

Changes in SARS CoV-2 Seroprevalence Over Time in Ten Sites in the United States, March – August, 2020

Travis Lim, DrPH¹; Mark Delorey, PhD; Nicolette Bestul, MPH¹; Michael Johannsen, PhD¹; Carrie Reed, PhD¹; Aron J. Hall, DVM¹; Alicia M. Fry, MD¹; Chris Edens, PhD¹; Vera Semenova, PhD¹; Han Li, PhD¹; Peter Browning¹, BS; Rita Desai¹, BS; Monica Epperson, PhD¹; Tao Jia, MS¹; Natalie J. Thornburg, PhD¹; Jarad Schiffer, MS¹; Fiona P. Havers, MD¹

1: U.S. Centers for Disease Control and Prevention COVID-19 Response Team, Atlanta, Georgia, United States

Corresponding author: Fiona Havers, MD, MHS

1600 Clifton RD, MS H24-6

Atlanta, GA 30329, United States

Phone: 404-639-5062; Email: fhavers@cdc.gov

Summary: The majority of people in the U.S. have not been infected with SARS-CoV-2. Reported cases continue to be less than estimated infections. Some statistically significant increases in seroprevalence followed spikes in reported cases. Some statistically significant decreases may be indirect evidence of waning of detectable antibodies at the population level.

Abstract:

Background. Monitoring of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibody prevalence can complement case reporting to inform more accurate estimates of SARS-CoV-2 infection burden, but few studies have undertaken repeated sampling over time on a broad geographic scale.

Methods. We performed serologic testing on a convenience sample of residual sera obtained from persons of all ages, at ten sites in the United States from March 23 through August 14, 2020, from routine clinical testing at commercial laboratories. We age-sex-standardized our seroprevalence rates using census population projections and adjusted for laboratory assay performance. Confidence intervals were generated with a two-stage bootstrap. We used Bayesian modeling to test whether seroprevalence changes over time were statistically significant.

Results. Seroprevalence remained below 10% at all sites except New York and Florida, where it reached 23.2% and 13.3%, respectively. Statistically significant increases in seroprevalence followed peaks in reported cases in New York, South Florida, Utah, Missouri and Louisiana. In the absence of such peaks, some significant decreases were observed over time in New York, Missouri, Utah, and Western Washington. The estimated cumulative number of infections with detectable antibody response continued to exceed reported cases in all sites.

Conclusions. Estimated seroprevalence was low in most sites, indicating that most people in the U.S. have not been infected with SARS-CoV-2 as of July 2020. The majority of infections are likely not reported. Decreases in seroprevalence may be related to changes in healthcare-seeking behavior, or evidence of waning of detectable anti-SARS CoV-2 antibody levels at the population level. Thus, seroprevalence estimates may underestimate the cumulative incidence of infection.

Keywords: COVID-19, Seroprevalence, SARS-CoV-2, serology

Introduction

In the United States (U.S.), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection causing coronavirus disease 2019 (COVID-19) has resulted in over 12 million reported cases and 250,000 deaths as of November 25, 2020 [1]. However, the true number of infections is thought to be far greater, as an estimated 15–40% of infections are asymptomatic [2-4], and many persons with symptoms either do not seek medical care or are not tested. Furthermore, case counts reported by state and territorial health departments are affected by changes in testing and reporting capacity and practices, all of which can distort the apparent extent of infections and disease in the population at a given time.

Given that the majority of persons that are infected with SARS-CoV-2 develop antibodies following infection [5-9], the prevalence of anti-SARS-CoV-2 antibodies has been used as an indicator of cumulative infections in a given population since the start of the pandemic [10-14]. Seroprevalence estimates can capture mild or asymptomatic infections that may be missed in reported case counts. They therefore allow better estimates of the true burden of infection, informing public health interventions. In addition, seroprevalence estimates can indicate population-level immunity, and the coverage of SARS-CoV-2 vaccines when they become available, both of which have implications for future transmission [15, 16]. While seroprevalence rates have been used to estimate the proportion of the population that has never been infected, they are not necessarily a perfect indicator. If antibody levels wane and become undetectable over time, as has been suggested by recent reports [5, 7, 8], this could affect the accuracy of these estimates as the pandemic progresses.

We and others have reported on seroprevalence estimates throughout in the pandemic [10, 11, 13, 17, 18], with many seroprevalence studies focused on assessing a narrow time window or single geographic location. In this study, the availability of a large number of residual sera collected from patients during the course of routine blood tests allowed us to estimate seroprevalence in 10 sites across the U.S. at multiple time points. In each of the sites, we compare the seroprevalence

trajectories to reported cases for a pandemic that has been heterogeneous with respect to time and geography.

Methods

Sampling

The U.S. Centers for Disease Control and Prevention (CDC) partnered with two commercial laboratory companies to select a random sample of deidentified sera remnants from a population of convenience, as previously described [17]. To be included, specimens must have been from routine diagnostic tests, such as metabolic panels or cholesterol levels, blinded to COVID-19 symptoms or diagnosis, that were collected from patients between March 23 and August 14, 2020, and, at minimum, included demographic information on age, sex, zip code of patient, and date of specimen collection.

