SARS-CoV-2 Cumulative Incidence and Period Seroprevalence: Results from a Statewide Population-Based Serosurvey in California

*Katherine Lamba MPH¹, *Heather Bradley PhD², Kayoko Shioda PhD DVM MPH³, Patrick S. Sullivan DVM PhD⁴, Nicole Luisi MPH MS⁴, Eric W. Hall PhD⁴, Megha L. Mehrotra PhD MPH¹, Esther Lim MPH¹, Seema Jain MD¹, Amanda Kamali MD¹, Travis Sanchez DVM MPH⁴, Benjamin A. Lopman PhD⁴, Mansour Fahimi PhD⁵, Aaron J. Siegler PhD⁴

*Katherine Lamba and Heather Bradley contributed equally to this manuscript

¹California Department of Public Health, Richmond, California, United States

²Department of Population Health Sciences, Georgia State University School of Public Health, Atlanta, Georgia, United States

³Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States

⁴Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States

⁵Marketing Systems Group, Philadelphia, Pennsylvania, United States

Summary: A population-based serosurvey conducted in California found approximately one in three SARS-CoV-2 infections had been reported by November 2020. Disparities in seroprevalence by race/ethnicity and social determinants of health were consistent with those seen among PCR-confirmed cases.
Corresponding Author:

Katherine Lamba, MPH
Epidemiologist, Infectious Diseases Branch
California Department of Public Health
850 Marina Bay Parkway
Richmond, California 94804
United States
Katherine.Lamba@cdph.ca.gov
(510) 620-3429

Alternate Corresponding Author:

Heather Bradley, PhD
Assistant Professor
Department of Population Health Sciences
Georgia State University School of Public Health
P.O. Box 3995, Atlanta GA, 30302-3995
United States
hbradley@gsu.edu
(404) 413-1400
ABSTRACT

Background: California has reported the largest number of COVID-19 cases of any U.S. state, with more than 3.5 million confirmed as of March 2021. However, the full breadth of SARS-CoV-2 transmission in California is unknown since reported cases only represent a fraction of all infections.

Methods: We conducted a population-based serosurvey, utilizing mailed, home-based SARS-CoV-2 antibody testing along with a demographic and behavioral survey. We weighted data from a random sample to represent the adult California population and estimated period seroprevalence overall and by participant characteristics. Seroprevalence estimates were adjusted for waning antibodies to produce statewide estimates of cumulative incidence, the infection fatality ratio (IFR), and the reported fraction.

Results: California’s SARS-CoV-2 weighted seroprevalence during August–December 2020 was 4.6% (95% CI: 2.8–7.4%). Estimated cumulative incidence as of November 2, 2020 was 8.7% (95% CrI: 6.4%–11.5%), indicating 2,660,441 adults (95% CrI: 1,959,218–3,532,380) had been infected. The estimated IFR was 0.8% (95% CrI: 0.6%–1.0%), and the estimated percentage of infections reported to the California Department of Public Health was 31%. Disparately high risk for infection was observed among persons of Hispanic/Latinx ethnicity and people with no health insurance and who reported working outside the home.
**Conclusions:** We present the first statewide SARS-CoV-2 cumulative incidence estimate among adults in California. As of November 2020, approximately one in three SARS-CoV-2 infections in California adults had been identified by public health surveillance. When accounting for unreported SARS-CoV-2 infections, disparities by race/ethnicity seen in case-based surveillance persist.

**Keywords:** SARS-CoV-2, seroprevalence, cumulative incidence
Introduction

Since SARS-CoV-2 virus infection was first detected in California in January 2020, more than 3.5 million reported cases (8,900 per 100,000 population) and 54,000 related deaths have been reported [1], making California the state with the largest number of reported cases in the U.S. Like other U.S. states, California relies on data from diagnosed infections, hospitalizations, and deaths that are reported to state and local health departments (e.g., California Department of Public Health [CDPH]) to monitor the burden of SARS-CoV-2 infections. These data are critical for monitoring, but because undiagnosed infections are not reported, have limited utility for estimation of key epidemiologic indicators such as cumulative incidence, the infection fatality ratio (IFR), and the reported fraction of infections. These indicators require population-based estimates of the seroprevalence of antibodies to SARS-CoV-2.

