

## **Are SARS-CoV-2 seroprevalence estimates biased?**

Saki Takahashi, PhD<sup>1</sup>, Bryan Greenhouse, MD, MA<sup>1,2</sup>, Isabel Rodríguez-Barraquer, MD, PhD<sup>1</sup>

<sup>1</sup>EPPIcenter Program; Division of HIV, ID, and Global Medicine; Department of Medicine, University of California, San Francisco, San Francisco, CA, USA

<sup>2</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA

### **Corresponding author:**

Isabel Rodríguez-Barraquer

University of California, San Francisco

Division of HIV, ID, and Global Medicine

995 Potrero Avenue, Ward 84

San Francisco, CA 94110

[isabel.rodriguez@ucsf.edu](mailto:isabel.rodriguez@ucsf.edu)

(628) 206-8844

**Brief summary:** Biased seroprevalence estimates can result from using serological assays optimized using validation sets that do not represent disease spectrum in the general population. Correct interpretation of SARS-CoV-2 serosurveys requires quantifying how assay sensitivity varies with disease severity and over time.

**Key words:** COVID-19; SARS-CoV-2 seroprevalence; antibody testing; test performance; spectrum bias

Accepted Manuscript

Serosurveys are needed to understand how many people have been infected by SARS-CoV-2 and where we are in the epidemic curve. Unprecedented serosurveillance efforts have been launched (e.g., Solidarity II, NIH) to generate local and global infection estimates to guide social distancing policies. Recently published results suggest that the proportion of the population that has been infected, even in places with explosive outbreaks such as Spain or New York City, is low and far from the levels required for herd immunity [1–3].

Multiple serological assays and rapid tests are now available and, as of July 16<sup>th</sup>, the Food and Drug Administration has authorized the use of twenty-nine, which report a range of test performance characteristics (i.e., sensitivity and specificity, as well as positive and negative predictive values) [4]. Assay validation requires samples from individuals with known infection status in order to determine test performance characteristics. Due to potential cross-reactivity of antibody responses to seasonal coronaviruses, much of the focus of assay development has been on ensuring near perfect specificity, to minimize the risk of false positive results. This is particularly important during early stages of the epidemic, when the number of true positives is expected to be very low. However, if the purpose of deploying a serological assay is to quantify the proportion of the population that has been infected by SARS-CoV-2 (i.e., serosurveillance), adequate characterization of assay sensitivity to detect prior infection in the general population is important as well. We raise this issue because in the absence of such characterization, it will not be possible to generate accurate estimates of population-level exposure to this novel pathogen [5].

A growing body of evidence suggests that asymptomatic and mild SARS-CoV-2 infections, together making up over 95% of all infections, may be associated with lower antibody titers than more severe infections [6–16]. Similarly, it is known that antibody levels peak a few weeks after infection and

then decay gradually [17–19]. Yet, positive controls used for assay optimization and validation are usually limited to early convalescence samples from hospitalized patients with severe disease, leading to what is commonly known as *spectrum bias*. Sensitivities estimated from these sample sets may therefore overestimate the actual sensitivity that the assay would have when applied to the general population, leading to underestimates of the true seroprevalence.

To illustrate this point, we quantified the amount of bias in estimating population seroprevalence potentially introduced by the choice of positive controls used to evaluate assay sensitivity. We evaluated the impact of: (a) using validation sets with different proportions of severe, mild, and asymptomatic infections (Figure 1A), and (b) using validation sets with samples from recent infections only (Figure 2A and 2B). While the relationship between disease severity and assay sensitivity has yet to be quantified for most assays, we assumed that sensitivity was highest for severe infections and considered a range of values for asymptomatic and mild infections. Similarly, we assumed that sensitivity peaked early after infection and then decayed over time (Figure 2C). For simplicity, we fixed test specificity at 100%.

