This is an Accepted Manuscript for Epidemiology & Infection. Subject to change during the editing and production process.

DOI: 10.1017/S095026882500010X

Using SARS-CoV-2 nucleoprotein antibodies to detect (re)infection

C.E. Hoeve, N. Neppelenbroek, E.R.A. Vos, A.J. Huiberts, S.P. Andeweg, G. den Hartog, Robert van Binnendijk, Hester de Melker, S. van den Hof, Mirjam Knol

Affiliations: National Institute for Public Health and the Environment, Center for Infectious Disease Control, the Netherlands

Corresponding author:

MJ Knol, mirjam.knol@rivm.nl

Abstract

We assessed the validity of serum total anti-Nucleoprotein Immunoglobulin (N-antibodies) to identify SARS-CoV-2 (re)infections by estimating persistence of N-antibody seropositivity and boosting following infection. From a prospective Dutch cohort study (VASCO), we included adult participants with ≥2 consecutive self-collected serum samples, 4-8 months apart, between May 2021-May 2023. Sample pairs were stratified by N-seropositivity of the first sample and by self-reported infection within the sampling interval. We calculated proportions of participants with N-seroconversion and fold-increase (1.5, 2, 3, 4) of N-antibody concentration over time since infection and explored determinants. We included 67,632 sample pairs. Pairs with a seronegative first sample (70%) showed 89% N-seroconversion after reported infection and 11% when no infection was reported. In pairs with a seropositive first sample (30%), 82%-65% showed a 1.5- to 4-fold increase with a reported reinfection, and 19%-10% without a reported reinfection, respectively. After one year, 83% remained N-seropositive post-first infection and 93%-61% showed 1.5-fold to 4-fold

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

increase post-reinfection. Odds for seroconversion/fold increase were higher for symptomatic infections and Omicron infections. In the current era with limited antigen or PCR testing, N-serology can be validly used to detect SARS-CoV-2 (re)infections at least up to a year after infection, supporting the monitoring of COVID-19 burden and vaccine effectiveness.

Introduction

Monitoring SARS-CoV-2 infections in the endemic phase is important to estimate the incidence in the population, e.g. to identify risk groups (1)or to estimate real-world vaccine effectiveness of COVID-19 vaccines (2-4). At the start of the pandemic wide-scale community testing was available to monitor the incidence of infections. In the Netherlands, SARS-CoV-2 community testing facilities have been scaled down and eventually closed in March 2023. In addition, the commitment to self-testing declined. As wide-scale testing is no longer available, alternative methods are necessary to keep monitoring SARS-CoV-2 infections.

One method to detect SARS-CoV-2 infections is through detecting antibodies induced by infection but not by vaccination. COVID-19 vaccines authorized in the Netherlands induce Spike S1-antibodies following vaccination, but these are also induced by infection. Spike S1 antibodies are therefore less suitable to detect SARS-CoV-2 infections in a highly vaccinated population, leaving antibodies against the Nucleoprotein (N), one of the structural proteins of SARS-CoV-2, a more specific marker to identify (re)infection (5). While a relative increase in antibodies can be used as a marker for reinfection (6-8), it is currently unknown how long N-antibodies persist after infection and which increase in N-antibodies can reliably detect reinfection.

The aim of this study is to describe the persistence of N-antibody seropositivity and boosting after infection, to estimate the most suitable fold increase to detect reinfection and whether there are factors that affect the sensitivity of detecting (re)infection.

Methods

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Study design and population

VASCO (VAccine Study COvid-19) is an ongoing 5-year prospective cohort study and has included approximately 45,000 community-dwelling participants aged 18-85 years in the Netherlands (9). The study started in May 2021, a few months after COVID-19 vaccines were introduced in the Netherlands, and the primary objective is to assess the real-world vaccine effectiveness against SARS-CoV-2 infection. Participants are asked to take a self-collected fingerprick blood sample every 6 months, and one month after primary vaccination, for detection of serum antibodies. In addition, participants are asked to complete monthly digital questionnaires in the first year and three-monthly in years 2-5, including questions on sociodemographic factors, health status, COVID-19 vaccination, SARS-CoV-2 related symptoms, positive SARS-CoV-2 tests (PCR or (self-administered) rapid antigen test), and willingness to test when having symptoms. Participants can also notify positive SARS-CoV-2 tests or COVID-19 vaccinations in the study app at any moment. After April 2022, regular testing has been scaled down in the Netherlands. Since then, participants receive self-tests free of charge and are encouraged to test when having COVID-19-like symptoms.

Antibody measurements

Self-collected fingerprick blood samples were collected in 0.5 µl Minicollect tubes (Greiner, #450533) and returned in a pre-printed and addressed safety-bag envelope, and centrifuged immediately upon arrival. Serum is subsequently separated, aliquoted and stored at -80°C. Serum samples were analyzed for total immunoglobulin (Ig) levels against N-antibodies on the Cobas e801 (Roche Diagnostics, Mannheim, Germany) using batch-specific, linear calibration-lines obtained with a dilution range of the NIBSC 20/136 WHO standard (NIBSC) or an internal pool of 125 N-antibody positive, anonymized patient sera calibrated against the WHO standard. The cut-off for N-positivity ranged from 2.05 to 3.79 BAU/mL between batches(10). The clinical sensitivity of the Roche assay is 99.5% (95% confidence interval: 97.0% – 100%) and the clinical specificity is 99.80 % (99.69% - 99.88%)(11).

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

To evaluate reproducibility, a total of 278 samples, ranging from 0.4 BAU/mL to 477 BAU/mL, were measured twice using different assay batches. Concordance of seropositivity was 99% (275/278). To evaluate reproducibility between assay batches, we selected all twice measured seropositive samples (201/278). Proportions of samples with a 1.5-, 2-, 3- and 4-fold difference between measurements were calculated for these samples. Of the 201 samples, 6 (3.0%) samples, ranging from 12 BAU/mL to 107 BAU/mL, had a difference in N-antibody levels of at least 1.5-fold, and 1 (0.5%) of 2-fold (12 BAU/mL). There were no samples with a difference in N-antibody levels of 3-fold or more between measurements. The coefficient of variation ranged between 0.01% and 20.4%.

Determinants

The following potential determinants were considered: age group (18-59 years vs. 60-85 years), vaccination status (unvaccinated, partly vaccinated [one primary series dose + 7 days], primary series [two primary series doses + 14 days, or one dose JCovden + 28 days], first booster, second booster, third booster [booster doses + 7 days]), calendar time (quarters), and log N-antibody concentration of first sample (continuous in BAU/mL). In case of a reported infection between the first and second sample the following potential determinants were also included: COVID-19 related symptoms (yes, no, unknown), severity of infection (local, systemic, other, not reported).