Serosurveys, which pair SARS-CoV-2 antibody testing with surveys about demographic, behavioral, and clinical characteristics, are an effective tool for estimating burden of diagnosed and undiagnosed infections alongside risk factors for infection [2,3]. When conducted in population-based samples, serosurveys can provide relatively unbiased estimates of the burden of disease in a geographic area and of disparities in infection across population groups. To date, most California-based serosurveys have had limited geographic reach or were conducted in specific populations, such as blood donors or essential workers, that are likely not representative of the general population of the state [4–7]. Given California's diverse population of approximately 40 million, a representative statewide study was needed to capture data from people less likely to be included in convenience samples and to understand the full extent of SARS-CoV-2 burden in California and disparities that might exist by population characteristics and behaviors.
We conducted a population-based serosurvey in California as part of the COVIDVu study, a longitudinal probability survey of U.S. households using mailed at-home specimen collection for PCR and serology testing [8]. In addition to period seroprevalence during August–December, 2020, we estimated cumulative incidence, the IFR, and the percentage of infections that were reported. Given the common finding in serosurveys of waning detectable antibodies over time [9-11], we applied a model [12] to adjust the seroprevalence estimate for antibody waning.

Methods

Sampling

We randomly sampled 8,726 households in California from a frame derived from the USPS Computerized Delivery Sequence File, which has been used extensively for survey research [13–15] and represents nearly all housed, non-institutionalized persons in the state. To account for sub-optimal response among under-represented racial and ethnic groups in a pilot survey, we over-sampled census blocks with >50% Black residents and households associated with surnames likely to signify Hispanic/Latinx ethnicity at a rate of three times the California general population.

Survey and laboratory procedures

Survey and laboratory procedures have been described elsewhere [8]. Briefly, during August–December, 2020, we mailed an introductory letter followed a few days later by a kit for self-collecting specimens for SARS-CoV-2 PCR and antibody testing. The kit included an anterior nares (AN) swab, a dried blood spot (DBS) card and single use lancet, and instructions for specimen self-collection using text and illustrations [16,17]. An adult resident
provided a list of all persons living in the household along with each person’s age, and one adult \( \geq 18 \) years was randomly selected for study participation by the study’s electronic platform. Persons who consented were asked to respond to an online survey and to provide AN and DBS specimens, which were mailed back prepaid to a central laboratory.

PCR testing was performed as previously described [18] on AN specimens using the Thermo Emergency Use Authorization (EUA) Version 2 kit (Thermo Fisher Scientific, Waltham, MA). DBS specimens were tested using the BioRAD Platelia Total Antibody test (i.e., IgA, IgM, IgG; BioRad, Hercules, California; sensitivity 92.2%, specificity 99.6%), which was validated as a Laboratory Developed Test (LDT) under CLIA/CAP protocols. Because the BioRAD test detects SARS-CoV-2 IgG targeting the nucleocapsid protein, which may wane more quickly than IgG targeting the spike protein [9,19], we performed a sensitivity analysis by testing non-reactive samples with the EUROIMMUN IgG assay (EUROIMMUN, Lubeck, Germany) that targets the spike protein. Test results were returned to participants. Participants were compensated $60-100 (depending on sampling group) for completing the survey and submitting specimens.

Patient Consent Statement

COVIDVu was approved by the Emory University Institutional Review Board (STUDY00000695) and was deemed exempt public health surveillance by the California Committee for the Protection of Human Subjects (2020-124). Written consent was obtained from participants.
Sample weights

We computed sample weights to estimate key epidemiologic parameters representing the non-institutionalized, housed adult (aged ≥18) population of California. The first step was to ensure that participants had complete data for weighting variables, a process accomplished with hierarchical hot deck imputation for gender (replacing 0.1% missing data), education (1.6% missing), race (7.1% missing), ethnicity (2.5% missing), marital status (2.9% missing), and income (14.4% missing). The second step was to develop design weights, the reciprocal of the probability of being selected, that were adjusted for differential non-response using Classification and Regression Tree analysis. This analysis identified characteristics distributed differently across responding and non-responding households, which included homeownership status (rent vs. own), residence in a household located in a census tract with >50% Black residents, presence of likely Hispanic/Latinx surname, and presence of household information about income or number of adults on the address-based sampling frame. The third step was to calibrate nonresponse-adjusted design weights to characteristics of adults residing in California, using an iterative proportional fitting (raking) procedure to align the weights of California respondents simultaneously to bivariate distributions of gender nested with age, race-ethnicity, education, income, and marital status from U.S. Census estimates for California [20]. In the final step, we examined weights to detect extreme outliers and trimmed at the 1st and 99th percentiles of the weight distribution.

Data analysis

We used standard survey analytic procedures in SAS v9.4 (PROC SURVEY) to estimate statewide seroprevalence during August–December, 2020. We estimated weighted seroprevalence overall and by participant characteristics and sampling month with accompanying 95% Modified Wilson score confidence intervals (CI). We assessed
differences in seroprevalence across participant characteristics using prevalence ratios (PRs), which were computed using average marginal predictions from logistic regression in SUDAAN and associated confidence intervals.

