Clinical Infectious Diseases

EDITORIAL COMMENTARY

A Tale of 2 Studies: Study Design and Our Understanding of Severe Acute Respiratory Syndrome Coronavirus 2 Seroprevalence

Ross M. Boyce,[1] Bonnie E. Shook-Sa,[2] and Allison E. Aiello[3]

[1]Division of Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA, [2]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA, and [3]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

(See the Brief Report by Bajema et al on pages e3120–3.)

Keywords: COVID-19; coronavirus; seroprevalence; epidemiology.

The coronavirus disease 2019 (COVID-19) pandemic is arguably the most important public health crisis of the last century. To date, infections with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus have caused nearly 300 000 deaths in the United States alone [1], while also contributing to substantial excess morbidity and mortality from delayed and deferred care [2]. In addition to the direct and indirect health effects, policies intended to limit the spread of the disease have resulted in large-scale disruptions to education systems, economic activity, and social networks. Put simply, the COVID-19 pandemic has affected the daily lives of nearly all Americans in a way that no other health crisis has in our lifetimes.

One of the most challenging aspects of SARS-CoV-2 surveillance and prevention is the high proportion of individuals who remain asymptomatic or experience only mild symptoms and therefore never seek care [3]. Frequent testing can help identify individuals who harbor asymptomatic infections, but there are numerous logistical barriers associated with screening a large population [4]. An alternative method for quantifying the number of asymptomatic infections and characterizing the overall prevalence of SARS-CoV-2 in the population is the use of seroprevalence studies that measure antibody response to infection. A positive serological test provides evidence of past exposure to the virus and can be coupled with data on symptoms to assess whether an individual experienced symptomatic or asymptomatic infection.

Thus, we should not be surprised that serological studies of SARS-CoV-2 prevalence among the general population have assumed an outsized role in the public square and are frequently highlighted in major news outlets, often in support of disparate policy aims. For example, early in the course of the US pandemic, the now widely critiqued Santa Clara County study [5] was cited as evidence that the "COVID-19 death toll would be closer to that of seasonal flu" and used to question the wisdom of public health mandates [6]. In contrast, the authors of a more rigorous study in Indiana, which was conducted in the same month and found a similarly low proportion of seropositive individuals as the Santa Clara County study, concluded that "many persons in Indiana remain susceptible" and that "adherence to evidence-based public health mitigation measures is needed to reduce surge in hospitalizations and prevent morbidity and mortality from COVID-19" [7].

As the pandemic has evolved, more seroprevalence studies have been published. Most of these studies have used convenience samples that reflect very different populations across a wide range of study designs, serological assays, and statistical methods. Undoubtedly, these studies have provided critical data and substantially advanced our understanding of transmission dynamics. Yet, results across these differing study designs are frequently interpreted interchangeably and extrapolated to the general population, despite important questions regarding representativeness, generalizability, and methodological consistency [8]. Comparisons across studies are made more challenging by the rapidly changing dynamics of the pandemic and geographic differences in infection patterns. What is urgently needed,

Received 8 December 2020; editorial decision 10 December 2020; published online 18 December 2020.

Correspondence: R. M. Boyce, Division of Infectious Diseases University of North Carolina at Chapel Hill CB# 7030, Bioinformatics Building 130 Mason Farm Rd, 2nd Floor Chapel Hill, NC 27599-7030 (ross_boyce@med.unc.edu).

Clinical Infectious Diseases® 2021;73(9):e3124–6

DOI: 10.1093/cid/ciaa1868

Table 1. Advantages and Disadvantages of Seroprevalence Study Designs

Type of Study	Recruitment Challenges	Nonparticipation/ Loss to Follow-up	Level of Representativeness	Efficiency & Timeliness	Costs
Remnant studies (eg, blood donation, outpatients, inpatients)	None or low	Low	Unclear or unknown	High	Low
Community studies (eg, convenience samples, targeted community samples without weighting or sampling methods for ensuring representativeness)	Moderate	Moderate to high	Moderate	Moderate	Moderate/ high
Population-based studies with rigorous population weighted designs and adequate participation	High	Moderate to high	High	Low	High

however, are studies that compare seroprevalence estimates derived from multiple approaches covering the same geographic areas and timeframes while using similar laboratory techniques. Such "studies of studies" can highlight the potential tradeoffs between factors such as accuracy, efficiency, and cost (Table 1). Moreover, these comparison studies can provide the critical data to allow modelers to more accurately estimate the direction of the pandemic and subsequent effects on the health system.

Therefore, the article presented here by Bajema et al works toward filling a key methodological gap in the SARS-CoV-2 seroprevalence literature. The authors attempted to estimate and compare the prevalence of SARS-CoV-2 antibodies among residents of the Atlanta metropolitan area using 2 approaches: (1) a representative community survey and (2) a convenience sample of remnant blood from a commercial laboratory. Samples were obtained over roughly the same period, whereas serological testing was performed in a single laboratory using a validated assay; minimizing the potential causes of variability. The 2 methods, after statistical adjustment, arrived at similar estimates of seroprevalence: 4.94% (95% confidence interval [CI], 3.34-6.64) for the clinical study versus 3.18% (95% CI, 1.49-6.67) for the community study, with a difference in seroprevalence of 1.76% (95% CI, -1.52 to 5.06).

