

Comparison of Estimated SARS-CoV-2 Seroprevalence through Commercial Laboratory Residual Sera Testing and a Community Survey

Kristina L. Bajema,¹ F. Scott Dahlgren,¹ Travis W. Lim,¹ Nicolette Bestul,¹ Holly M. Biggs,¹ Jacqueline E. Tate,¹ Claudio Owusu,¹ Christine M. Szablewski,^{1,2} Cherie Drenzek,² Jan Drobeniuc,¹ Vera Semenova,¹ Han Li,¹ Peter Browning,¹ Rita Desai,¹ Monica Epperson,¹ Lily T. Jia,¹ Natalie J. Thornburg,¹ Chris Edens,¹ Alicia M. Fry,¹ Aron J. Hall,¹ Jarad Schiffer,¹ Fiona P. Havers¹

¹CDC COVID-19 Response, Centers for Disease Control and Prevention, Atlanta, GA

²Georgia Department of Public Health, Atlanta, GA

Corresponding author:

Kristina L. Bajema

Centers for Disease Control and Prevention

CDC COVID-19 Response

1600 Clifton Road NE, Mailstop H24-6

Atlanta, GA 30329

pgz0@cdc.gov

Abstract

We compared severe acute respiratory syndrome–related coronavirus-2 seroprevalence estimated from commercial laboratory residual sera and a community household survey in metropolitan Atlanta during April-May 2020 and found these two estimates to be similar (4.94% versus 3.18%). Compared with more representative surveys, commercial sera can provide an approximate measure of seroprevalence.

Keywords:

Severe acute respiratory syndrome–related coronavirus-2 (SARS-CoV-2)

Seroprevalence

Coronavirus disease 2019 (COVID-19)

Convenience sampling

Accepted Manuscript

Introduction

In the setting of widespread community transmission of severe acute respiratory syndrome–related coronavirus-2 (SARS-CoV-2), serological surveys are important to inform the public health response. They can help to estimate infections at the population-level and inform transmission dynamics [1]. By detecting evidence of past infection, serological surveys can identify more infections than reported cases alone. This is particularly useful where testing is limited and for persons with mild or asymptomatic infections who may not be tested.

Many SARS-CoV-2 serological surveys are being conducted, but vary by participant and sample selection methods, assay type, period of study, and regional differences in the burden of disease. Carefully designed studies to select participants representative of the underlying population may provide the most accurate seroprevalence estimates but require considerable investment of time and resources [2, 3]. Conversely, convenience sampling can facilitate large, efficient studies and be used to follow seroprevalence trends over time, though the population from which specimens are obtained may not be representative of the general population [4-8].

To better understand how sampling strategies impact estimates of coronavirus disease 2019 (COVID-19) burden and to assess the ability of commercial laboratory sampling to accurately estimate SARS-CoV-2 population seroprevalence, we compared the seroprevalence estimate from a convenience sample of commercial laboratory residual sera (“commercial sampling”) with the estimate from a community survey designed to randomly select a representative sample of metropolitan Atlanta residents (“community survey”) [3]. For this comparison, we used the same serologic assay and matched overlapping geographical areas and time periods. With the commercial sampling, we also estimated seroprevalence at multiple timepoints from April-July 2020 in the Atlanta, Georgia, Metropolitan Statistical Area (MSA).

Methods

Survey Populations

In commercial sampling, persons of all ages throughout the Atlanta MSA, which includes DeKalb and Fulton Counties, had blood testing through a commercial laboratory from April 27-May 1, 2020 for routine screening or clinical management. Additional testing was also performed in the Atlanta MSA from May 23-30, June 15-17, July 6-10, and July 27-28. A convenience sample of deidentified specimens was selected by the laboratory for further SARS-CoV-2 testing. We targeted at least 300 specimens per age group [5]. The community survey was conducted in Dekalb and Fulton Counties from April 28-May 3, 2020. Three hundred ninety-four households and 696 participants of all ages were enrolled using a stratified two-stage cluster sampling design as previously described [3].

This activity was reviewed by CDC and was conducted consistent with applicable federal law and CDC policy.¹

SARS-CoV-2 Antibody Testing

For both investigations, SARS-CoV-2 antibody testing was performed by the Centers for Disease Control and Prevention (CDC) using a CDC-developed and validated enzyme-linked immunosorbent assay (ELISA) with pan-immunoglobulin against the anti-spike protein of SARS-CoV-2. Through methods previously described, specificity was 99.3% (95% CI: 98.3-99.9%) and sensitivity was 96.0% (95% CI: 90.0-99.9%) [5, 9]. To compare seroprevalence estimates using the same assay, serum samples in the community survey originally tested with a different assay [3] were re-tested using the CDC ELISA.