For each sample pair, vaccination status was determined at the sampling date of the second blood sample, as described before (10). Calendar time was determined by the sampling date of the second blood sample. Occurrence and type of COVID-19 symptoms were collected if participants reported a new positive test in the study app and in monthly follow-up questionnaires after infection. Infections were defined as systemic when at least one of the following symptoms was reported: fever, general malaise, extreme fatigue, joint pain, muscle pain, irritability or confusion, nausea or vomiting, diarrhea, stomach pain, pain while breathing, and shortness of breath. Infections were defined as

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

local if at least one of the following symptoms was reported without any of the systemic symptoms present: cough, sore throat, runny nose, loss of smell and/or taste and headache.

Data analysis

First, N-antibody levels 200 days pre- and post-infection were explored and visualized using a generalized additive model for first, second and third reported infections separately. Models were fitted with time since infection as a base-spline with 6 knots. The most favorable number of knots and knot positions were determined for each subsequent infection by AIC value of the models.

Then, we included participants with at least two blood samples with an interval of 4-8 months between May 2021 and May 2023 for which antibody assessment was available. Subsequently, the data was organized and analyzed in sample pairs consisting of two consecutive samples of the same participant. A participant could contribute more than one sample pair if the participant had submitted more than two samples, i.e. the second sample of the first sample pair could serve as a first sample in the second sample pair (Figure 1, panel A).

All sample pairs were stratified by N-seropositivity of the first sample. Subsequently, the sample pairs were stratified by whether an infection was reported within the sampling interval. In case an infection was reported, time since infection was calculated as the time between the infection and the second sample. In case of a reported infection, the third sample, if available, was added to the sample pair, provided that between the second and third sample no infection was reported and no increase in N-levels of >=2 fold was observed (Figure 1, panel B).

For sample pairs where the first sample was seronegative, we calculated the percentage and 95% confidence interval (CI) of N-seroconversion overall and by categories of the determinants, stratified by reported infection within the sample pair interval. To estimate the association between potential determinants and N-seroconversion we performed univariable and multivariable logistic regression using generalized estimating equation (GEE) models with exchangeable correlation structure to

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

account for dependencies within participants. Odds ratios with 95% CI and p-values were provided. Variables included in the univariable and multivariable models were age group (18-59 years, 60-85 years), vaccination status, time period (year-quarter), and severity of infection (only for sample pairs with a reported infection in the sampling interval). A sensitivity analysis was performed on participants who reported to (almost) always test in case of symptoms. For sample pairs with a reported infection the duration of seropositivity after infection was assessed by calculating the proportion (and 95% CI) of seropositive samples over time.

For sample pairs where a first sample was seropositive, the percentage of sample pairs with a 1.5-, 2-, 3- and 4-fold was calculated overall and by categories of the determinants, stratified by reported infection within the sample pair interval. The fold increase levels were chosen arbitrarily to explore different relative increase measures as a marker for unreported reinfections. A univariable and multivariable GEE model with exchangeable correlation structure were used to estimate the association between potential determinants and n-fold increase. Odds ratios with 95% CI and p-values were provided. Variables included in the univariable models were age group, vaccination status, time period, concentration of first sample in the sample pair, occurrence of symptoms, and severity of infection (only for sample pairs with a reported infection in the sampling interval). In the multivariable models the same variables were included except occurrence of symptoms due to collinearity with the variable severity of infection. A sensitivity analysis was performed on participants who reported to (almost) always test in case of symptoms. For sample pairs with a reported infection, the fold increase over time was assessed by calculating the percentage (and 95% CI) of samples with a 1.5-, 2-, 3- and 4-fold increase relative to the first sample of the sample pair.

All analysis were performed using R version 4.4.1, including the tidyverse, geepack and mgcv packages (12-15).

Ethics

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

The VASCO study is conducted in accordance with the principles of the Declaration of Helsinki and the study protocol was approved by the not-for-profit independent Medical Ethics Committee of the Stichting Beoordeling Ethiek Biomedisch Onderzoek (BEBO), Assen, the Netherlands (NL76815.056.21). VASCO was registered in the online Dutch clinical trials register (trialregister.nl, registration number NL9279). Written informed consent was obtained from all participants prior to enrollment into the study.

Results

N-antibody concentrations following reported infection

Between May 2021 and May 2023 a total of 131,791 samples were collected from 44,407 participants. N-antibody levels 200 days before and after the first three reported infections are plotted in Figure 2. N-antibody levels post-infection showed a high degree of variation between participants (Figure 2). Overall, after a first reported infection N-antibody levels showed an increase during the first weeks after which a geometric mean concentration (GMC) of 39 BAU/mL was observed. This was followed by a gradual decline to 30 BAU/mL 200 days post-infection. After a second reported infection we observed a peak GMC of 162 BAU/mL, after which N-antibody levels declined to 91 BAU/mL 200 days post-infection. After a third reported infection N-antibody levels reached a peak GMC of 215 BAU/mL, which then declined to 150 BAU/mL 200 days post-infection.

Selection of sample pairs

In total 67,632 sample pairs of 33,283 participants were included. 8,627 participants contributed one sample pair, 14,963 participants two sample pairs and 9,693 three sample pairs. Figure 3 shows the number of sample pairs by seropositivity of the first sample. An infection was reported in the sampling interval for 31% of sample pairs with the first sample N-seronegative and for 14% of sample pairs with the first sample N-seropositive.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Among sample pairs where the first sample was seronegative, the median N-antibody level of the second samples was 32.6 BAU/mL for sample pairs with a reported infection in the sampling interval and 1.3 BAU/mL for sample pairs without a reported infection. Additionally, the distribution of N-antibody levels was visually distinct between the two groups (Figure 4, panel A). Among sample pairs where the first sample was seropositive the median N-antibody level of the second samples was 237.4 BAU/mL for sample pairs with a reported infection in the sampling interval and 30.6 BAU/mL for sample pairs without a reported infection (Figure 4, panel B).

First infections

Among the sample pairs with a seronegative first sample, 14,685 infections were reported and an additional 3,495 sample pairs for which no infection was reported showed seroconversion. Since the specificity of the Roche antibody test for N-seropositivity is 99.5%, this indicates that 19% of first infections would be missed in this study population when only considering reported infections.