Assays with imperfect sensitivity lead to underestimates of the true seroprevalence (Figure 1B, grey line), but can be easily corrected for if the actual sensitivity of the assay in the sampled population is known (Figure 1B, purple line) [20]. However, if test sensitivity has been determined from positive control sets skewed towards those with severe clinical outcomes (high antibody levels), the estimated prevalence, even after correction, will still underestimate the true prevalence (Figure 1B, cyan and gold lines). The magnitude of the underestimate will depend on how biased the distribution of positive controls is relative to the population, and on how much assay sensitivity varies with disease severity (Figure 1C). Similarly, corrected estimates of prevalence will only equal

the true prevalence if decreases in sensitivity due to waning antibody responses over time can be accounted for (Figure 2C). If spectrum bias stemming from clinical outcomes as well as times since infection are both present, underestimation of the true prevalence will be greater.

These results have important implications for assay development and for the interpretation of SARS-CoV-2 seroprevalence studies. First, they highlight the need to quantify the extent to which the sensitivity of the assays used in ongoing serosurveillance studies varies with disease severity and over time, and possibly differentially over time by disease severity. Incorporating loss of sensitivity with increasing time since infection will gain importance as the pandemic progresses. More importantly, these results highlight the need for long term studies characterizing kinetics of antibody responses to SARS-CoV-2 across the severity spectrum. If antibody responses are significantly lower in milder cases, or if there is significant waning in the months following infection, assays for seroprevalence studies should be optimized to detect these lower titers. Finally, these results caution against accepting aggregate sensitivities and specificities reported by assay manufacturers at face value. Ideally, sensitivities and specificities should be stratified by disease severity and time since infection, and the characteristics of the validation set should be reported at a minimum.

Based on these results, we propose the following recommendations to improve the optimization, validation, and interpretation of SARS-CoV-2 seroprevalence studies. First, we emphasize the need for studies measuring antibody levels longitudinally in individuals who have had confirmed SARS-CoV-2 infection. Such studies should include individuals from a range of disease severity, as discussed here, as well as other factors that may affect assay sensitivity such as age and immunodeficiency. Second, a set of standards should be established for assay validation. Importantly, positive controls should include samples across the spectrum of severity and times

since infection. To these ends, a SARS-CoV-2 reference serum bank of well-characterized and representative positive control sets could be created. This reference panel would also facilitate direct comparability of test performance across assays. Third, we recommend developing a set of criteria for reporting the results of assay validation. In particular, characteristics of validation samples used to determine test performance should be reported, as well as any heterogeneities in performance identified in the validation process.

Correct interpretation of SARS-CoV-2 seroprevalence studies will not be possible until we know the sensitivity of serological assays to detect mild and asymptomatic infections and infections occurring months beforehand. We urge the FDA to revisit reporting requirements for performance characteristics of SARS-CoV-2 serological assays with this in mind.

Accepted Manuscript

**Author contributions:** ST, BG, and IRB conceived the study, performed the simulations, and wrote the manuscript.

**Declaration of interests:** We declare no competing interests.

**Role of the funding source:** ST was supported by the Schmidt Science Fellows, in partnership with the Rhodes Trust. The funders had no role in the writing of the manuscript or the decision to submit it for publication.

**Corresponding author:**

Isabel Rodríguez-Barraquer

University of California, San Francisco

Division of HIV, ID, and Global Medicine

995 Potrero Avenue, Ward 84

San Francisco, CA 94110

[isabel.rodriguez@ucsf.edu](mailto:isabel.rodriguez@ucsf.edu)

