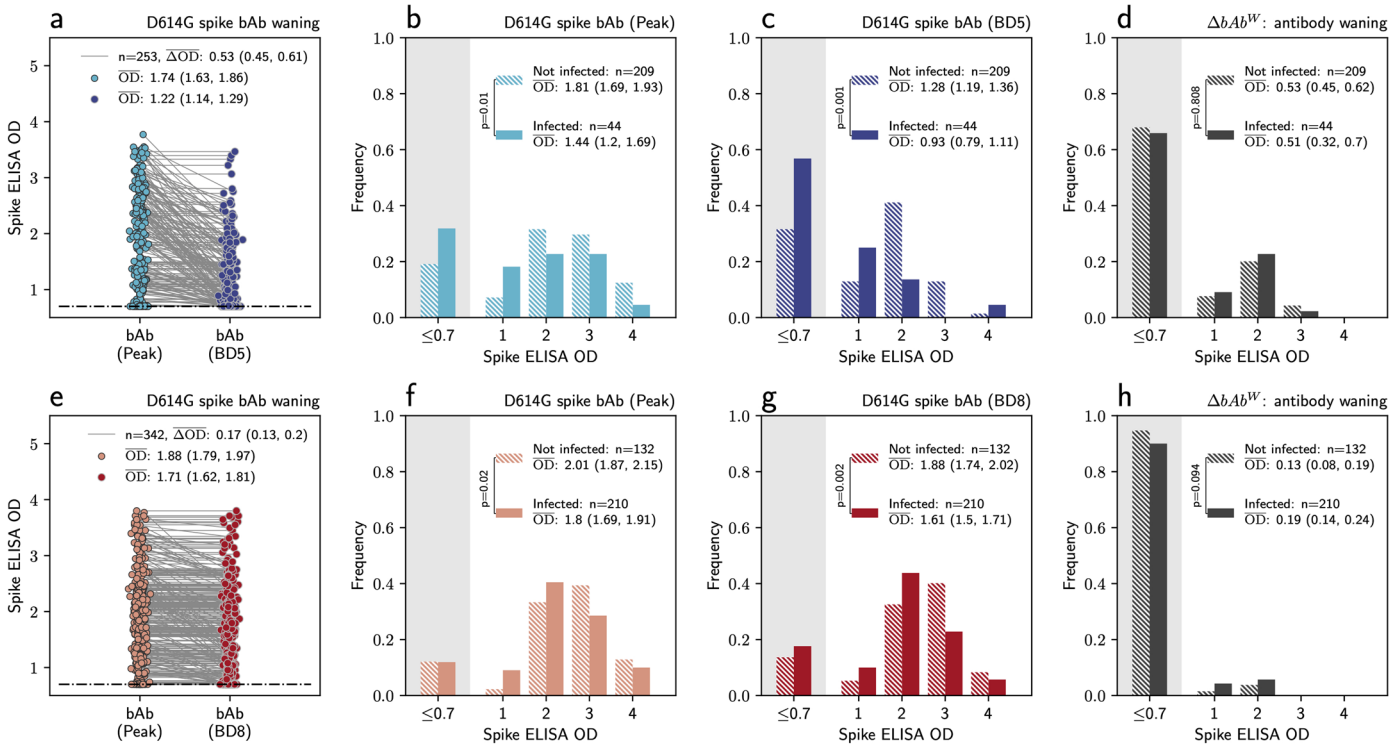
Extended Data Fig. 1 | Flowchart of participants included in the Delta-wave subgroup analysis. Grey boxes represent participants excluded from the Delta-wave subgroup analysis. *Based on a previously published study[31]. †Household with more than 6 infected individuals would be computationally intractable to track all possible transmission chain configurations (Methods Section 3).

https://doi.org/10.1038/s41591-024-03131-2



Extended Data Fig. 2 | Flowchart of participants included in the Omicron-wave subgroup analysis. Grey boxes represent participants excluded from the Omicron-wave subgroup analysis. *Based on a previously published study[31]. †Household with more than 6 infected individuals would be computationally intractable to track all possible transmission chain configurations (Methods Section 3).