Among participants who reported an infection with a seronegative first sample, the overall seroconversion rate was 89%. The seroconversion rate varied around 90% for the different age groups, and levels of vaccination status and severity of infection (Figure 5). For samples collected during Q4 of 2021 and Q1 of 2022 (mostly Delta infections) the N-seroconversion rate was significantly lower at 71% compared to ~90% in samples taken later (mostly Omicron infections) ($p<0.01$). The N-seroconversion rate was significantly higher for infections with local and systemic symptoms (both 90%) than for asymptomatic infections (81%, $p < 0.001$). In the multivariable model, N-seroconversion was more likely for participants over 60 years compared to those under 60 years (OR: 1.25 [1.10-1.41]), unvaccinated participants compared to vaccinated participants (e.g. 0.47 [0.30-0.73] for primary series), and for local infections (1.92 [1.40-2.65]) and systemic infections (1.81 [1.34-2.45]) compared to asymptomatic infections (Supplementary table S1A). The odds of N-seroconversion increased with calendar-time (8.04 [5.41-11.93] in 2022 Q2 to 24.39 [6.65-89.45] in 2023 Q2, compared to 2021 Q4).

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

The overall N-seroconversion rate was 11% among participants with a seronegative first sample, who did not report an infection in the sample interval. The N-seroconversion rate was significantly higher among participants aged 18-59 years (12%) compared to 60-85 years (10%, p<0.001). The N-seroconversion rate was 18% for unvaccinated participants and was significantly lower for vaccinated participants (4-16%, p<0.05). N-seroconversion rates increased significantly with calendar-time from 1% in 2021 Q4 to 26% in 2023 Q2 (p<0.001). In the multivariable analysis only vaccination status and calendar time remained significant factors (Supplementary table S1A).

N-seroconversion increased to 96% 8-9 weeks after a reported infection and started declining after 20-29 weeks (Figure 6). At 50-59 weeks after infection the N-seroconversion rate was still 83%.

Reinfections

In sample pairs where the first sample was seropositive, the median fold increase in N-antibody levels was 7.6 (IQR: 2.3-22.9) for sample pairs with a reported infection in the sample interval (n= 2934) and 0.6 (IQR: 0.4 – 1.1) for sample pairs without reported infection (n = 17,512).

Among the sample pairs with a seropositive first sample, a total of 2,934 infections were reported and an additional 3,297 sample pairs for which no infection was reported showed a 1.5-fold increase, 2,676 a 2-fold increase, 2,040 a 3-fold increase, and 1,723 a 4-fold increase. Using a 1.5, 2-, 3-, or 4-fold increase suggests that, respectively, 53%, 48%, 41%, 37% of infections are missed when only using reported infections for detection.

Figure 7 shows the percentage of sample pairs with a 1.5-, 2-, 3- or 4-fold increase in N-antibody level by determinants for sample pairs with and without a reported infection between samples. Among sample pairs with a reported infection in the sampling interval 82% showed a 1.5-fold increase, 77% a 2-fold increase, 71% a 3-fold increase and 65% a 4-fold increase. A fold increase was more likely among vaccinated participants, following local and systemic infections, following infections after the first quarter of 2022 and less likely with increasing antibody concentrations of

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

the first sample of a sample pair (Supplementary tables S1B-S1E). Age group did not affect fold increase. In the multivariable model most of these effects were sustained, but there was no significant association between vaccination status and fold increase.

Among sample pairs without a reported infection in the sampling interval the proportion of sample pairs with a 1.5-, 2-, 3- and 4-fold increase were respectively 19%, 15%, 12% and 10%. Among these sample pairs a fold increase was more likely in the 18–59 year age group, in unvaccinated participants, for infections after 2022 Q1, and less likely with increasing antibody concentrations of the first sample of a sample pair. These results were maintained in the multivariable model (Supplementary tables S1B-S1E).

The percentage of sample pairs with a 1.5-fold increase rose from 57% 0-1 weeks after infection to 89% 30-39 weeks after infection (Figure 8). A similar pattern is seen for the 2-, 3- and 4-fold increase in N-antibody levels, but with lower proportions with a higher fold increase. The 3-fold and 4-fold increase in N-level appeared to decline 20-29 weeks after infection.

Sensitivity analysis

A sensitivity analysis on participants who report to (almost) always test in case of symptoms revealed slightly lower proportions (approximately 2 percent points) of participants with seroconversion or fold increase among those that did not report an infection, particularly in the 18-59 year age group (Supplementary tables S2A-S2E).

Discussion

In this study we aimed to describe the persistence of N-antibody seropositivity after infection and to evaluate the sensitivity of different fold increases in order to detect reinfection(s) and whether this differed by several potential determinants. We found that seroconversion rate to detect primary infections (using reported positive tests as a gold standard) is 89% and using fold-increases to detect reinfections provided a sensitivity of 82% for 1.5-fold to 65% for 4-fold. We found that in our cohort,

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

by not using serology data but only reported infections, we would have missed 19% of primary infections and 37%-53% of reinfections depending on the fold increase used. Among participants with a reported infection, seroconversion, 1.5-fold and 2-fold increases were maintained in more than 80% of participants, 3-fold increases in 73% of participants and 4-fold increases in 67% after 6 months. This allows detection of infections well for at least half a year after the occurrence of infection, making the method suitable for sampling intervals such as in the VASCO study (a 6 month interval).

We observed several determinants that affected seroconversion/fold increase. First, we observed that seroconversion/fold increase was more likely among participants with symptomatic infections. Higher antibody responses can be expected with more severe infections, resulting in a higher chance of seroconversion or a fold increase. Since widescale testing has discontinued it is likely that participants only test when being symptomatic. This is supported by our observation that the proportion with a fold increase among those without a test increases over time. Furthermore, our sensitivity analysis among participants who (almost) always test when having symptoms showed marginal differences in proportions for samples pairs without reported infections, suggesting that unreported infections are indeed mostly asymptomatic. Second, higher seroconversion rates were seen among unvaccinated compared to vaccinated participants. After (multiple) doses, vaccinated individuals usually have higher levels of antibodies against the Spike protein (16). As a consequence, infections may be cleared earlier before substantial levels of N-antibodies are induced, resulting in less fold increase and a lower rate of seroconversion in vaccinated participants and lower sensitivity of the assay, similar to what was shown by Bazin et al (17). Furthermore, due to absence of protective antibodies unvaccinated participants may experience more severe infections and therefore induce higher antibody concentrations, resulting in higher seroconversion rates (5, 18). Fold increase rates were not significantly different between vaccinated and unvaccinated participants, suggesting that vaccination status does not affect detection of reinfections. Third, among sample pairs without reported infections, a fold increase was more likely among the younger