Our target sample size was 1,800 specimens per site per round with 450 specimens in each of four age groups (0-18, 19-49, 50-64, and 65+ years), based on calculations of statistical power needed to estimate a prevalence within +/- 2% in a population with 5% prevalence. The laboratories generated de-duplicated, de-identified sample lists from their internal electronic laboratory information systems of all specimens with sufficient serum volume. Specimens were sampled by stratified random sampling by age group, with the aim of collecting equal numbers of specimens for each of the four age groups.

Each round of specimens covered a one-week period of collections, except for Utah and Minnesota, which covered a two-week period, as fewer specimens were available per week. Specimen collection rounds were a minimum of three weeks apart for each site, and sites were not synchronized.

Between three and five specimen collection rounds occurred in ten sites, one from each of the 10 Health and Human Services regions (Fig e1): New York City Metro Area (NY), Western Washington State (WA), South Florida (FL), Missouri (MO), Connecticut (CT), Utah (UT), Louisiana (LA), San

Francisco Bay Area (CA), Philadelphia Greater Metro Area (PA), and Minneapolis-St. Paul-St. Cloud combined statistical area (MN).

This activity was reviewed by CDC and was conducted consistent with applicable federal law and CDC policy.[§] Informed consent was waived, as de-identified data were used. The results from individual timepoints were previously released on CDC's website [19].

Laboratory Assay

Sera were tested with a CDC assay as previously described [17]. Briefly, sera specimens were tested with an enzyme-linked immunosorbent assay (ELISA) against pan-immunoglobulin (pan-Ig) to the anti-spike protein of SARS-CoV-2. Signal-to-threshold ratios were calculated from optical density results with a reactive cutoff of >1.0. A receiver operator curve was created to maximize overall accuracy. The pan-Ig assay performed with a sensitivity of 96.0% (CI: 89.98 – 98.89%) and a specificity of 99.3% (CI: 98.32 – 99.88%) against a PCR-confirmed test set [20].

Analysis

Age-sex-standardized estimates of seroprevalence with confidence intervals that account for serology assay performance were generated using a two-stage bootstrap method described previously [17].

For each site and time point, estimated cumulative infections were calculated as age-sex standardized seroprevalence multiplied by catchment population. A case ascertainment ratio was calculated by dividing estimated infections by cumulative reported cases as of the last day of specimen collection for that time period and catchment area. Catchment populations were mostly

[§] See e.g., 45 C.F.R. part 46.102(l)(2), 21 C.F.R. part 56; 42 U.S.C. §241(d); 5 U.S.C. §552a; 44 U.S.C. §3501 et seq.

consistent over time but were updated from the second timepoint onwards in NY, WA, and PA, based on counties that provided the majority of specimens (Supplement).

To assess the relative magnitude and significance of changes in seroprevalence over time, while accounting for uncertainty in sampling and lab assay accuracy, we used Bayesian hierarchical modeling to obtain the joint posterior distribution of our parameters, optimized by Gibbs sampling [21]. Equations are listed in the statistical Appendix; briefly, we modeled the odds of having detectable SARS-CoV-2 antibodies at each site, assuming a binomial distribution for antibody detection and using a logit link, as a function of the following parameters: age, sex, county, and time period (an indicator variable for change between sequential collection rounds). Relative risks were calculated from the posterior probabilities of positivity as described in the appendix. For UT, model diagnostics indicated a poor fit, so we instead used a non-parametric permutation method (Appendix).

Estimated cases and catchment populations

Publicly reported daily case data at the county level were obtained from USAFacts [22]. County-level population projections by age and sex were downloaded from the U.S. Census Bureau [23].

Statistical analyses were programmed in R 3.6.1 and RStudio v1.2.1335 (R Foundation for Statistical Computing).

Results

Sample characteristics

A total of 78,990 specimens were collected from patients between March 23 and August 14, 2020 across ten sites. Of those, 43.7% (34,516) were male. The median age was 55, with 59.0% (46,628) of all specimens from patients over 50 years of age. Table 1 shows the demographic characteristics of patient specimens at each site. A comparison of demographic and geographic information over each

timepoint shows that our samples remained generally consistent within sites for known characteristics and are coming from the same source counties except for the first timepoint in NY, WA, and PA (Table e1).

Results by site

In the northeastern U.S., including NY, PA, and CT, reported cases increased markedly in April 2020, around or just before the first time point. New York had the most dynamic changes in seroprevalence and the highest seroprevalence, increasing from 6.9% by April 1 to 23.2% by April 16, as estimated by the bootstrap analysis (adjusted Risk Ratio (aRR) from Bayesian analysis: 2.3, 95% Confidence Interval (CI): 1.4 – 4.9, Table 2 and Figure 1a). Daily reported cases in NY then steadily declined; subsequent seroprevalence measurements indicated a small decrease by the first week of May (aRR 0.7, CI: 0.6-0.9). In PA and CT, cases also began to decline gradually starting in May (Figure e2b-c), and the aRR of detectable antibodies was not statistically different from 1.0 in CT and PA.