https://doi.org/10.1038/s41591-024-03131-2



Extended Data Fig. 3 | D614G spike binding antibody (bAb) level for the Delta wave and the Omicron wave analysis. a, for Delta wave subgroup, the distribution of the peak bAb level to BD5 (light blue dots) and the D614G spike bAb level at BD5 (dark blue dots), among individuals who had one prior SARS-CoV-2 infection before blood draw 5. Each dot represents one individual, with two measurements of the same individual connected through a gray line. OD: absorbance at 450 nm, measured in optical density; $\overline{OD}$: the average of OD; $\overline{OD}$: the average drop of OD. **b**, for Delta wave subgroup, the distribution of the peak D614G spike bAb up to BD5, stratified by individuals who were infected during the Delta wave (solid bar) vs those who were not infected (dashed bar). Independent samples t-test (two-sided) is used to determine the statistical significance (anti reported on the legend) of difference between the $\overline{OD}$ of the two groups. **c**, same as **b** but for D614G spike bAb level at BD5. **d**, same as **b** but for ΔbAb^W. **e**, for Omicron wave subgroup, the distribution of the peak bAb level to BD8 (light red dots) and the D614G spike bAb level at BD8 (dark red dots), among individuals who had one prior SARS-CoV-2 infection before BD8. Each dot represents one individual, with two measurements of the same individual connected through a gray line. **f**, for the Omicron wave subgroup, the distribution of the D614G spike bAb level at BD8, stratified by individuals who were infected during the Omicron wave (solid bar) vs those who were not infected (dashed bar). Independent samples t-test (two-sided) is used to determine the statistical significance (p-value reported on the legend) of difference between the $\overline{OD}$ s of the two groups. **g**, same as **f** but for D614G spike bAb level at BD8. **h**, same as **f** but for ΔbAb^W.

https://doi.org/10.1038/s41591-024-03131-2

Extended Data Table 1 | Positivity rate of different serologic assays by the variant type of prior exposure for the Delta and Omicron wave subgroup

Delta wave subgroup			
Seropositivity	**Prior D614G infection**	**Prior Beta infection**	**Prior Delta exposure**
Anti-nucleocapsid assay were positive in at least one of the first 5 blood draws	109/113 (97%)	133/140 (95%)	–
Anti-nucleocapsid assay were positive at BD5	104/113 (92%)	129/140 (92%)	–
Anti-D614G nAb assay were positive for peak nAb response.	87/113 (77%)	60/140 (43%)	
Anti-D614G nAb were positive for nAb response at BD5	81/113 (72%)	59/140 (42%)	–
Omicron wave subgroup			
Seropositivity	**Prior D614G exposure**	**Prior Beta exposure**	**Prior Delta exposure**
Anti-nucleocapsid assay were positive in at least one of the first 8 blood draws	60/61 (98%)	116/120 (97%)	160/161 (99%)
Anti-nucleocapsid assay were positive at BD8	58/61 (95%)	108/120 (90%)	159/161 (99%)
Anti-D614G nAb were positive for nAb response at BD8	57/61 (93%)	71/120 (59%)	140/161 (87%)
Anti-BA.1 nAb were positive for nAb response at BD8	29/61 (48%)	36/120 (30%)	50/161 (31%)

https://doi.org/10.1038/s41591-024-03131-2

Extended Data Table 2 | Mediation analysis for nAbs as CoPs against serologically ascertained Delta and Omicron wave infections, with a variant-specific model for direct effect

