

# UC Irvine

## UC Irvine Previously Published Works

### Title

Estimated seroprevalence of SARS-CoV-2 antibodies among adults in Orange County, California.

### Permalink

<https://escholarship.org/uc/item/57t831b5>

### Journal

Scientific reports, 11(1)

### ISSN

2045-2322

### Authors

Bruckner, Tim A  
Parker, Daniel M  
Bartell, Scott M  
et al.

### Publication Date

2021-02-01

### DOI

10.1038/s41598-021-82662-x

Peer reviewed



OPEN

## Estimated seroprevalence of SARS-CoV-2 antibodies among adults in Orange County, California

Tim A. Bruckner<sup>1</sup>✉, Daniel M. Parker<sup>1</sup>, Scott M. Bartell<sup>1,2</sup>, Veronica M. Vieira<sup>1</sup>, Saahir Khan<sup>3</sup>, Andrew Noymer<sup>1</sup>, Emily Drum<sup>1</sup>, Bruce Albala<sup>4</sup>, Matthew Zahn<sup>5</sup> & Bernadette Boden-Albala<sup>1</sup>✉

Clinic-based estimates of SARS-CoV-2 may considerably underestimate the total number of infections. Access to testing in the US has been heterogeneous and symptoms vary widely in infected persons. Public health surveillance efforts and metrics are therefore hampered by underreporting. We set out to provide a minimally biased estimate of SARS-CoV-2 seroprevalence among adults for a large and diverse county (Orange County, CA, population 3.2 million). We implemented a surveillance study that minimizes response bias by recruiting adults to answer a survey without knowledge of later being offered SARS-CoV-2 test. Several methodologies were used to retrieve a population-representative sample. Participants (n = 2979) visited one of 11 drive-thru test sites from July 10th to August 16th, 2020 (or received an in-home visit) to provide a finger pin-prick sample. We applied a robust SARS-CoV-2 Antigen Microarray technology, which has superior measurement validity relative to FDA-approved tests. Participants include a broad age, gender, racial/ethnic, and income representation. Adjusted seroprevalence of SARS-CoV-2 infection was 11.5% (95% CI: 10.5–12.4%). Formal bias analyses produced similar results. Prevalence was elevated among Hispanics (vs. other non-Hispanic: prevalence ratio [PR] = 1.47, 95% CI 1.22–1.78) and household income < \$50,000 (vs. > \$100,000: PR = 1.42, 95% CI: 1.14 to 1.79). Results from a diverse population using a highly specific and sensitive microarray indicate a SARS-CoV-2 seroprevalence of ~ 12 percent. This population-based seroprevalence is seven-fold greater than that using official County statistics. In this region, SARS-CoV-2 also disproportionately affects Hispanic and low-income adults.

Active and routine surveillance of infectious diseases serves a critical public health function by permitting estimation of overall prevalence and incidence of new infections<sup>1,2</sup>. California reported the first case of local spread of SARS-CoV-2 causing COVID-19 disease in the US on February 26, 2020 and over the following six months has recorded > 586,000 cases and > 10,600 deaths from this disease<sup>3</sup>. Owing in part to limited testing capacity, no state in the US (including California) has enacted routine population-based surveillance of SARS-CoV-2. Instead, as a proxy, states typically collect SARS-CoV-2 polymerase chain reaction (PCR) testing for individual clinical diagnostic purposes. SARS-CoV-2 infection, however, may cause minimal or no symptoms in some individuals<sup>4,5</sup>. Since persons without (or with minor) symptoms may not seek care, and because testing capacity has frequently been limited, clinic-based estimates may considerably undercount the true incidence of SARS-CoV-2 infections in the population<sup>6</sup>.