In MN, reported cases first increased in early May, and remained elevated throughout our study period without clear peaks or declines. Seroprevalence climbed from 2.4% by May 12 to 6.1% by July 18. Increases during late June and early July were statistically significant (Figure e2d).

In the southern U.S. (FL, LA), seroprevalence rose gradually in April and May; FL increased to 2.8% and LA to 5.8% by April 23 (Figure e2g). Both sites experienced a marked increase in cases starting in late June through late July. By July 23, seroprevalence in FL rose to 13.3% (aRR 2.5, CI: 1.7 – 3.9), and by August 8, seroprevalence in LA rose to 9.8% (aRR 2.4, CI: 1.6 - 3.7) (Figure 1b).

In the western U.S. (CA, WA), seroprevalence in the study remained low. Western Washington State stayed between 1.0% and 2.1%, with a statistically significant decrease between May 4 and June 15 (aRR 0.3, CI: 0.1 - 0.7) (Figure 1c). The San Francisco Bay Area remained below 1.0% through May 27, and increased to 1.7% by July 23 (aRR 1.3, CI: 0.7 – 1.2).

In the central U.S., in MO and UT, reported cases were at steady low levels from April through early June, then cases in both sites began to increase sharply toward the third and fourth study timepoint. Seroprevalence in MO decreased from 2.8% in late May to 0.8% (aRR 0.6, CI: 0.4-0.8) in late June, then an increase to 1.4% (aRR 1.6, CI: 1.0 – 2.3) in July (Figure 1d). UT seroprevalence decreased from 2.2% in early May to 1.1% in early June (aRR 0.3, CI: 0.1-0.8) then increased to 2.7% in early July (aRR 3.3, CI: 1.2 – 12.0) and again to 4.8% in late July (aRR 2.4, CI: 1.3 – 4.6) (Figure e2e).

In a multivariable model that adjusted for age, sex, county, and time, persons 65 and older had an overall lower risk of infection (aRR < 1.0) in MO, CT, PA, and MN, compared to persons 19-49 (table 3). In FL, persons 50 and older had a higher risk of infection and in LA, persons 50 and older had a lower risk of infection. No statistically significant difference in risk of infection was observed between males and females (table 3).

Case ascertainment ratio

At the first time point, estimated actual infections exceeded reported cases by at least ten-fold in seven of 10 sites (range 6 to 24-fold, Table 2); by the fourth time point, the ratio of estimated infections to reported cases was between two-fold in Missouri to seven-fold in NY.

Discussion

In a large-scale SARS-CoV-2 seroprevalence study, we collected and tested 78,990 specimens from ten U.S. sites that experienced different epidemic curves, over multiple timepoints between March and August 2020. Our use of convenience samples of commercial laboratory residual sera allowed serial measurements of seroprevalence from a large number of specimens drawn from similar sampling frames. During this period, less than 10% of the population had detectable antibodies to SARS-CoV-2 in all sites except NY and FL. Through Bayesian modeling, we determined that seroprevalence increases in NY, FL, UT, and MO, which were observed following an increase in reported cases, were statistically significant. We also observed statistically significant declines in

seroprevalence in MO, UT, NY, and WA, following periods of stable, low numbers of cases. This decrease in seroprevalence, even in settings with ongoing transmission, may suggest that waning of assay-detectable antibodies may be occurring; the effects of waning antibodies on overall seroprevalence may be more apparent during periods with fewer new cases, possibly in combination with changes in healthcare-seeking behaviors and changing adherence to stay-at-home orders.

Since the beginning of the COVID-19 pandemic, seroprevalence estimates have been used to estimate cumulative incidence of infection [10-14, 18]. These studies can capture a broader range of infections than case reporting, including mild or asymptomatic cases that may go undetected and unreported. Our seroprevalence estimates at specific time points are similar to those of other studies undertaken in similar locations, including in San Francisco, Utah, and Miami-Dade [24-26]. In New York, seroprevalence in a state-wide survey of grocery store patrons was 14% in late March [12], and serial samples in a health care setting found a peak of seroprevalence of 19.3% by mid-April [27], both approximating our estimates. In a representative survey in Connecticut, seroprevalence was 3.1%, lower than our estimate of 6.2% at a comparable timepoint, although lower assay sensitivity and different sampling methods may explain some of the difference in results [28].