Wave			Delta	Omicron
Protection against reinfection	**Direct effect** (Protection absent of nAbs)	**Prior D614G exposure** (odds ratio, absent of waning)	0.76 (0.36, 1.61)	1.23 (0.63, 2.38)
		Prior Beta exposure (odds ratio, absent of waning)	0.47 (0.30, 0.76)	0.78 (0.50, 1.21)
		Prior Delta exposure (odds ratio, absent of waning)	–	0.47 (0.29, 0.76)
	Mediators effect (Protection from nAbs)	**Anti-D614G nAb** (odds ratio, per 10-fold increase)	0.59 (0.43, 0.83)	–
		ΔnAb^{W} (odds ratio, per 10-fold increase)	1.00 (0.72, 1.39)	–
		Anti-BA.1 nAb (odds ratio, per 10-fold increase)	–	0.73 (0.56, 0.95)
		ΔnAb^{E} (odds ratio, per 10-fold increase)	–	0.94 (0.76, 1.15)
	Total protection (relative risk compared to naïve individuals)		0.40 (0.38, 0.42)	0.70 (0.68, 0.72)
	Proportion of protection mediated by nAbs		37% (34%, 40%)	11% (9%, 12%)
Protection against onward transmission (Odds ratio compared to naïve individuals)			0.20 (0.05, 0.72)	1.11 (0.62, 2.00)

Average and 95% CIs are provided for each of the model parameters. ΔnAb^{W}: the quantity of D614G nAbs waned from peak level to that at BD5. ΔnAb^{E}: the quantity of antibodies that can neutralize D614G but fail to neutralize Omicron BA.1 at BD8 due to Omicron's immune escape.

https://doi.org/10.1038/s41591-024-03131-2

Extended Data Table 3 | Mediation analysis for nAbs as CoPs against Delta (ascertained by both serology and PCR) and Omicron wave infections, with a waning model for direct effect

Wave			Delta	Omicron
Protection against reinfection	**Direct effect** (Protection absent of nAbs)	**Effect size** (odds ratio, absent of waning)	0.34 (0.17, 0.64)	0.29 (0.17, 0.51)
		Waning half-life (days)	128 (77, 261)	
	Mediators effect (Protection from nAbs)	**Anti-D614G nAb** (odds ratio, per 10-fold increase)	0.65 (0.49, 0.86)	–
		ΔnAb^W (odds ratio, per 10-fold increase)	1.02 (0.78, 1.36)	–
		Anti-Omicron BA.1 nAb (odds ratio, per 10-fold increase)	–	0.73 (0.56, 0.95)
		ΔnAb^E (odds ratio, per 10-fold increase)	–	1.01 (0.84, 1.21)
	Total protection (relative risk compared to naïve individuals)		0.41 (0.40, 0.43)	0.62 (0.61, 0.64)
	Proportion of protection mediated by nAbs		33% (30%, 35%)	11% (9%, 12%)
Protection against onward transmission (Odds ratio compared to naïve individuals)			0.23 (0.08, 0.71)	1.19 (0.66, 2.13)

Average and 95% CIs are provided for each of the model parameters. ΔnAb^W: the quantity of D614G nAbs waned from peak level to that at BD5. ΔnAb^E: the quantity of antibodies that can neutralize D614G but fail to neutralize Omicron BA.1 at BD8 due to Omicron's immune escape.

https://doi.org/10.1038/s41591-024-03131-2

Extended Data Table 4 | Mediation analysis for D614G spike binding antibody as CoPs against serologically ascertained Delta and Omicron wave infections, with a variant-specific model for direct effect

Protection against reinfection			Delta (serology)	Omicron (serology)
	Direct effect (Protection absent of nAbs)	**Prior D614G exposure** (odds ratio, absent of waning)	0.60 (0.24, 1.48)	1.38 (0.67, 2.84)
		Prior Beta exposure (odds ratio, absent of waning)	0.51 (0.32, 0.83)	0.91 (0.58, 1.45)
		Prior Delta exposure (odds ratio, absent of waning)	–	0.61 (0.38, 0.97)
	Mediators effect (Protection from nAbs)	**D614G binding Ab** (odds ratio, per 10-unit increase)	0.26 (0.12, 0.59)	0.60 (0.46, 0.77)
		ΔbAb^W (odds ratio, per 10-unit increase)	1.02 (0.55, 1.91)	1.02 (0.55, 1.87)
	Total protection (relative risk compared to naïve individuals)		0.40 (0.38, 0.42)	0.65 (0.63, 0.67)
	Proportion of protection mediated by spike binding Ab		35% (32%, 38%)	27% (25%, 29%)
Protection against onward transmission (Odds ratio compared to naïve individuals)			0.22 (0.06, 0.74)	1.18 (0.65, 2.13)

