Michele Santacatterina, Brian Burke, Mihili Gunaratne, William S. Weintraub, Mark A. Espeland, Adolfo Correa, DeAnna Friedman-Klabanoff, Michael Gibbs, David Herrington, Kristen E. Miller, John W. Sanders, Austin L. Seals, Diane Uschner, Thomas F. Wierzba and Morgana Mongraw-Chaffin*

Using repeated antibody testing to minimize bias in estimates of prevalence and incidence of SARS-CoV-2 infection

https://doi.org/10.1515/em-2023-0012
Received April 26, 2022; accepted July 17, 2023; published online August 9, 2023

Abstract

Objectives: The prevalence and incidence of SARS-CoV-2, the virus which causes COVID-19, at any given time remains controversial, and is an essential piece in understanding the dynamics of the epidemic. Cross-sectional studies and single time point testing approaches continue to struggle with appropriate adjustment methods for the high false positive rates in low prevalence settings or high false negative rates in high prevalence settings, and post-hoc adjustment at the group level does not fully address this issue for incidence even at the population level.
Methods: In this study, we use seroprevalence as an illustrative example of the benefits of using a case definition using a combined parallel and serial testing framework to confirm antibody-positive status. In a simulation study, we show that our proposed approach reduces bias and improves positive and negative predictive value across the range of prevalence compared with cross-sectional testing even with gold standard tests and post-hoc adjustment. Using data from the North Carolina COVID-19 Community Research Partnership, we applied the proposed case definition to the estimation of SARS-CoV-2 seroprevalence and incidence early in the pandemic.
Results: The proposed approach is not always feasible given the cost and time required to administer repeated tests; however, it reduces bias in both low and high prevalence settings and addresses misclassification at the individual level. This approach can be applied to almost all testing contexts and platforms.
Conclusions: This systematic approach offers better estimation of both prevalence and incidence, which is important to improve understanding and facilitate controlling the pandemic.

Keywords: antibody test; bias; repeated testing; SARS-Cov-2; simulation

***Corresponding author: Morgana Mongraw-Chaffin**, Department of Epidemiology & Prevention, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA, E-mail: mmongraw@wakehealth.edu. https://orcid.org/0000-0001-6541-7198
Michele Santacatterina, Department of Population Health, New York University Grossman School of Medicine, New York, NY, USA; and The Biostatistics Center, George Washington University, Rockville, MD, USA
Brian Burke, **Mihili Gunaratne and Diane Uschner**, The Biostatistics Center, George Washington University, Rockville, MD, USA
William S. Weintraub and Kristen E. Miller, Georgetown University, Washington, DC, USA; and MedStar Health Research Institute, Washington, DC, USA
Mark A. Espeland, **David Herrington**, **John W. Sanders**, **Austin L. Seals and Thomas F. Wierzba**, Wake Forest School of Medicine, Winston-Salem, NC, USA
Adolfo Correa, University of Mississippi Medical Center, Jackson, MS, USA
DeAnna Friedman-Klabanoff, Center for Vaccine Development and Global Health, University of Maryland School of Medicine, Baltimore, MD, USA
Michael Gibbs, Department of Emergency Medicine, Atrium Health, Charlotte, NC, USA

Introduction

Accurately estimating the prevalence and incidence of a disease underpins public health and policy decisions about disease and is the foundation of all successful control measures. Throughout the SARS-CoV-2 pandemic, determining prevalence and incidence from testing has been challenging and controversial, and failure to understand the basic role of prevalence in misclassification has exacerbated these issues. While the importance of antibody testing to determine SARS-CoV-2 seroprevalence has decreased as the pandemic continued, the early challenges in this area and their contribution to confusion and uncertainty provide an illustrative example to avoid these same problems with other testing platforms for future decision making. Right from the beginning, cross-sectional SARS-CoV-2 seroprevalence study groups struggled to find appropriate adjustment methods for the high false positive rates in low prevalence settings early in the pandemic. This remains one of the primary criticisms of early work in the field, (Infectious Diseases Society of America 2020; Offord 2020) and has led to confusion and misunderstanding about the scope and risk of the emerging COVID-19 pandemic. Approaches to date have been post-hoc and focused on adjustment for test performance at the group level (Bajema et al. 2021; Basto-Abreu et al. 2022; Bryant et al. 2020; Pollán et al. 2020). These include adjusting the prevalence point estimate using a simple calibration equation (Basto-Abreu et al. 2022; Bryant et al. 2020), adjusting for time-varying test performance (Perez-Saez et al. 2021), or providing a range of estimates using IgM AND IgG as the case definition for the lower bound, which optimizes for specificity, and IgM OR IgG as the case definition for the upper bound, which optimizes sensitivity (Pollán et al. 2020). While these approaches have the benefit of simplicity and enhance the accuracy of cross-sectional serosurveys, they are limited by failing to address misclassification at the individual level.

Individual level misclassification remains a major problem for every type of SARS-CoV-2 test regardless of improved test performance throughout the pandemic with implications for invidual decsion making, clinical care, and policy. For instance, heterogeneity in test performance by time since exposure or symptom onset continues to be a source of confusion for the public and a source of misdiagnosis for clinical care and bias in research. Worse, a study of pooled estimates from Reverse Transcriptase Polymerase Chain Reaction tests suggests that in some contexts the mean false negative rate may not decrease below 20 % regardless of the timing (Kucirka et al. 2020). Both over- and under-estimation of prevalence and incidence from such misclassification may negatively impact clinical and public health planning, including preparing for medical capacity and staff needs, estimating demand and access to testing, vaccines, and other supplies, and maintaining public trust by providing clear and accurate information. This misclassification may have wide ranging impacts on estimates of asymptomatic cases, seroconversion rates, impact on long-term outcomes, determinants and outcomes of incident infections, and questions of immunity and the possibility of reinfection.

In this paper our goal is to use the illustrative example of seroprevalence early in the pandemic as a framework for improving the accuracy of estimates from all types of SARS-CoV-2 testing in the future. We offer proof-of-concept of an approach that includes parallel and sequential testing (PSeq) to increase positive and negative predictive values that will substantially reduce overestimation of prevalence due to false positives and preserve high sensitivity, when compared to using a single test approach (IgM OR IgG on single test). We use simulation analysis to show how our proposed case definition reduces bias. We further expand our simulation to mimic test performance of current gold standard we also showcase the difference in results using data from a large, ongoing surveillance program. Hypothesizing that estimation of incidence would be improved by reducing misclassification at the individual level, we show that the IgM OR IgG definition overestimates the cumulative incidence and incidence rates compared to our proposed approach. This type of repeated testing approach is becoming more common for clinical and personal decisions making protocols, and we draw connections to their effective use in these contexts to argue for this approach to be expanded to improve research, public health, and policy.