At all ten sites, across all time periods during the study, the estimated number of infections was much higher than the number of reported cases. This case ascertainment ratio changed over the course of the study, from at least 10-fold in seven sites [17], to between two- to seven-fold. This decrease may be due to several factors, including improved testing availability, changing testing patterns [29] and increased contact tracing, thus potentially resulting in the detection and reporting of a larger proportion of infections. Changes in healthcare-seeking behavior may also have affected both case reporting and seroprevalence estimates in unpredictable ways. The case ascertainment ratio based on seroprevalence is a conservative estimate, however, as some individuals infected near the end of our specimen collection period may have not yet developed antibodies and not all

infected persons develop antibodies. We would also be unable to detect antibodies lost due to waning in previously infected persons, which would underestimate the population that has been infected.

The observation that seroprevalence stayed flat, or even decreased, in some sites, while cumulative case counts increased at all ten sites, offers indirect evidence of waning levels of assay-detectable antibodies in a certain proportion of the population. In the absence of waning, repeated serological testing of similar samples from a population in which the virus continues to circulate should find that the number of persons with past infection increases monotonically. Reports on humoral response and antibody kinetics found waning anti-SARS-CoV-2 antibody levels among a subset of recovered patients along timelines that our study would have captured, with declines 40-60 days after the onset of symptoms [7, 8, 14]. IgG antibodies typically have a half-life of 7-21 days [30], but IgG to the SARS-CoV-2 spike protein, which our ELISA assay can detect, may persist longer [5, 31]. Antibody levels following infection also appear positively correlated with disease severity [5, 7]. Many infections identified through seroprevalence studies likely had mild or asymptomatic disease, and thus potentially lower initial antibody titers than reported cases. These persons may be more likely to have antibody levels that fall below detectable levels when sampled at later timepoints. If the rate of new seroconversion from new infections falls below the rate of antibody waning in previously infected individuals, seroprevalence may decrease, possibly explaining the observed seroprevalence changes in New York following the spring surge of new infections.

The implications of seroprevalence estimates for herd immunity are not fully known, since the presence of antibodies at levels detectable in most ELISAs does not necessarily correlate with protection against reinfection and disease [32]. Conversely, there may be residual immune protection from past infections even if antibodies fall below the threshold of detection, and immunity may be conferred by T-cell mediated mechanisms independently of antibody levels [33, 34].

Limitations

Our study has several limitations. While use of residual clinical specimens allows for large sample sizes spanning multiple timepoints, the population from whom these are drawn could differ from the population in multiple ways discussed previously [17]. Although the laboratory sera were sampled from routine tests not directly associated with COVID-19, increases in symptomatic infections may have driven a disproportionate number of COVID-19-infected patients to clinical settings, and the clinical specimens could have been taken as part of their acute or follow-up care. Pediatric specimens sometimes did not meet our sample size targets in some sites, especially Utah, likely because healthy children infrequently have blood drawn.

In some instances, there was geographic variability within a site over time, with variability especially in suburban counties at some sites. Given local geographic variation in SARS-CoV-2 clusters, this may have affected our seroprevalence estimates, although our multivariable model for relative odds of seropositivity controlled for county. In addition, as noted above, healthcare-seeking behavior likely also changed over time for the general population [35]. This could potentially affect the study population from whom residual clinical samples are available, potentially biasing these seroprevalence estimates in unpredictable ways. Since our data were drawn from serial independent samples and not cohorts, the effects of antibody waning and potential sampling biases are inextricable. In addition, not all persons infected with SARS-CoV-2 mount an antibody response [5-9].

Conclusions

This study shows that as of August 2020, most of the predominantly adult population sampled from clinical laboratories in the 10 sites studied had no evidence of having past infection with SARS-CoV-2. Moreover, among people who become infected, a majority—many of whom likely had mild or asymptomatic infections—are not captured through case reporting, although the gap between estimated infections based on seroprevalence and reported case has decreased over time. Our

findings also suggest that antibodies may wane over time. Thus, both seroprevalence rates themselves, and the estimated number of infections based on seroprevalence, are likely most accurate during the early months of each site's epidemic, whereas by summer 2020 and later, seroprevalence estimates likely underestimate cumulative infections until appropriate methods are identified to account for waning.

Despite potential limitations, large-scale geographic seroprevalence surveys will continue to be an important tool in our understanding of potential herd immunity and may be useful for understanding the changing epidemiology of SARS-CoV-2 as vaccines or other pharmaceutical or non-pharmaceutical interventions are introduced.

SARS-CoV-2 continues to circulate across the U.S., and our study suggests both that the vast majority of the population do not have evidence of previous SARS-CoV-2 infection and that many people likely do not know when they are infected and able to transmit the virus to others. These findings reinforce the critical importance of continued precautions at the individual and population level, such as wearing cloth face masks, avoiding large gatherings, handwashing, and maintaining social distancing when possible.