Multiple studies have used serologic testing to measure prevalence of COVID-19 in different populations<sup>4–8</sup>, but the accuracy of these studies has been questioned due to insufficient specificity of the serologic assay for a low prevalence population and due to potential sampling bias<sup>9,10</sup>. Due to variability in performance of the serologic assay and sampling methodology, these studies have estimated anywhere from a tenfold to 85-fold higher

<sup>1</sup>Program in Public Health, University of California, Irvine, 653 E. Peltason Dr, Irvine, CA 92697, USA. <sup>2</sup>Department of Statistics, University of California, Irvine, Bren Hall 2019, Irvine, CA 92697-1250, USA. <sup>3</sup>School of Medicine, University of California, Irvine, 1001 Health Sciences Rd, Irvine, CA 92697, USA. <sup>4</sup>Center for Clinical Research, School of Medicine, University of California, Irvine, 1001 Health Sciences Rd, Irvine, CA 92617, USA. <sup>5</sup>Orange County Health Care Agency, 405 W. 5th St., Santa Ana, CA 92701, USA. ✉email: tim.bruckner@uci.edu; BBODENAL@hs.uci.edu

prevalence as compared to case counts. An ideal seroprevalence study design would recruit a representative sample from the population and detect SARS-CoV-2 (via antibodies from a blood test). To minimize selection bias and skewing of seroprevalence to people with suspected SARS-CoV-2 infection, serum samples would be obtained outside of a clinic *without* oversampling those who are more likely to have been infected (e.g., due to self-selection into a study based on knowledge of symptoms and being offered an antibody test). As of our study date, we know of only one population-based prevalence estimate in the US that meets these criteria<sup>7</sup>.

Orange County (OC) California includes a large, ethnically diverse (34.0% Hispanic, 21.7% Asian) metropolitan region and is the sixth most populous county in the US<sup>11</sup>. Despite mandating individual and community-wide physical distancing measures, school and business closure, and confinement measures, OC has reported 46,057 cumulative SARS-CoV-2 cases and 957 deaths as of 8/16/20<sup>12</sup>. We set out to provide a minimally biased estimate of SARS-CoV-2 seroprevalence among all adults in OC.

Our analysis improves upon previous US estimates in several ways<sup>7,9,13,14</sup>. First, we recruited subjects outside of a clinical setting and employed strategies to minimize bias of recruiting mostly symptomatic cases. Second, we applied a robust SARS-CoV-2 Antigen Microarray technology which has superior sensitivity and specificity relative to what were currently available FDA-approved tests used by others<sup>15</sup>. Third, we recruited a sufficiently large sample of adults to calculate seroprevalence by race/ethnicity, age, and gender, which may uncover important differences across these groups. Fourth, we recruited subjects by administering a questionnaire *without* initially disclosing an offer for an antibody test. In addition to serving a critical surveillance function, we intend for our results to yield a more accurate measure of the infection fatality risk.

## Methods

**Recruitment.** This study represents a joint effort between University of California, Irvine (UCI) and the Orange County Health Care Agency (OCHCA). We received human subjects approval from the UCI Institutional Review Board (HS# 2020-5952) and obtained informed consent from all study participants. All methods were performed in accordance with the relevant guidelines and regulations.

We focused our study on adults 18 years or older residing in OC on July 1st 2020. We know of no full roster of the adult population in OC that would permit sampling based on complete enumeration. Absent this information, we used a proprietary database, reflecting the age, income, and racial/ethnic diversity of OC and maintained by SoapBoxSample, an LRW Group Company, to recruit participants. The proprietary database is targeted towards certain demographic groups and contains contact information for 800,000 adults—almost one-third of all adults in OC. The database, moreover, contains sociodemographic information which allows us to assess potential non-response bias.

The database contains contact information of individuals rather than households. Using this database, we invited (via email [36.4%] or random-digit telephone dialing [63.6%]) one resident per household to participate in a study about their opinions of COVID-19, without initial mention of SARS-CoV-2 antibody testing. We did not allow the opportunity to defer the survey to another household member. Participants for this analysis were not enrolled if other members of the household already were enrolled. Participants completed a survey regarding socio-demographics (e.g., age, gender, race/ethnicity, household income), daily work and social activities related to SARS-CoV-2, any known previous infection with SARS-CoV-2, and history of SARS-CoV-2 symptoms in the last few months (see details in Supplemental Material, Question about Symptoms). Once the respondent provided these answers, we then asked if they were willing to participate in a drive-thru finger-prick blood test for SARS-CoV-2 antibodies. Participants received a \$10 gift card as compensation for completion of the survey and blood test.