(628) 206-8844

## References

1. Gobierno de España. Estudio ENE-COVID19: Primera Ronda [Internet]. Available from: [https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID\\_Informe\\_preliminar\\_cierre\\_primera\\_ronda\\_13Mayo2020.pdf](https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID_Informe_preliminar_cierre_primera_ronda_13Mayo2020.pdf)
2. Stringhini S, Wisniak A, Piumatti G, et al. Repeated seroprevalence of anti-SARS-CoV-2 IgG antibodies in a population-based sample from Geneva, Switzerland [Internet]. *Infectious Diseases (except HIV/AIDS)*. medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.05.02.20088898v1.abstract>
3. Amid Ongoing COVID-19 Pandemic, Governor Cuomo Announces Results of Completed Antibody Testing Study of 15,000 People Showing 12.3 Percent of Population Has COVID-19 Antibodies. **2020** [cited 2020 May 16]; . Available from: <https://www.governor.ny.gov/news/amid-ongoing-covid-19-pandemic-governor-cuomo-announces-results-completed-antibody-testing>
4. Center for Devices, Radiological Health. EUA Authorized Serology Test Performance [Internet]. U.S. Food and Drug Administration. 2020 [cited 2020 Aug 5]. Available from: <https://www.fda.gov/medical-devices/emergency-situations-medical-devices/eua-authorized-serology-test-performance>
5. Krammer F, Simon V. Serology assays to manage COVID-19. *Science* [Internet]. American Association for the Advancement of Science; **2020** [cited 2020 May 15]; . Available from: <https://science.sciencemag.org/content/early/2020/05/14/science.abc1227.full>
6. Whitman JD, Hiatt J, Mowery CT, et al. Test performance evaluation of SARS-CoV-2 serological assays [Internet]. *Infectious Diseases (except HIV/AIDS)*. medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.04.25.20074856v1.abstract>
7. Okba NMA, Muller MA, Li W, et al. SARS-CoV-2 specific antibody responses in COVID-19 patients [Internet]. *Infectious Diseases (except HIV/AIDS)*. medRxiv; 2020. Available from: <http://dx.doi.org/10.1101/2020.03.18.20038059>
8. Wu F, Wang A, Liu M, et al. Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications [Internet]. *Infectious Diseases (except HIV/AIDS)*. medRxiv; 2020. Available from: <http://dx.doi.org/10.1101/2020.03.30.20047365>
9. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Available from: <http://dx.doi.org/10.25561/77482>
10. Cervia C, Nilsson J, Zurbuchen Y, et al. Systemic and mucosal antibody secretion specific to SARS-CoV-2 during mild versus severe COVID-19 [Internet]. *bioRxiv*. 2020 [cited 2020 May 30]. p. 2020.05.21.108308. Available from: <https://www.biorxiv.org/content/10.1101/2020.05.21.108308v1>
11. Solbach W, Schiffner J, Backhaus I, et al. Antibody profiling of COVID-19 patients in an urban low-incidence region in Northern Germany [Internet]. *Infectious Diseases (except HIV/AIDS)*. medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.05.30.20111393v1.abstract>
12. Bartolini A, Scapaticci M, Bioli M, Lazzarotto T, Re MC, Mancini R. Immunochromatographic



assays for COVID-19 epidemiological screening: our experience [Internet]. Pathology. medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.05.28.20116046v1.abstract>

13. Long Q-X, Tang X-J, Shi Q-L, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nat Med [Internet]. **2020**; . Available from: <http://dx.doi.org/10.1038/s41591-020-0965-6>
14. Long Q-X, Liu B-Z, Deng H-J, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. Nat Med [Internet]. **2020**; 26(6):845–848. Available from: <http://dx.doi.org/10.1038/s41591-020-0897-1>
15. Lynch KL, Whitman JD, Lacanienta NP, et al. Magnitude and kinetics of anti-SARS-CoV-2 antibody responses and their relationship to disease severity [Internet]. Infectious Diseases (except HIV/AIDS). medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.06.03.20121525v1.abstract>
16. Lei Q, Li Y, Hou H, et al. Antibody dynamics to SARS-CoV-2 in Asymptomatic and Mild COVID-19 patients [Internet]. Infectious Diseases (except HIV/AIDS). medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.07.09.20149633v1.abstract>
17. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2. JAMA [Internet]. **2020**; . Available from: <http://dx.doi.org/10.1001/jama.2020.8259>
18. Rosado J, Cockram C, Merklings S, et al. Serological signatures of SARS-CoV-2 infection: Implications for antibody-based diagnostics. **2020** [cited 2020 May 14]; . Available from: <https://hal-pasteur.archives-ouvertes.fr/pasteur-02569149/document>
19. Iyer AS, Jones FK, Nodoushanian A, et al. Dynamics and significance of the antibody response to SARS-CoV-2 infection [Internet]. Infectious Diseases (except HIV/AIDS). medRxiv; 2020. Available from: <https://www.medrxiv.org/content/10.1101/2020.07.18.20155374v1.abstract>
20. Diggle PJ. Estimating Prevalence Using an Imperfect Test [Internet]. Epidemiology Research International. 2011. p. 1–5. Available from: <http://dx.doi.org/10.1155/2011/608719>