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

age group (18-59 years). Participants in the younger age group are expected to experience less symptoms when infected and are less likely to test when having symptoms. This is supported by our findings in the sensitivity analysis among participants who (almost) always test when having symptoms, where the age effect was no longer present. In contrast, seroconversion among sample pairs with a reported test was more likely among the older age group. The older age group possibly experienced more severe first infections compared with the younger age group, which may have led to enhanced antibody induction in the older age group. While we adjust for symptoms after infections, we may not be able to fully adjust for severity as categorization is only based on type of symptom (local vs systemic) and does not include a measure for seriousness of the symptoms. Fourth, we observed lower rates of seroconversion or fold increase among those with and without a reported infection during the Delta prevalent period. The Omicron variant has deviated more from the vaccine strain than the Delta variant. This may in turn result in a less adequate immune response among vaccinated participants during an Omicron infection than during a Delta infection due to immune escape. This in turn might result in more severe symptoms and thus higher rates of seroconversion following Omicron infections. Finally, we observed that a fold increase is less likely when the N-antibody levels in the first sample were already high. Recent research has shown that higher serum N-antibody levels are associated with higher mucosal immunity, thus reducing the risk of reinfection (19). However, in addition to observed boosting of N-antibody levels following each additional infection (at least up to three infections), we also observed a ceiling effect. The higher the N-antibody levels in the first sample, the lower the chance that a fold increase will be found as high levels cannot always accurately be measured due to saturation of the assay. This limits the possibility to use the fold increase as a proxy for reinfection when multiple infections occur in a relatively short period of time. Overall, using serology in addition to reported positive tests in population-based studies will help identify more infections, specifically in unvaccinated participants and younger age groups.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

294 A 4-fold increase in antibody concentrations is considered a gold standard for various pathogens, but
295 has been posed as too strict for population studies and resulting in under detection of reinfections
296 (6, 7). Traditionally, a 2-fold increase was considered to carry a risk of measurement error causing
297 false-positives (6). False-positives may be caused by differences in sample concentrations by other
298 factors than infection such as variation of sample distribution and storage conditions, or variations in
299 lab measurement. This may cause small variations in concentrations that may be picked up as a fold
300 increase (e.g. a difference of 1.5 BAU/mL between a sample with 3 BAU/mL and 4.5 BAU/mL is a 1.5-
301 fold increase, but not necessarily due to an infection). We found 3% false-positives when using a 1.5-
302 fold increase and 0.5% false-positives when using a 2-fold increase in samples measured in duplicate.
303 The extend of other sources of false-positives, such as conditions during transport of samples, could
304 not be measured in the current study. Where false-positives might not be wanted in a clinical
305 setting, in the context of cohort studies such as the VASCO study, a small proportion of false-
306 positives may be more acceptable. For example, if one would like to exclude participants with a
307 recent infection, this would result in the inappropriate exclusion of only a small group, but
308 simultaneously identify a large number of true positives.

309 Not all infections can be identified using seroconversion or fold increases. One factor that affects the
310 possibility of detection is the time of sampling in relation to the moment of infection. N-antibody
311 levels are still increasing during the first 4-6 weeks after infection and then decrease over time. For
312 example, if the infection occurred just before the second sample, an infection could be missed. For
313 reinfections, if the first sample is shortly after a prior infection or the second sample is too long after
314 the reinfection, a fold increase may not be observed. We showed that fold increases remained high
315 for at least half a year, making the fold increase method suitable for serial sampling with an interval
316 of 6 months and possibly longer. Other approaches have been suggested, for example to only look at
317 fold increases after a downward trend in antibodies has been observed (20). However, this approach
318 requires sampling with smaller intervals as one would need to identify the downward trend before
319 the reinfection occurs.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Recommendation

When choosing the most suitable fold increase to detect reinfections a trade-off needs to be made between having more false positives with a lower fold increase or having less true positives with a higher fold increase. In our study we gained 82% true positives and 3% false positives when using a 1.5-fold increase, whereas we gained 77% true positives and 0.5% false positives when using a 2-fold increase, thus decreasing the proportion of false positives to near-zero. Using a 3-fold increase the proportion of true positives decreased to 71% whereas only a marginal reduction in false positives was found compared to using a 2-fold increase (0%). Furthermore, due to increasing antibody levels and the assay limit of detection, a higher fold increase is hard to maintain over time as larger fold increases become less practical to measure for commercial kits. At this time a 2-fold increase therefore appears most favorable. In the end, the most favorable fold increase may depend on the research question.

Strengths and limitations

VASCO has a large study population with blood sampling at a regular interval allowing us to identify (re)infections in addition to reported infections based on (self-)testing. However, the interval of 6 months makes it more difficult to study trends (e.g. downward trends) within individuals. It should be noted that the duration of N-seropositivity, but also the sensitivity and specificity is dependent on the assay used (21, 22). When using an assay with lower sensitivity over time, this may limit the applicability of seroconversion or a fold increase as a proxy for (re)infection. Our findings should therefore be read in the context of the Roche assay measuring total Ig. In addition, the Roche assay has a high specificity (99.5%), making it suitable for this application. Finally, our method assumes that sample collection is complete, samples are collected between 4-8 months, and all have a measurable result. In reality, this is not always the case, which leads to less detection of infections than presented here.

Conclusion

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Seroconversion and fold-increase are suitable methods to detect (re)infections in population based prospective research at least up to a year after infection. Which fold increase to use requires a trade-off to include either more false positives or less true positives and this may depend on the research question. Overall, using a 2-fold increase resulted in the detection of a large proportion of additional infections in our data, with only a small share of false-positives.

Funding source

This study is funded by the Dutch ministry of Health, Welfare and Sport.

Ethical approval

Medical Ethics Committee of the Stichting Beoordeling Ethiek Biomedisch Onderzoek, Assen, the Netherlands, gave ethical approval for this work.

Data availability

Anonymized data reported from this study can be obtained from the corresponding author upon request. The dataset may include individual data and a data dictionary will be provided. Data requests should include a proposal for the planned analyses. Data transfer will require a signed data sharing agreement.

Competing interests

The authors declare none.

References

1. Vos ERA, van Hagen CCE, Wong D, Smits G, Kuijer M, Wijmenga-Monsuur AJ, et al. SARS-CoV-2 Seroprevalence Trends in the Netherlands in the Variant of Concern Era: Input for Future Response. Influenza and other respiratory viruses. 2024;18(6):e13312.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

2. Huiberts AJ, de Gier B, Hoeve CE, de Melker HE, Hahne SJM, den Hartog G, et al. Vaccine effectiveness of primary and booster COVID-19 vaccinations against SARS-CoV-2 infection in the Netherlands from July 12, 2021 to June 6, 2022: A prospective cohort study. International journal of infectious diseases. 2023;133:36-42.

3. Huiberts AJ, de Gier B, Hoeve CE, de Melker HE, Hahne SJ, den Hartog G, et al. Effectiveness of bivalent mRNA booster vaccination against SARS-CoV-2 Omicron infection, the Netherlands, September to December 2022. Euro surveillance. 2023;28(7).