Methods

Case definitions for OR, AND, and PSeq

All case definitions are included in Box 1. For the OR case definition we consider participants to be a positive case if they have either a positive IgM or a positive IgG response on a single test. This definition optimizes sensitivity. For the AND case definition we consider participants to be a positive case only if they have both a positive IgM and a positive IgG on the same test. This definition optimizes specificity. The OR and AND definitions use only parallel testing. We define PSeq using both parallel and sequential testing as:

IgM AND IgG on first test OR [(*IgM OR IgG on first test) AND (IgM OR IgG on second test*)]

The first condition uses the information contained in the first test in both IgG and IgM. The second condition deals with the case of having either one of IgG or IgM positive on the first test and a confirmatory IgG or IgM positive tests on the second. In this paper, we consider only two tests. The PSeq case definition can be extended to more than two tests.

Simulation

In this Section, we describe the design of a simulation study designed to evaluate and compare estimated prevalence based on PSeq with that of the OR and AND case definitions, and that of the OR with adjustment with respect to absolute bias and positive predictive value, across levels of true prevalence (π). We considered a sample size of n = 1,000 and 1,000 iterations. To compute prevalence based on the OR (π^{OR}), AND (π^{AND}) case definitions, and OR with adjustment (π^{aOR}), we assumed that the vector containing the number for each combination of positive and negative IgG and IgM was sampled from a multinomial distribution with probabilities defined as in Eqs. (2)–(5) of the Supplementary Material. In other words, we considered a 2-by-2 table containing the number of negative tests for both IgG and IgM (n−−), IgG positive and IgM negative (n+−), IgG negative and IgM positive (n−+) and both positive (n++) (Table 1 in the Supplementary Material). For each iteration we then computed π^{OR} and π^{AND} as the proportion between the sum of n+−, n−+ and n++, and between n++ and the total number of participants n respectively. We computed $\pi^{\text{aOR}} = \frac{\pi^{\text{OR}}+\text{sp}^{\text{OR}}-1}{\text{se}^{\text{OR}}+\text{sp}^{\text{OR}}-1}$ where se^{OR} and sp^{OR} were defined as in Eqs. (8) and (9) in the Supplementary Material. To compute prevalence based on the PSeq case definition (π^{PSeq}), we assumed the vector of positive and negative tests was sampled from a binomial distribution with probability $\pi(se^{\text{PSeq}}) + (1-\pi)(1-sp^{\text{PSeq}})$, where se^{PSeq} and sp^{PSeq} are defined in Eqs. (16)–(22) and (20) of the Supplementary Materials. Sensitivity and specificity of the first and second test for IgG and IgM were set equal to 0.8 and 0.9, respectively based roughly on actual test performance (Frederick National Laboratory for Cancer Research 2020). For our primary analysis, we considered the tests to be uncorrelated, i.e., we assumed conditional independence (Hui and Walter 1980). For each iteration, we computed π^{PSeq} as the sum of positive tests divided by n. We considered 81 values of the true prevalence π, from 0.1 to 0.9 and computed absolute bias of π^{OR}, π^{AND}, π^{aOR}, and π^{PSeq} as the absolute difference between their means and the true value of prevalence. Positive and negative predictive values for OR, AND, and PSeq were calculated using formulas presented in Eqs. (10), (11), (14), (15) and (35) in the Supplementary Materials, respectively.

We further extended our simulation to represent the current gold standard in COVID diagnosis, Nucleic Acid Amplification Tests (NAATs). Using the same principles as for our antibody test simulation, we calculated the absolute bias, negative

Box 1: Equations for different case definitions used in this study.

Case definition[a]	Equation	Parallel, sequential, or both	Goal
OR	IgM OR IgG on any single test	Parallel	Maximize sensitivity
AND	IgM AND IgG on any single test	Parallel	Maximize specificity
PSeq	IgM AND IgG on first test OR [(IgM OR IgG on first test) AND (IgM OR IgG on second test)]	Both	Reduce bias across the range of prevalence

[a]aOR is a parallel approach that uses the OR definition and uses the following equation to adjust the prevalence for testing performance: $\pi^{\text{aOR}} = \frac{\pi^{\text{OR}}+\text{sp}^{\text{OR}}-1}{\text{se}^{\text{OR}}+\text{sp}^{\text{OR}}-1}$.

predictive value, and positive predictive value for a single test and for a repeated test approach, respectively. We used estimates of sensitivity (0.98) and specificity (0.97) for NAATs taken from the higher range of test performance in ideal testing situations (Hanson et al. 2023). Estimates for NAAT performance range widely and many in real world settings are substantially lower than those used here (Hanson et al. 2023). We do not provide a simulation of the impact of these case definitions on incidence, given that the additional assumptions would be substantial and may be depend on heterogeneity of risk in specific populations.

Real world case-study

In this Section, we describe an empirical application of our proposed case definition. We used data between April 16, 2020 and December 20, 2020 from the North Carolina COVID-19 Community Research Partnership study (NC-CRP), a health system-based longitudinal syndromic and sero-surveillance study (COVID-19 Community Research Partnership 2022; Munawar et al. 2021). Briefly, participants were initially recruited through participating healthcare systems and healthcare workers were oversampled. Inclusion required an active email address and data capable cell phone. Additional details on the NC-CRP study design can be found on the study website and in the design paper (HYPERLINK "http://www.covid19communitystudy.org/index.html" \o "http://www.covid19communitystudy.org/index.html" http://www.covid19communitystudy.org/index.html) (COVID-19 Community Research Partnership 2022).

Participants were tested for antibodies against SARS-CoV-2 using two different at-home lateral flow assays from Syntron and Innovita. Both tests use the Scanwell platform and test for IgM and IgG to both SARS-CoV-2 nucleocapsid and spike proteins. The sensitivity and specificity of the Syntron test were IgG 0.73 and 0.99, and IgM 0.93 and 0.975, respectively, based on test performance results from the National Cancer Institute (Frederick National Laboratory for Cancer Research 2021). Similarly, the sensitivity and specificity of the Innovita test were IgG 0.87 and 0.99, and IgM 0.90 and 0.99, respectively. For the primary analysis and unless otherwise specified we used the Syntron test performance characteristics. Because we only included tests through December 2020, which was before access to COVID-19 vaccinations was widespread, a positive test represents evidence of prior infection with SARS-CoV-2. From the initial cohort of 9,830 participants, 39 had an inconclusive test, and 3,859 had only one test. The final cohort was comprised of 5,932 participants with more than one test, ranging from 2 to 6, with a median of 4 tests. Sixty-one percent were female, 89 % were non-Hispanic white and 46, 41, and 13 % were of age 18–45, 46–65, and 65+, respectively. We estimated prevalence using four different approaches: π^{OR}, π^{AND}, π^{PSeq}, and π^{aOR}. The first three target the OR, AND, and PSeq case definition described above where prevalence is estimated by calculating the number of positive cases for each case definition and dividing them by 5,932 (the total cohort). The fourth approach, π^{aOR}, estimates prevalence adjusting it by test characteristics. Confidence intervals were constructed by using the normal approximation (Blyth and Still 1983).