NOTES

Acknowledgements

We would like to acknowledge the following members of the Microbial Pathogenesis and Immune Response (MPIR) Laboratory at CDC for their contributions to the testing of specimens for this project: Bailey Alston, MS; Muiyiwa Ategbale, MPH; Shanna Bolcen, MSPH; Darbi Boulay, BS; Li X Cronin, MS; Ebenezer David, PhD; Yamini Gorantla, PhD; Panagiotis Maniatis, MS; Kimberly Moss, BS; Kristina Ortiz, MS; So Hee Park; Palak Patel, MS; Yunlong Qin, PhD; Evelene Steward-Clark, MS; Heather Tatum; Andrew Vogan, MS; Briana Zellner, PhD. We would also like to thank Michelle Lewis, PhD; David Blum, PhD; and the University of Georgia Bioexpression and Fermentation Facility for protein expression and purification for helping to create and optimize the pan-IG ELISA assay. We thank Marla Petway, MPH, and the CDC STATT laboratory for specimen handling support. Finally, we would like to thank our collaborators at the State Health Departments in all participating study sites, including Carina Blackmore, DVM; Debra Blog, MD; Shua J. Chai, MD; Angela Dunn, MD; Julie Hand, MSPH; Seema Jain, MD; Scott Lindquist, MD; Ruth Lynfield, MD; Scott Pritchard, MPH; Theresa Sokol, MPH; Lynn Sosa, MD; George Turabelidze, MD; Sharon M. Watkins, PhD; John Wiesman, DrPH; Randall W. Williams, MD; and Stephanie Yendell, DVM. The de-identified data and the codebook from the study can be requested from the corresponding author to further public health research or practice, at the time of publication. Additional documents such as the study protocol are available upon request. Data sharing will be subject to applicable laws, regulations, and rules established in U.S. Code as well as agency policies of the U.S. Centers for Disease Control and Prevention. Data sharing may require a data access agreement.

Funding

This work was supported by the U.S. Centers for Disease Control and Prevention. The authors did not receive external funding for this study.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the U.S. Centers for Disease Control and Prevention.

Conflict of interest disclosures: The authors have no conflicts of interest to report.

Accepted Manuscript

References

1. U.S. Centers for Disease Control and Prevention. Cases in the U.S. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html>. Accessed 8/27/2020.
2. Nishiura H, Kobayashi T, Miyama T, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis* **2020**; 94: 154-5.
3. Wells PM, Doores KM, Couvreur S, et al. Estimates of the rate of infection and asymptomatic COVID-19 disease in a population sample from SE England. **2020**: 2020.07.29.20162701.
4. Yousaf AR, Duca LM, Chu V, et al. A prospective cohort study in non-hospitalized household contacts with SARS-CoV-2 infection: symptom profiles and symptom change over time. *Clinical Infectious Diseases* **2020**.
5. Grandjean L, Saso A, Ortiz A, et al. Humoral Response Dynamics Following Infection with SARS-CoV-2. **2020**: 2020.07.16.20155663.
6. Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clinical Infectious Diseases* **2020**; 71(15): 778-85.
7. Long QX, Tang XJ, Shi QL, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* **2020**.
8. Seow J, Graham C, Merrick B, et al. Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection. **2020**: 2020.07.09.20148429.
9. Wang Y, Zhang L, Sang L, et al. Kinetics of viral load and antibody response in relation to COVID-19 severity. *The Journal of Clinical Investigation* **2020**; 130(10).
10. Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. **2020**: 2020.04.14.20062463.
11. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *The Lancet* **2020**.

12. Rosenberg ES, Tesoriero JM, Rosenthal EM, et al. Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. *Annals of Epidemiology* **2020**; 48: 23-9.e4.
13. Stringhini S, Wisniak A, Piumatti G, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *The Lancet* **2020**; 396(10247): 313-9.
14. Xu X, Sun J, Nie S, et al. Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. *Nat Med* **2020**.
15. Brevia LP, Domingo JD, Martínez Beneito MÁ, Barberà JP. Vaccine coverage estimation using a computerized vaccination registry with potential underreporting and a seroprevalence study. *Vaccine* **2015**; 33(18): 2183-8.
16. Winter AK, Martinez ME, Cutts FT, et al. Benefits and Challenges in Using Seroprevalence Data to Inform Models for Measles and Rubella Elimination. *J Infect Dis* **2018**; 218(3): 355-64.
17. Havers FP, Reed C, Lim T, et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. *JAMA Internal Medicine* **2020**.
18. Bajema KL, Wiegand RE, Cuffe K, et al. Estimated SARS-CoV-2 Seroprevalence in the US as of September 2020. *JAMA Internal Medicine* **2020**.
19. U.S. Centers for Disease Control and Prevention. Commercial Laboratory Seroprevalence Survey Data. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/commercial-lab-surveys.html>. Accessed 9/23/2020.
20. Freeman B, Lester S, Mills L, et al. Validation of a SARS-CoV-2 spike protein ELISA for use in contact investigations and sero-surveillance. **2020**: 2020.04.24.057323.
21. Mitchell PK, Mier-y-Teran-Romero L, Biggerstaff BJ, et al. Reassessing Serosurvey-Based Estimates of the Symptomatic Proportion of Zika Virus Infections. *American Journal of Epidemiology* **2018**; 188(1): 206-13.