To estimate cumulative incidence of SARS-CoV-2 infection through the median date of sampling (November 2, 2020), we adjusted the statewide seroprevalence estimate to account for antibody waning below a detectable level. In this analysis, we used a Bayesian model that has been previously described [12]. Briefly, it estimates timing of infections based on 1) an external estimate of time from symptom onset to seroconversion [21], 2) estimated time from seroconversion to seroreversion from New York City [12], 3) time series data on COVID-19-related deaths reported to CDPH through February 10, 2021, and 4) the distribution of timing of symptom onset to deaths in California. The model is calibrated with the statewide seroprevalence data estimated from this analysis. The model allowed us to directly estimate the IFR and derive a cumulative incidence estimate using the total number of modeled infections since the beginning of the epidemic in California. We estimated the reported fraction (the number of COVID-19 cases reported divided by the estimated number of total infections) as the ratio of PCR-confirmed COVID-19 cases reported to CDPH to the estimated cumulative incidence as of November 2, 2020. Credible intervals (CrI) for this ratio were derived using the Bayesian 95% credible intervals for the cumulative incidence estimate.

Results

Of 8,726 California households sampled, 357 (4.1%) were ineligible as evidenced by letters or kits returned undeliverable. Of 8,369 eligible households, 1,188 (14.2%) completed household enumeration, consent, and the online survey. Of those, 983 (83%) completed specimen self-collection and had a valid immunoglobulin (Ig) result, representing 11.7% of
eligible households (Figure 1). Unweighted and weighted distributions of the sample by characteristics are described in Table 1 with comparison to the adult California population.

Overall, weighted statewide SARS-CoV-2 seroprevalence among adults during August–December, 2020 was 4.6% (95% CI: 2.8%–7.4%) with a median specimen collection date of November 2, 2020 (Table 2). Ninety-eight non-reactive samples were re-tested using the EUROIMMUN IgG assay targeting the spike protein; none were reactive for seropositivity. Seroprevalence was 7.5 times as high among Hispanic/Latinx persons compared to non-Hispanic/Latinx white persons (PR 95% CI: 2.8–20.2) and higher among 35 - 44 year olds compared to 55 - 64 year olds (PR=3.3; 95 CI: 1.0–10.1%). Seroprevalence was also higher among people with no health insurance compared to people with private insurance (PR=4.5, 95% CI: 1.2–16.9%) and among people who left home for work versus those who did not (PR=3.9, 95% CI: 1.1–14.0%). Having contact with someone with a confirmed COVID-19 infection and having a prior COVID-19 diagnosis were both associated with higher seroprevalence, but having COVID-19 symptoms was not associated with higher seroprevalence.

The estimated statewide cumulative incidence of SARS-CoV-2 infection among California adults as of November 2, 2020, adjusted for waning antibodies was 8.7% (95% CrI: 6.4%–11.5%); an estimated 2,660,441 total infections (95% CrI: 1,959,218–3,532,380) had occurred among adults by that date. Based on these estimates and the number of PCR-confirmed COVID-19 cases reported to CDPH by November 2, 2020, the estimated reported fraction was 31%, meaning about one in three SARS-CoV-2 infections among adults were diagnosed and reported to CDPH as COVID-19 cases through early November. The estimated IFR among California adults was 0.8% (95% CrI: 0.6%–1.0%).
Discussion

We report the first representative statewide estimate of cumulative SARS-CoV-2 incidence using a population-based probability sampling approach for California, adjusted for waning antibodies. By early November 2020, nearly 9% of California adults had been infected with SARS-CoV-2, with about one in three infections diagnosed and reported to the state. By accounting for unreported infections, we estimated an IFR of 0.8% among adults in California.

Results from this study indicated large disparities in burden of infection among Californians, particularly among Hispanic/Latinx persons compared to non-Hispanic/Latinx white persons. These data support other findings in California, which have documented seroprevalence up to 3 times as high among Hispanic/Latinx persons compared to non-Hispanic/Latinx white persons [6,22,23]. We also reported differences in seroprevalence by insurance status and workplace. This is consistent with previous reports of varying seroprevalence by social determinants of health such as household income and housing status [2,22], and parallels inequities seen among PCR-confirmed cases [24]. Socioeconomic factors including essential worker status and ability to physically distance from others while at work or home, are associated with risk for SARS-CoV-2 infection; therefore a continued focus on health equity in California’s vaccine distribution is essential [23].