As a secondary analysis to estimate prevalence and its uncertainty while controlling for correlations between tests and uncertain sensitivity and specificity, we repeated our analysis using Bayesian models (Gelman and Carpenter 2020). Specifically, we chose a truncated 0–1 normal distribution with mean 0.73, 0.99, 0.93, 0.975 and standard deviation 0.01 as informative priors for sensitivity and specificity of IgG and IgM, respectively. We used the same priors for first and second tests. Also, we fitted models using priors based on the Innovita IgG and IgM performance, *i.e.*, a truncated 0–1 normal distribution with mean 0.87 (sensitivity IgG), 0.99 (specificity IgG), 0.90 (sensitivity IgM), 0.99 (specificity IgM), and standard deviation 0.01. We chose a truncated 0–1 normal distribution with mean 0.50 and standard deviation 0.10 as informative prior for covariances between IgG and IgM and first and second tests. These values were chosen to reflect correlations around 0.1. Similarly as in our simulation scenarios, we assumed a multinomial model for the OR and AND case definitions and a binomial model for the PSeq case definition. The posteriors of prevalence distributions were computed using a Markov-chain Monte Carlo with 2,000 iterations and 4 chains. When we considered correlation, as opposed to π^{aOR} for which we assumed conditional independence, adjusted prevalence estimation, π^{Adj}, was obtained by solving Eqs. (2)–(5) in the Supplementary Material.

We also estimated cumulative incidence over calendar time and incidence rates for 2-month intervals. We excluded those who reported a COVID-19 diagnosis or previous positive test for SARS-CoV-2 at enrollment. The Kaplan–Meier estimator was used to estimate cumulative incidence for each case definition across calendar time. Follow-up was from study enrollment to date of event or date of last entry, whichever came first. Incidence rates were calculated as the ratio between number of events and person-years across calendar time. Person-years were calculated from the beginning of each 2-month interval until the date of event or interval. Intervals for incidence rates were as follows: 1st interval 4/6/2020 to 6/12/2020, 2nd interval 6/13/2020 to 8/8/2020, 3rd interval 8/9/2020 to 10/4/2020, and 4th interval 10/5/2020 to 11/30/2020 and estimates are presented per 10 person-years. R version 4.0.1 was used for all analyses (R Project for Statistical Computing, https://www.r-project.org) (R Core Team 2021).

Results

Simulation results

As shown in Figure 1, π^{PSeq} outperformed π^{OR} and π^{AND} with respect to absolute bias, positive predictive value, and negative predictive value across values of prevalence from 0.1 to 0.9. Absolute bias ranged from 0 to 0.3 depending on the prevalence with the highest bias for the OR definition at low prevalence and the highest bias for the AND definition at high prevalence. Absolute bias of π^{PSeq} was similar to that of π^{aOR} across prevalence values. Positive and negative predictive values are not calculable for the π^{aOR} approach as it adjusts the group level prevalence without accounting for misclassification at the individual level. Using sensitivity and specificity of current gold standard NAATs, absolute bias ranged from 0 to 1.5 and repeated testing of lower performance tests still exhibited reduced bias compared to a single test across the range of prevalence (Figure 2). PPV and NPV for NAATs were also similarly poor at low and high prevalence respectively within the range concordant with current waves of transmission and in the context of reduced but more selective clinical testing.

Real world case-study results

Table 1 shows that the highest and lowest estimated prevalences were obtained by using the π^{OR} and π^{AND}, respectively, e.g., $\hat{\pi}^{\text{OR}} = 11.5$ (95 % confidence interval: 9.2–13.7) while $\hat{\pi}^{\text{AND}} = 3.6$ (95 % confidence interval: 2.3–4.9). Estimates of π^{aOR} and π^{PSeq}, were between those of π^{OR} and π^{AND}. In addition, similar results were obtained using Bayesian models and accounting for correlation between tests (Table 2). Similarly as for prevalence estimation, Kaplan–Meier cumulative incidence based on the OR and AND case definitions was the highest and the lowest, respectively, with estimates based on the PSeq case definition falling in between (Figure 2). The overestimation of cumulative incidence based on the OR case definition at the beginning of the study period is due to the presence of false positives and that overestimation continues across the entire timeline.

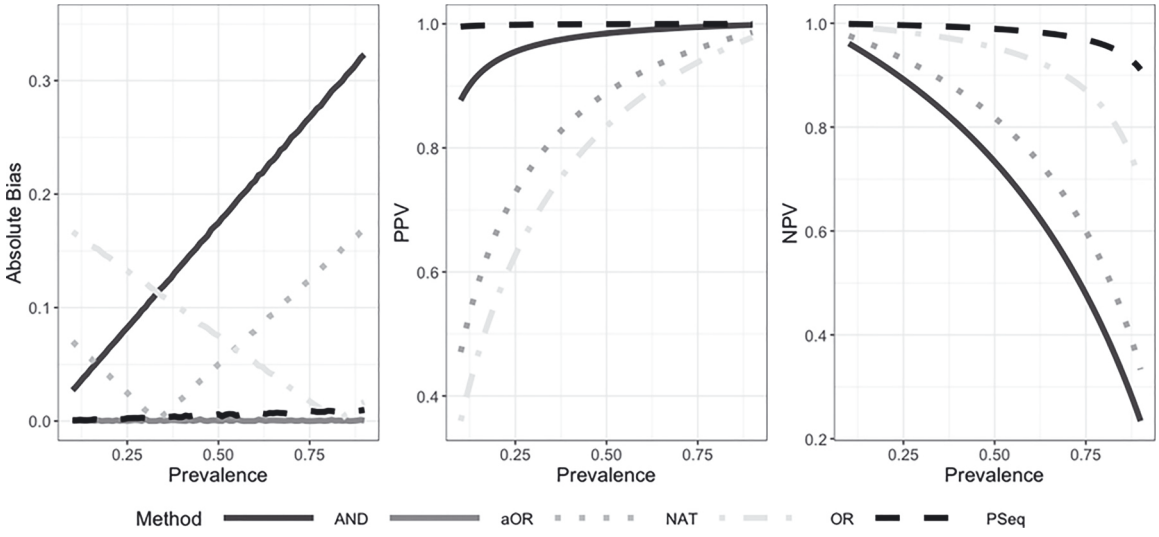

Figure 1: Absolute bias (left panel), positive predictive value (PPV – middle panel) and negative predictive value (NPV – right panel) of π^{OR}, π^{AND}, π^{aOR}, and π^{PSeq} from antibody tests (sensitivity = 0.80 and specificity = 0.90) and from a single NAAT test (sensitivity = 0.98 and specificity = 0.97) across values of SARS-CoV-2 prevalence from 0.1 to 0.9 under conditional independence from simulation model. Positive predictive values are provided only for π^{OR}, π^{AND}, π^{aOR}, and π^{PSeq} and NAAT.