22. USAFacts. US Coronavirus Cases and Deaths - Track COVID-19 data daily by state and county. Available at: https://usafactsstatic.blob.core.windows.net/public/data/covid-19/covid_confirmed_usafacts.csv. Accessed 8/11/2020.
23. U.S. Census Bureau. County Population by Characteristics: 2010-2018. Available at: <https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html>. Accessed 8/11/2020.
24. Miami-Dade County. Second round of COVID-19 community testing completed; Miami-Dade County and the University of Miami Miller School of Medicine announce initial findings. Available at: <https://www.miamidade.gov/releases/2020-04-24-sample-testing-results.asp>. Accessed 9/9/2020.
25. Ng D, Goldgof G, Shy B, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood from the San Francisco Bay Area. **2020**: 2020.05.19.20107482.
26. University of Utah. Utah Hero Project Announces Phase One Findings. Available at: <https://healthcare.utah.edu/publicaffairs/news/2020/07/utah-hero-project-phase-1.php>. Accessed 9/9/2020.
27. Stadlbauer D, Tan J, Jiang K, et al. Seroconversion of a city: Longitudinal monitoring of SARS-CoV-2 seroprevalence in New York City. **2020**: 2020.06.28.20142190.
28. Mahajan S, Srinivasan R, Redlich CA, et al. Seroprevalence of SARS-CoV-2-Specific IgG Antibodies Among Adults Living in Connecticut Between March 1 and June 1, 2020: Post-Infection Prevalence (PIP) Study. **2020**: 2020.08.04.20168203.
29. Tromberg BJ, Schwetz TA, Pérez-Stable EJ, et al. Rapid Scaling Up of Covid-19 Diagnostic Testing in the United States — The NIH RADx Initiative. **2020**; 383(11): 1071-7.
30. Vidarsson G, Dekkers G, Rispens T. IgG Subclasses and Allotypes: From Structure to Effector Functions. **2014**; 5(520).
31. Iyer AS, Jones FK, Nodoushania A, et al. Dynamics and significance of the antibody response to SARS-CoV-2 infection. **2020**: 2020.07.18.20155374.

32. To KK-W, Hung IF-N, Ip JD, et al. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. *Clinical Infectious Diseases* **2020**.
33. Grifoni A, Weiskopf D, Ramirez SI, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **2020**; 181(7): 1489-501.e15.
34. Sekine T, Perez-Potti A, Rivera-Ballesteros O, et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell* **2020**.
35. Kocher KE, Macy ML. Emergency Department Patients in the Early Months of the Coronavirus Disease 2019 (COVID-19) Pandemic—What Have We Learned? *JAMA Health Forum* **2020**; 1(6): e200705-e.

Accepted Manuscript

Figure 1. Changes in seroprevalence point estimates and 95% CI from age-sex standardized two-stage bootstrapping method described previously¹⁷, measured at sequential time points from March – July 2020 (left axis, line graph), at four of ten sites in the United States. Overlay: daily reported cases per 10,000 population (right axis, area graph). * = adjusted relative risk of infection compared to the preceding timepoint was statistically different from 1.0, based on model coefficients and 95% credible intervals of Bayesian hierarchical modeling that adjusts for sex, age group, and county. Figures for remaining six sites are in the online supplement, figure e2.

Accepted Manuscript

Table 1. Demographic characteristics of total patient specimens selected for commercial laboratory seroprevalence survey across up to five time points at ten sites in the United States from March – July 2020.

		WA	NY	FL	PA	MO	UT	CA	CT	MN	LA
Age group (yrs)	0-18	342 (4.5%)	1435 (18.5%)	313 (4.6%)	1392 (16.1%)	920 (9.8%)	244 (2.8%)	159 (3.2%)	1068 (12.4%)	386 (4.9%)	103 (3.6%)
	19-49	2100 (27.8%)	2393 (30.9%)	1692 (25%)	2222 (25.7%)	2633 (28.1%)	3637 (41.5%)	1782 (35.8%)	2433 (28.2%)	3155 (39.7%)	1306 (45.8%)
	50-64	2097 (27.8%)	1827 (23.6%)	1349 (19.9%)	2480 (28.6%)	2638 (28.2%)	2558 (29.2%)	1223 (24.6%)	2424 (28.1%)	2377 (29.9%)	784 (27.5%)
	65+	3005 (39.8%)	2091 (27%)	3424 (50.5%)	2563 (29.6%)	3165 (33.8%)	2330 (26.6%)	1816 (36.5%)	2704 (31.3%)	2026 (25.5%)	656 (23%)
Sex	Female	4268 (56.6%)	4341 (56%)	3746 (55.3%)	4852 (56%)	5289 (56.5%)	5050 (57.6%)	2747 (55.2%)	4795 (55.6%)	4379 (55.1%)	1747 (61.3%)
	Male	3276 (43.4%)	3405 (44%)	3032 (44.7%)	3805 (44%)	4067 (43.5%)	3719 (42.4%)	2233 (44.8%)	3834 (44.4%)	3565 (44.9%)	1102 (38.7%)
Serology result ¹	non-reactive	7396 (98%)	6414 (82.8%)	6455 (95.2%)	8286 (95.7%)	9123 (97.5%)	8528 (97.3%)	4887 (98.1%)	8136 (94.3%)	7552 (95.1%)	2630 (92.3%)
	Reactive	148 (2%)	1332 (17.2%)	323 (4.8%)	371 (4.3%)	233 (2.5%)	241 (2.7%)	93 (1.9%)	493 (5.7%)	392 (4.9%)	219 (7.7%)