**Figure 1 Title: Impact of assay validation set and symptom-dependent sensitivity on estimation of seroprevalence**

**Figure 1 Legend:** **(A)** True distribution of clinical outcomes observed among SARS-CoV-2 infected individuals (left bar), classified as asymptomatic (blue), mild (green), or severe (red), and three simulated sets of positive controls used in assay validation. **(B)** For a range of true prevalence (x-axis), we calculated the prevalence that would be estimated in the population (y-axis) under the scenarios of no correction for test characteristics (grey line), and corrections for test sensitivity as determined from each validation set (purple, gold, and cyan lines). We assumed values of symptom-dependent sensitivities ( $Se$ ) as shown in the panel and a specificity of 100%, leading to crude sensitivities of 53.15% (set 1), 76.5% (set 2), and 95% (set 3) in the validation sets. For all of these scenarios, the true sensitivity in the general population would be 53.15%. **(C)** For a range of sensitivities in mild (x-axis) and asymptomatic infections (assuming sensitivity in asymptomatic infections is  $\frac{2}{3}$  of the sensitivity in mild infections), we calculated the ratio of estimated to true prevalence (y-axis) for each validation set used for correction. The scenarios in Figure 1B are highlighted with the corresponding color.

Accepted Manuscript

**Figure 2 Title: Impact of assay validation set and time-dependent sensitivity on estimation of seroprevalence**

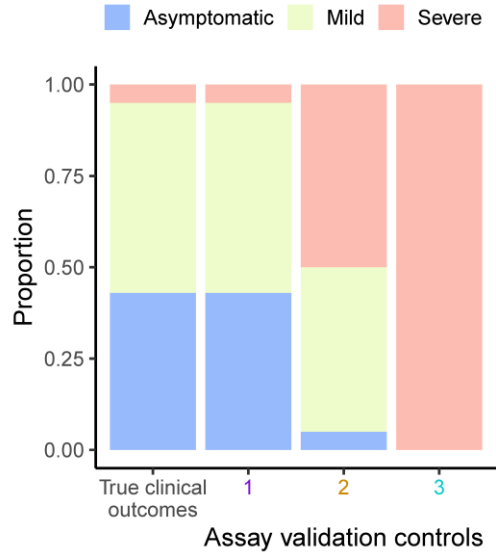
**Figure 2 Legend: (A)** Epidemic curve from an SEIR transmission model with SARS-CoV-2-like parameters in a completely susceptible population, with mean latent period of 3.5 days, mean infectious period of 7 days, and  $R_0$  of 2.1. We simulated cross-sectional seroprevalence surveys performed at 60, 180, and 300 days into the epidemic (labeled points on the curve). **(B)** The proportion of the ever-infected individuals that experienced infection within 0-60 days (black), 60-180 days (medium grey), and 180+ days (light grey) of the serosurveys performed at days 60, 180, and 300. The last bar shows a simulated set of positive controls used in assay validation, which includes only recent samples taken within 0-60 days after infection. We assume for these simulations that the validation controls, and all infections, represent severe infections with a baseline test sensitivity of 95%. **(C)** For a range of true prevalence (x-axis), we calculated the prevalence values that would be estimated in the population at days 60, 180, and 300 (y-axis) (assuming sensitivity ( $Se$ ) is reduced to 80% of the baseline level in infections that are 60-180 days old, and to 60% of the baseline level in infections that are 180+ days old, and a specificity of 100%), correcting for test characteristics using the positive control set from the last bar in Figure 2B. We used the procedure described in the Extended Methods section to calculate the estimated prevalences. We also considered prevalence values that would be estimated in the population if the two sources of spectrum bias (both clinical outcomes and time since infection) are present. For this simulation (filled squares), we calculated the prevalence that would be estimated at day 180, now assuming that the distribution of true clinical outcomes and their test sensitivities is equal to that in Figure 1B (i.e., in terms of severity: 43% asymptomatic, 52% mild, 5% severe, and in terms of sensitivity: 95% in severe, 60% in mild, and 40% in asymptomatic), and that the positive controls are all severe, recent infections.