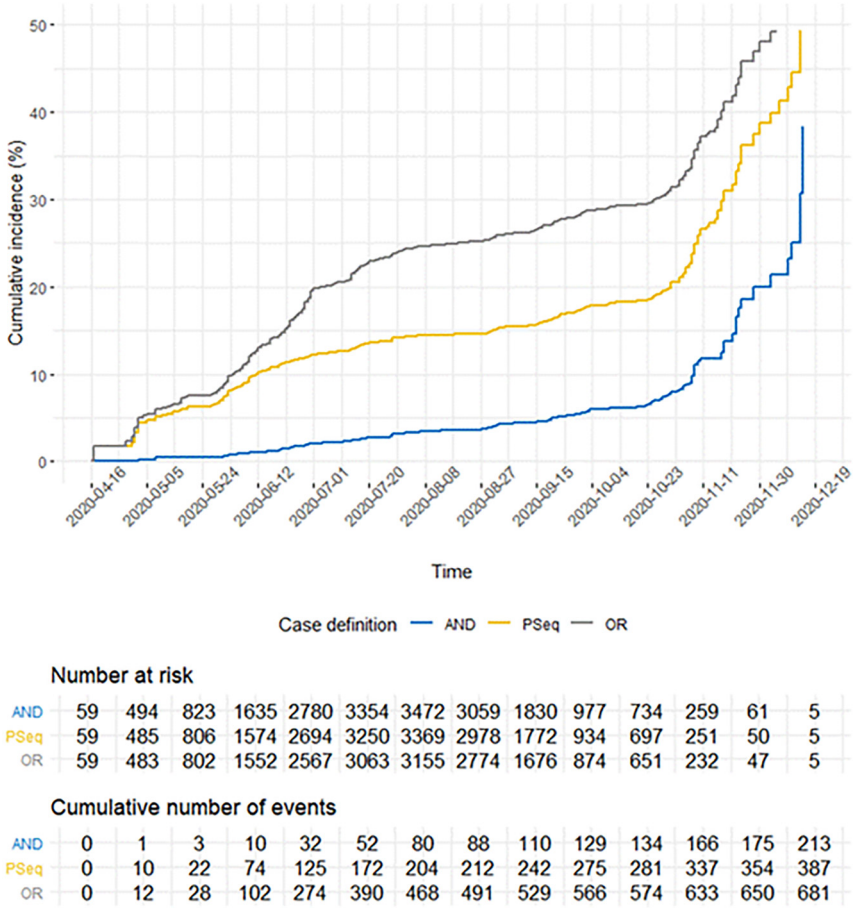

Figure 2: Kaplan–Meier estimates of cumulative incidence of SARS-CoV-2 in a subset of 5,932 participants from the NC-CRP study, North Carolina, April–December 2020, number at risk and cumulative number of events over calendar time using the OR, AND, and PSeq case definitions.

Table 1: Prevalence of SARS-CoV-2 infection and 95 % confidence interval in a subset of 5,932 participants from the NC-CRP study, North Carolina, April–December 2020 using π^{OR}, π^{AND}, $\pi^{aORSyntron}$, $\pi^{aORInnovita}$ and π^{PSeq}.

	Prevalence % (95 % CI)				
	$\hat{\pi}^{AND}$	$\hat{\pi}^{PSeq}$	$\hat{\pi}^{aORSyntron}$	$\hat{\pi}^{aORInnovita}$	$\hat{\pi}^{OR}$
Overall	3.6 (2.3–4.9)	6.5 (4.8–8.3)	8.4 (6.5–10.4)	9.8 (7.9–11.6)	11.5 (9.2–13.7)

Incidence rates were higher for the OR definition than the PSeq definition across the entire timeline (Table 3). Similarly, the AND definition underestimated incidence rates compared to the PSeq definition during every interval.

Table 2: Prevalence (%) of SARS-CoV-2 infection and 95 % uncertainty interval in a subset of 5,932 participants from the NC-CRP study, North Carolina, April – December 2020 using π^{OR}, π^{AND}, π^{Adj}, and π^{PSeq} estimated using Bayesian models.

Antibody test	Case definition			
	$\hat{\pi}^{AND}$	$\hat{\pi}^{PSeq}$	$\hat{\pi}^{Adj(a)}$	$\hat{\pi}^{OR}$
Syntron[b]	3.8 (3.4; 4.2)	6.5 (6.0; 7.1)	5.1 (4.4; 5.8)	11.2 (10.5; 11.8)
Innovita[c]	3.7 (3.3; 4.2)	6.5 (6.0; 7.1)	4.3 (3.7; 5.0)	11.1 (10.4; 11.7)

[a]From Eqs. (2)–(5) in the Supplementary Material. We chose a truncated 0–1 normal distribution with mean 0.50 and standard deviation 0.10 as informative prior for covariances between IgG and IgM and first and second tests. [b]Truncated 0–1 normal distributions with mean 0.73 (sensitivity IgG), 0.99 (specificity IgG), 0.93 (sensitivity IgM), 0.975 (specificity IgM) and standard deviation 0.01 were used as priors. [c]Truncated 0–1 normal distributions with mean 0.87 (sensitivity IgG), 0.99 (specificity IgG), 0.90 (sensitivity IgM), 0.99 (specificity IgM) and standard deviation 0.01 were used as priors.

Table 3: Incidence Rate per 10 person-years[a] by case definition across 2-month time intervals from 4/6/2020 to 11/30/2020 in a subset of 5,932 participants from the NC-CRP study, North Carolina, April – December 2020.

Case definition	Two-month time intervals			
	1st interval 4/6/2020 to 6/12/2020	2nd interval 6/13/2020 to 8/8/2020	3rd interval 8/9/2020 to 10/4/2020	4th interval 10/5/2020 to 11/30/2020
AND	0.9	1.3	1.4	5.9
PSeq	7.2	3.0	2.1	10.6
OR	9.9	9.0	3.0	12.0

[a]Incidence rates were calculated as the ratio between number of events and person-years across calendar time. Person-years were calculated from the beginning of each 2-month interval until the date of event or interval.