N=total numbers (% = percent within each subgroup at the same site). ¹ Sera specimens were tested with an enzyme-linked immunosorbent assay (ELISA) against pan-immunoglobulin (pan-Ig) to the anti-spike protein of SARS-CoV-2. Signal-to-threshold ratios were calculated from optical density results with a reactive cutoff of >1.0. A receiver operator curve was created to maximize overall accuracy. The pan-Ig assay performed with a sensitivity of 96.0% (CI: 89.98 – 98.89%) and a specificity of 99.3% (CI: 98.32 – 99.88%) against a PCR-confirmed test set¹⁹.

Table 2. Seroprevalence estimates, infection estimates, and case reporting ascertainment ratios from sampling over multiple serial cross-sectional time points at ten sites in the United States.

Site	Timepoint	Last date of collection	Seroprevalence* (Point estimate, 95% CI)	Cases Reported†	Estimated infections (N, upper – lower range)	Ratio of estimated to reported cases**
CA	2	5/27/2020	0.7% (0.4-1.8)	11913	46637 (26650 - 119924)	4
	3	7/23/2020	1.8% (1.2-3)	38603	119924 (79949 - 199874)	3
CT	2	5/26/2020	5.2% (3.9-6.6)	41234	185275 (138957 - 235157)	4
	3	6/17/2020	6.3% (5.0-7.8)	45347	224468 (178149 - 277913)	5
	4	7/6/2020	5.2% (4-6.7)	46975	185275 (142520 - 238720)	4
	5	7/27/2020	6.7% (5.3-8.4)	48983	238720 (188838 - 299291)	5
FL	2	4/27/2020	2.8% (1.9-4.3)	19310	177670 (120562 - 272850)	9
	3	5/27/2020	4.2% (2.8-6.2)	30055	266504 (177670 - 393411)	9
	4	7/23/2020	13.3% (10.5-16.2)	171659	843931 (666261 - 1027946)	5
LA	2	4/27/2020	6.7% (3.2-12.4)	27068	312219 (149119 - 577837)	12
	3	8/14/2020	9.8% (7.4-12.7)	136735	455517 (343962 - 590313)	3
MN	2	6/6/2020	2.2% (1.4-3.3)	21066	84865 (54005 - 127297)	4
	3	6/27/2020	4.3% (3.4-5.8)	26528	165872 (131154 - 223734)	6
	4	7/18/2020	6.1% (4.4-8.3)	34379	235306 (169729 - 320171)	7
	5	8/8/2020	8.8% (6.6-11.2)	46045	339458 (254594 - 432038)	7
MO	2	5/30/2020	2.8% (1.7-4.1)	12956	171102 (103884 - 250543)	13
	3	6/20/2020	0.8% (0.6-1.8)	17588	48886 (36665 - 109994)	3
	4	7/9/2020	1.4% (0.9-2.5)	25762	85551 (54997 - 152770)	3
	5	7/30/2020	3.2% (2.2-4.5)	48816	195546 (134438 - 274986)	4
NY	2	4/16/2020	20.9% (16.5 -25.9)	196928	2551011 (2013956 - 3161301)	13
	3	5/6/2020	23.2% (20-26.3)	281670	2831745 (2441159 – 3210124)	10
	4	6/21/2020	19.5% (17.2-21.9)	329418	2380130 (2099397 – 2673069)	7
	5	7/11/2020	17.6% (15.5-20)	338224	2148220 (1891898 – 2441159)	6
	6	7/30/2020	18.4% (16.1-20.7)	347104	2245866 (1965133 - 2526600)	6
PA	2	5/30/2020	3.6% (2.5-5.1)	56318	242681 (168529 – 343798)	4
	3	6/20/2020	3.8% (2.8-5.1)	64095	256163 (188752 – 343798)	4
	4	7/11/2020	5% (3.8-6.5)	71746	337057 (256163 – 438174)	5
	5	8/1/2020	6.1% (4.8-7.6)	81832	411210 (323575 - 512327)	5
UT‡	2	6/5/2020	1.1% (0.6-2.1)	10310	24013 (13098 - 45844)	2
	3	6/27/2020	1.5% (0.9-2.6)	18836	32745 (19647 - 56759)	2