Average and 95% CIs are provided for each of the model parameters. ΔbAb^W: the quantity of D614G spike binding antibodies waned from peak level to that at BD5 for Delta (at BD8 for Omicron).

nature portfolio

Corresponding author(s): Kaiyuan Sun, Cheryl Cohen

Last updated by author(s): 05/28/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were entered during visits on tablet computers with use of the Research Electronic Data Capture application (REDCap).
Data analysis	Software packages used: python version 3.8.11; scipy version 1.7.1; R version 4.1.2; R package "brglm" version 0.7.2; R package "mice" version 3.16.0. Code to reproduce the figures will be made available at a public repository once the paper is accepted and published.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data availability

Aggregate data to reproduce the figures are available at Zenodo (DOI: 10.5281/zenodo.11375487). Individual-level data cannot be publicly shared because of ethical restrictions and the potential for identifying included individuals. Accessing individual participant data and a data dictionary defining each field in the dataset

would require provision of protocol and ethics approval for the proposed use. To request individual participant data access, please submit a proposal to C.C. who will respond within 1 month of request. Upon approval, data can be made available through a data sharing agreement.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Information on sex were collected based on self-reporting.
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	Participant characteristics are summarized in Table 1 of the manuscript.
Recruitment	Households were randomly selected, from the HDSS database for the rural site and with use of GPS coordinates for the urban site (appendix p 1 of https://doi.org/10.5281/zenodo.7260083). Households with more than two members and where at least 75% of members consented to participate were eligible. In brief, we first approached households previously enrolled in PHIRST, and then prospectively approached new potentially eligible households using the site-specific sampling frame used for PHIRST until the required number of households were enrolled. For households that failed to recruit all members the infection status of household members that were not recruited remained unknown, which could underestimate the force-of-infection within the household setting for the transmission model. Exclusion of household with <2 members could have biased estimates of attack rates.
Ethics oversight	The PHIRST-C protocol was approved by the University of Witwatersrand Human Research Ethics Committee (Reference 150808) and the U.S. Centers for Disease Control and Prevention's Institutional Review Board relied on the local review (#6840). Informed consent was obtained from all adult participants (aged ≥18 years), assent from children aged 7 to 17 years, and consent from a parent or guardian for children younger than 18 years before data collection. Participants receive grocery store vouchers of ZAR50 (USD 3) per visit to compensate for time required for specimen collection and interview.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	NA (The original sample size calculation for the primary endpoints of the PHIRST-C cohort is described at https://doi.org/10.1016/S1473-3099(22)00069-X, this study represents analysis of the cohort data that was not planned during the original study design)
Data exclusions	The analysis was performed on a subset of 797 for the Delta wave analysis and 535 for the Omicron wave analysis of the total 1,200 cohort participants, detail inclusion criteria are described in the manuscript.
Replication	NA (This is a prospective household observational study during the COVID-19 pandemic, circulating variants and population immunity have fundamentally changed throughout the study thus it is impossible to replicate the real-world epidemiological situations in a similar population)
Randomization	NA (covariates were controlled by fitting a multi-variable chain-binomial model
Blinding	NA (This is a prospective household observational study, no blinding performed)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: monoclonal antibodies (mAbs) CR3022 (source: produced by SAMRC AIRU) were used as the positive control. Palivizumab (source: Medimmune; Identifier: Synagis; RRID: AB_2459638) was used as negative control.

Validation: All antibodies were validated for the species and application by the relevant manufacturer and in a previous study: https://doi.org/10.1016/j.xcrm.2022.100510

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): 293T/ACE2.MF cells modified to overexpress human ACE2 (kindly provided by M. Farzan (Scripps Research)), 293T/17 cells purchased from ATCC

Authentication: Cell lines were authenticated by the providers

Mycoplasma contamination: All cell lines tested negative for mycoplasma

Commonly misidentified lines (See ICLAC register): None used in this study

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration: NCT02519803; NCT05277298

Study protocol: https://clinicaltrials.gov/ct2/show/NCT02519803; https://clinicaltrials.gov/ct2/show/NCT05277298