4. Huiberts AJ, Hoeve CE, de Gier B, Cremer J, van der Veer B, de Melker HE, et al. Effectiveness of Omicron XBB.1.5 vaccine against infection with SARS-CoV-2 Omicron XBB and JN.1 variants, prospective cohort study, the Netherlands, October 2023 to January 2024. Euro surveillance. 2024;29(10):2400109.

5. van den Hoogen LL, Smits G, van Hagen CCE, Wong D, vos ERA, van Boven M, et al. Seropositivity to Nucleoprotein to detect mild and asymptomatic SARS-CoV-2 infections: A complementary tool to detect breakthrough infections after COVID-19 vaccination? Vaccine. 2022;40:2251-7.

6. Cauchemez S, Horby P, Fox A, Mai le Q, Thanh le T, Thai PQ, et al. Influenza infection rates, measurement errors and the interpretation of paired serology. PLoS pathogens. 2012;8(12):e1003061.

7. Zhao X, Siegel K, Chen MI, Cook AR. Rethinking thresholds for serological evidence of influenza virus infection. Influenza and other respiratory viruses. 2017;11(3):202-10.

8. den Hartog G, Andeweg SP, Hoeve CE, Smits G, Voordouw B, Eggink D, et al. Assessment of hybrid population immunity to SARS-CoV-2 following breakthrough infections of distinct SARS-CoV-2 variants by the detection of antibodies to nucleoprotein. Scientific reports. 2023;13(1):18394.

9. Huiberts AJ, Kooijman MN, de Melker HE, Hahne SJ, Grobbee DE, Hoeve C, et al. Design and baseline description of an observational population-based cohort study on COVID-19 vaccine effectiveness in the Netherlands - The VAccine Study COvid-19 (VASCO). Preprint. 2022.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

10. Huiberts AJ, Hoeve CE, Kooijman MN, de Melker HE, Hahne SJ, Grobbee DE, et al. Cohort profile: an observational population-based cohort study on COVID-19 vaccine effectiveness in the Netherlands – The VAccine Study COvid-19 (VASCO). Research Square (preprint). 2024.

11. Roche Diagnostics. Elecsys® Anti-SARS-CoV-2: Immunoassay for the qualitative detection of antibodies (incl. IgG) against SARS-CoV-2 2024 [Available from: https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html.

12. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

13. Wickham H AM, Bryan J,, Chang W ML, François R,, Grolemund G HA, Henry L,, Hester J KM, Pedersen TL,, Miller E BS, Müller K, Ooms, J RD, Seidel DP, Spinu V,, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1686.

14. Højsgaard S, Halekoh, U. & Yan J. The R Package geepack for Generalized Estimating Equations. Journal of Statistical Software. 2006;15(2):1-11.

15. Wood SN. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. Journal of the Royal Statistical Society Series B (Statistical Methodology). 2011;73(1):3-36.

16. Hoeve CE, Huiberts AJ, de Gier B, Andeweg SP, den Hartog G, de Melker HE, et al. COVID-19 vaccination-induced antibody responses and waning by age and comorbidity status in a large population-based prospective cohort study. Vaccine. 2024.

17. Bazin R, Rochette S, Perreault J, Fournier MJ, Gregoire Y, Boivin A, et al. Evaluation of anti-nucleocapsid level variation to assess SARS-CoV-2 seroprevalence in a vaccinated population. Infectious diseases. 2023;55(6):425-30.

18. Van Elslande J, Oyaert M, Lorent N, Vande Weygaerde Y, Van Pottelbergh G, Godderis L, et al. Lower persistence of anti-nucleocapsid compared to anti-spike antibodies up to one year after SARS-CoV-2 infection. Diagnostic microbiology and infectious disease. 2022;103(1):115659.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

19. Verheul MK, Kaczorowska J, Hofstee MI, Schepp RM, Smits GP, Wessels Beljaars D, et al. Protective mucosal SARS-CoV-2 antibodies in the majority of the general population in the Netherlands. Mucosal immunology. 2024.

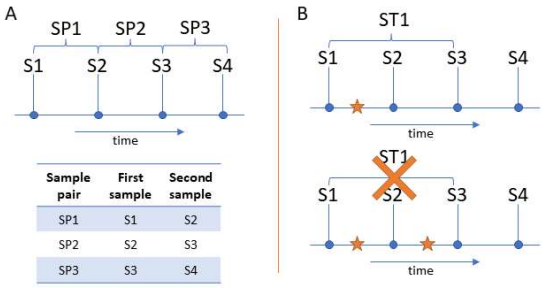
20. Richard L, Nisenbaum R, Colwill K, Mishra S, Dayam RM, Liu M, et al. Enhancing detection of SARS-CoV-2 re-infections using longitudinal sero-monitoring: demonstration of a methodology in a cohort of people experiencing homelessness in Toronto, Canada. BioMed Central infectious diseases. 2024;24(1):125.

21. Alfego D, Sullivan A, Poirier B, Williams J, Adcock D, Letovsky S. A population-based analysis of the longevity of SARS-CoV-2 antibody seropositivity in the United States. EClinicalMedicine. 2021;36:100902.

22. Stone M, Grebe E, Sulaeman H, Di Germanio C, Dave H, Kelly K, et al. Evaluation of Commercially Available High-Throughput SARS-CoV-2 Serologic Assays for Serosurveillance and Related Applications. Emerging infectious diseases. 2022;28(3):672-83.

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

433 **Figure 1. Organization of sample pairs. A)** Subsequent samples (S) form a sample pair (SP). A second sample may serve as a
434 first sample in the next sample pair of the same individual (e.g. sample 2 is considered a second sample in sample pair 1,
435 but a first sample in sample pair 2). **B)** In case of an infection (star symbol) between a first and second sample of a sample
436 pair, a third sample may be added to evaluate a longer time interval since infection (sample triple (ST)). A third sample can
437 only be added to the sample pair if there is no infection between the second and third sample. Fourth samples were not
438 included due to limited numbers. S3 may therefore form an ST sample pair with S1 in the upper figure, and a sample pair
439 with S2 in panel A.