Population-level seroprevalence and cumulative incidence are critical indicators for monitoring the course of epidemics in populations. These indicators have been particularly challenging to estimate for SARS-CoV-2 because of the large number of asymptomatic infections and barriers to testing and diagnosis, particularly early in the pandemic [25]. Estimates of SARS-CoV-2 burden of disease are primarily derived from cases reported to health departments, and the Centers for Disease Control and Prevention’s (CDC) most recent
estimate suggests approximately 22% of infections are diagnosed and reported nationally [26]. Infections among racial or ethnic subgroups may be less likely to be detected through routine surveillance because of limited access to, or usage of, testing services [27]. This was evident in our findings, which show racial and ethnic disparities observed among PCR-confirmed cases may be larger when accounting for undiagnosed infections. California surveillance data suggest Hispanic/Latinx persons have nearly 3 times the PCR-confirmed infections (per population size) of white persons [28]. We estimated that seroprevalence was 7.5 times as high among Hispanic/Latinx persons compared to non-Hispanic/Latinx white persons. Equitable access to SARS-CoV-2 testing will aid in identification and reporting of cases across racial and ethnic groups [29]. Because there is no evidence that antibody waning is differential by race, cross-sectional serosurveys will continue to be an important tool for monitoring disparities.

With nearly 40 million residents representing a geographically and demographically diverse population, conducting representative serosurveys generalizable to California’s population has been an ongoing challenge. Seroprevalence estimates from several local and population-specific serosurveys in California have ranged from <1% to over 21% and have varied greatly depending on sampling period, geographic location, and population sampled [4–7,22–24,30–33]. CDC estimates for California using clinical laboratory residual specimens ranged from 4.1% in September 2020 to 18.1% in January 2021, with an estimated seroprevalence from mid-November 2020 of 6.6% [34]. This CDC estimate may be lower than our estimated cumulative incidence for a similar time period because it does not account for waning antibodies and excludes specimens specifically collected for COVID-19 testing. Statewide estimates based on electronic laboratory reporting to CDPH from clinical laboratories and blood banks indicate 38% seropositivity during February 2021; these data
include persons seeking clinical care, donating blood, and those that may have been vaccinated and may not be representative of all Californians [35].

This study has limitations. First, while our response rate was within expected range for an address-based survey, it was suboptimal, with approximately 12% of sampled households providing a valid specimen for SARS-CoV-2 antibody testing. This study was the first in the U.S. to mail out self-collection kits to a randomly selected probability sample, but even household surveys employing door-to-door outreach methods have only reached response rates of approximately 24% [36,37]. The need for on-going monitoring of population-level infection burden and vaccine coverage will require improved and innovative methods for recruiting participants, particularly those who may be under-represented in surveillance data due to issues with testing access or usage. Monetary compensation for participation may have contributed to sampling bias by making it more likely that persons of lower socioeconomic status would participate. An advantage of the address-based approach is that we were able to account for differential non-response by comparing responding to non-responding households by characteristics available at the census-tract level (e.g., racial distribution of census tract, Hispanic/Latinx surname, and home ownership status), and using predictors of non-response in weighting computation. However, we were not able to adjust for variables associated with individual non-response, such as variables not assessed by or substantially associated with our weighting schema, such as essential worker status or prior infection history. Differential non-response according to factors associated with seropositivity, if unadjusted for, would contribute to bias in our estimates.

There are also two potential limitations regarding our cumulative incidence estimates. First, due to waning SARS-CoV-2 antibodies, previous infections may be undetected by antibody assays [9,10]. This issue may be exacerbated when using laboratory assays targeting the nucleocapsid, versus spike, protein [19,38]. We addressed this limitation in two ways: 1)
by using a modeling approach to estimate cumulative incidence given the observed period seroprevalence and death data and 2) by retesting a sample of non-reactive specimens with an assay targeting the spike protein. Nevertheless, there may be some degree of misclassification in antibody positivity. Second, we used an estimate for duration of seropositivity from published data from New York City to parameterize the model. The New York estimate was generated using the CDC ELISA kit that detects total SARS-CoV-2 Ig targeting the spike protein, while the assay used for our study detects total Ig targeting the nucleocapsid protein. However, the New York City estimate is the only available approximation of the timeline for population-level waning antibodies at this point. Comparisons to external estimates allay laboratory and modeling-related concerns to some extent. For example, the estimated reported fraction from our study is similar to a previous CDC estimate for California suggesting 24% of infections were reported statewide during July-August 2020 [34].

As part of the COVIDVu study, we will perform two rounds of follow-up with study participants: once in the first quarter of 2021 and again in the second quarter of 2021. We will collect DBS samples to estimate incident infections and to assess antibody waning among baseline participants. We will test for IgA, IgG and IgM to nucleocapsid (BioRad Platelia Total Antibody test) and for potentially vaccine-associated IgG (i.e., antibodies to Spike, EuroIMMUN IgG assay). We will also administer follow-up surveys at both time points, focusing on vaccination and on-going infection risk. Our estimates provide an important baseline for ongoing efforts to monitor seroprevalence statewide using laboratory surveillance data and additional population-based serosurveys.