We estimated the total sample size needed for a two-sample binomial test, powered at 80% to detect differences in seroprevalence between two equally sized groups (e.g., male vs. female). Based on this sample size calculation, we then arrived at targeted quotas for each sociodemographic stratum, according to the population distribution of each stratum in Orange County. Based on US Census-derived population distributions in OC of age, gender, race/ethnicity, and household income<sup>11</sup>, we targeted recruitment to ensure adequate sample size of subjects from the following strata: age-by-gender (18–34 years, 35–54 years, 55 years or above; by male, female), race/ethnicity (Hispanic, Asian non-Hispanic, and other non-Hispanic including white and Black, and household income (< \$50,000, \$50,000–99,999, and \$100,000 or above).

We encouraged participation of racial/ethnic minorities and/or lower-income adults in several ways. First, for four of the six weeks of the survey, we offered the survey in the five most commonly spoken languages in OC (i.e., English, Spanish, Mandarin, Vietnamese, and Korean). Second, we invited all household members of underrepresented groups to receive testing at the drive-thru site. Third, for the last two weeks of the survey we targeted particular zip codes with a large proportion of racial/ethnic minorities. To increase participation among persons considered vulnerable (e.g., aged > 65 years), we offered in-home visits at which licensed phlebotomists obtained the finger-prick blood sample at the participant's home.

For participants who agreed to the antibody test, we invited them to one of eleven drive-thru test sites that spanned the geography of OC. These sites were distributed across OC to permit ease of access (by public transit and/or less than 15 min driving time) and previously identified by OC as Points of Dispensing sites as allowing ease of drive-thru testing given existing security and experience with large numbers of people. We assigned each participant a unique alphanumeric code, to be presented at the drive-thru test site, to ensure that only survey participants (and not others interested in receiving the antibody test) were processed. To increase the participation rate, we sent appointment reminders via email and telephone at regular intervals leading up to the scheduled date and time. We conducted SARS-CoV-2 antibody testing from July 10th to August 16th, 2020 on Thursdays through Sundays in an attempt to accommodate a range of participant work schedules and availability. We used UCI student and alumni volunteers to staff the sites overseen by trained faculty.

**Laboratory methods.** We assessed past SARS-CoV-2 infection using a coronavirus antigen microarray (CoVAM). This microarray approach has been widely used for multiple pathogens<sup>16</sup>. Unlike then available FDA authorized tests which detect antibodies against one or two antigens<sup>17</sup>, the CoVAM quantitatively measures IgG and IgM antibodies against 12 antigens from SARS-CoV-2, including spike (S) as whole protein or separated domains including S1, S2, and receptor-binding domain (RBD), nucleocapsid protein (NP), and membrane (M) protein (Supplemental Material Fig. S1). The CoVAM also measures IgG and IgM antibodies against 53 antigens from other viruses that cause acute respiratory illness, including SARS-1, MERS, and the four seasonal coronaviruses, which can be used for assessment of antibody cross-reactivity. The antigens are printed onto microarrays in quadruplicate, probed with blood specimens and secondary antibodies for IgM and IgG, and imaged to determine background-subtracted median fluorescence intensity averaged across quadruplicate spots as previously described<sup>15,16,18,19</sup>. The protein microarray requires less than 100  $\mu$ l of blood, which we retrieved from participants using a finger-prick for blood collection into capillary tubes. We processed blood specimens for plasma separation and microarray analysis in a high throughput manner within four days of collection.