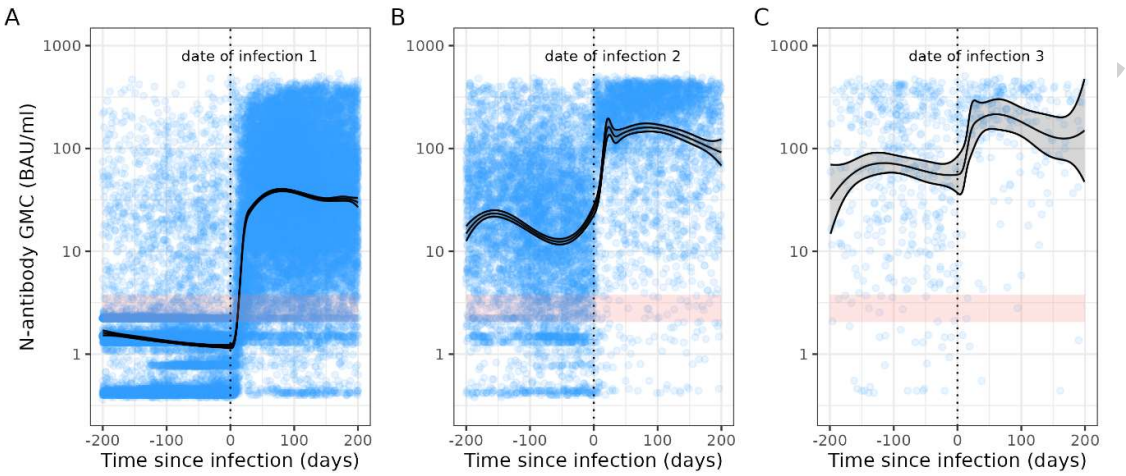
Sample pair	First sample	Second sample
SP1	S1	S2
SP2	S2	S3
SP3	S3	S4

440

441

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

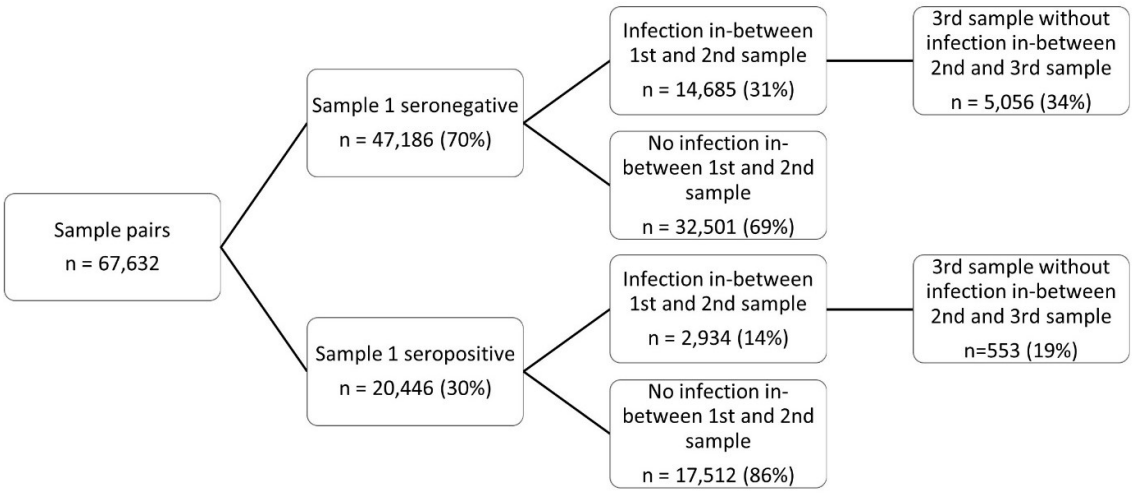
Figure 2. Generalized additive model showing N-antibody geometric mean concentration (GMC) over time before and after a first (A, n = 46,090 samples), second (B, n = 9,607 samples) and third (C, n = 719 samples) reported infection. Black lines represent N-antibody GMC and 95% confidence interval, blue scatter represents all individual samples used in the model, dotted vertical lines represent the moment of infection; red area corresponds with the cut-off range for seropositivity for different assay batches.



https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

449 **Figure 3. Flowchart of included sample pairs.** Sample pairs are grouped based on seropositivity of the first sample of a

450 sample pair and a reported infection in the sampling interval.

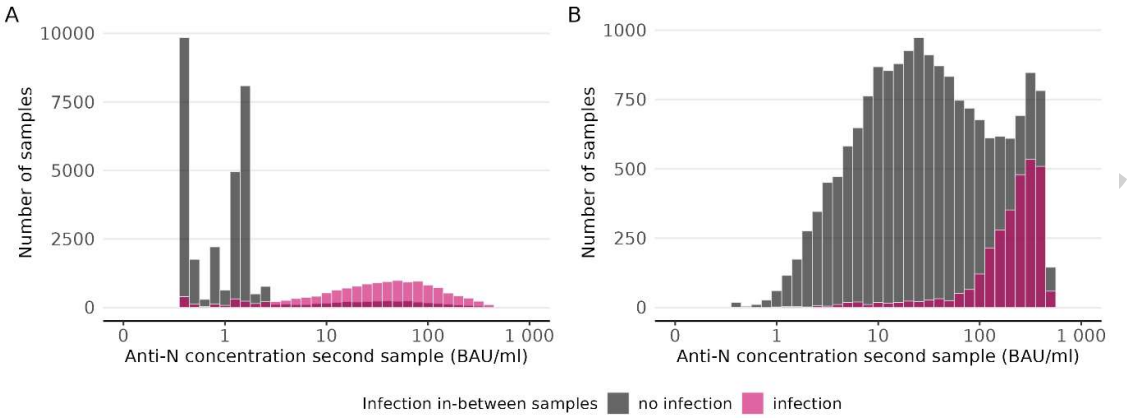


451

452

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Figure 4. Histogram of N-antibody levels of the second sample of sample pairs. A) sample pairs with the first sample seronegative (n=47,186) stratified by reported infection in the sampling interval, B) sample pairs with the first sample seropositive (n=20,446) stratified by reported infection in the sampling interval. Bars are plotted with overlap.