¹See e.g., 45 C.F.R. part 46, 21 C.F.R. part 56; 42 U.S.C. §241(d); 5 U.S.C. §552a; 44 U.S.C. §3501 et seq.

Analysis

We compared age, sex, and available race/ethnicity data in the two populations with demographics of the corresponding catchment areas using U.S. Census data [10]. For commercial sampling, we identified residents of the Atlanta MSA by ZIP codes contained in the 29-county catchment area using population density mapping. Seroprevalence results were age- and sex-standardized to the underlying population and adjusted for assay characteristics using previously described methods [5]. For this analysis, we report the final adjusted seroprevalence estimate as the mean of the bootstrap distribution.

We repeated analysis of the community survey using test results from the CDC assay by methods previously described [3]. For this analysis, we also adjusted the seroprevalence estimates to account for the assay characteristics [9], uncertainty in the test positivity from the community survey, and uncertainty of the sensitivity and specificity of the assay [11].

We examined whether there was a significant difference in seroprevalence estimates between commercial sampling and the community survey conducted in similar geographic areas during the same time period. We calculated the effective sample sizes using a binomial distribution with the corresponding mean and variance from each study and used a two-sample z-test to compare the estimates.

Finally, we calculated seroprevalence estimates for commercial specimens collected in the Atlanta MSA during three additional time periods from May through July and compared changes to reported cases over the same time period [12]. SAS (version 9.4), RStudio (version 4.0.2), and ArcGIS (version 10.7.1) were used to perform analyses.

Results

Among 1,343 persons in the Atlanta MSA with commercial residual sera collected from April 27-May 1, the median age was 57 [interquartile range (IQR) 36-73] and 601 (44.8%) were men (Table 1). Most resided in the northeastern part of the Atlanta MSA and in areas with predominantly non-Hispanic White residents (Figure 1). About 26% (353 persons) resided in DeKalb and Fulton Counties. Of the 696 community survey participants providing specimens from April 28-May 3 in the

original analysis, testing with the CDC ELISA could not be performed in 27 persons due to insufficient sample volume. Among 669 persons with a CDC ELISA result, the median age was 46 [IQR 36-73] and 302 (45.1%) were men. The sample sizes allowed 70% power to detect a 5% difference in seroprevalence (see Supplementary Methods).

The adjusted seroprevalence estimate from commercial sampling in the Atlanta MSA was 4.94% (95% confidence interval [CI], 3.34-6.64%). The adjusted seroprevalence estimate in DeKalb and Fulton Counties in the community survey was 3.18% (95% CI, 1.49-6.67%). These estimates did not differ significantly [$P = .40$; difference in seroprevalence 1.76% (95% CI, -1.52-5.06%)].

Seroprevalence estimates from commercial sampling declined in the Atlanta MSA from 4.94% in April 27-May 1 to 3.23% (95% CI, 1.86-4.72%) during June 15-17 and increased to 6.82% (5.04-8.71%) during July 27-28 (Supplementary Figure S1). The incidence of reported COVID-19 in the Atlanta MSA was relatively stable between late March and early June and increased after mid-June [12]. Demographic characteristics and geographic distribution of persons selected for SARS-CoV-2 testing did not change significantly over sampling periods.

Discussion

The validity of large-scale SARS-CoV-2 seroprevalence surveys utilizing clinical blood samples relative to survey methods traditionally more representative of the population has not been evaluated. We compared SARS-CoV-2 seroprevalence estimates from commercial laboratory residual sera testing with a community survey in metropolitan Atlanta and found estimates to be similar. The general trend in seroprevalence estimated from commercial sampling was also consistent with the pattern of reported cases in the Atlanta MSA.

Seroprevalence surveys involving commercial residual sera testing can estimate the total number of SARS-CoV-2 infections across many different geographic areas. Furthermore, the ability to efficiently conduct repeated sampling over time has allowed CDC to track jurisdiction-level spread of SARS-CoV-2 across the U.S. [8]. Despite these important advantages, such convenience sampling may be subject to bias, as persons who have their blood drawn for routine clinical purposes may differ systematically from the general population with regard to demographics, geographic distribution,

underlying health status, exposure risk, care-seeking behavior, and healthcare access [5]. Even with these biases, we found that commercial sampling provided an efficient means to estimate community SARS-CoV-2 seroprevalence relative to the community survey conducted during the same period in approximately overlapping geographic areas.

Many factors contributed to overall differences in the seroprevalence estimates, but we could not quantify their relative contribution or direction of bias in all cases. Local variation in SARS-CoV-2 outbreaks can impact seroprevalence estimates, particularly when measured over small geographic areas. While we observed early COVID-19 incidence to be higher in DeKalb and Fulton Counties compared with the entire Atlanta MSA, we were not able to compare seroprevalence in the two counties alone as the commercial sampling was not powered to examine county-level seroprevalence. This study was also not powered to detect less than a 5% difference between commercial sampling and the community survey. In addition, risk of SARS-CoV-2 infection and seroprevalence have been shown to vary across age and race/ethnicity groups [2, 3, 5, 8, 13] and can therefore impact seroprevalence estimates. Though individual-level race information was not available for the commercial sampling, the geographic distribution of the residual specimens within the Atlanta MSA (Supplementary Figure S2) suggests that commercial sampling resulted in underrepresentation of racial and ethnic minorities. Minority underrepresentation was also observed in the community survey [3], but without individual-level race information, we could not directly compare differences between the two surveys. Differences in the geographic areas as well as the populations sampled could account for the difference in seroprevalence estimates in the two studies.

In conclusion, we found estimation of SARS-CoV-2 seroprevalence from commercial laboratory residual sera testing to be a reasonable alternative approach to a community survey in metropolitan Atlanta. Given the limitations of commercial testing, more resource-intensive approaches may still be required to ascertain more accurate seroprevalence estimates in special populations, including minority groups, children, and healthcare workers [14]. Large-scale geographic seroprevalence surveys offer important advantages to allow tracking of SARS-CoV-2 spread through the U.S.

NOTES

Acknowledgments

The authors thank Quest for supplying specimens and the Molecular Pathogenesis and Immune Response Laboratory team for testing specimens: Bailey Alston, Muyiwa Ategbale, Shanna Bolcen, Darbi Boulay, Li Cronin, Ebenezer David, Yamini Gorantla, Panagiotis Maniatis, Kimberly Moss, Kristina Ortiz, So Hee Park, Palak Patel, Yunlong Qin, Evelene Steward-Clark, Heather Tatum, Andrew Vogan, and Briana Zellner; Samuel Lerma for assistance data management; the CDC STATT Laboratory for assistance with handling of specimens: Marla Petway, Brandi Smith, Wallace Rumph, Consuelo Hopkins, Brandy Rider, Brenda Upshaw, Daniel Perry, Novelle Smith, and Janelle Moore; the Georgia Department of Public Health: Lynn Paxton, Sandra Ford, Eric Jens, and Kathleen Toomey; and the CDC Community Serosurvey Team: Jennifer Harris, Lucy Breakwell, Glen Abedi, Nirma Bustamante, Olivia Almendares, Amy Schnall, Zunera Gilani, Tiffany Smith, Laura Gieraltowski, Kelsey McDavid, Ilana Schafer, Raiza Amiling, Claire Mattison, Margaret Cortese, Nicole Brown, Karen Change, Nicolas Deputy, Rodel Desamu-Thorpe, Chase Gorishek, Arianna Hanchey, Michael Melgar, Benjamin Monroe, Carrie Nielson, Gerald Pellegrini, Mays Shamout, Laura Tison, Sara Vagi, and Rachael Zacks.

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

Funding

This work was supported by the Centers for Disease Control and Prevention, Atlanta, Georgia.

Potential Conflicts of Interest

The authors have no reported conflicts of interest.