	4	7/15/2020	2.7% (1.8-3.9)	28303	58942 (39295 - 85138)	2
	5	8/6/2020	4.8% (3.7-6.2)	39061	104308 (80404 - 134731)	3
WA	2	5/11/2020	2.1% (1.2-3.5)	13098	124063 (70893 - 206772)	9
	3	6/20/2020	1.7% (0.9-3.2)	17185	100432 (53170 - 189049)	6
	4	7/7/2020	1.3% (0.8-2.4)	21299	76801 (47262 - 141787)	4
	5	7/31/2020	1.3% (0.9-2.4)	30416	76801 (53170 - 141787)	3

* = seroprevalence point estimates calculated using bootstrap methods described previously¹⁷. † = publicly reported cases by county and date were compiled by

USAfacts, a non-profit consortium including University of Pennsylvania and Stanford University. ‡ = insufficient samples received from pediatric strata from Utah,

estimates refer to seroprevalence among adults ≥ 19 and assumes publicly reported cases were 91% adult based on CDC case surveillance inquiries. **These estimates

compare the estimated infections based on the proportion of the population with antibodies to SARS-CoV-2 and compare it to the number of reported cases as of the last

date of specimen collection. These numbers are conservative, as some persons who are recently infected may not have developed detectable antibodies at the time of

specimen collection, or antibodies could wane over time.

Table 3. Coefficients from Bayesian hierarchical models for adjusted relative risk of SARS CoV-2 infection at ten sites.

	NY	WA	MO	CT	PA	MN	CA	LA	FL	UT
Sex										
Female	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Male	1.1 (1.0, 1.1)	1.4 (0.8, 2.1)	1.0 (0.7, 1.3)	1.0 (0.9, 1.2)	0.9 (0.6, 1.1)	0.9, (0.7, 1.1)	1.5 (0.7, 2.7)	1.1 (0.9, 1.5)	1.0 (0.8, 1.1)	0.9 (0.5, 1.5)
Age group (yrs)										
0-18	1.0 (1.0, 1.0)	0.5 (0.1, 1.7)	1.1 (0.5, 1.7)	0.5 (0.3, 0.8)	1.0 (0.7, 1.6)	1.0 (0.5, 1.7)	0.3 (0.8, 1.2)	0.7 (0.4, 1.5)	1.1 (0.8, 1.4)	0.8 (0.1, 2.2)
19-49	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
50-64	1.0 (1.0, 1.1)	0.7 (0.4, 1.3)	0.7 (0.4, 1.0)	0.8 (0.6, 1.1)	0.7 (0.5, 1.1)	0.6 (0.4, 0.8)	1.0 (0.9, 1.3)	0.7 (0.5, 0.9)	1.1 (1.0, 1.4)	0.7 (0.4, 1.1)
65+	1.0 (1.0, 1.1)	0.6 (0.3, 1.1)	0.7 (0.5, 1.1)	0.4 (0.3, 0.5)	0.2 (0.1, 0.5)	0.3 (0.2, 0.5)	1.1 (0.3, 3.1)	0.5 (0.3, 0.7)	1.5 (1.2, 2.0)	0.6 (0.3, 1.1)
Timepoint										
1	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
2 vs 1	2.3 (1.4, 4.9)	1.7 (0.8, 5.1)	0.7 (0.5, 1.2)	1.0 (0.7, 1.6)	1.1 (0.5, 2.4)	1.2 (0.9, 1.6)	0.9 (0.6, 1.6)	0.7 (0.4, 1.0)	0.9 (0.5, 1.3)	0.3 (0.1, 0.8)
3 vs 2	1.9 (1.5, 2.6)	0.3 (0.1, 0.7)	0.4 (0.2, 0.7)	1.1 (0.8, 1.5)	1.3 (0.7, 2.4)	2.3 (1.4, 4.3)	1.3 (0.8, 2.2)	2.4 (1.6, 3.7)	1.0 (0.7, 1.6)	1.1 (0.6, 1.9)
4 vs 3	0.7 (0.6, 0.9)	1.6 (0.6, 4.0)	1.4 (0.6, 3.3)	0.8 (0.5, 1.1)	1.2 (0.8, 2.0)	1.3 (1.0, 1.8)			2.5 (1.7, 3.9)	3.3 (1.2, 12.0)
5 vs 4	0.9 (0.8, 1.1)	1.3 (0.6, 3.1)	2.8 (1.5, 5.3)	1.3 (0.9, 2.0)	1.2 (0.8, 1.8)	1.0 (0.8, 1.4)				2.4 (1.3, 4.6)
6 vs 5	1.0 (0.8, 1.2)									

Model estimates the risk ratio of an individual having detectable anti-SARS-CoV-2 antibodies, as a function of sex, age group, and timepoint of collection as categorical variables, adjusted for county. Ref = reference category. *For Utah only, convergence and resulting posterior distributions were highly dependent upon the priors and the starting values for the chains. Because of this, we have little confidence in the estimates of model parameters. However, for completeness, we report the results using non-informative priors [Normal (0, 0.0001)] for the model coefficients and randomly generated starting values and emphasize the sensitive nature of the model fit.

Figure 1