**Statistical analysis.** SARS-CoV-2 antibody reactivity was determined as previously described<sup>15</sup>. Briefly, the CoVAM data for each specimen was compared to 99 positive PCR-confirmed cases with blood collected at least 7 days (range 7–50 days, median 11 days) post symptom onset and 88 negative pre-pandemic controls with blood collected prior to November 1, 2019. From the 12 SARS-CoV-2 antigens on the array, a linear regression model was trained on positive and negative controls to determine optimal weighted combinations of reactive antigens to calculate composite SARS-CoV-2 IgG and IgM antibody levels that discriminate these two groups. Based on Receiver Operating Characteristic curve analyses, reactivity thresholds for IgG and IgM were chosen to classify unknown specimens demonstrating 100% specificity and 94% sensitivity for either IgM or IgG<sup>15,18</sup>. These test characteristics compare favorably to current FDA-authorized test kits<sup>17</sup> and indicate a very low likelihood of false positive detection of SARS-CoV infection.

We first calculated crude seroprevalence for SARS-CoV-2 antibodies (excluding household members of participants who were tested after study recruitment) by dividing the number of reactive cases by the total number of participants. Next, we estimated population-representative SARS-CoV-2 seroprevalence by post-stratifying results by the age, gender, race/ethnicity, and income distribution for each of the 88 zip codes in OC. This process, known as raking<sup>20</sup>, uses age-gender-race/ethnicity population estimates and income population estimates from the US Census<sup>11</sup> to determine appropriate weights for estimating the overall seroprevalence for OC as a whole. Participants who did not provide complete demographic information were excluded from the weighted analyses, as were empty strata. In secondary analyses, we used multiple imputation to assess non-response bias and Bayesian methods to adjust for antibody test accuracy (Table S2, Supplemental Material).

We estimated prevalence ratios (i.e., relative risk across subgroups of a seropositive SARS-CoV-2 test) and 95% uncertainty intervals using both unweighted and population-weighted<sup>21</sup> log-binomial multiple regression, comparisons across age, gender, income, and race/ethnicity subgroups, mutually adjusting for these covariates as well as sampling week, to limit confounding by any changes in recruitment and seroprevalence over time. We determined reference groups based on the stratum with the largest sample size. Finally, we estimated the infection fatality risk among adults in OC by dividing the number of cumulative excess SARS-CoV-2 deaths among adults by the number of implied past infections (i.e., seroprevalence  $\times$  adult population) at the time of the survey that could have led to a recorded death by the time of our analysis. Excess deaths were calculated as the sum of actual deaths minus expected deaths, in which both are for all-causes, for 1 March through 16 August 2020. Source data for the deaths were daily death counts provided OCHCA<sup>12</sup>. We compared this infection fatality risk to that using published statistics from OCHCA<sup>12</sup>.