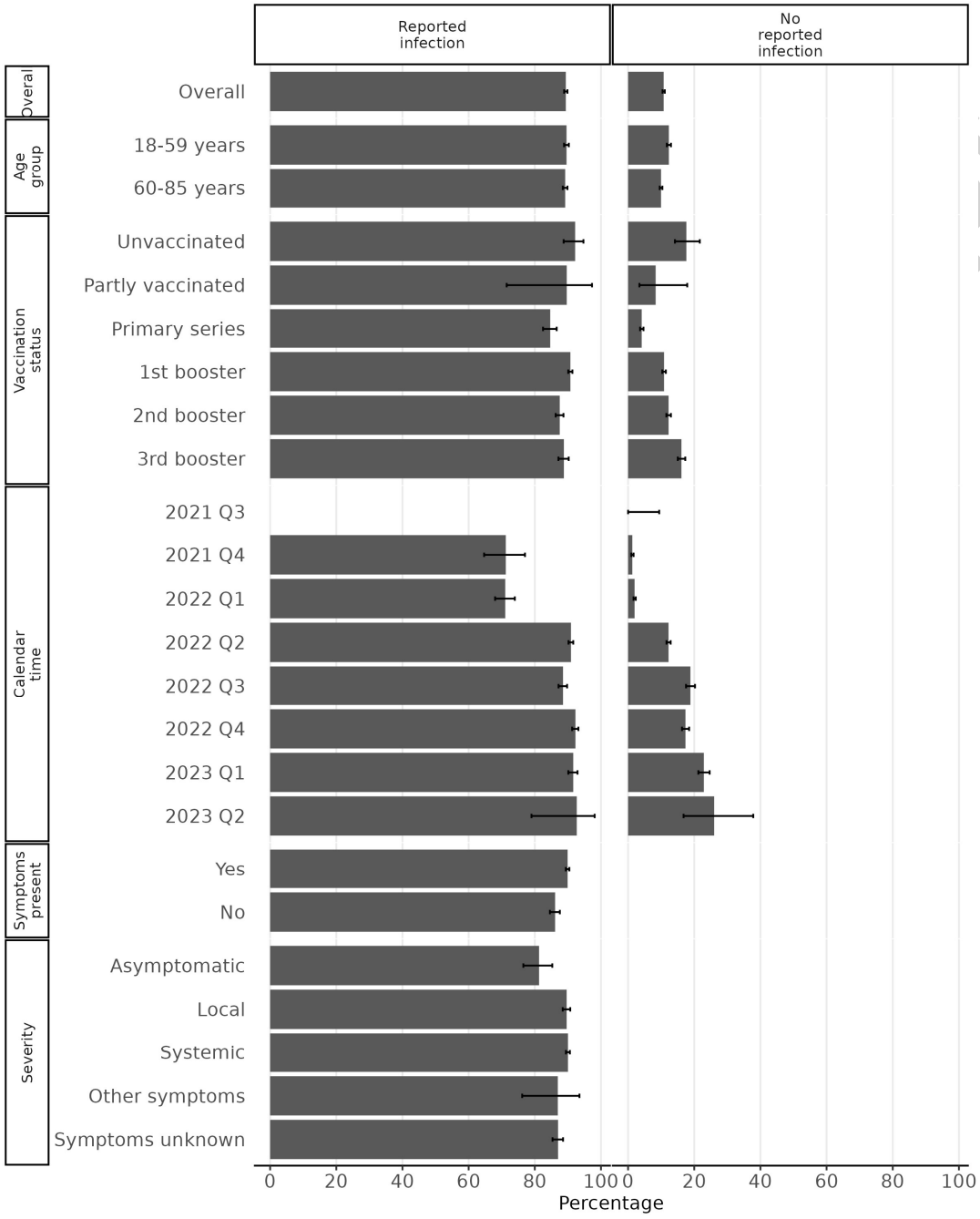


https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

458 **Figure 5. N-seroconversion of sample pairs with (left) and without (right) reported infection.** When there were less than

459 10 datapoints for a determinant, data is not shown in the figure. Calendar time was determined by the sampling date of

460 the second blood sample.



