SARS-CoV-2 Serosurveys: How Antigen, Isotype and Threshold Choices Affect the Outcome

Raquel A. Binder,[1,a] Gavin F. Fujimori,[1,a] Catherine S. Forconi,[1,a] George W. Reed,[1] Leandro S. Silva,[1] Priya Saikumar Lakshmi,[1] Amanda Higgins,[2] Lindsey Cincotta,[2] Protiva Dutta,[2] Marie-Claire Salive,[2] Virginia Mangolds,[2] Otuwe Anya,[2] J. Mauricio Calvo Calle,[3] Thomas Nixon,[4] Qiushi Tang,[4] Mireya Wessolossky,[1] Yang Wang,[5] Dominic A. Ritacco,[1] Courtney S. Bly,[1] Stephanie Fischinger,[6] Caroline Atyeo,[6] Peter O. Oluoch,[7] Boaz Odwar,[1] Jeffrey A. Bailey,[8] Ana Maldonado-Contreras,[7] John P. Haran,[2,7] Aaron G. Schmidt,[6,9] Lisa Cavacini,[5] Galit Alter,[6] and Ann M. Moormann[1]

[1]Department of Medicine, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA; [2]Department of Emergency Medicine, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA; [3]Department of Pathology, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA; [4]Horae Gene Therapy Center, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA; [5]MassBiologics, University of Massachusetts Medical School, Boston, Massachusetts, USA; [6]Ragon Institute of MGH, MIT, and Harvard, Cambridge, Massachusetts, USA; [7]Department of Microbiology and Physiological Systems, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA; [8]Department of Pathology and Laboratory Medicine, Warren Alpert Medical School, Brown University, Providence, Rhode Island, USA; and [9]Department of Microbiology, Harvard Medical School, Boston, Massachusetts, USA

Background. Evaluating the performance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) serological assays and clearly articulating the utility of selected antigens, isotypes, and thresholds is crucial to understanding the prevalence of infection within selected communities.

Methods. This cross-sectional study, implemented in 2020, screened PCR–confirmed coronavirus disease 2019 patients (n = 86), banked prepandemic and negative samples (n = 96), healthcare workers and family members (n = 552), and university employees (n = 327) for anti–SARS-CoV-2 receptor-binding domain, trimeric spike protein, and nucleocapsid protein immunoglobulin (Ig)G and IgA antibodies with a laboratory-developed enzyme-linked immunosorbent assay and tested how antigen, isotype and threshold choices affected the seroprevalence outcomes. The following threshold methods were evaluated: (i) mean + 3 standard deviations of the negative controls; (ii) 100% specificity for each antigen-isotype combination; and (iii) the maximal Youden index.

Results. We found vastly different seroprevalence estimates depending on selected antigens and isotypes and the applied threshold method, ranging from 0.0% to 85.4%. Subsequently, we maximized specificity and reported a seroprevalence, based on more than one antigen, ranging from 9.3% to 25.9%.

Conclusions. This study revealed the importance of evaluating serosurvey tools for antigen-, isotype-, and threshold-specific sensitivity and specificity, to interpret qualitative serosurvey outcomes reliably and consistently across studies.

Keywords. ELISA; Massachusetts; SARS-CoV-2; antigen; isotype; serosurvey; threshold; Youden index.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) seroprevalence studies (i.e., serosurveys) are essential public health tools that can be incorporated into disease spread models to estimate the SARS-CoV-2 effective reproductive number, SARS-CoV-2 transmission potentials, disease dynamic forecasts, and assess the impact of public health and clinical interventions, which are of particularly importance in the beginning of an outbreak [1].

However, given the number of SARS-CoV-2 serosurvey tools, evaluating test performance and clearly articulating their utility across different study populations has become crucial. This is particularly true for laboratory-developed serological assays that are not designed for clinical use, were not given emergency use authorization by the US Food and Drug Administration, and whose makeup and performance vary [2]. Ideally, serosurvey tools are evaluated for their reproducibility, sensitivity, and specificity based on PCR–confirmed coronavirus disease 2019 (COVID-19) clinical specimens and well-characterized prepandemic negative controls [3]. Furthermore, thresholds for each detected antigen-isotype combination and potential compound measurements (e.g., final positive call based on more than one SARS-CoV-2 antigen) must be evaluated to allow for qualitative assessments (i.e., providing results that are either positive or negative for the antibodies of interest). While the present study focuses on SARS-CoV-2, the results can be leveraged for any infectious disease that requires the development and standardization of serological tools and respective threshold methods for qualitative

outcomes, especially (re)emerging infectious diseases that require the rapid development of reliable serological assays.

Here, we describe the evaluation and implementation of a laboratory-developed SARS-CoV-2 enzyme-linked immunosorbent assay (ELISA) [4] based on samples from hospitalized COVID-19 patients (CHCs), healthcare workers (HCWs) and their family members (HCW Family), and return-to-work employees (RTWs) during the first COVID-19 wave in Worcester, Massachusetts (MA), from April to August 2020. Overall, this study aims to demonstrate the importance of (i) rapidly funding academic medical centers to design, evaluate, and deploy laboratory-developed tests in an outbreak setting, and (ii) evaluating serosurvey tools for antigen-, isotype-, and threshold-specific sensitivity and specificity, in order to interpret seroprevalence measurements reliably and consistently across study populations.

METHODS

Study Participant Recruitment and Enrollment
Study participants were enrolled at the University of Massachusetts Chan Medical School (UMass Chan) and the University of Massachusetts Memorial Medical Center (UMass Memorial) in Worcester, MA between April and August of 2020 under the institutional review board–approved consolidated COVID-19 Clinical and Observational Pathogenesis and Epidemiology (COVID-COPE) study protocol (H00020145; see the Supplementary Methods for more details). Briefly, once consent was obtained, participants were asked to complete a survey capturing demographic and symptomatic information, and blood samples were collected.

The HCWs consisted of front-line workers from the UMass Memorial emergency department and their family members (HCW Family). The RTW employees consisted of UMass Chan employees who returned to campus in August after the state-mandated stay-at-home order. In addition, patients with COVID-19–like symptoms who were being evaluated at the UMass Memorial emergency department were approached for study participation (COVID-19 Hospitalized Cases [CHC]). Once the participants consented, the patients or their healthcare proxy completed a survey and blood samples were collected. COVID-19–related clinical information, including symptoms, were extracted from the participants' medical charts. Finally, de-identified banked prepandemic and prescreened blood samples negative for anti–SARS-CoV-2 antibody served as negative controls (total n = 96; see Supplementary Materials). RTW employees, HCWs, and HCW Family participants received antibody results (Supplementary Figure 1), with the caveat that this was a research assay and not a diagnostic test.

SARS-CoV-2 ELISA
The SARS-CoV-2 ELISA developed at the Ragon Institute was implemented at UMass Chan [4], using receptor-binding domain (RBD), spike trimer (S; truncated to 1208 amino acids, deleted transmembrane domain and C-terminal domain), and nucleocapsid protein (N; full-length) SARS-CoV-2 antigens provided by MassBiologics of UMass Chan, Mattapan, MA [5]. We also used SARS-CoV2 RBD antigen provided by the Ragon Institute, Cambridge, MA. (See Supplementary Methods for more details and validation.)

Statistical Analysis
Statistical calculations and graphs were done using Prism (version 8.2.1), R (version 3.6.1 or 4.1.1), and Stata (version 17) software. The heat map was generated by the pheatmap package in R software (version 4.1.1). Three thresholds were evaluated; see the Supplementary Methods for more details. First, we applied the threshold formula "mean + 3 standard deviations of negative controls" (3 SD) threshold [6]. Next, we applied a maximum specificity (Max Spec) threshold by setting the cutoff at 100% specificity for each antigen-isotype combination (i.e., with the cutoff value being the highest optical density [OD] value among the negative control group for each antigen-isotype combination). Finally, we determined a threshold based on the maximal Youden index (sensitivity + specificity − 1) and the receiver operating characteristic (ROC) curve (Supplementary Figure 2), which balances sensitivity and specificity for each antigen-isotype combination (Youden threshold). Experimental samples with OD values equal to or higher than the determined threshold value for each antigen-isotype combination were considered serologically positive (Supplementary Table 1).

RESULTS

Study Participant Recruitment
A total of 212 suspected COVID-19 patients (CHCs) were approached for study participation when they were being evaluated in the emergency department, and 134 (63.2%) were subsequently enrolled. Of those, 86 (64.2%) were PCR-confirmed COVID-19 cases (Table 1). The majority of the CHCs were white (60.4% [n = 81]) and non-Hispanic (65.7% [n = 88]), aged ≥60 years (59.7% [n = 80]; average age, 67 years), and male (47.8% [n = 64]). The latter two determinants (increased age and being male) are common COVID-19 risk factors and consistent with the demographics of those severely affected during the first COVID-19 wave in MA [7].

A total of 253 HCWs volunteered to participate in the study, and 299 of their family members (HCW Family) were also enrolled. Based on self-reported data, the majority of HCWs were white (76.3% [n = 193]) and non-Hispanic (80.2% [n = 203]). The average age of the HCWs was 41 years (range, 21–68 years), and 49.8% (n = 126) were female. Among the HCW family members who reported demographic information, 54.2% (n = 162) were white, 57.5% (n = 172) non-Hispanic, and 33.4% (n = 100) female. A total of 327 RTW employees

Table 1. Demographics, SARS-CoV-2 PCR Results and Symptom Distribution Among Study Participants (n = 1109)

Characteristic	Participants, No. (%)				
	CHC	HCW	HCW Family	RTW	Negatives[a]
Total	134 (12.1)	253 (22.8)	299 (27.0)	327 (29.5)	96 (8.7)
Age(years)					
0–17	0 (0.0)	0 (0.0)	26 (8.7)	0 (0.0)	0 (0.0)
18–30	4 (3.0)	45 (17.8)	49 (16.4)	85 (26.0)	19 (19.8)
31–40	5 (3.7)	69 (27.3)	36 (12.0)	103 (31.5)	7 (7.3)
41–50	12 (9.0)	59 (23.3)	33 (11.0)	52 (15.9)	3 (3.1)
51–60	19 (14.2)	32 (12.7)	23 (7.7)	66 (20.2)	4 (4.2)
61–70	18 (13.4)	13 (5.1)	13 (4.3)	18 (5.5)	4 (4.2)
≥71	58 (43.3)	0 (0.0)	3 (1.0)	2 (0.6)	0 (0.0)
Missing	18 (13.4)	35 (13.8)	116 (38.8)	1 (0.3)	59 (61.5)[b]
Gender					
Male	64 (47.8)	91 (36.0)	87 (29.1)	141 (43.1)	30 (31.3)
Female	52 (38.8)	126 (49.8)	100 (33.4)	184 (56.3)	27 (28.1)
Non-binary	0 (0.0)	0 (0.0)	1 (0.3)	1 (0.3)	0 (0.0)
Missing	18 (13.4)	36 (14.2)	111 (37.1)	1 (0.3)	39 (40.6)
Race					
White	81 (60.4)	193 (76.3)	162 (54.2)	229 (70.0)	51 (53.1)
Black	10 (7.5)	6 (2.4)	2 (0.7)	4 (1.2)	3 (3.1)
Asian	4 (3.0)	13 (5.1)	13 (4.3)	71 (21.7)	3 (3.1)
Other	19 (14.2)	2 (0.8)	3 (1.0)	18 (5.5)	0 (0.0)
Missing	20 (14.9)	39 (15.4)	119 (39.8)	5 (1.5)	39 (40.6)
Ethnicity					
Hispanic	20 (14.9)	11 (4.4)	5 (1.7)	19 (5.8)	5 (5.2)
Non-Hispanic	88 (65.7)	203 (80.2)	172 (57.5)	304 (93.0)	52 (54.2)
Missing	26 (19.4)	39 (15.4)	122 (40.8)	4 (1.2)	39 (40.6)
SARS-COV-2 PCR test result					
Positive	86 (64.2)	84 (33.2)	34 (11.4)	8 (2.4)	NA
Negative	44 (32.8)	135 (53.4)	149 (49.8)	318 (97.3)	NA
Missing	4 (3.1)	34 (13.4)	116 (38.8)	1 (0.3)	NA
COVID-19 symptoms[c]					
≥1	134 (100.0)	87 (34.4)	80 (26.8)	29 (8.9)	NA
None	0 (0.0)	105 (41.5)	81 (27.1)	1 (0.3)	NA
Missing or NA	0 (0.0)	61 (24.1)	138 (46.2)	297 (90.8)	NA

Abbreviations: CHCs, COVID-19 hospitalized case; COVID-19, coronavirus disease 2019; HCWs, healthcare workers; HCW Family, family members of enrolled HCWs; NA, not applicable. RTW, return-to-work employees who had been working remotely from March to August 2020, during the first COVID-19 wave in Worcester, Massachusetts; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

[a]SARS-CoV-2–negative samples: prepandemic and pre-screened SARS-CoV-2 negative samples.

[b]Among these 59 participants who provided negative samples but were missing age-specific information, 17 were adults (aged ≥18 years), one was in his 20s, and two who were in their 30s, but with no exact ages available.

[c]The RTW, HCW, and HCW Family members were asked to report symptoms only if they had a positive SARS-COV-2 test result or had reason to believe they were infected. The following COVID-19 symptoms were reported: cough, fever or chills, shortness of breath or difficulty breathing, loss of taste or smell, sore throat, headache, unexplained muscle or body aches, unusual weakness or fatigue, chilblains, and/or diarrhea.

volunteered for the study, of whom the majority were white (70.0% [n = 229]) and non-Hispanic (93.0% [n = 304]). About half of them were female (56.3% [n = 184]), and the average age was 40 years (range, 22–73 years). Hence, a total of 1013 study participants were successfully enrolled; demographic and symptomatic information was recorded, blood samples were collected, and serum or plasma samples were screened for anti–SARS-CoV-2 antibodies.

Among these study participants, the missing data ranged from 0.3% to 40.8% (Table 1), depending on the demographic category or study population subgroup. Note that this does not include the COVID-19 symptom reports, as

RTWs, HCWs, and HCW Family participants were asked to report COVID-19 symptoms only if they had a positive SARS-COV-2 PCR test or had reason to believe they were infected. HCW Family members were the least likely to answer demographic questions and therefore had the most missing data, despite repeated contact attempts. RTWs were most likely to fill out all the full questionnaires.

Threshold Methods and Sensitivity/Specificity

PCR-confirmed SARS-CoV-2–positive CHCs (n = 86) and negative samples (n = 96) were screened by means of ELISA for anti-RBD, anti-S, and anti-N SARS-CoV-2

immunoglobulin (Ig)A and IgG antibodies. We further evaluated (i) merging the three antigen-isotype combinations (RBD IgG, N IgG, and RBD IgA) with the largest area under the ROC curve (i.e., with maximum sensitivity and specificity —"Top Three"); (ii) combining all IgA isotypes (N, RBD and S IgA—"all IgA"); and (iii) combining all IgG isotypes (N, RBD and S IgG—"all IgG") as serological outcome measures (Figure 1). We used isotype-based compound measurements because a heat map of CHC, HCW, and RTW-derived OD values with at least one positive antigen-isotype combination clustered by isotype, rather than antigen (Figure 2). The clustering was confirmed with a correlogram comparing antigen-isotype combinations (Supplementary Figure 3). Again, when analyzing PCR-confirmed CHCs or all seropositive participants, we found the strongest positive correlation between the same isotypes rather than the same antigens (Supplementary Figure 3). This clustering of IgG and IgA responses to SARS-CoV-2 infection has been described by others [8, 9] and is likely a function of time since infection [10] and class switching [8].

Applying the 3 SD threshold resulted in 100% specificity for all isotype-antigen combinations across IgG and IgA (Table 2). Conversely, sensitivity was low, ranging from 25.6% to 76.7%.

Sensitivities for RBD IgA and IgG—25.6% and 55.8%, respectively—were particularly low. Sensitivities for IgA and IgG S trimer and IgA and IgG N were similar, ranging from 72.1% to 76.7%. The RBD IgM antigen-isotype combination resulted in relatively low sensitivity (range, 41.9%–93.0% across threshold methods), with specificities ranging from 93.7% to 100.0% across threshold methods, similar to previously published results [4, 11]. Given the short duration of IgM production, while appearing concomitantly with IgG after SARS-CoV-2 infection [3], IgM was not deemed a useful serosurvey tool and was omitted from the rest of the analyses.

Based on the Max Spec threshold (fixing the specificity at 100%), sensitivities for RBD IgA and IgG increased compared with the 3 SD threshold but remained relatively low, at 74.4% and 64.0%, respectively. Again, sensitivities for IgA and IgG S trimer and IgA and IgG N were similar and ranged from 81.4% to 84.9%

Based on the Youden threshold, robust sensitivity was obtained across all antigen-isotype combinations, ranging from 84.9% to 97.7%. Sensitivities for RBD IgA and IgG were 96.5% and 97.7%, respectively. Sensitivities for IgA and IgG S trimer and IgA and IgG N ranged from 84.9% to 94.2%. Conversely, specificities were slightly lower than with the other two methods and ranged from 92.7% to 99.0% (Table 2).
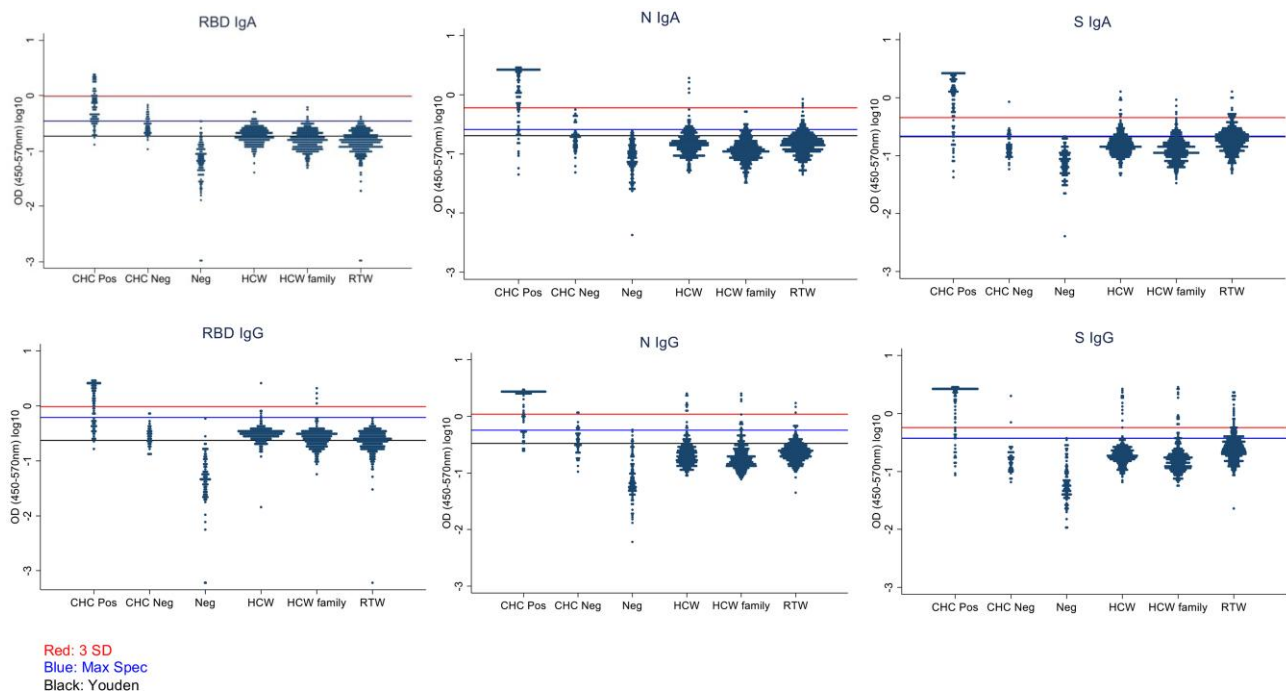

Red: 3 SD
Blue: Max Spec
Black: Youden

Figure 1. **Optical density (OD) value distribution among each measured subgroup**. OD values (450-570nm) were plotted according to each antigen-isotype combination and subgroup (y-axis in log10 scale). The top horizontal line in the graph indicates the cut off for the 3 standard deviation above the mean threshold method (3 SD). The second line indicates the cut off when the threshold was chosen at the highest value of the negative controls (Max Spec). The third line indicates the cut off when the threshold was chosen based on the Youden threshold (Youden). Note, the last two lines converge for S IgA and S IgG. Abbreviations: CHC Neg, PCR-negative CHCs (COVID-19 Hospitalized Case); CHC Pos, PCR confirmed hospitalized COVID-19 patients; HCW, health care workers; HCW Family, family members of health care worker listed under HCW; N, SARS-CoV-2 nucleocapsid protein; Neg, negative controls; RBD, SARS-CoV-2 receptor-binding domain; RTW, return to work employees who had been working remotely from March to August of 2020, during the first COVID-19 wave in Worcester, MA; S Trimer; SARS-CoV-2 spike trimer.

Serological Results for HCWs, HCW Family, and RTW Employees

Next, we screened HCW (n = 253), HCW Family (n = 299), and RTW (n = 327) participants for anti-RBD, anti-S, and anti-N SARS-CoV-2 IgA and IgG antibodies with the ELISA method. The overall OD value spread for the HCW, HCW Family, and RTW subgroups was similar (Figure 1). However, the overlap between the PCR-positive CHCs (positive controls) and negative control values varied notably depending on the antigen-isotype combination, affecting the cutoffs for each method and therefore the subsequently determined seroprevalence for each subgroup. Furthermore, most negative control OD values were notably lower than the experimental subgroups (HCW, HCW Family, and RTW) for RBD, S, and N IgG but not so much for the IgA equivalents. Hence, when applying the three different thresholds to the experimental subgroups the seroprevalence outcomes varied widely, not just across threshold methods (Youden, 3 SD, and Max Spec) but also across antigen-isotype combinations (Figure 1, Figure 3, and Supplementary Figure 4).

When we compared the overall seroprevalences for the antigen-isotype combinations, RBD IgA, RBD IgG, and the Top Three compound estimate (RBD IgG, RBD IgA, and N IgG) yielded the highest seroprevalences, especially when applying the Youden threshold, ranging from 29.7% to 85.4% (Table 3 and Figure 3). The S IgA– and All IgA–based seroprevalences ranged from 12.5% to 38.6% when applying the Youden threshold (Table 3). Overall, the N IgA–, N IgG–, and S IgG–based seroprevalences were lowest, ranging from 0.0% to 20.6%, independent of applied threshold; note the wide range of seroprevalence outcomes (Table 3).

When we compared the three participant subgroups based on the Youden threshold, the HCWs had the highest seroprevalence followed by the HCW Family and RTW groups when measuring RBD IgA (51.4%, 34.1%, and 29.7%, respectively) and RBD IgG (85.4%, 62.2%, and 52.9%, respectively) (Table 3 and Figure 3). For the other antigen-isotype combinations, the order was less pronounced, except for S IgA, where the RTW subgroup had the highest seroprevalence based on the Youden and Max Spec thresholds, compared to the HCW and HCW Family subgroups.

When we compared the threshold methods, applying the Youden cutoff, which was the most sensitive (highest average sensitivity across all thresholds and isotype combinations), yielded the highest seroprevalences across all antigen-isotype combinations, though with a wide range (6.0%–85.4%; Table 3 and Figure 3). Furthermore, while the RBD IgA–, RBD IgG–, and the Top Three compound–based seroprevalence estimates were highest when applying the Youden threshold method, they also represented the most notable discrepancy when comparing the Youden with the Max Spec and 3 SD thresholds. Based on the Youden threshold method, the RBD IgG–based seroprevalences among all three subgroups ranged

from 52.9% to 85.4%, as compared to the 3 SD and Max Spec thresholds, where the seroprevalences ranged from 0.0% to 2.0%.

Given the notable variation in OD spread among the antigen-isotype combinations and the resulting wide range of seroprevalence estimates (including the overinflated seroprevalence based on RBD IgA/IgG when applying the Youden threshold; Figure 3), maximizing the specificity and including
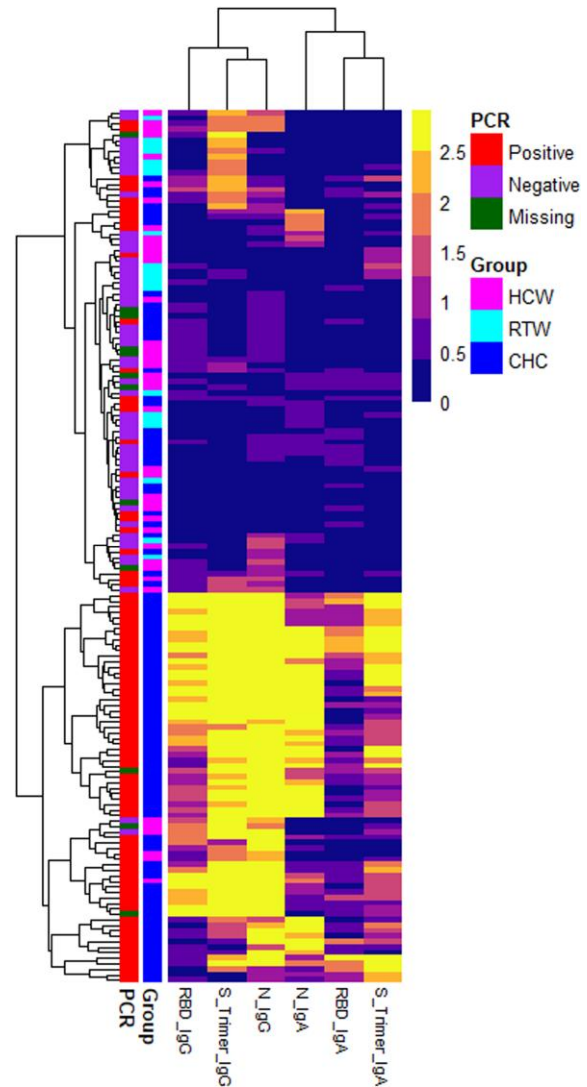

Figure 2. Heat map of optical density (OD) values. A heat map was built based on unbiased clustering of the OD values of the experimental subgroups with at least one positive antigen-isotype combination. The OD values are represented in the color scale, ranging from blue to yellow. The results clustered by isotype, rather than antigen (see top branching into immunoglobulin [Ig]G and IgA from left to the right). Abbreviations: CHC, COVID-19 hospitalized case; HCW, healthcare workers and their family members; N, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid protein; RBD, SARS-CoV-2 receptor-binding domain; RTW, return-to-work employees who had been working remotely from March to August 2020, during the first COVID-19 wave in Worcester, Massachusetts; S trimer, SARS-CoV-2 spike trimer.

Table 2. Sensitivity and Specificity for Each Threshold Method and Isotype-Antigen Combination based on PCR-confirmed Hospitalized Coronavirus Disease 2019 Patients (n = 86) and Negative Samples (n = 96)

Threshold Method[a]	Antigen	IgA		IgG		IgM	
		Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
3 SD	RBD	25.6	100.0	55.8	100.0	67.4	98.4
	S trimer	75.6	100.0	76.7	100.0	NA	NA
	N	74.4	100.0	72.1	100.0	NA	NA
Max Spec	RBD	74.4	100.0	64.0	100.0	41.9	100.0
	S trimer	84.9	100.0	84.9	100.0	NA	NA
	N	83.7	100.0	81.4	100.0	NA	NA
Youden	RBD	96.5	93.7	97.7	97.9	93.0	93.7
	S trimer	84.9	99.0	84.9	99.0	NA	NA
	N	89.5	94.8	94.2	92.7	NA	NA

Abbreviations: Ig, immunoglobulin; Max Spec, maximum specificity; N, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid protein; NA, not applicable; RBD, SARS-CoV-2 receptor-binding domain; S trimer; SARS-CoV-2 spike trimer; SD, standard deviation.

[a]The 3SD threshold was determined by calculating the mean + 3 SDs for negative controls; the Max Spec threshold, by setting the cutoff at the highest value of the negative controls for each antigen-isotype combination; and the Youden threshold, by determining the maximum Youden index on the ROC curve.

more than one antigen (compound measurements) for the overall serosurvey measurement was prioritized.

Based on the Top Three compound estimate and the Max Spec threshold, we estimated that the seroprevalences among the HCW, HCW Family, and RTW study participants were 14.7%, 10.1%, and 9.3%, respectively (HCWs ranking highest; Table 3 and Figure 3). Based on the All IgG compound estimate and the Max Spec threshold, we estimated that the seroprevalences among the HCW, HCW Family, and RTW study participants were 15.9%, 11.5%, and 25.9%, respectively (with RTW employees ranking highest). Hence, based on these two compound estimates and the Max Spec threshold, we estimated that the anti–SARS-CoV-2 antibody seroprevalences among the HCW, HCW Family, and RTW study participants ranged from 9.3% to 25.9% during the first COVID-19 wave in the spring of 2020, while the ranking was unclear owing to relatively low overall seroprevalence. Note that at the time no COVID-19 vaccines were available, so detection of any SARS-CoV-2 antibodies was considered an indication of infection.

Finally, participant COVID-19 symptoms and test reports were analyzed to validate the serological results. As part of the participant surveys, HCWs, HCW Family, and RTW participants were asked to report COVID-19–like symptoms if they either had a positive SARS-CoV-2 PCR result or were thought to have been infected (Table 1). Among HCW, HCW Family, and RTW participants, HCWs reported the most positive SARS-CoV-2 PCR results (33.2% [n = 84]) and COVID-19 symptoms (34.4% [n = 87]), followed by HCW Family and RTW participant groups (Table 1), thus confirming a potential HCW, HCW Family, and RTW group ranking. Interestingly, among the RTW employees, HCWs, and HCW Family members who reported SARS-CoV-2–positive PCR results, only 54.8% (n = 69) reported COVID-19 symptoms,

while 32.5% (n = 41) reported no symptoms, and 12.7% (n = 16) did not answer that question. Note that molecular SARS-CoV-2 tests were not widely available outside hospital settings during the study period.

DISCUSSION

Given the number of serological SARS-CoV-2 assays, evaluating test performance and ideal cutoffs, along with clearly articulating their utility across different study populations is crucial. In this study, we validated a laboratory developed SARS-CoV-2 ELISA along with three different methods to transform quantitative OD values into qualitative seroprevalence estimates.

The Youden index is a commonly applied threshold method and conventional ROC summary measure designed to provide optimal separation of negative and positive values and therefore maximize sensitivity and specificity [4, 12, 13]. Fixing the specificity at 100% (i.e., Max Spec threshold) and using 3 SDs above the mean of the negative controls are also commonly applied serological threshold methods and are designed to maximize specificity [13–16]. Hence, we chose methods that have been evaluated by others and either maximize the sensitivity-specificity balance or maximize specificity.

The present study found notable variation in seroprevalence estimates among healthy study populations when comparing the three threshold methods, with estimates ranging from 0.0% to 85.4%. The seroprevalence outcomes were lowest and varied least across antigen-isotype combinations based on the 3 SD and Max Spec thresholds. Although applying the Youden threshold increased sensitivity, which is important in a clinical setting, this threshold yielded unrealistically high seroprevalence estimates within our community-based serosurvey, up to 85.4% (RBD IgG in HCWs), and these varied significantly among antigen-isotype combinations. For
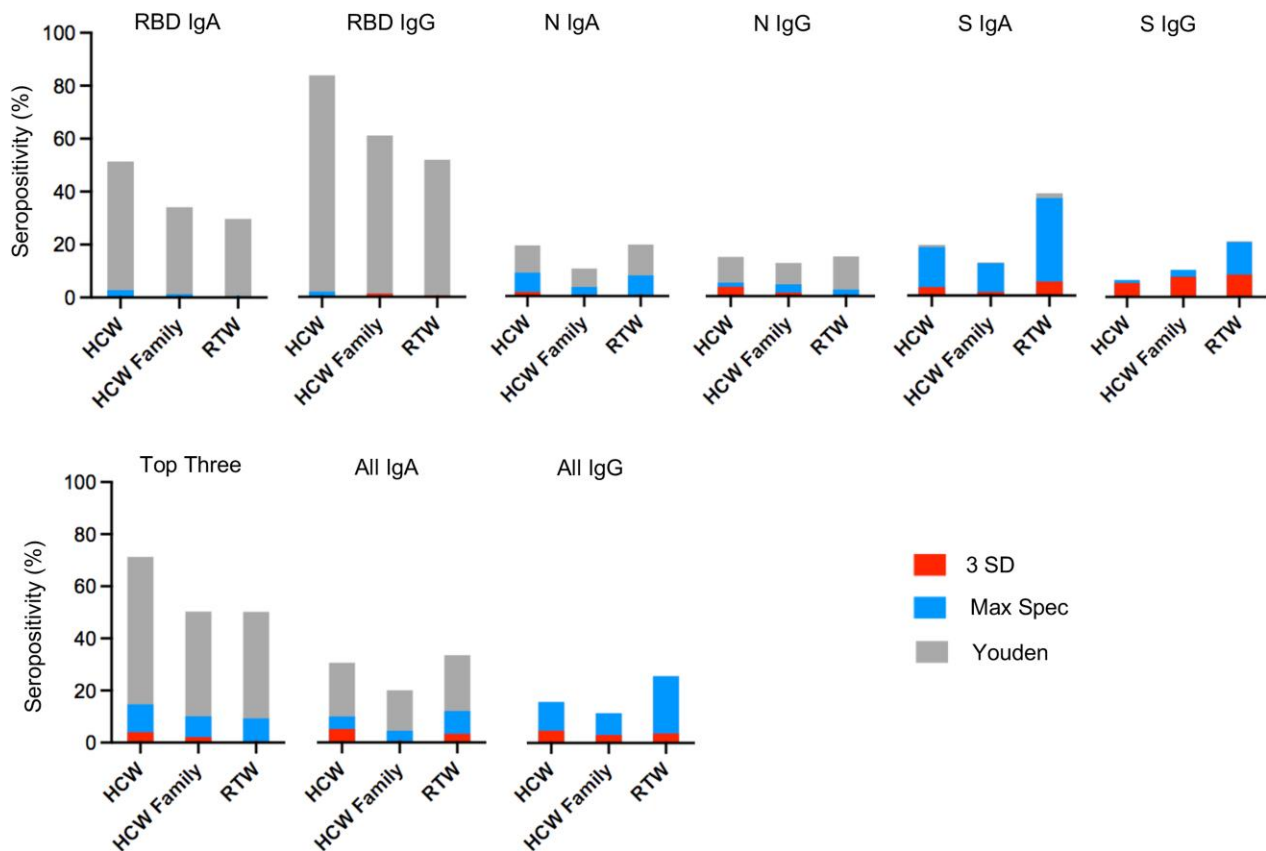
Figure 3. Seropositivity of each subgroup, according to the antigen-isotype combination. The red bars represent the seroprevalence based on the threshold defined as the mean + 3 SDs among negative controls (3 SD). The blue bars represent the seroprevalence based on the maximum specificity threshold (Max Spec). The black bars represent the seroprevalence based on the Youden threshold (Youden). Abbreviations: HCWs, healthcare workers; HCW Family , family members of participating HCWs; Ig, immunoglobulin; N, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid protein; RBD, SARS-CoV-2 receptor-binding domain; RTW, return-to-work employees who had been working remotely from March to August of 2020, during the first coronavirus disease 2019 wave in Worcester, Massachusetts; S trimer, SAR-S-CoV-2 spike trimer; Top Three, combination of the three antigen/isotype combinations (RBD IgG, N IgG, and RBD IgA) with the largest area under the ROC curve (i.e., with maximum sensitivity and specificity).

example, the RBD IgG–based seroprevalence among the HCW, HCW Family, and RTW groups ranged from 52.9% to 85.4% when applying the Youden threshold and from 0.0% to 2.0% when applying the Max Spec threshold (Table 3).

While such discrepancies were most drastic among RBD-based seroprevalence estimates and were unrealistically high for both RBD IgA and IgG when applying the Youden threshold, the seroprevalence outcomes were not as different when comparing the three thresholds for N IgA, N IgG, and S IgG. These results emphasized the importance of maximizing assay specificity, as generally done in population-based studies, and using compound measurements for our final serosurvey outcome (i.e., population-based exposure status as determined by seroprevalence). The Top Three (93.0% sensitivity, 100.0% specificity) and the all IgG (94.2% sensitivity, 100.0% specificity) compound estimates, combined with the Max Spec threshold, exhibited the least antigen- or isotype-dependent variation and were considered the most reliable outcome measurements to determine seroprevalence in our study population

(Table 3). Based on these two compound estimates and the Max Spec threshold, we estimated that the anti–SARS-CoV-2 antibody seroprevalence among the HCW, HCW Family, and RTW study participants ranged from 9.3% to 25.9% during the first COVID-19 wave in the spring of 2020, while the ranking was unclear owing to relatively low overall seroprevalence.

A similar hospital-based study found seroprevalence estimates of 5.5% among HCWs in Boston, MA, in July of 2020 [17]. Another Massachusetts population-based study found seroprevalence estimates of 31.5% in April 2020 [18]. However, the difference could be due to the serosurvey tool, as the study used an IgM-IgG point-of-care lateral flow immunoassay. A US-wide blood donation-based SARS-CoV-2 seroprevalence study found that infection-induced seroprevalence estimates increased from 3.5% to 11.5% between July and December 2020, and it estimated the northeast of the US (including MA) to have reached a seroprevalence of 19.3% by May of 2021 when the study ended [19]. Hence, our HCW, HCW Family, and

Table 3. Seroprevalence for Each Participant Group, Antigen-Isotype Combination, and Threshold Method

Isotype-Antigen Combination	Threshold Method[a]	Seroprevalence, % (No.)		
		HCW (n = 253)	HCW Family (n = 299)	RTW (n = 327)
RBD IgA	Youden	51.4 (130)	34.1 (102)	29.7 (97)
	Max Spec	2.8 (7)	1.3 (4)	0.9 (3)
	3 SD	0.0 (0)	0.0 (0)	0.0 (0)
RBD IgG	Youden	85.4 (216)	62.2 (186)	52.9 (173)
	Max Spec	2.0 (5)	1.3 (4)	0.0 (0)
	3 SD	0.4 (1)	1.3 (4)	0.0 (0)
S trimer IgA	Youden	19.1 (48)	12.5 (36)	38.6 (124)
	Max Spec	18.3 (46)	12.2 (35)	36.8 (118)
	3 SD	3.2 (8)	1.4 (4)	5.3 (17)
S trimer IgG	Youden	6.0 (15)	7.3 (21)	20.6 (66)
	Max Spec	6.0 (15)	7.3 (21)	20.2 (65)
	3 SD	4.8 (12)	4.5 (13)	7.2 (23)
N IgA	Youden	19.9 (50)	10.8 (31)	20.2 (65)
	Max Spec	9.2 (23)	3.5 (10)	8.1 (26)
	3 SD	1.6 (4)	0.0 (0)	0.9 (3)
N IgG	Youden	15.1 (38)	12.8 (37)	15.3 (49)
	Max Spec	5.2 (13)	4.5 (13)	2.5 (8)
	3 SD	3.6 (9)	1.4 (4)	0.9 (3)
Top Three	Youden	71.3 (179)	50.3 (145)	50.2 (161)
	Max Spec	14.7 (37)	10.1 (29)	9.3 (30)
	3 SD	4.0 (10)	2.1 (6)	0.9 (3)
All IgA	Youden	30.7 (77)	20.1 (58)	33.6 (108)
	Max Spec	10.0 (25)	4.5 (13)	12.1 (39)
	3 SD	5.2 (13)	1.0 (3)	3.4 (11)
All IgG	Youden	15.9 (40)	11.5 (33)	25.9 (83)
	Max Spec	15.9 (40)	11.5 (33)	25.9 (83)
	3 SD	4.8 (12)	3.1 (9)	3.7 (12)

Abbreviations: HCW, healthcare workers; HCW Family, family members of participating HCWs; Ig, immunoglobulin; Max Spec, maximum specificity; N, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid protein; RBD, SARS-CoV-2 receptor-binding domain; RTW, return-to-work employees who had been working remotely from March to August 2020, during the first COVID-19 wave in Worcester, Massachusetts; S trimer; SARS-CoV-2 spike trimer; SD, standard deviation; Top Three, combination of the three antigen/isotype combinations (RBD IgG, N IgG, and RBD IgA) with the largest area under the ROC curve (i.e., with maximum sensitivity and specificity).

[a]The 3SD threshold was determined by calculating the mean + 3 SDs for negative controls; the Max Spec threshold, by setting the cutoff at the highest value of the negative controls for each antigen-isotype combination; and the Youden threshold, by determining the maximum Youden index on the ROC curve.

RTW populations fall into the estimated ranges in similar time periods.

Longitudinal studies have found that SARS-CoV-2 IgA and IgG antibodies are first detected within similar time frames post-symptom onset (PSO), but IgA production ceases earlier. [20, 21] IgA seroreversion has been reported between two and six months PSO, while IgG antibodies are detected up to eight months PSO. [20, 21] Despite the short time frame of the study during the early phase of the pandemic, within five months of the first COVID-19 wave in MA, the discrepancies between antigen-isotype combinations may have been due to individual variation in kinetics of natural antibody production among study participants or the respective half-lives of circulating antigen-specific antibodies. For example, SARS-CoV-2 IgG persistence and titer have been found to depend on disease severity. [22] Another study found a substantial decline in anti-N SARS-CoV-2 antibodies over a five-month period as part of a Canadian serosurveillance program [23]. Furthermore, others have shown that S-specific serum IgA levels decay significantly

($P < .002$) faster than S-specific IgG after COVID-19 messenger RNA vaccination [10]. Others have found differences in post-vaccination antibody production depending on sex or age [24]. Hence, the choice of (i) measured antigen-isotype combination, (ii) threshold method, and (iii) presumably positive and negative controls can affect the positive threshold determination and therefore the qualitative outcomes of a serological assays.

Our study had limitations. The study participants were mostly white and non-Hispanic. Others have observed significant differences in the breadth and strength of the humoral immune response in relation to ethnicity [25]. Hence, our results may not apply to more diverse populations. Furthermore, our results may not be applied to currently circulating SARS-CoV-2 variants since the study samples were collected during the early phase of the pandemic. Overall, there was a general lack of molecular tests during active infection for non-hospitalized cases. Hence, we were not able to confirm whether our study included asymptomatic infections or

potential non-seroconverters and whether the lack of antibodies may have been a function of time since infection. Similarly, there was only limited information on days since symptom onset for the CHCs, which would have allowed determining sensitivity and specificity in terms of time since symptom onset.

Furthermore, there could have been self-selection bias among HCW, HCW Family, and RTW groups whereby people who thought they had been infected or exposed to SARS-CoV-2 were more likely to enroll in our study and therefore artificially inflate the seroprevalence. Finally, ELISAs are being increasingly replaced by multiplex bead assays. However, methods to define seropositivity still need to be developed [26]. Note that in the COVID-19 vaccine era, serological assays are designed to distinguish between vaccination (i.e., anti-S or anti-RBD only antibodies) and infection (i.e., anti-S, anti-RBD and/or anti-N antibodies), even though one has to consider potential cross-reactivity with endemic human coronaviruses, especially OC43 and HKU1, which are most closely related to SARS-CoV-2 [27].