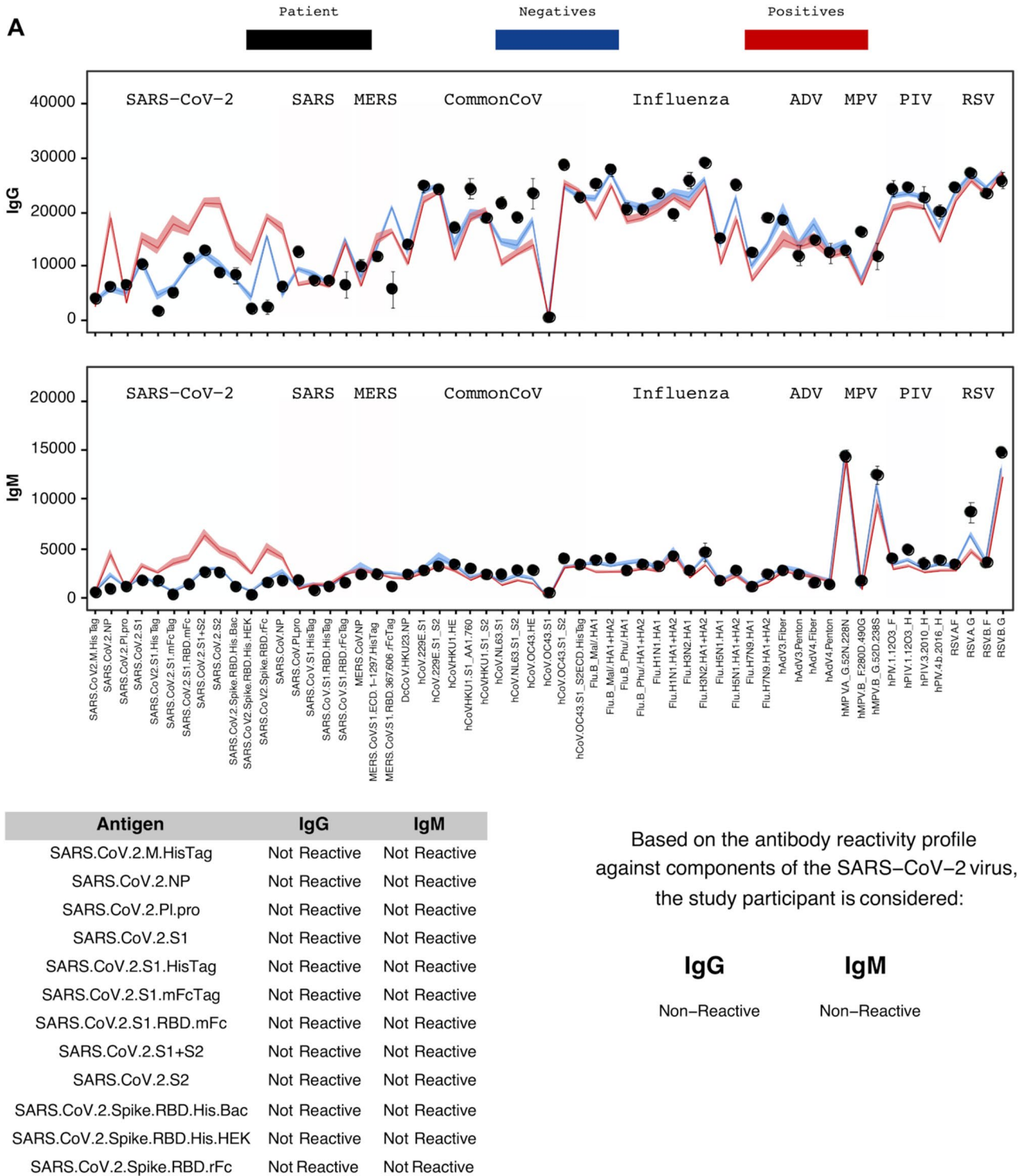
## Results

Over the study period, 4555 adults completed the survey and consented to participate in the blood test. Of these participants, 2,979 (65.4%) showed up to provide a viable blood sample for CoVAM analysis from July 10th to August 16th, 2020. Figure 1A, and B illustrate negative and positive SARS-CoV-2 test results using CoVAM. Comparison of the black circles (i.e., patient IgG and IgM values) across these figures show clear differences in the SARS-CoV-2 antibody reactive profile between a negative and positive test. A full heat map of all specimens, with classification and positive and negative controls, appears in Supplemental Material (Fig. S1).