In this population-based study, we estimated California statewide SARS-CoV-2 seroprevalence in adults. Accounting for waning antibody response, we estimated a cumulative incidence of 8.7% as of November 2, 2020. The estimated IFR was 0.8% for adults, and we found that only one in three SARS-CoV-2 infections in adults were reported to
CDPH through early November 2020. Disparities documented in our study by ethnicity and insurance status may be larger than previously suggested by local seroprevalence studies [6,23] and surveillance data [28]. Taken together, these data underscore the continued need to focus public health interventions, including access to testing and vaccination in socioeconomically vulnerable communities most heavily impacted. Serosurveys are an important tool for understanding the full extent of SARS-CoV-2 infections and will be critical for on-going monitoring of population-level immunity in the vaccine era.
NOTES

Disclaimer

The findings and conclusions in the article are those of the authors and do not necessarily represent the official position of the California Department of Public Health.

Funding

This work was supported by the Epidemiology and Laboratory Capacity for Prevention and Control of Emerging Infectious Diseases cooperative agreement CK19-1904 (NU50CK000539) and the US National Institute of Allergy and Infectious Diseases (3R01AI143875-02S1).

Potential conflicts of interest

All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. Dr. Sullivan reported receiving grants from the US National Institutes of Health during the conduct of the study. Dr. Lopman reported grant support from NSF/RAPID 2032084, NIH/NIAID RO1 AI143875, NIH/NIGMS R01 GM124280 during the conduct of the study; grants and personal fees from Takeda Pharmaceuticals, and personal fees from World Health Organization outside the submitted work. Dr. Siegler reported receiving grants from the US National Institutes of Health and the Woodruff Foundation during the conduct of the study. Dr. Fahimi reported receiving payments to his institution for statistical support and consulting for this work. No other potential conflicts of interest were disclosed.
References

1. State of California. Tracking COVID-19 in California. Available at: https://covid19.ca.gov/state-dashboard/. Accessed 21 March 2021.

2. Clapham H, Hay J, Routledge I, et al. Seroepidemiologic Study Designs for Determining SARS-COV-2 Transmission and Immunity. Emerg Infect Dis 2020; 26:1978–1986.

3. Peeling RW, Olliaro PL. The time to do serosurveys for COVID-19 is now. Lancet Respir Med. 2020; 8:836–838.

4. Appa A, Takahashi S, Rodriguez-Barraquer I, et al. Universal PCR and antibody testing demonstrate little to no transmission of SARS-CoV-2 in a rural community. medRxiv 2020; Available at: http://dx.doi.org/10.1101/2020.08.15.20175786.

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 Intern Med 2020; 180:1576–1586.

6. Yamaki J, Peled H, Mathews S, et al. Seroprevalence of Novel Coronavirus SARS-CoV-2 at a Community Hospital Emergency Department and Outpatient Laboratory in Northern Orange County, California. J Racial Ethn Health Disparities 2020; Available at: http://dx.doi.org/10.1007/s40615-020-00918-0.

7. Riverside University Health System Public Health. Antibody study shows coronavirus spread wider in Riverside County. Available at: https://rivcoph.org/Portals/0/Documents/CoronaVirus/July/News/7.27.20%20antibody%20testing%20results.pdf?ver=2020-07-27-144931-703&timestamp=1595886602504.

8. Siegler AJ, Sullivan PS, Sanchez T, et al. Protocol for a national probability survey using
home specimen collection methods to assess prevalence and incidence of SARS-CoV-2 infection and antibody response. Ann Epidemiol 2020; 49:50–60.

9. Choe PG, Kang CK, Suh HJ, et al. Waning Antibody Responses in Asymptomatic and Symptomatic SARS-CoV-2 Infection. Emerg Infect Dis 2021; 27. Available at: http://dx.doi.org/10.3201/eid2701.203515.

10. Perreault J, Tremblay T, Fournier M-J, et al. Waning of SARS-CoV-2 RBD antibodies in longitudinal convalescent plasma samples within 4 months after symptom onset. Blood 2020; 136:2588–2591.

11. Stephens DS, McElrath MJ. COVID-19 and the path to immunity. JAMA 2020; 324:1279–1281.

12. Shioda K, Lau MS, Kraay AN, et al. Estimating the cumulative incidence of SARS-CoV-2 infection and the infection fatality ratio in light of waning antibodies. medRxiv 2020; Available at: http://dx.doi.org/10.1101/2020.11.13.20231266.

13. Lavange LM, Kalsbeek WD, Sorlie PD, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010; 20:642–649.

14. Chido-Amajuoyi OG, Yu RK, Agaku I, Shete S. Exposure to Court-Ordered Tobacco Industry Antismoking Advertisements Among US Adults. JAMA Netw Open 2019; 2:e196935.