The study does have limitations related to seroprevalence comparisons between the 2 cohorts. First, there was a mismatch in target populations between the clinical and community-based studies: the Atlanta Metropolitan Statistical Area, comprising a 29-county region (2019 population estimate: 6 million), was used for the clinical study, whereas the community-study focused on DeKalb and Fulton counties (2019 population estimate: 1.8 million). The authors note in a previous publication that DeKalb and Fulton counties had the highest number of reported COVID-19 cases in Georgia at the time the community study began [9], creating some doubt about the comparability of these estimates. Second, the clinical sample was only standardized by age and sex and did not account for differences in socioeconomic status or race/ethnicity between the participants and the population, even as disparities in SARS-CoV-2 infection and outcomes have been associated with these factors [10]. Third, although the authors find no significant difference between the seroprevalence estimates, we must remind ourselves that a failure to reject a null hypothesis of no difference is not an acceptance of the null. In other words, the finding of no statistical difference between approaches does not indicate that no differences exist. We would also highlight that seemingly small differences can have substantial effects when applied to a large population. For example, the 1.76% absolute difference reported between the approaches represent more than 100 000 infections in the Atlanta metropolitan area, a number that certainly could affect contingency plans for hospital beds.

Despite the limitations of the study, it is generally reassuring that the seroprevalence estimates based on community and clinical samples were "pretty close." Unfortunately, this is no guarantee that community and clinical estimates will align for different time periods or geographic areas because factors driving selection bias can vary across populations and can change over time. Thus, we cannot say that the 2 methods should be used interchangeably. Instead, we should leverage the advantages of each approach to draw a more complete picture of the current state of SARS-CoV-2 transmission in the community. Convenience samples, and particularly those that use remnant clinical specimens, provide timely estimates at relatively low cost compared with community seroprevalence studies. Clinical studies may be particularly useful when repeated in the same target population over multiple time points. In contrast, more resource-intensive, prospective studies that seek to enroll a representative sample of the population are likely the only way to accurately estimate seroprevalence. This is particularly the case for subgroups such as historically marginalized populations that may not be well represented in samples relying on engagement with existing health systems. Given the well-established disparities related to socioeconomic status and race/ethnicity, getting the numbers right is critically important to achieving health equity and ultimately ending the pandemic.

Notes

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Financial support. This work was supported by funding for seroprevalence studies from the North Carolina Department of Health and Human Services (R. M. B., B. E. S., A. E. A.). R. M. B. is supported by the National Institutes of Health (K23AI141764). B. E. S. receives support from the Bahama University of North Carolina at Chapel Hill Center For AIDS Research, a National Institutes of Health–funded program P30 AI050410. A. E. A. received research support from Eunice Kennedy Shriver National Institute of Child Health and Human Development and National Institute on Aging of the National Institutes of Health under award numbers P2C HD050924 and P30 AG066615, respectively.

Potential conflicts of interest. The authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References

1. Johns Hopkins University Medicine. Coronavirus Resource Center. Available at: https://coronavirus.jhu.edu. Accessed 7 December 2020.
2. Czeisler MÉ, Marynak K, Clarke KEN, et al. Delay or avoidance of medical care because of COVID-19-related concerns - United States, June 2020. MMWR Morb Mortal Wkly Rep **2020**; 69:1250–7.
3. Oran DP, Topol EJ. Prevalence of asymptomatic SARS-CoV-2 infection: a narrative review. Ann Intern Med **2020**; 173:362–7.
4. Binnicker MJ. Challenges and controversies to testing for COVID-19. J Clin Microbiol **2020**; 58:e01695–20.
5. Bendavid E, Mulaney B, Sood N, et al. COVID-19 antibody seroprevalence in Santa Clara County, California. medRxiv **2020**; 2020.04.14.20062463.
6. Bogan A. New data suggest the coronavirus isn't as deadly as we thought; a study finds 50 to 85 times as many infections as known cases—meaning a far lower fatality rate. Wall Street Journal (Online) **2020**. Available at: http://libproxy.lib.unc.edu/login?url=https://www.proquest.com/newspapers/new-data-suggest-coronavirus-isnt-as-deadly-we/docview/2390912277/se-2?accountid=14244.
7. 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**; 69:960–4.
8. 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**; 222:1086–9.
9. Biggs HM, Harris JB, Breakwell L, et al.; CDC Field Surveyor Team. Estimated community seroprevalence of SARS-CoV-2 antibodies - two Georgia counties, April 28-May 3, 2020. MMWR Morb Mortal Wkly Rep **2020**; 69:965–70.
10. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial disparities in incidence and outcomes among patients with COVID-19. JAMA Netw Open **2020**; 3:e2021892.