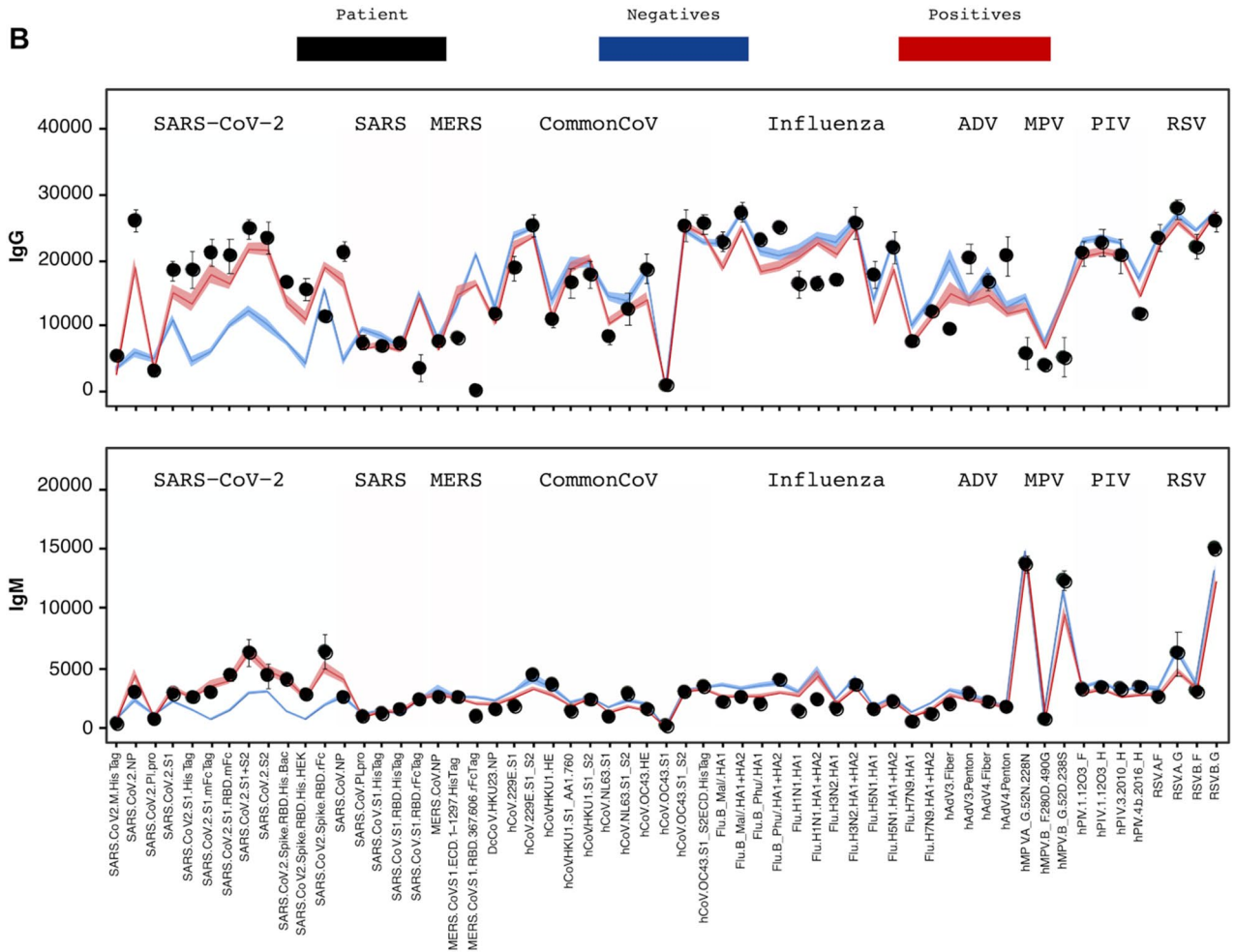
Our recruitment strategy successfully garnered participation from a broad range of socioeconomic, age, and racial/ethnic backgrounds (Table 1). The overall unadjusted seroprevalence of SARS-CoV-2 is 11.8% (351/2979). Table 1 also shows elevated crude SARS-CoV-2 seroprevalence among Hispanic and low-income (<\$50,000) adults.

The fact that our sample skews towards higher-income earners and older adults (Supplemental Material Table S3) indicates that post-stratification weighting may provide a more representative estimate of SARS-CoV-2 seroprevalence for OC. This weighting produced an adjusted seroprevalence of 11.5% (95% Confidence Interval: 10.5–12.4%). We also performed sensitivity analyses to ascertain potential bias induced by differential non-response of adults with respect to SARS-CoV-2 positivity status. We assessed non-response bias in several ways (Supplemental Material, Bias Analysis). Depending on the level of differential non-response by persons suspected to be SARS-CoV-2 positive and the adjustment strategy used, adjusted point estimates for seroprevalence range from 11.3 to 14.4%.