Data collection: The study aims to characterize the community burden (including the clinical features) and transmissibility of SARS-CoV-2 within the context of a functional antibody response. In addition,the study will assess the effect of the interaction of SARS-CoV-2 with influenza virus and RSV on disease severity and transmission dynamics. A household-level prospective cohort study will be conducted in one rural and one urban community located in Mpumalanga Province and North West Province, respectively. The study will be conducted for 12 months of intensive follow up (July 2020 to August 2021) with a post-intensive follow-up continuing for a further 16 months (until December 2022). Two hundred households; 1,000 study participants of all ages; will be randomly selected from a list of 327 hoseholds that participated and successfully completed a 10-months follow-up period in a study similar to that currently proposed, but directed at community burden and transmission dynamics of influenza, respiratory syncytial virus and other respiratory pathogens. Each household and household member will be enumerated and the HIV infection status and the level of immunosuppression of HIV-infected individuals will be assessed. Each household member will be followed twice per week during the intense follow-up period (12 months) of the study. During this period upper respiratory tract samples will be collected irrespective of presence of symptoms and data on key symptoms, healthcare seeking, hospitalization and death will be captured at each follow up visit. Respiratory samples will be tested by reverse transcriptase real-time polymerase chain reaction (rRT-PCR) for SARS-CoV-2, influenza and RSV, and selected samples will be cultured and sequenced. An infection risk questionnaire will be administered to all study participants at enrollment and every month thereafter. Sera will be collected at enrollment and every 2 months during the 12-month intense follow-up period from all participants. In addition, sera will be collected every 2 months for a further 6 months following the 12-month intense follow-up period from study participants that tested positive for SARS-CoV-2 by rRT-PCR on respiratory specimens at 14, 16 and 18 months and from all study participants at 18 months. Sera will be tested for the presence of SARS-CoV-2, influenza and RSV antibodies. Wearable proximity sensors will be deployed for 8-12 days in each household over the 6-month intense follow-up period.

Outcomes: Primary outcomes (see also at https://clinicaltrials.gov/study/NCT05277298):
• Outcome Measure: Describe the symptomatic fraction of SARS-CoV-2 infections among individuals in household cohort in a rural and urban setting, South Africa 2020/2021. [Time Frame: 14 months]. Measure Description: The proportion of SARS-CoV2 infection that are symptomatic. This will be stratified by age, underlying conditions, HIV infection.
• Outcome Measure: Describe the household secondary infection risk of SARS-CoV-2 infection among individuals in a household

cohort in a rural and urban setting, South Africa 2020/2021. [Time Frame: 14 months]. Measure Description: The number of secondary cases in a household divided by the total number of susceptible gives the secondary attack risk during the 14months of follow-up. This will be explored by underlying conditions, age and HIV status

• Outcome Measure: Describe the serial interval for SARS-CoV-2 in households over 14 months of follow-up in an urban and rural setting, South Africa 2020/2021 [Time Frame: 14 months]. Measure Description: Calculate the time between successive cases in each household as documented by onset of symptoms and/or PCR positive test.

• Outcome Measure: Describe the duration of shedding of SARS-CoV-2 in a household cohort in a rural and urban setting, South Africa 2020/2021. [Time Frame: 14 months]. Measure Description: The duration of PCR positive SARS-CoV-2 infection in individual over 14 months of follow-up, including describing the interval by age, HIV status and underlying conditions.

• Outcome Measure: Describe the incidence of infection by PCR and serology in a household cohort in a rural and urban setting, South Africa 2020/21 [Time Frame: 14 months]. Measure Description: The number of new cases of symptomatic illness recorded by symptoms reported cohort over the 14 months of follow-up. Including estimating a person time to follow-up incident rate

• Outcome Measure: Describe the incidence illness by PCR and/or serology in a household cohort in a rural and urban setting, South Africa 2020/21 [Time Frame: 14 months]. Measure Description: Describe the incidence of infection by PCR and serology in a household cohort over 14 months of follow-up, in a rural and urban setting, South Africa 2020/21

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA