Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in *All of Us* Research Program Participants, 2 January to 18 March 2020

Keri N. Althoff,[1,2] David J. Schlueter,[3] Hoda Anton-Culver,[4] James Cherry,[5] Joshua C. Denny,[2] Isaac Thomsen,[6] Elizabeth W. Karlson,[7] Fiona P. Havers,[8] Mine S. Cicek,[9] Stephen N. Thibodeau,[9] Ligia A. Pinto,[10] Douglas Lowy,[5] Bradley A. Malin,[6] Lucila Ohno-Machado,[11] Carolyn Williams,[12] David Goldstein,[13] Aymone Kouame,[6] Andrea Ramirez,[2] Adrienne Roman,[6] Norman E. Sharpless,[5] Kelly A. Gebo,[14,a] and Sheri D. Schully[2,a]; on behalf of the *All of Us* Research Program

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [2]*All of Us* Research Program, National Institutes of Health, Bethesda, Maryland, USA; [3]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA; [4]Department of Medicine, School of Medicine, University of California, Irvine, Irvine, California, USA; [5]National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; [6]Vanderbilt University Medical Center, Nashville, Tennessee, USA; [7]Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts, USA; [8]Centers for Disease Control and Prevention, Atlanta, Georgia, USA; [9]Mayo Clinic, Rochester, Minnesota, USA; [10]Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, Maryland, USA; [11]School of Medicine, University of California, San Diego, San Diego, California, USA; [12]National Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda, Maryland, USA; [13]Columbia University Medical Center, New York, New York, USA; and [14]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Background. With limited severe acute respiratory syndrome coronavirus (SARS-CoV-2) testing capacity in the United States at the start of the epidemic (January–March 2020), testing was focused on symptomatic patients with a travel history throughout February, obscuring the picture of SARS-CoV-2 seeding and community transmission. We sought to identify individuals with SARS-CoV-2 antibodies in the early weeks of the US epidemic.

Methods. *All of Us* study participants in all 50 US states provided blood specimens during study visits from 2 January to 18 March 2020. Participants were considered seropositive if they tested positive for SARS-CoV-2 immunoglobulin G (IgG) antibodies with the Abbott Architect SARS-CoV-2 IgG enzyme-linked immunosorbent assay (ELISA) and the EUROIMMUN SARS-CoV-2 ELISA in a sequential testing algorithm. The sensitivity and specificity of these ELISAs and the net sensitivity and specificity of the sequential testing algorithm were estimated, along with 95% confidence intervals (CIs).

Results. The estimated sensitivities of the Abbott and EUROIMMUN assays were 100% (107 of 107 [95% CI: 96.6%–100%]) and 90.7% (97 of 107 [83.5%–95.4%]), respectively, and the estimated specificities were 99.5% (995 of 1000 [98.8%–99.8%]) and 99.7% (997 of 1000 [99.1%–99.9%]), respectively. The net sensitivity and specificity of our sequential testing algorithm were 90.7% (97 of 107 [95% CI: 83.5%–95.4%]) and 100.0% (1000 of 1000 [99.6%–100%]), respectively. Of the 24 079 study participants with blood specimens from 2 January to 18 March 2020, 9 were seropositive, 7 before the first confirmed case in the states of Illinois, Massachusetts, Wisconsin, Pennsylvania, and Mississippi.

Conclusions. Our findings identified SARS-CoV-2 infections weeks before the first recognized cases in 5 US states.

Keywords: SARS-CoV-2; United States; Epidemic; Immunoglobulin G antibodies; *All of Us* Research Program.

Among the first 12 known cases of severe acute respiratory syndrome coronavirus (SARS-CoV-2) infection in the United States, the earliest recognized symptoms onset date was 14 January 2020, and all 12 case patients had recently traveled to mainland China or were close contacts of recent travelers [1]. Domestic testing for SARS-CoV-2 began in mid-January 2020, and the Food and Drug Administration granted emergency use authorization for real-time reverse-transcription polymerase chain reaction (RT-PCR) testing at the Centers for Disease Control and Prevention (CDC) on 4 February 2020 [2]. With limited testing capacity, testing throughout February was focused on symptomatic patients with a positive travel history. Many states recognized their first confirmed cases in the last week of February or the first week of March 2020.

Principles infectious disease epidemiology indicate that low-level circulation of the pathogen before the recognized outbreak was likely [3]. Phylogenetic analyses suggested evolution of the SARS-CoV-2 virus between October and December 2019 [4–6]. The deaths of 2 California residents infected with SARS-CoV-2 virus on 6 and 17 February 2020, and an infected passenger or crew member on a cruise ship that left San Francisco on 11 February, suggest that the virus was present in California in February, although a signal in syndromic surveillance was not detected until the end of February [7]. A 2020 study in blood donation specimens collected among residents of 9

states between 13 December and 17 January found antibodies to SARS-CoV-2 in the United States as early as mid-December 2019 [8].

Determining the presence and location of SARS-CoV-2 in the earliest days of the US pandemic, together with other information on the spread and severity of coronavirus disease 2019 (COVID-19) illness, is important for understanding of the emergence of the virus, the epidemiology of this virus, and informing simulation models used to predict cases, deaths, and healthcare utilization and subsequently guide future pandemic planning, policy development, and resource allocation.

Serologic evidence of SARS-CoV-2 antibodies can identify individuals who have been infected with SARS-CoV-2, including those who were asymptomatic or had subclinical illness, who are typically undetected at the start of an infectious disease epidemic. Serologic assays differ in their targets and in their sensitivity and specificity; previous studies have compared commercially available assays using the same positive and negative control specimens [9]. The CDC recommends minimizing false-positive antibody test results using a sequential testing approach ("employing two independent tests in sequence when the first test yields a positive result"), particularly when the prevalence of SARS-Cov-2 is expected to be low [10]. The objective of our study was to determine seroprevalence of SARS-CoV-2 immunoglobulin G (IgG) antibodies among participants in the *All of Us* cohort, from whom a blood specimen was collected during study visits occurring at the start of 2020 (from 2 January to 18 March 2020).

METHODS AND MATERIALS

The *All of Us* Research Program is an observational cohort study that, beginning in May 2018, enrolls diverse adults (>18 years of age) in the United States from study site locations in all 50 states, with a goal of enrolling at least 1 million participants, and has been described elsewhere [11]. Briefly, the program includes genomic measurements and the large sample size needed for precision medicine research. Participants are enrolled after an informed consent process at clinics and regional medical centers that compose the *All of Us* Research Program network. Biospecimens collected from participants during *All of Us* study visits occurring from the first day in 2020 (2 January) to the day when in-person visits were paused owing to the SARS-CoV-2 public health emergency (18 March 2020). Participants who had not withdrawn from the *All of Us* study and had a serum specimen deemed acceptable by the biobank were eligible to participate in our nested study. Alaskan Native and American Indian participants were excluded from this analysis, at the request of tribal leaders. The *All of Us* internal review board approved our nested study.

We withdrew frozen serum specimens from the *All of Us* biorepository and prepared for serologic testing, starting with specimens collected in approximately 2-week intervals from 18 March 2020, moving backward in time until there was a week with no positive specimens or 2 January 2020, whichever came first. This strategy was agreed on before testing of any specimens and was necessary owing to the costs associated with testing this large sample size. The testing was conducted through a contract by Quest Diagnostics in an environment that meets the Clinical Laboratory Improvement Amendments regulations. We selected the Abbott Architect SARS-CoV-2 IgG enzyme-linked immunosorbent assay (ELISA) (target: nucleocapsid; manufacturer's sensitivity, 89.3% [95% confidence interval (CI) 82.6%–93.7%]; manufacturer's specificity, 99.6% [99.0%–99.9%]) and the EUROIMMUN SARS-CoV-2 ELISA (target: spike protein; manufacturer's sensitivity, 90.0% [74.4%–96.5%]; manufacturer's specificity, 100% [95.4%–100%]). Laboratory testing began in May 2020 and continued through January 2021; we analyzed the results as they were completed.

We estimated the sensitivity, specificity, and Clopper-Pearson exact binomial 95% CIs for the Abbott and EUROIMMUNE ELISAs, using blinded positive control specimens from individuals who were hospitalized or discharged and convalescing with RT-PCR–confirmed SARS-CoV-2 infection (median duration [interquartile range] since symptom onset, 14 [11–18] days for 18 positive controls from Brigham and Women's Hospital, Boston, Massachusetts, 45 [41–55] days for 44 positive controls from Vanderbilt University Medical Center, Nashville, Tennessee, and unknown for 45 positive controls from the Mayo Clinic, Rochester, Minnesota). Negative control specimens came from *All of Us* participants who enrolled and had blood specimens collected in the year before the emergence of SARS-CoV-2, specifically from January to March 2019 (n = 1000). Up to 8 duplicate positive control specimens and 2 duplicate negative control specimens were used to ensure there were at least 1 positive and 1 negative control specimen on each plate.

Given the anticipated low prevalence of SARS-CoV-2–seropositive participants from 2 January to 18 March 2020, our definition of seropositive was that an individual must be seropositive with the Abbott assay and subsequently seropositive with the EUROIMMUN assay, consistent with the CDC guidance for sequential testing; for this testing algorithm, net sensitivity and specificity and Clopper-Pearson exact binomial 95% CIs were estimated [10]. In a sensitivity analysis of potential false-positive results, SARS-CoV-2 IgG nucleocapsid and spike protein titers were quantified at a National Cancer Institute research, non–Clinical Laboratory Improvement Amendments laboratory, using ELISAs developed in house. In addition, to estimate the probability of false-positives, we simulated 1 million participants and applied the lower 95% CI bounds of the sensitivity and specificity for the Abbott and EUROIMMUN assays and mimicked the sequential testing algorithm to estimate the maximum probability of a false-positive result with

our definition of seropositive (see Supplementary Materials for details of the simulation).

We report the date of specimen collection, state of residence, and characteristics (patient age at specimen collection, sex at birth, race, and ethnicity) of those meeting our definition of seropositive using the sequential testing algorithm. An exception was granted to the *All of Us* program's data and statistics dissemination policy to report individual test results [12]. We also reviewed the electronic health records data and the participant-reported survey data for those who met our seropositive definition for information regarding recognized respiratory illness.

RESULTS

The study included a total of 24 079 *All of Us* participants who had a biospecimen collected in our study period, had not withdrawn from the study, and had an acceptable specimen for testing. Participants were predominantly female sex at birth (57%); 49% were non-Hispanic white, 24% non-Hispanic Black/African American, and 17% Hispanic, Latino, or Spanish (Table 1). Participants resided in 50 states, with the largest number of

Table 1. Characteristics of *All of Us* Research Program Participants With Blood Specimens Collected From 2 January to 18 March 2020 and Available for Serologic Testing

Characteristic	Participants, No. (%)[a] (N = 24 079)
Age, median (IQR), y	53 (37–65)
Sex at birth	
Female	13 692 (57)
Male	10 100 (42)
Other sex at birth	8 (0)
Skip, prefer not to say, or no answer	279 (1)
Race and ethnicity[b]	
Asian	630 (3)
Non-Hispanic Black/African American	5712 (24)
Hispanic Black/African American	81 (0)
Non-Hispanic white	11 896 (49)
Hispanic white	279 (1)
Pacific Islander	14 (0)
Hispanic, Latino or Spanish	4059 (17)
Other race	785 (3)
Skip, prefer not to say, or no answer	623 (3)
Geographic area of US residence	
Northeast	6953 (29)
Southeast	6092 (25)
Midwest	5612 (23)
West	5408 (22)
US territories	14 (0)

Abbreviation: IQR, interquartile range.

[a]Data represent no. (%) of participants unless otherwise specified.

[b]Alaskan Native and American Indian participants were excluded from this analysis, at the request of tribal leaders.

participants residing in California (14%), Massachusetts (11%), Alabama (10%), Illinois (10%), Pennsylvania (8%), Arizona (8%), New York (8%), Wisconsin (7%), Florida (6%), or Michigan (5%, Figure 1).

Using our positive and negative controls, the estimated sensitivities of the Abbott and EUROIMMUN assays were 100% (107 of 107 [95% CI: 96.6%–100%]) and 90.7% (97 of 107 [83.5%–95.4%]), respectively, and the respective specificities were 99.5% (995 of 1000 [98.8%–99.8%]) and 99.7% (997of 1000 [99.1%–99.9%]), respectively. The net sensitivity and specificity of our sequential testing algorithm were 90.7% (97 of 107 [83.5%–95.4%]) and 100% (1000 of 1000 [99.6%–100%]), suggesting a low probability of false-positives. Of the 1000 negative controls, 5 were false-positives with the Abbott assay and 3 with the EUROIMMUN assay; no negative control samples were false-positive with our sequential method of testing (that is, no negative controls were false-positive with both assays).

Of the 147 *All of Us* participants with a positive Abbott result, 9 had samples that were subsequently positive with the EUROIMMUNE assay and met our definition of seropositive (Table 2). Seven of the 9 seropositive individuals were detectable before the first confirmed cases in the states of their residence. These included individuals with specimens collected on 7 January in Illinois, 8 January in Massachusetts, 3 February in Wisconsin, 15 February in Pennsylvania, and 6 March in Mississippi (Figure 2). Of the 2 seropositive participants who responded to the *All of Us* COVID-19 Participant Experience (COPE) survey, 1 reported experiencing fever, cough, sore or painful throat, and a belief they had COVID-19 within a reasonable time frame (2 weeks after) specimen collection, given the survey was administered in May 2020. Review of electronic health record data during the relevant time frame revealed that 2 seropositive participants had illnesses compatible with mild COVID-19 (eg, fatigue and mild respiratory symptoms), but additional testing was limited and no diagnosis was confirmed. The other 7 seropositive participants had no evidence of healthcare use in their electronic health record data.

In addition, in sensitivity analyses of SARS-CoV-2 IgG antibody titers, 7 of the 9 seropositive participants had titers above the limit of detection, including 1 with nucleocapsid IgG titers, 4 with spike IgG titers, and 3 with both nucleocapsid and spike IgG titers. Finally, we estimated the mean probability of false-seropositive results, given the low prevalence of SARS-CoV-2 during the study period through simulations. Using our prevalence of 0.00037 (9 of 24 079), the estimated mean probability that all 9 results were false-positive was 0.00001 across 1000 replications of the simulation study; there was a 0.00019, 0.00210, 0.01405, 0.06251, 0.19487, 0.43841, 0.72959, and 0.93492 probability that there were at least 8, 7, 6, 5, 4, 3, 2, and 1 false-positive results, respectively.

Figure 1. Numbers of *All of Us* participants with blood specimens collected from 2 January to 18 March 2020 and available for serologic testing, by state (N = 24 079). (Alaska and Hawaii each had <20 participants with specimens available).

DISCUSSION

Our retrospective study of blood specimens from *All of Us* participants collected 2 January to 18 March 2020 suggests evidence of SARS-CoV-2 infection weeks before recognition of the virus in Illinois, Wisconsin, Pennsylvania, Mississippi, and Massachusetts. As recommended by the CDC, our study used a rigorous sequential testing definition of seropositive to minimize false-positive results.

Table 2. Characteristics of 9 *All of Us* Participants With Severe Acute Respiratory Syndrome Coronavirus 2–Seropositive Blood Specimens Collected From 2 January to 18 March 2020

| Participant | Race/Ethnicity | Age Group, y | Sex at Birth | SARS-CoV-2 Assay | |
				Abbott (Cut-off, 1.4)	EUROIMMUN (Cut-off, 1.1)
1	Black/African American	45–49	Female	3.09	2.681
2	Black/African American	50–54	Male	3.99	1.416
3	Black/African American	55–59	Male	3.17	1.846
4	Black/African American	55–59	Female	1.83	2.487
5	Black/African American	60–64	Female	4.13	1.984
6	Hispanic, Latino, or Spanish	60–64	Female	4.01	1.544
7	Hispanic, Latino, or Spanish	70–74	Male	2.43	1.125
8	White	50–54	Female	3.76	1.449
9	White	65–69	Male	8.09	2.891

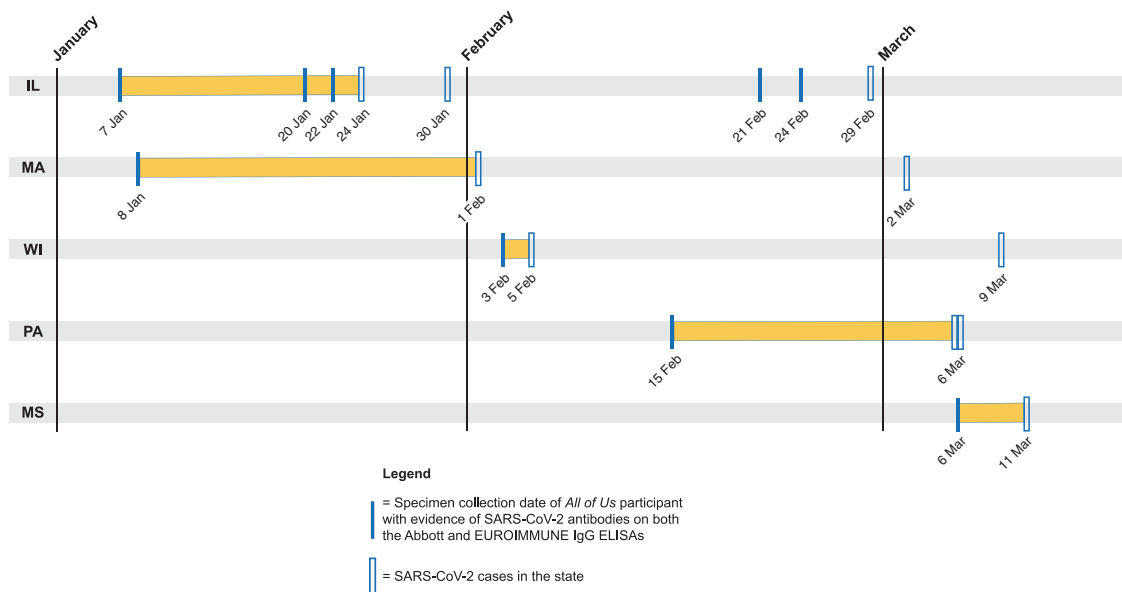Abbreviation: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

Figure 2. Date of blood specimen collection compared with date of the state's initial confirmed severe acute respiratory syndrome coronavirus (SARS-CoV-2) cases for 9 SARS-CoV-2–seropositive *All of Us* participants with specimens collected from 2 January to 18 March 2020. The first SARS-Cov-2 case in Illinois (IL) was confirmed on 24 January 2020 (https://www.dph.illinois.gov/news/city-chicago-announces-first-local-patient-travel-related-case-2019-novel-coronavirus), and the spouse of the first confirmed case patient had the second confirmed case, confirmed on 30 January (https://www.dph.illinois.gov/news/second-illinois-2019-novel-coronavirus-case-identified); the third case was a presumptive positive case announced on 29 February while awaiting confirmation by the Centers for Disease Control and Prevention (CDC) (https://www.dph.illinois.gov/news/state-illinois-public-health-officials-announce-new-presumptive-posi- tive-covid-19-case-illinois). The first SARS-CoV-2 case in Massachusetts (MA) was confirmed on 1 February 2020 (https://www.mass.gov/news/man-returning-from-wuhan-china-is-first-case-of-2019-novel-coronavirus-confirmed-in); the second case was confirmed on 2 March (https://www.mass.gov/news/first-presumptive-positive-case-of-covid-19-identified-by-massachusetts-state-public-health). The first 2 SARS-CoV-2 cases in Pennsylvania (PA) were presumptive positive (while awaiting CDC confirmation), on 6 March 2020 (https://www.governor.pa.gov/newsroom/wolf-administration-confirms-2-presumptive-positive-cases-of-covid-19/). The first SARS-Cov-2 case in Wisconsin (WI) was confirmed on 5 February 2020 (https://www.dhs.wisconsin.gov/news/releases/020520.htm), with the second case confirmed on 9 March (https://www.dhs.wisconsin.gov/news/releases/030920.htm). Finally, the first SARS-CoV-2 case in Mississippi (MS) was presumptive positive (while awaiting CDC confirmation), on 11 March 2020 (https://msdh.ms.gov/msdhsite/_static/23,21819,341.html).

The median time from SARS-CoV-2 infection to IgG antibody presence is 14 days [13–15]. Assuming that individuals who were seropositive according to our sequential definition were infected ≥2 weeks before biospecimen collection, our findings suggests that the virus may have been present in Illinois as early as 24 December 2019. Our findings of suspected SARS-CoV-2 infection in January in Wisconsin (n = 1 participant) and Massachusetts (n = 1) corroborate findings of a retrospective study of antibodies that were reactive in microneutralization with live SARS-CoV-2 (USA-WA1/2020) in Wisconsin (n = 3) and Massachusetts (n = 16) in blood donations made between 30 December 2019 and 17 January 2020 [8]. We found evidence of SARS-CoV-2 antibodies in Mississippi in early March 2020, before the state's first case on 11 March. A study of blood donors from California, Connecticut, Iowa, Massachusetts, Michigan, Oregon, Rhode Island, Washington, and Wisconsin used other laboratory approaches, including microneutralization tests and a receptor binding domain/angiotensin-converting enzyme 2 blocking activity assay. Our estimates of seropositive individuals using a sequential testing approach with commercial assays triangulate the findings from the blood donors, suggesting the robustness of the findings of seropositive individuals in both studies before the first state-recognized cases.

Furthermore, our findings expand the knowledge of undetected SARS-CoV-2 infections likely occurring in early January in Illinois and early February in Pennsylvania. These data suggest that SARS-CoV-2 infection was present in states far from the initial hot spots considered to be points of entry to the United States (Seattle, Washington, and New York City, New York). Although the virus was presumed to be circulating in New York City, Seattle, and the state of California, none of the *All of Us* participants in these states tested positive, perhaps due to the low (albeit increasing) and highly localized transmission from January through mid-March (and the smaller numbers of *All of Us* participants from the state of Washington).

Given the disproportionate burden of the subsequent US COVID-19 epidemic in minority populations, it is noteworthy that 7 of the 9 seropositive individuals were older minority participants [16]. Although the *All of Us* study is enriched with populations that are underrepresented in biomedical research, there was a disproportionate increased burden of seropositivity in Black/African Americans (5 of 9) and Hispanic, Latino, or Spanish participants (2 of 9), compared with the race and ethnicity distribution in the *All of Us* study population. Although the numbers are limited, these findings reinforce scientific hypotheses of the impact of social factors on viral circulation,

including structural discrimination against racial and ethnic minority groups.

The current study contributes to the evidence of low-level circulation of SARS-CoV-2 in many states at the start of the US epidemic. Federal testing recommendations included travel to a geographic area with known SARS-CoV-2 transmission or contact with a confirmed SARS-CoV-2 case patient until cases were confirmed in most states; the 13 June 2020 consolidation of testing recommendations did not include the travel epidemiologic link [17]. These epidemiologic links were particularly important at the beginning of the US pandemic when testing capacity was limited. Although our data suggest that antibody evidence of SARS-CoV-2 infection weeks before the first confirmed cases in 5 states, the CDC's first report of community transmission (ie, a confirmed infection in a person without a travel history or exposure to a confirmed SARS-CoV-2) did not occur until 26 February 2020 [18]. The epidemiologic links in the testing recommendations may have been in place too long, obscuring the geographic spread of SARS-CoV-2 found in our results. Future pandemic management should carefully consider the impact of epidemiologic links in testing recommendations and reduce testing restrictions as early as possible.

There are limitations to our study. First, *All of Us* participants were not confirmed to be infected with SARS-CoV-2 via molecular diagnostic tests or paired acute-convalescent serum specimens. The strength of serologic studies, however, is the capture of potentially asymptomatic individuals and those with subclinical illness who do not seek or were unable to obtain testing. Second, it is possible that we have detected preexisting, non-SARS coronavirus antibodies that bind to SARS-CoV-2 nucleocapsid and spike protein in these 9 individuals, rendering these individuals false-positive [9, 19, 20]. Previous studies have shown low cross-reactivity of the S1 domain of the spike protein (target for the EUROIMUNE assay) and low-level cross-reactivity of the S2 domain of the spike protein and the nucleocapsid protein (target for the Abbott assay) and a strong correlation between antibodies against spike protein and neutralization; owing to this evidence, neutralization experiments were not performed on specimens from the 9 seropositive individuals [9, 19, 21, 22].

Our sensitivity analyses of nucleocapsid and spike protein IgG titer quantification provided additional supporting evidence that it is unlikely all 9 of these seropositive individuals were false-positives (estimated probability, 0.00001) and was aligned with the probability of 2 false-positives (7 of the 9 seropositive individuals had quantified IgG titers, and the estimated probability of 2 false-positives was 0.72959; we did not believe further depleting the specimen for neuralization results was warranted. The sequential testing algorithm net specificity was 100% (95% CI: 99.6%–100%); none of the negative control specimens were positive with both Abbott and EUROIMMUNE assays. It may be more likely that we have misclassified those with low titers as seronegative; the positive controls used to estimate Abbott and EUROIMMUN sensitivity were individuals who were hospitalized, or recently discharged, with RT-PCR–confirmed SARS-CoV-2 infection. Third, we do not yet know if these SARS-CoV-2 antibodies were the result of infections acquired during travel or within the participant's community. Fourth, the findings likely cannot be extrapolated to broader seroprevalence estimates, because participants in a longitudinal study enriched for underrepresented minorities in biomedical research may not be generalizable, particularly in the setting of a low prevalence of infection. A clear strength of our study is that it is nested in the *All of Us* Research Program, which provides the opportunity to investigate a critical time in the SARS-CoV-2 US epidemic within a single-protocol-driven, longitudinal study with large-scale geographic, race, and ethnicity breadth, and ongoing follow-up with participants.

Our findings highlight the importance of diverse longitudinal cohort studies that collect biospecimens for conducting retrospective studies to expand the understanding of the epidemiology of SARS-CoV-2, particularly spread of the virus before recognized cases in 5 US states, and informing public health surveillance testing strategies, computational models of the entrance of the novel virus into susceptible populations, and subsequent intervention and mitigation efforts.

Supplementary Data

Supplementary materials are available at *Clinical 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. Kujawski SA, Wong KK, Collins JP, et al. Clinical and virologic characteristics of the first 12 patients with coronavirus disease 2019 (COVID-19) in the United States. Nat Med **2020**; 26:861–8.

2. Hinton D. 2019-Novel coronavirus (2019-nCoV) real-time reverse transcriptase (RT)-PCR diagnostic panel emergency use authorization for the Centers for Disease Control and Prevention. **2020**. Available at: https://www.fda.gov/media/134919/download. Accessed 11 December 2020.

3. Lourenço J, Paton R, Ghafari M, et al. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. medRxiv [Preprint: not peer reviewed]. 26 March 2020. Available from: https://www.medrxiv.org/content/10.1101/2020.03.24.20042291v1.

4. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med **2020**; 26:450–2.

5. Li X, Zai J, Zhao Q, et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. J Med Virol **2020**; 92:602–11.

6. van Dorp L, Acman M, Richard D, et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. Infect Genet Evol **2020**; 83:104351.

7. Jorden MA, Rudman SL, Villarino E, et al. Evidence for limited early spread of COVID-19 within the United States, January–February 2020. MMWR Morb Mortal Wkly Rep **2020**; 69:680–4.

8. Basavaraju SV, Patton ME, Grimm K, et al. Serologic testing of U.S. blood donations to identify SARS-CoV-2-reactive antibodies: December 2019-January 2020. Clin Infect Dis **2020**; 72. doi:10.1093/cid/ciaa1785.

9. Patel EU, Bloch EM, Clarke W, et al. Comparative performance of five commercially available serologic assays to detect antibodies to SARS-CoV-2 and identify individuals with high neutralizing titers. J Clin Microbiol **2021**; 59:e02257-20.

10. US Centers for Disease Control and Prevention. Interim guidelines for COVID-19 antibody testing in clinical and public health settings. Available at: https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html. Accessed 10 December 2020.

11. The All of Us Research Program investigators. The "All of Us" Research Program. N Engl J Med **2019**; 381:668–76.

12. The *All of Us* Research Program. The *All of Us* Research Program: data and statistics dissemination policy. Available at: https://www.researchallofus.org/wp-content/themes/research-hub-wordpress-theme/media/2020/05/AoU_Policy_Data_and_Statistics_Dissemination.pdf. Accessed 16 December 2020.

13. Guo L, Ren L, Yang S, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). Clin Infect Dis **2020**; 71:778–85. doi:10.1093/cid/ciaa310.

14. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis **2020**; 71:2027–34. Available at: https://doi.org/10.1093/cid/ciaa344.

15. To KKW, Tsang OTY, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis **2020**; 20:565–74.

16. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. JAMA **2020**; 323:2466–7.

17. US Centers for Disease Control and Prevention. CDC releases consolidated COVID-19 testing recommendations. Available at: https://www.cdc.gov/media/releases/2020/s0613-covid19-testing-recommendations.html. Accessed 16 December 2020.

18. US Centers for Disease Control and Prevention. CDC confirms possible instance of community spread of COVID-19 in the US. Available at: https://www.cdc.gov/media/releases/2020/s0226-Covid-19-spread.html. Accessed 16 December 2020.

19. Okba NMA, Müller MA, Li W, et al. Severe acute respiratory syndrome coronavirus 2-specific antibody responses in coronavirus disease patients. Emerg Infect Dis **2020**; 26:1478–88.

20. Ng K, Faulkner N, Cornish G, et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. Science **2020**; 370:1343.

21. de Assis RR, Jain A, Nakajima R, et al. Analysis of SARS-CoV-2 antibodies in COVID-19 convalescent blood using a coronavirus antigen microarray. Nat Commun **2021**; 12:6.

22. Khan S, Nakajima R, Jain A, et al. Analysis of serologic cross-reactivity between common human coronaviruses and SARS-CoV-2 using coronavirus antigen microarray. bioRxiv [Preprint: not peer reviewed]. 25 March 2020. Available from: https://www.biorxiv.org/content/10.1101/2020.03.24.006544v1.