15. Cerel J, Maple M, van de Venne J, Moore M, Flaherty C, Brown M. Exposure to suicide in the community: Prevalence and correlates in one U.s. state. Public Health Rep 2016; 131:100–107.

16. Sullivan PS, Sailey C, Guest JL, et al. Detection of SARS-CoV-2 RNA and antibodies in
diverse samples: protocol to validate the sufficiency of provider-observed, home-collected blood, saliva, and oropharyngeal samples. JMIR public health and surveillance 2020; 6:e19054.

17. Suitability and Sufficiency of telehealth clinician-observed participant-collected samples for SARS-CoV2 testing: the iCollect Cohort Pilot Study. Available at: https://www.jmir.org/preprint/19731. Accessed 24 May 2020.

18. Sullivan PS, Sailey C, Guest JL, et al. Detection of SARS-CoV-2 RNA and Antibodies in Diverse Samples: Protocol to Validate the Sufficiency of Provider-Observed, Home-Collected Blood, Saliva, and Oropharyngeal Samples (Preprint). Available at: http://dx.doi.org/10.2196/preprints.19054.

19. Lumley SF, Wei J, O’Donnell D, et al. The duration, dynamics and determinants of SARS-CoV-2 antibody responses in individual healthcare workers. Clin Infect Dis 2021; Available at: http://dx.doi.org/10.1093/cid/ciab004.

20. US Census Bureau. QuickFacts. 2019.

21. Iyer AS, Jones FK, Nodoushani A, et al. Dynamics and significance of the antibody response to SARS-CoV-2 infection. medRxiv 2020; Available at: http://dx.doi.org/10.1101/2020.07.18.20155374.

22. Bruckner TA, Parker DM, Bartell SM, et al. Estimated seroprevalence of SARS-CoV-2 antibodies among adults in Orange County, California. Sci Rep 2021; 11:3081.

23. Routledge I, Epstein A, Takahashi S, et al. Citywide serosurveillance of the initial SARS-CoV-2 outbreak in San Francisco. Res Sq 2021; Available at: http://dx.doi.org/10.21203/rs.3.rs-180966/v1.
24. Chamie G, Marquez C, Crawford E, et al. Community transmission of severe acute respiratory syndrome coronavirus 2 disproportionately affects the Latinx population during shelter-in-place in San Francisco. Clin Infect Dis 2020; Available at: https://academic.oup.com/cid/advance-article-abstract/doi/10.1093/cid/ciaa1234/5895337.

25. Dyer O. Covid-19: US testing ramps up as early response draws harsh criticism. BMJ 2020; 368:m1167.

26. Centers for Disease Control and Prevention. Estimated Disease Burden of COVID-19. 2021. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html. Accessed 21 March 2021.

27. Centers for Disease Control and Prevention. Health Equity Considerations and Racial and Ethnic Minority Groups. 2021. Available at: https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html. Accessed 21 March 2021.

28. California Department of Public Health. COVID-19 Race and Ethnicity Data. Available at: https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/COVID-19/Race-Ethnicity.aspx. Accessed 21 March 2021.

29. Blueprint For a Safer Economy: Equity Focus. Available at: https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/COVID-19/CaliforniaHealthEquityMetric.aspx. Accessed 21 March 2021.

30. Bendavid E, Mulaney B, Sood N, et al. COVID-19 antibody seroprevalence in Santa Clara County, California. Int J Epidemiol 2021; Available at: http://dx.doi.org/10.1093/ije/dyab010.

31. Sood N, Simon P, Ebner P, et al. Seroprevalence of SARS-CoV-2–Specific Antibodies
Among Adults in Los Angeles County, California, on April 10-11, 2020. JAMA 2020; 323:2425–2427.

32. Ng DL, Goldgof GM, Shy BR, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood from the San Francisco Bay Area. medRxiv 2020; Available at: http://dx.doi.org/10.1101/2020.05.19.20107482.

33. Lewnard JA, Mora AM, Nkwocha O, et al. Prevalence and clinical profile of severe acute respiratory syndrome coronavirus 2 infection among farmworkers, California, June–November 2020. Emerg Infect Dis. 2021 May; Available at: https://doi.org/10.3201/eid2705.204949

34. Centers for Disease Control and Prevention. COVID Data Tracker: Nationwide Commercial Laboratory Seroprevalence Survey. Available at: https://covid.cdc.gov/covid-data-tracker/#national-lab. Accessed 11 February 2021.

35. Porter GS. California State Department of Public Health. Cal West Med 1933; 38:335–336.

36. 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; 69:965–970.

37. 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; 69:960–964.