Discussion

Using simulations, we demonstrate that using a case definition that takes advantage of repeated testing minimizes bias in prevalence estimation across the prevalence range. Our proposed approach both minimizes the number of false positives in the low prevalence setting and minimizes the false negatives in the high prevalence context, highlighting how this approach can reduce bias during low prevalence at the beginning of the pandemic. Concerns over false negatives also apply to groups with high pretest probabilities such as in those with symptoms or known exposures, and concerns over high false positives apply more generally to estimates of incidence in the general population given the expected low positive predictive value from the prevalence of current infection remaining under 10 % at any given time even during pandemic wave peaks. The repeated testing approach outperformed the other case definitions and reduced bias even in the context of high test performance concordant with current gold standard tests (NAATs). We further demonstrate the implications of choosing a repeated testing case definition compared to various case definitions that rely on a single test using real world data.

Context

Some of the earliest estimates of SARS-CoV-2 prevalence in the US suffered from fatal methodological flaws and may have damaged the response to the pandemic by encouraging the public to embrace incorrect conclusions about its nature (Offord 2020). These attempts were followed by the elegant work of investigators who provided seroprevalence estimates in multiple countries (Bajema et al. 2021; Pollán et al. 2020; Stringhini et al. 2020). These authors make use of gold standard approaches to account for false positives and low positive predictive values,

but overwhelmingly rely on a cross-sectional design with a single antibody test. Very few studies on SARS-CoV-2 with large sample sizes and broad recruitment have included repeated antibody testing in their design, and current estimates of active infection often rely on a single testing framework. Further, few studies investigating the risk factors for serious COVID outcomes or the impact of SARS-CoV-2 infection have attempted to address misclassification. Our findings speak to the importance of repeated testing and the improvement in prevalence and incidence estimates that can be achieved with a confirmatory test. Given that the positive and negative predictive values differ depending on the underlying prevalence, comparisons between groups will likely be underestimated using results from a single test. The bias is likely to be the highest when the difference between two groups is the greatest. While *post hoc* adjustment generally performs well for estimates of prevalence, caveats about the disconnect between the population of interest and the study sample used to estimate test characteristics have been previously noted (Accorsi et al. 2021; Takahashi, Greenhouse, and Rodríguez-Barraquer 2020). Moreover, similar approaches for measures of incidence or time-to-event analyses are lacking and may not be effective if they do not address misclassification at the individual level. This may prove particularly challenging in the context of increasing reinfection rates. In addition, test performance may vary among individuals, and thus extensions to individual-level misclassification approaches may be important. Our results and basic testing theory suggest that: 1. Additional use of single cross-sectional test samples will continue to generate biased incidence estimates and hinder prevention and control measures, and 2. The proposed approach will be just as successful at minimizing bias throughout the pandemic as it has been in the low prevalence setting of the early pandemic.

Repeated testing has already been implemented or recommended to improve accuracy in SARS-CoV-2 testing in a variety of ways: 1. To rule out false negatives even when using gold standard NAAT tests for diagnosis given the heterogeneity in test performance by timing since exposure or symptom onset, 2. For improved sensitivity of rapid antigen tests in both symptomatic and asymptomatic individuals (Soni et al. 2023), 3. To reduce false positives from screening of asymptomatic groups (Connor et al. 2022), and 4. To improve negative predictive value and rule out false negatives in high pretest probability settings when using rapid antigen tests at home for those with symptoms, exposure, or risk of infecting others (Centers for Disease Control 2022; United States Food & Drug Administration 2022). While these are just some examples, our results confirm the potential of a repeated testing approach to benefit clinical care, research, and policy across a wide range of contexts. Extending a repeated testing approach to more research studies and public health surveillance efforts would greatly benefit our understanding of this pandemic.

Limitations and strengths

The primary limitation of this work is that we do not have a gold standard test of SARS-CoV-2 infection with which to confirm our results. This issue is not unique to our study and continues to hamper the field as a whole; however, the misclassification problems we investigate here remain a serious concern for all test types, including those currently considered the gold standard. Second, not all Wake Forest NC-CRP participants had more than one antibody test and cadence between tests was highly variable. Further, those with and without more than one antibody test are demographically different based on differential timing of recruitment. The surveillance data from this cohort is likewise not informative with regards to ideal testing cadence or optimal testing protocols for the general population. Similarly, this study sample includes results obtained using two different tests as the test was changed to improve test performance. The main impact of this is on the *post hoc* adjustments. Given the limitations of the NC-CRP cohort as a data source, we extracted a subset of the NC-CRP as a clean cohort to best demonstrate the proposed approach. As such, the real world data we use for this analysis may not be fully generalizable to other populations, although we expect that the performance of these case definitions will not differ substantially across populations. Addressing other issues that impact prevalence and incidence estimation including sampling strategy and access to testing and care is beyond the scope of this paper. Finally, the improved estimates from our approach come at the expense of the cost and time required for participant follow-up and repeated testing.

Despite these limitations, our study used the illustrative example of SARS-CoV-2 seropositivity with antibody testing to demonstratesing the utility of a case definition that leverages repeated testing. With a combination of simulations and real world data, we demonstrate the importance of choosing a case definition that reduces bias across the range of prevalence as the COVID-19 pandemic progresses. While none of the methods used in this paper are new, the application of this approach during the COVID-19 pandemic is novel and it is expected that these results could impact our future public health response. By going back to basics and building on traditional testing theory, our proposed approach could be expanded to any type of testing to reduce the bias of SARS-CoV-2 prevalence and incidence estimates. This approach will also offer the benefit of addressing misclassification at the individual level. While *post hoc* prevalence adjustments that account for single test characteristics perform well, they are not available for adjusting estimates of incidence which are vulnerable to misclassification of both the numerator and denominator. Further, this approach can be applied to other types of test, and specifically may reduce the uncertainty around the ideal time frame for testing in many situations.

Conclusions

We have proposed and evaluated an approach using sequential testing to increase positive and negative predictive value compared to other case definitions using an illustrative example. Additional applications for this approach include estimates of SARS-CoV-2 incidence, changing incidence rates, antibody duration, reinfection rates, and all other metrics that depend on SARS-CoV-2 testing. A repeated testing approach will generate better estimates of SARS-CoV-2 prevalence and incidence from which to build evidence-based recommendations and guide more effective public health policy.