In summary, the present study found notable variation among seroprevalence outcomes depending on the antigen-isotype combination and the chosen threshold method, and we found that compound estimate-based outcomes calculated with the Max Spec threshold were most reliable. Robust serosurveys are powerful public health tools that can determine the extent of previous SARS-CoV-2 infections or vaccination rates among populations with different exposure risks. Knowing the fraction of susceptible individuals within a specific population assists with (i) establishing effective risk assessments, (ii) implementing specific infection control measures, and (iii) determining the effectiveness of those measures over time. This is especially important in regions outside the United States, where vaccinations are not widely implemented and longitudinal serosurveys will help monitor changes in transmission patterns. However, to obtain reliable seroprevalence estimates, serosurvey tools need to be evaluated for antigen-, isotype-, and threshold-specific sensitivity and specificity, in order to interpret qualitative serosurvey outcomes reliably and consistently across study populations.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

References

1. Larremore DB, Fosdick BK, Bubar KM, et al. Estimating SARS-CoV-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys. Elife **2021**; 10:e64206.
2. Farnsworth CW, Anderson NW. SARS-CoV-2 serology: much hype, little data. Clin Chem **2020**; 66:875–7.
3. Centers for Disease Control and Prevention. Interim guidelines for COVID-19 antibody testing. https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html. Accessed 1 July 2022.
4. Roy V, Fischinger S, Atyeo C, et al. SARS-CoV-2-specific ELISA development. J Immunol Methods **2020**; 484–485: 112832.
5. Ejemel M, Li Q, Hou S, et al. A cross-reactive human IgA monoclonal antibody blocks SARS-CoV-2 spike-ACE2 interaction. Nat Commun **2020**; 11:4198.
6. Lardeux F, Torrico G, Aliaga C. Calculation of the ELISA's cut-off based on the change-point analysis method for detection of *Trypanosoma cruzi* infection in Bolivian dogs in the absence of controls. Mem Inst Oswaldo Cruz **2016**; 111: 501–4.
7. Archive of COVID-19 cases in Massachusetts: Massachusetts Department of Public Health, Boston, MA, **2020**.
8. Zohar T, Loos C, Fischinger S, et al. Compromised humoral functional evolution tracks with SARS-CoV-2 mortality. Cell **2020**; 183:1508–19e12.
9. Ivanov A, Semenova E. Long-term monitoring of the development and extinction of IgA and IgG responses to SARS-CoV-2 infection. J Med Virol **2021**; 93:5953–60.
10. Wisnewski AV, Campillo Luna J, Redlich CA. Human IgG and IgA responses to COVID-19 mRNA vaccines. PloS One **2021**; 16:e0249499.

11. Landry ML. Immunoglobulin M for acute infection: true or false? Clin Vaccine Immunol **2016**; 23:540–5.

12. Plebani M, Padoan A, Negrini D, Carpinteri B, Sciacovelli L. Diagnostic performances and thresholds: the key to harmonization in serological SARS-CoV-2 assays? Clin Chim Acta **2020**; 509:1–7.

13. Plebani M, Padoan A, Sciacovelli L, Basso D. Towards the rational utilization of SARS-CoV-2 serological tests in clinical practice. Clin Chem Lab Med **2020**; 58:e189–e91.

14. Lew TTS, Aung KMM, Ow SY, et al. Epitope-functionalized gold nanoparticles for rapid and selective detection of SARS-CoV-2 IgG antibodies. ACS Nano **2021**; 15(7): 12286–97.

15. Halpern MD, Jain S, Jewett MW. Enhanced detection of host response antibodies to Borrelia burgdorferi using mmune-PCR. Clin Vaccine Immunol **2013**; 20:350–7.

16. Yap G, Pok KY, Lai YL, et al. Evaluation of chikungunya diagnostic assays: differences in sensitivity of serology assays in two independent outbreaks. PloS Negl Trop Dis **2010**; 4:e753.

17. Kataria Y, Cole M, Duffy E, et al. Seroprevalence of SARS-CoV-2 IgG antibodies and risk factors in health care workers at an academic medical center in Boston, Massachusetts. Sci Rep **2021**; 11:9694.

18. Naranbhai V, Chang CC, Beltran WFG, et al. High seroprevalence of anti-SARS-CoV-2 antibodies in Chelsea, Massachusetts. J Infect Dis **2020**; 222:1955–9.

19. Jones JM, Stone M, Sulaeman H, et al. Estimated US infection- and vaccine-induced SARS-CoV-2 seroprevalence based on blood donations, July 2020–May 2021. JAMA **2021**; 326:1400–9.

20. Cohen KW, Linderman SL, Moodie Z, et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. Cell Rep Med **2021**; 2:100354.

21. Ortega N, Ribes M, Vidal M, et al. Seven-month kinetics of SARS-CoV-2 antibodies and role of pre-existing antibodies to human coronaviruses. Nat Commun **2021**; 12:4740.

22. Yan X, Chen G, Jin Z, et al. Anti-SARS-CoV-2 IgG levels in relation to disease severity of COVID-19. J Med Virol **2022**; 94:380–3.

23. Bolotin S, Tran V, Osman S, et al. SARS-CoV-2 seroprevalence survey estimates are affected by anti-nucleocapsid antibody decline. J Infect Dis **2021**; 223:1334–8.

24. Ward H, Whitaker M, Flower B, et al. Population antibody responses following COVID-19 vaccination in 212,102 individuals. Nat Commun **2022**; 13:907.

25. Smith M, Abdesselem HB, Mullins M, et al. Age, disease severity and ethnicity influence humoral responses in a multi-ethnic COVID-19 cohort. Viruses **2021**; 13::786.

26. Chan Y, Fornace K, Wu L, et al. Determining seropositivity-a review of approaches to define population seroprevalence when using multiplex bead assays to assess burden of tropical diseases. PLoS Negl Trop Dis **2021**; 15: e0009457.

27. Liu DXL, Liang JQ Fung TS. Human coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). Encycl Virol **2021**:428–40. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7204879/