38. Gamble A, Garnier R, Chambert T, Gimenez O, Boulinier T. Next-generation serology: integrating cross-sectional and capture-recapture approaches to infer disease dynamics. Ecology 2020; 101:e02923.
Table 1. Demographic characteristics of serosurvey participants (N=983) and weighted sample size compared to the California population aged ≥ 18 years

| Characteristic      | Sample | Weighted Sample | California Population (≥18 years)¹ |
|---------------------|--------|-----------------|-----------------------------------|
|                     | N      | Column %        | N                                | Column % |
| Overall             | 98     | 3 100.0         | 29,446,494                      | 100.0    |
| Sex                 |        |                 | 30,617,58                       | 100.0    |
| Male                | 41     | 6 42.3          | 14,160,171                      | 48.1     |
| Female              | 56     | 7 57.7          | 15,286,322                      | 51.9     |
| Race/ethnicity      |        |                 |                                  |          |
| Hispanic/Latinx     | 26     | 1 26.6          | 10,356,872                      | 35.2     |
| Non-Hispanic/Latinx| 54     | 9 55.8          | 12,079,332                      | 41.0     |
| White               | 9      |                 | 10,947,32                       |          |
| Non-Hispanic/Latinx| 35     | 3 3.6           | 1,424,280                       | 4.8      |
| Black               | 10     | 7 10.9          | 4,537,655                       | 15.4     |
| Asian               | 7      |                 | 1,903,134                       |          |
| Non-Hispanic/Latinx| 31     | 3 3.2           | 1,048,355                       | 3.6      |
| Other               |        |                 | 161,754                         | 0.5      |
| Age                 |        |                 |                                  |          |
| 18-34 years         | 23     | 6 24.0          | 8,930,380                       | 30.3     |
| 35-44 years         | 15     |                 | 5,322,270                       | 18.1     |
| 45-54 years         | 16     | 7 17.0          | 4,737,818                       | 16.1     |
| 55-64 years         | 18     | 8 19.1          | 4,776,294                       | 16.2     |
| 65+ years           | 23     | 6 24.0          | 5,679,731                       | 19.3     |

¹2019 Bridged-Race Estimates (NVSS)
Table 2. Unweighted and weighted SARS-CoV-2 antibody prevalence among serosurvey participants (N=983) and prevalence ratios by demographic and epidemiologic characteristics, California, August–December, 2020

| Characteristic                  | Unweighted                              | Weighted                              |       |       |       |       |
|--------------------------------|-----------------------------------------|---------------------------------------|-------|-------|-------|-------|
|                                | n  | N   | %   | n    | N   | %    | 95% CI | PR  | 95% CI | P value |
| Overall                        | 33 | 983 | 3.4 | 1,338,730 | 29,446,494 | 4.5 | 2.8–7.4 | n/a |
| Sex                            |    |     |     |       |       |     |       |       |     |
| Male                           | 13 | 416 | 3.1 | 562,771  | 14,160,171 | 4.0 | 1.8–8.7 |       |       |
| Female                         | 20 | 567 | 3.5 | 775,959  | 15,286,322 | 5.1 | 2.7–9.3 | 1.3  | 0.5–3.6 |
| Race/ethnicity                 |    |     |     |       |       |     |       |       | <0.001|
| Hispanic/Latinx                | 21 | 261 | 8.0 | 1,069,853 | 10,356,872 | 10.3 | 5.8–17.8 | 7.5  | 2.8–20.2 |
| Non-Hispanic/Latinx White      | 10 | 549 | 1.8 | 166,116  | 12,079,332 | 1.4 | 0.6–3   |       |       |
| Non-Hispanic/Latinx Black      | 1  | 35  | 2.9 | 46,436   | 1,424,280  | 3.3 | 0.6–15.7 | 2.4  | 0.3–19.7 |
| Non-Hispanic/Latinx Asian      | 0  | 107 | 0.0 | .        | 4,537,655  | .   | .       | n/a  |       |
| Non-Hispanic/Latinx Other      | 1  | 31  | 3.2 | 56,325   | 1,048,355  | 5.4 | 1–23.7  | 3.9  | 0.5–31.2 |
| Age                            |    |     |     |       |       |     |       |       | 0.07  |
| 18-34 years                    | 14 | 236 | 5.9 | 461,821  | 8,930,380  | 5.2 | 2.8–9.5 | 2.7  | 0.9–8  |
| 35-44 years                    | 9  | 156 | 5.8 | 329,858  | 5,322,270  | 6.2 | 3–12.3  | 3.3  | 1.1–10.2 |
| 45-54 years                    | 3  | 167 | 1.8 | 410,604  | 4,737,818  | 8.7 | 2.5–25.6 | 4.6  | 1–21.5 |
| 55-64 years                    | 6  | 188 | 3.2 | 90,012   | 4,776,294  | 1.9 | 0.7–5   |       |       |