References

1. 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**.
2. Menachemi N, Yiannoutsos CT, Dixon BE, et al. Population Point Prevalence of SARS-CoV-2 Infection Based on a Statewide Random Sample - Indiana, April 25-29, 2020. *MMWR Morb Mortal Wkly Rep* **2020**; 69(29): 960-4.
3. Biggs HM, Harris JB, Breakwell L, et al. Estimated Community Seroprevalence of SARS-CoV-2 Antibodies - Two Georgia Counties, April 28-May 3, 2020. *MMWR Morb Mortal Wkly Rep* **2020**; 69(29): 965-70.
4. Rosenberg E, Tesoriero J, Rosenthal E, et al. Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. *Annals of Epidemiology* **2020**.
5. 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**.
6. Dodd RY, Xu M, Stramer SL. Change in Donor Characteristics and Antibodies to SARS-CoV-2 in Donated Blood in the US, June-August 2020. *Jama* **2020**.
7. Anand S, Montez-Rath M, Han J, et al. Prevalence of SARS-CoV-2 antibodies in a large nationwide sample of patients on dialysis in the USA: a cross-sectional study. *The Lancet* **2020**.
8. Bajema KL, Wiegand RE, Cuffe K, et al. Estimated SARS-CoV-2 Seroprevalence in the US as of September 2020. *JAMA internal medicine* **2020**.
9. 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. *bioRxiv* **2020**: 2020.04.24.057323.
10. United State Census Bureau. County Population by Characteristics: 2010-2019. Available at: <https://www.census.gov/content/census/en/data/datasets/time-series/demo/popest/2010s-counties-detail.html>. Accessed September 12.
11. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol* **1978**; 107(1): 71-6.

12. USAFacts. US Coronavirus Cases and Deaths. Available at:
<https://usafacts.org/visualizations/coronavirus-covid-19-spread-map>. Accessed September 15.
13. Centers for Disease Control and Prevention. CDC COVID Data Tracker: Demographic Trends of COVID-19 cases and deaths in the US reported to CDC. Available at:
<https://covid.cdc.gov/covid-data-tracker/#demographics>. Accessed September 23.
14. Shook-Sa BE, Boyce RM, Aiello AE. Estimation Without Representation: Early Severe Acute Respiratory Syndrome Coronavirus 2 Seroprevalence Studies and the Path Forward. *J Infect Dis* **2020**; 222(7): 1086-9.

Accepted Manuscript

Table 1. Demographic characteristics of persons tested for SARS-CoV-2 antibodies through commercial residual sera sampling (Atlanta, Georgia Metropolitan Statistical Area) and community surveillance (DeKalb and Fulton Counties, Georgia) compared with 2019 postcensal estimates for the respective catchment areas

	Commercial Residual Sera Testing N=1343	Community Surveillance N=669	Atlanta Metropolitan Statistical Area ⁴ N= 6,020,364	DeKalb and Fulton Counties ⁴ N= 1,823,234
Age, years (N, %)				
0-17	225 (16.8)	43 (6.4)	1,451,061 (24.1)	402,362 (22.1)
18-49	316 (23.5)	332 (49.6)	2,672,314 (44.4)	867,468 (47.6)
50-64	254 (18.9)	185 (27.7)	1,135,455 (18.9)	327,713 (18.0)
≥65	548 (40.8)	109 (16.3)	761,534 (12.6)	225,691 (12.4)
Male gender (N, %)	601 (44.8)	302 (45.1)	2,907,383 (48.3)	872,939 (47.9)
Race/Ethnicity	--			
White, non-Hispanic		315 (47.1)	2,779,924 (46.2)	643,896 (35.3)
Black or African American, non-Hispanic		260 (38.9)	2,068,532 (34.4)	870,869 (47.8)
Hispanic		43 (6.4)	660,674 (11.0)	141,530 (7.8)
Asian/Pacific Islander, non-Hispanic		24 (3.6)	379,325 (6.3)	129,375 (7.1)
Multiple race/Other/Unknown ¹		27 (4.0)	131,909 (2.2)	37,564 (2.1)

Dates of specimen collection	April 27-May 1	April 28-May 3	--	--
Seroprevalence (Estimate, 95% CI)	4.94% (3.34-6.64%) ²	3.18% (1.49-6.67%) ³		

¹Includes people with more than one race/ethnicity; people identifying as American Indian, Alaskan Native, or other races/ethnicities; and people with missing data on race/ethnicity.

²Age- and sex-standardized to the Atlanta MSA population and adjusted for assay characteristics.

³Age-, sex-, and race/ethnicity-standardized to the DeKalb and Fulton population and adjusted to account for assay characteristics, uncertainty in test positivity from the community survey, and uncertainty of the sensitivity and specificity of the assay.

⁴U.S. Census Bureau, 2019 postcensal estimates: <https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/counties/asrh/>.

Abbreviations: IQR, interquartile range; SARS-CoV-2, severe acute respiratory syndrome-related coronavirus-2

Accepted Manuscript