Acknowledgments: The authors would like to thank other non-author collaborators of the North Carolina COVID-19 Community Research Partnership Study Group including: **Wake Forest School of Medicine:** Alain Bertoni, MD, Martha A. Alexander-Miller, PhD, Allison Mathews, PhD, Iqra Munawar, MS, Brian Ostasiewski, Christine Ann Pittman Ballard, MPH, Metin Gurcan, PhD, Alexander Ivanov, MD, Giselle Melendez Zapata, MD, Marlena Westcott, PhD, Karen Blinson, Laura Blinson, Douglas McGlasson, Mark Mistysyn, Donna Davis, Lynda Doomy, Perrin Henderson, MS, Alicia Jessup, Kimberly Lane, Beverly Levine, PhD, Jessica McCanless, MS, Sharon McDaniel, Kathryn Melius, MS, Christine O'Neill, Angelina Pack, RN, Ritu Rathee, RN, Scott Rushing, Jennifer Sheets, Sandra Soots, RN, Michele Wall, Samantha Wheeler, John White, Lisa Wilkerson, Rebekah Wilson, Kenneth Wilson, Deb Burcombe. **Atrium Health:** Michael S. Runyon, MD MPH, Lewis H. McCurdy, MD, Yhenneko J. Taylor, PhD, Lydia Calamari, MD, Hazel Tapp, PhD, Amina Ahmed, MD, Michael Brennan, DDS, Lindsay Munn, PhD RN, Keerti L. Dantuluri, MD, Timothy Hetherington, MS, Lauren C. Lu, Connell Dunn, Melanie Hogg, MS, CCRA, Andrea Price, Marina Leonidas, Laura Staton, Kenniesha Spencer, MPH, Melinda Manning, Whitney Rossman, MS, Frank X. Gohs, MS, Anna Harris, MPH, Bella Gutnik MS, Jennifer S. Priem, PhD, MA, Ryan Burns, MS. **MedStar Health Research Institute:** Chris Washington, Allison Moses, Sarahfaye Dolman, Julissa Zelaya-Portillo, John Erkus, Joseph Blumenthal, Ronald E. Romero Barrientos, Sonita Bennett, Shrenik Shah, Shrey Mathur, Christian Boxley, Paul Kolm, PhD, Eva Hochberger, Ella Franklin, Deliya Wesley, Naheed Ahmed. **Tulane**: Richard Oberhelman*, MD, Joseph Keating*, PhD, Patricia Kissinger, PhD, John Schieffelin, MD, Joshua Yukich, PhD, Andrew Beron, MPH, Devin Hayes, BS, Johanna Teigen, MPH. **University of Maryland School of Medicine:** Karen Kotloff*, MD, Wilbur H. Chen*, MD, MS, Andrea A. Berry, MD, Helen Powell, PhD, Lynnee Roane, MS, RN, Reva Datar, MPH. **University of Mississippi**: Leandro Mena*, MD, MPH, Bhagyashri Navalkele, MD, Yuan-I Min, PhD, Alexandra Castillo, MPH, Lori Ward, PhD, MS, Robert P. Santos, MD, Courtney Gomillia, MS-PHS, Pramod Anugu, Yan Gao, MPH, Jason Green, Ramona Sandlin, RHIA, Donald Moore, MS, Lemichal Drake, Dorothy Horton, RN, Kendra L. Johnson, MPH. **Wake Med Health and Hospitals:** William H. Lagarde*, MD, LaMonica Daniel, BSCR. **New Hanover:** Patrick D. Maguire*, MD, Charin L. Hanlon, MD, Lynette McFayden, RN, Isaura Rigo, MD, Kelli Hines, Lindsay Smith, Alexa Drilling, Monique Harris, Belinda Lissor, Vivian Cook, Maddy Eversole, Terry Herrin, Dennis Murphy, Lauren Kinney, Polly Diehl, Nicholas Abromitis, Tina St. Pierre, Judy Kennedy BSCS,

MBA, Lauren Kinney, BS, Bill Heckman, Denise Evans, Vivian Cook, Maddy Eversole, Julian March, Ben Whitlock, Wendy Moore. **Vidant Health:** Thomas R. Gallaher*, MD, Shakira Henderson, PhD, DNP, MS, MPH, Michael Zimmer, PhD, Danielle Oliver, Tina Dixon, Kasheta Jackson, Martha Reavis, Monica Menon, Brandon Bishop, Rachel Roeth, Mathew Johanson, Alesia Ceaser, Amada Fernandez, Carmen Williams, Jeremiah Hargett, Keeaira Boyd, Kevonna Forbes, Latasha Thomas, Markee Jenkins, Monica Coward, Derrick Clark, Omeshia Frost, Angela Darden, Lakeya Askew, Sarah Phipps, Victoria Barnes. **Campbell University School of Osteopathic Medicine**: Robin King-Thiele*,DO, Terri S. Hamrick*, Phd, Abdalla Ihmeidan, MHA (Campbell University School of Osteopathic Medicine), Chika Okafor, MD (Cape Fear Valley Medical Center), Regina B. Bray Brown, MD (Harnett Health System, Inc.), Poornima Vinod, MD (Southeastern Health), Amber Brewster, MD (Harnett Health System), Danius Bouyi, DO (Harnett Health System), Katrina Lamont, MD (Harnett Health System), Kazumi Yoshinaga, DO (Harnett Health System), A. Suman Peela, MD (Southeastern Health System), Giera Denbel, MD (Southeastern Health System), Jason Lo, MD: Southeastern Health System, Mariam Mayet-Khan, DO (Southeastern Health System), Akash Mittal, DO (Southeastern Health System), Reena Motwani, MD (Southeastern Health System), Mohamed Raafat, MD (Southeastern Health System), Evan Schultz, DO (Cumberland County Hospital System, Cape Fear Valley), Aderson Joseph, MD (Cumberland County Hospital System, Cape Fear Valley), Aalok Parkeh, DO (Cumberland County Hospital System, Cape Fear Valley), Dhara Patel, MD (Cumberland County Hospital System, Cape Fear Valley), Babar Afridi, DO (Cumberland County Hospital System, Cape Fear Valley). **George Washington University Data Coordinating Center:** Sharon L Edelstein, ScM, Greg Strylewicz, PhD, Meghan Turney, MA, Shirley Qin Zhou, MS, Ashley H. Tjaden, MPH, Matthew Bott. **Oracle Corporation:** Rebecca Laborde. **Sneez LLC:** Lucy Guill, MBA, Danielle Lamphier, MBA, Anna Schaefer, MSM, William M. Satterwhite, JD, MD. **Vysnova Partners:** Anne McKeague, PhD, Grace Tran, MPH, Johnathan Ward, Joyce Dieterly, MPH, Nana Darko, MPH, Kimberly Castellon, Isabella Malcolm, Ryan Brink, MS. **Javara Inc:** Atira Goodwin. **External Advisory Council:** Ruth Berkelman, MD, Emory, Kimberly Hanson, MD, U of Utah, Scott Zeger, PhD, Johns Hopkins, Cavan Reilly, PhD, U. of Minnesota, Kathy Edwards, MD, Vanderbilt, Helene Gayle, MD MPH, Chicago Community Trust (*Site Principal Investigator).
Ethical approval: The Wake Forest School of Medicine Institutional Review Board approved all study activiities.
Informed consent: Informed consent was obtained from all individuals included in this study.
Author contributions: MS and MMC designed the analyses and wrote the manuscript. MS conducted all analyses. All authors contributed to the interpretation of results and edited the manuscript for content. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: Authors state no conflict of interest.
Research funding: The North Carolina COVID-19 Community Research Partnership gratefully acknowledges the commitment and dedication of the study participants. Programmatic, laboratory, and technical support was provided by Vysnova Partners, Inc., Javara, Inc., and Oracle Corporation. This publication was supported by the CARES Act of the U.S. Department of Health and Human Services (HHS) [Contract # NC DHHS GTS #49927].
Data availability: All code for this project can be included in an online supplement. Data from the North Carolina COVID-19 Community Research Partnership is available through a standardized request process to the Coordinating Center and P & P Committee.