Downloaded from https://academic.oup.com/ofid/advance-article/doi/10.1093/ofid/ofab379/6329153 by Wallace Library Serials Dept. user on 28 July 2021
- 65+ years: 1 236 0.4 46,436 5,679,731 0.8 0.1–4.5 0.4 0.1–3.7

- Education
  - High School/GED or less: 9 121 7.4 797,929 10,161,577 7.9 3.7–16 2.3 0.8–6.6
  - Some college/associate's degree: 10 309 3.2 296,372 9,422,451 3.1 1.6–6.2 0.9 0.3–2.5
  - Bachelor's degree: 10 317 3.2 216,239 6,313,350 3.4 1.7–6.8 reference
  - Graduate degree: 4 236 1.7 28,190 3,549,116 0.8 0.2–3 0.2 0.1–0.8

- Annual Income
  - $0 to $49,999: 16 310 5.2 487,740 8,021,148 6.1 3.4–10.6 2.1 0.7–6.6
  - $50,000 to $99,999: 7 286 2.4 231,462 8,118,641 2.9 1.1–7 reference
  - $100,000+: 10 387 2.6 619,529 13,306,704 4.7 1.9–11 1.6 0.4–6.2

- Health Insurance
  - No health insurance: 3 30 10.0 98,187 791,167 12.4 4.1–31.9 4.5 1.2–16.9
  - Medicare/Medicaid/Other government plan: 7 281 2.5 223,129 7,902,141 2.8 1.2–6.3 1.0 0.4–2.9
  - Private insurance/parent's plan: 18 595 3.0 497,215 17,999,789 2.8 1.5–5 reference
  - Don't know: 5 77 6.5 520,200 2,753,397 18.9 7.3–40.9 6.8 2.2–21

- Household size
  - 1-2 persons: 12 585 2.1 451,298 14,808,778 3.0 1.2–7.5 reference
  - 3-5 persons: 17 359 4.7 714,877 13,063,378 5.5 2.8–10.4 1.8 0.6–5.8
  - >5 persons: 4 39 10.3 172,555 1,574,338 11.0 4–26.9 3.6 0.9–15.2

- Leave home for work^2
  - 14 357 3.9 598,873 11,766,211 5.1 2.4–10.6 3.9 1.1–14 0.04
|                                                                 | PR | CI           |
|------------------------------------------------------------------|----|-------------|
| COVID-19 symptoms since January 1st 2020<sup>2,3</sup>           | 25 | 551 4.5     |
| Contact to a confirmed case<sup>2</sup>                          | 20 | 139 14.4    |
| Prior COVID-19 diagnosis<sup>2,4</sup>                           | 13 | 20 65.0     |
| Month of sample collection                                       |    |             |
| August/September                                                 | 3  | 158 1.9     |
| October                                                          | 3  | 259 1.2     |
| November/December                                                | 27 | 566 4.8     |

PR: Prevalence ratio. CI: Confidence interval

1Confidence intervals are calculated using the modified Wilson method.

2Reference group is persons without characteristic

3Symptoms include: cold/flu, cough, shortness of breath or loss of taste or smell.

4Provider told them they likely had COVID-19 or a positive COVID-19 test result.
Figure 1. Flow diagram of probability sample of California households to estimate seroprevalence and cumulative incidence of SARS-CoV-2 infections among adults, August–December, 2020

Figure 1 Legend.

*Consent was required at the household level for household enumeration, and then at the individual level for the randomly selected member of an enumerated household.

†Test results considered invalid for the following reasons: Sample not sufficient to process, processing incomplete by study closeout, sample collection date outside of range 8/9/20-12/8/20.
Figure 1

COVIDVu Materials Mailed to CA Households
n = 8,726

Eligible Households
n = 8,369 (95.9%)

- Ineligible (Mail Undeliverable), n = 357 (4.1%)
- No Response from Household, n = 6,841 (81.7%)
  - Did Not Complete Household Consent*, n = 76 (0.9%)
  - Did Not Complete Individual Consent*, n = 219 (2.6%)
  - Did Not Complete Baseline Survey, n = 45 (0.5%)

Completed Enumeration, Consent*, and Baseline Survey
n = 1,188 (14.2%)

Primary Analytic Sample†
 n = 983 (11.7%)

- Valid Ig and AN Result, n = 967 (81.4%)
- Valid Ig Result, Invalid† AN Result, n = 19 (1.8%)
- Valid AN Result, Invalid† Ig Result, n = 91 (7.8%)

Did Not Return Specimens to Study Laboratory, n = 89 (7.5%)
- Invalid† Ig and AN Result, n = 27 (2.3%)