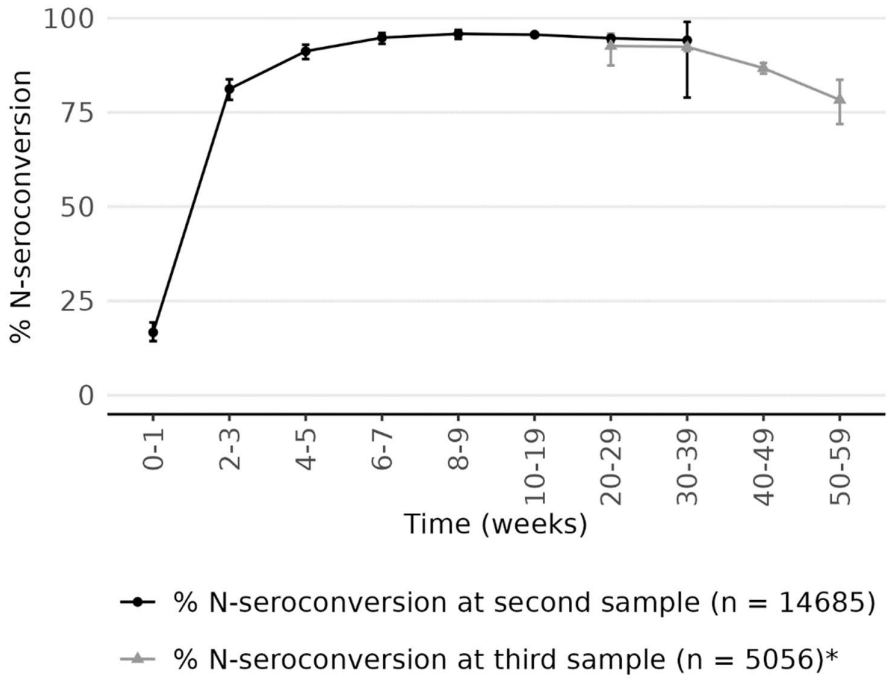
461

462

https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

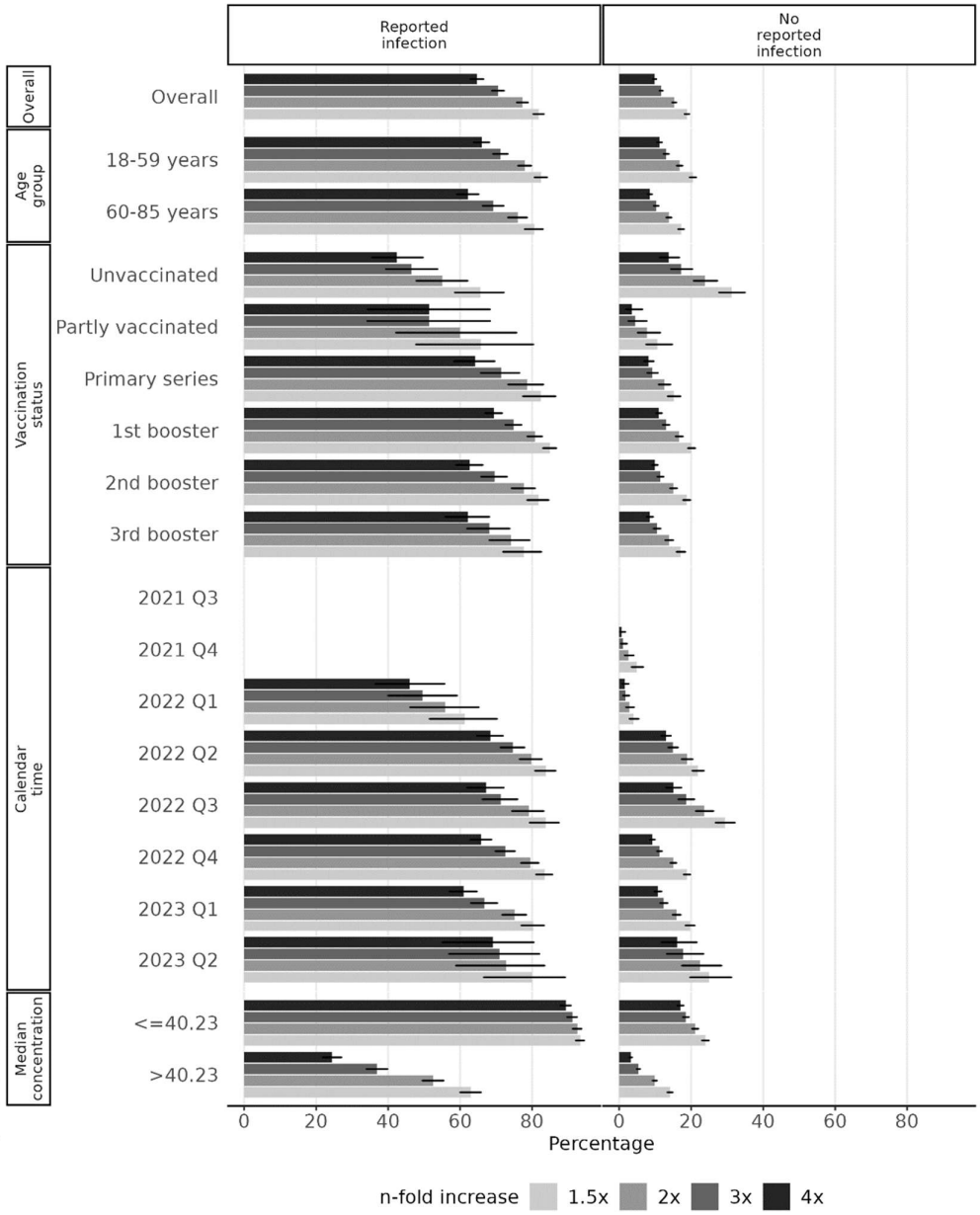
Figure 6. **N-seroconversion in sample pairs with or without third sample, with a reported infection only between 1st and 2nd sample, by time since infection.** Third samples were only included if there was no reported infection between the second and third sample and absence of a 2-fold increase. The error bars represent the 95% confidence interval around the percentage. When there were less than 10 datapoints for a period, data was excluded from the figure.

*time since infection equals time between infection and third sample.



https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

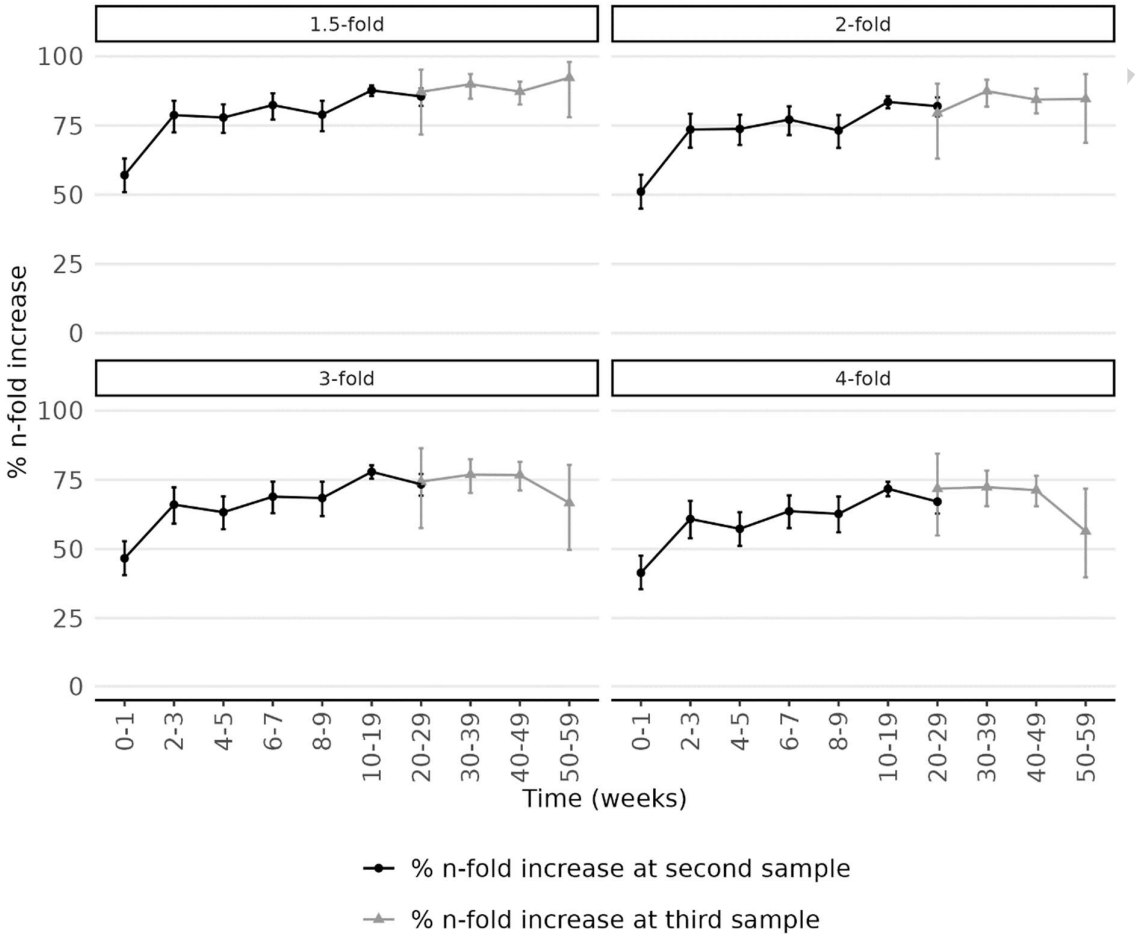
Figure 7. Percentage of sample pairs with 1.5-, 2-, 3- or 4-fold increase by determinant, stratified by reported infection in-between samples. When there were less than 10 datapoints for a determinant, data was excluded from the figure. Fold increase is presented by the saturation of the bars with the lightest bars representing samples with a 1.5-fold increase and the darker bars samples with a 4-fold increase. Calendar time was determined by the sampling date of the second blood sample.



https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press

Figure 8. Percentage of 1.5-, 2-, 3- or 4-fold increase in sample pairs with or without third sample, with a reported infection only between 1st and 2nd sample. Third samples were only included if there was no reported infection between the second and third sample and absence of a 2-fold increase. The error bars represent the 95% confidence interval around the percentage. When there were less than 10 datapoints for a period, data was excluded from the figure.

*time since infection equals time between infection and third sample.



https://doi.org/10.1017/S095026882500010X Published online by Cambridge University Press