References

Accorsi, E. K., X. Qiu, E. Rumpler, L. Kennedy-Shaffer, R. Kahn, K. Joshi, E. Goldstein, M. J. Stensrud, R. Niehus, M. Cevik, and M. Lipsitch. 2021. "How to Detect and Reduce Potential Sources of Biases in Studies of SARS-CoV-2 and COVID-19." *European Journal of Epidemiology* 36 (2): 179—96.

Bajema, K. L., R. E. Wiegand, K. Cuffe, S. V. Patel, R. Iachan, T. Lim, A. Lee, D. Moyse, F. P. Havers, H. Lee, A. M. Fry, A. J. Hall, K. Martin, M. Biel, Y. Deng, W. A. Meyer, M. Mathur, T. Kyle, A. V. Gundlapalli, N. J. Thornburg, L. R. Petersen, and C. Edens. 2021. "Estimated SARS-CoV-2 Seroprevalence in the US as of September 2020." *JAMA Internal Medicine* 181 (4): 450—60.

Basto-Abreu, A., M. Carnalla, L. Torres-Ibarra, M. Romero-Martínez, J. Martínez-Barnetche, I. López-Martínez, R. Aparicio-Antonio, T. Shamah-Levy, C. Alpuche-Aranda, J. A. Rivera, T. Barrientos-Gutierrez, L. Cuevas-Nasu, E. B. Gaona-Pineda, M. A. Ávila-Arcos, F. Reyes-Sánchez, R. Torres-Álvarez, N. López-Olmedo, D. Vidaña-Perez, R. González-Morales, D. Barrera-Nuñez, C. Perez-Ferrer, C. Gaspar-Castillo, D. Stern, F. Canto-Osorio, A. Sanchez-Pájaro, and Ensanut-Covid Collaborators. 2022. "Nationally Representative SARS-CoV-2 Antibody Prevalence Estimates after the First Epidemic Wave in Mexico." *Nature Communications* 13 (1): 589.

Blyth, C. R., and H. A. Still. 1983. "Binomial Confidence Intervals." *Journal of the American Statistical Association* 78 (381): 108—16.

Bryant, J. E., A. S. Azman, M. J. Ferrari, B. F. Arnold, M. F. Boni, Y. Boum, K. Hayford, F. J. Luquero, M. J. Mina, and I. Rodriguez-Barraquer. 2020. "Serology for SARS-CoV-2: Apprehensions, Opportunities, and the Path Forward." *Science Immunology.* 5 (47): eabc6347.

Centers for Disease Control. 2022. "Self-Testing at Home or Anywhere." Last Modified 9-6-2022. https://www.cdc.gov/coronavirus/2019-ncov/testing/self-testing.html (accessed October 10, 2022).

Connor, B. A., M. Rogova, J. Garcia, D. Patel, M. Couto-Rodriguez, D. Nagy-Szakal, and M. Rendel. 2022. "Comparative Effectiveness of Single vs Repeated Rapid SARS-CoV-2 Antigen Testing Among Asymptomatic Individuals in a Workplace Setting." *JAMA Network Open* 5 (3): e223073.

COVID-19 Community Research Partnership. 2022. "The COVID-19 Community Research Partnership: A Multistate Surveillance Platform for Characterizing the Epidemiology of the SARS-CoV-2 Pandemic." *Biology Methods and Protocols* 7 (1): bpac033.

Frederick National Laboratory for Cancer Research. 2020. *Federal Drug Administration Report. June 21, 2020 Serology Test Evaluation Report for "Quik Pac II COVID-19 IgG IgM Test"*. Tianjin New Bay Bioresearch C. https://www.accessdata.fda.gov/cdrh_docs/presentations/maf/maf3298-a001.pdf.

Frederick National Laboratory for Cancer Research. 2021. *Federal Drug Administration Report. March 8, 2021 Serology Test Evaluation Report for "Scanwell SARS-CoV-2 IgM IgG Test"*. Teco Diagnostics. https://www.accessdata.fda.gov/cdrh_docs/presentations/maf/maf3434-a001.pdf.

Gelman, A., and B. Carpenter. 2020. "Bayesian Analysis of Tests with Unknown Specificity and Sensitivity." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 69 (5): 1269—83.

Hanson, K. E., A. M. Caliendo, C. A. Arias, M. K. Hayden, J. A. Englund, M. J. Lee, M. Loeb, R. Patel, A. El Alayli, O. Altayar, P. Patel, Y. Falck-Ytter, V. Lavergne, R. L. Morgan, M. H. Murad, S. Sultan, A. Bhimraj, and R. A. Mustafa. "Infectious Diseases Society of America Guidelines on the Diagnosis of COVID-19." https://www.idsociety.org/practice-guideline/covid-19-guideline-diagnostics/ (accessed January 2, 2023).

Hui, S. L., and S. D. Walter. 1980. "Estimating the Error Rates of Diagnostic Tests." *Biometrics* 36 (1): 167—71.

Infectious Diseases Society of America. 2020. *IDSA COVID-19 Antibody Testing Primer*. Last Modified 5-4-2020. https://www.idsociety.org/globalassets/idsa/public-health/covid-19/idsa-covid-19-antibody-testing-primer.pdf (accessed June 24, 2020).

Kucirka, L. M., S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler. 2020. "Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure." *Annals of Internal Medicine* 173 (4): 262—7.