**Code Availability:** All R code used for this analysis is available at <https://github.com/sakitakahashi/COVID-sensitivity>

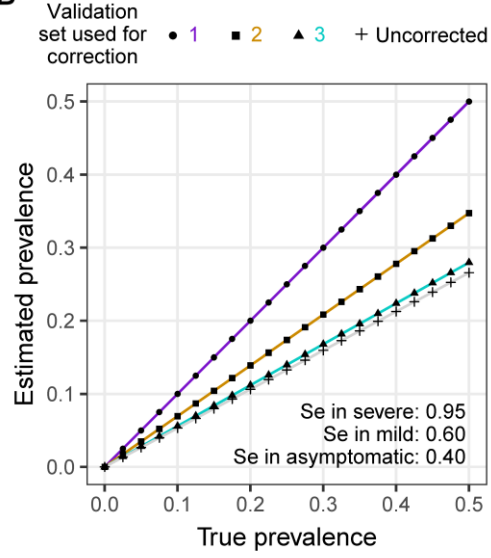
Accepted Manuscript

Figure\_1

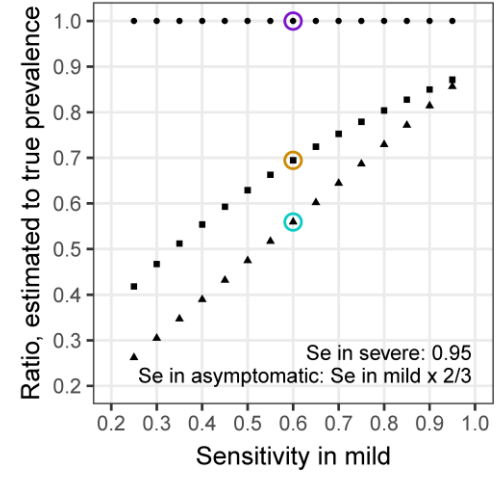
A



B

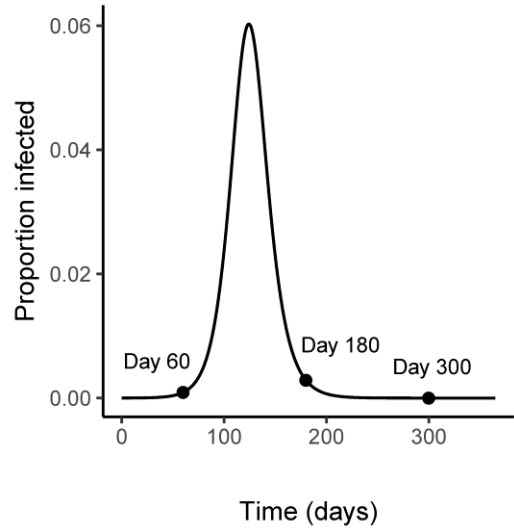


C

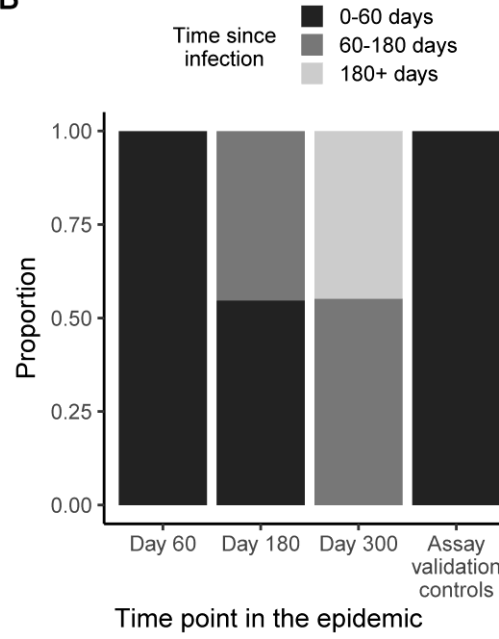


Figure\_2

A



B



C