Munawar, I., A. L. Seals, J. W. Sanders, D. M. Herrington, and T. F. Wierzba. 2021. "1213. Vaccine Uptake Amongst Participants in the North Carolina COVID-19 Community Research Partnership Who Were Initially Receptive or Hesitant to Receive a COVID-19 Vaccine." *Open Forum Infectious Diseases* 8 (Supplement_1): S696—7.

Offord, C. 2020. *How (Not) to Do an Antibody Survey for SARS-CoV-2*. The Scientist: News & Opinion. Last Modified 5-28-2020. https://www.the-scientist.com/news-opinion/how-not-to-do-an-antibody-survey-for-sars-cov-2-67488 (accessed June 24, 2020).

Perez-Saez, J., M.-E. Zaballa, S. Yerly, D. O. Andrey, B. Meyer, I. Eckerle, J.-F. Balavoine, F. Chappuis, D. Pittet, D. Trono, O. Kherad, N. Vuilleumier, L. Kaiser, I. Guessous, S. Stringhini, and A. S. Azman. 2021. "Persistence of Anti-SARS-CoV-2 Antibodies: Immunoassay Heterogeneity and Implications for Serosurveillance." *Clinical Microbiology and Infection* 27 (11): 1695.

Pollán, M., B. Pérez-Gómez, R. Pastor-Barriuso, J. Oteo, M. A. Hernán, M. Pérez-Olmeda, J. L Sanmartín, A. Fernández-García, I. Cruz, N. Fernández de Larrea, M. Molina, F. Rodríguez-Cabrera, M. Martín, P. Merino-Amador, J. León Paniagua, J. F. Muñoz-Montalvo, F. Blanco, R. Yotti, F. Blanco, R. Gutiérrez Fernández, M. Martín, S. Mezcua Navarro, M. Molina, J. F. Muñoz-Montalvo, M. Salinero Hernández, J. L. Sanmartín, M. Cuenca-Estrella, R. Yotti, J. León Paniagua, N. Fernández de Larrea, P. Fernández-Navarro, B. Pérez-Gómez, M. Pollán, A. Avellón, G. Fedele, A. Fernández-García, J. Oteo Iglesias, M. T. Pérez Olmeda, I. Cruz, M. E. Fernandez Martinez, F. D. Rodríguez-Cabrera, M. A. Hernán, S. Padrones Fernández, J. M. Rumbao Aguirre, J. M. Navarro Marí, B. Palop Borrás, A. B. Pérez Jiménez, M. Rodríguez-Iglesias, A. M. Calvo Gascón, M. L. Lou Alcaine, I. Donate Suárez, O. Suárez Álvarez, M. Rodríguez Pérez, M. Cases Sanchís, C. J. Villafáfila Gomila, L. Carbo Saladrigas, A. Hurtado Fernández, A. Oliver, E. Castro Feliciano, M. N. González Quintana, J. M. Barrasa Fernández, M. A. Hernández Betancor, M. Hernández Febles, L. Martín Martín, L. M. López López, T. Ugarte Miota, I. De Benito Población, M. S. Celada Pérez, M. N. Vallés Fernández, T. Maté Enríquez, M. Villa Arranz, M. Domínguez-Gil González, I. Fernández-Natal, G. Megías Lobón, J. L. Muñoz Bellido, P. Ciruela, A. Mas i Casals, M. Doladé Botías, M. A. Marcos Maeso, D. Pérez del Campo, A. Félix de Castro, R. Limón Ramírez, M. F. Elías Retamosa, M. Rubio González, M. S. Blanco Lobeiras, A. Fuentes Losada, A. Aguilera, G. Bou, Y. Caro, N. Marauri, L. M. Soria Blanco, I. del Cura González, M. Hernández Pascual, R. Alonso Fernández, P. Merino-Amador, N. Cabrera Castro, A. Tomás Lizcano, C. Ramírez Almagro, M. Segovia Hernández, N. Ascunce Elizaga, M. Ederra Sanz, C. Ezpeleta Baquedano, A. Bustinduy Bascaran, S. Iglesias Tamayo, L. Elorduy Otazua, R. Benarroch Benarroch, J. Lopera Flores, and A. Vázquez de la Villa. 2020. "Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A Nationwide, Population-Based Seroepidemiological Study." *The Lancet* 396 (10250): 535—44.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Soni, A., C. Herbert, H. Lin, C. Pretz, P. Stamegna, B. Wang, T. Orwig, C. Wright, S. Tarrant, S. Behar, T. Suvarna, S. Schrader, E. Harman, C. Nowak, V. Kheterpal, L. V. Rao, L. Cashman, E. Orvek, D. Ayturk, L. Gibson, A. Zai, S. Wong, P. Lazar, Z. Wang, A. Filippaios, B. Barton, C. J. Achenbach, R. L. Murphy, M. L. Robinson, Y. C. Manabe, S. Pandey, A. Colubri, L. O'Connor, S. C. Lemon, N. Fahey, K. L. Luzuriaga, N. Hafer, K. Roth, T. Lowe, T. Stenzel, W. Heetderks, J. Broach, and D. D. McManus. 2023. "Performance of Rapid Antigen

Tests to Detect Symptomatic and Asymptomatic SARS-CoV-2 Infection : A Prospective Cohort Study." *Annals of Internal Medicine* 176 (7): 975—82.

Stringhini, S., A. Wisniak, G. Piumatti, A. S. Azman, S. A. Lauer, H. Baysson, D. De Ridder, D. Petrovic, S. Schrempft, K. Marcus, S. Yerly, I. Arm Vernez, O. Keiser, S. Hurst, K. M. Posfay-Barbe, D. Trono, D. Pittet, L. Gétaz, F. Chappuis, I. Eckerle, N. Vuilleumier, B. Meyer, A. Flahault, L. Kaiser, and I. Guessous. 2020. "Seroprevalence of Anti-SARS-CoV-2 IgG Antibodies in Geneva, Switzerland (SEROCoV-POP): A Population-Based Study." *The Lancet* 396 (10247): 313—9.

Takahashi, S., B. Greenhouse, and I. Rodríguez-Barraquer. 2020. "Are Seroprevalence Estimates for Severe Acute Respiratory Syndrome Coronavirus 2 Biased?" *The Journal of Infectious Diseases* 222 (11): 1772—5.

United States Food & Drug Administration. 2022. "At-Home COVID-19 Antigen Tests-Take Steps to Reduce Your Risk of False Negative: FDA Safety Communication." Last Modified 8-11-2022. https://www.fda.gov/medical-devices/safety-communications/home-covid-19-antigen-tests-take-steps-reduce-your-risk-false-negative-fda-safety-communication (accessed October 10, 2022).

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/em-2023-0012).