Consistent with the unadjusted results, adjustment for post-stratification weighting indicates that prevalence is elevated among Hispanics (vs. other non-Hispanic: PR = 1.47, 95% CI 1.22–1.78), household income <\$50,000 (vs. >\$100,000: PR = 1.42, 95% CI 1.14–1.79), and symptomatic adults (vs. no symptoms: PR = 1.68, 95% CI 1.41–2.00; Table 2). Males also show a lower prevalence (PR = 0.85 95% CI 0.71–1.01) relative to females.



**Figure 1.** Coronavirus antigen microarray result report for representative negative (A) and positive (B) study participants. For each result, the IgG and IgM antibody reactivity quantified by mean fluorescence intensity is plotted with comparison to reference positives and negatives and classified as “Reactive” or “Non-Reactive” for each individual antigen and overall for IgG and IgM.



Antigen	IgG	IgM
SARS.CoV.2.M.HisTag	Not Reactive	Not Reactive
SARS.CoV.2.NP	<b>Reactive</b>	Not Reactive
SARS.CoV.2.Pl.pro	Not Reactive	Not Reactive
SARS.CoV.2.S1	<b>Reactive</b>	Not Reactive
SARS.CoV.2.S1.HisTag	<b>Reactive</b>	Not Reactive
SARS.CoV.2.S1.mFcTag	<b>Reactive</b>	Not Reactive
SARS.CoV.2.S1.RBD.mFc	<b>Reactive</b>	<b>Reactive</b>
SARS.CoV.2.S1+S2	<b>Reactive</b>	<b>Reactive</b>
SARS.CoV.2.S2	<b>Reactive</b>	<b>Reactive</b>
SARS.CoV.2.Spike.RBD.His.Bac	<b>Reactive</b>	Not Reactive
SARS.CoV.2.Spike.RBD.His.HEK	<b>Reactive</b>	Not Reactive
SARS.CoV.2.Spike.RBD.rFc	Not Reactive	<b>Reactive</b>

Based on the antibody reactivity profile against components of the SARS-CoV-2 virus, the study participant is considered:

**IgG**                      **IgM**  
**Reactive**                      **Reactive**

Figure 1. (continued)

We calculated the incidence fatality risk, separately for adults 18–64 years and those 65 years and above, given the strong age-related morbidity and mortality of SARS-CoV-2 (Supplemental Material, Incidence Fatality Risk)<sup>22</sup>. Using the seroprevalence results gives a range of infection fatality rates that are 0.94–1.9 deaths per 1000 persons for adults 18–64 years, and 10.5–11.6 deaths per 1,000 persons for adults 65 years and above. In comparing these values to those calculated using only reported statistics (through to 8/16/20, the end of our study period) from OCHCA, for 18–64 year olds the population-based SARS-CoV-2 infection fatality rate is an estimated 3.2–6.3 fold lower. Also, for 65 years and above, the population-based SARS-CoV-2 infection fatality rate is an estimated 9.0 to 10.0 fold lower than that reported by OCHCA.

Characteristic	Participants		Seroprevalence	
	N	(%)	Cases/total	(%)
Overall	2979	(100)	351/2979	(11.8)
<b>Age</b>				
18–34 years	673	(22.6)	83/673	(12.3)
35–54 years	1362	(45.7)	168/1362	(12.3)
≥ 55 years	944	(31.7)	100/944	(10.6)
<b>Gender</b>				
Female	1680	(56.4)	223/1680	(13.3)
Male	1296	(44.6)	128/1296	(9.9)
<b>Race/ethnicity</b>				
Hispanic (including multi-race)	1108	(37.2)	172/1108	(15.5)
Asian (non-Hispanic, non-African American)	469	(15.7)	50/469	(10.7)
Other non-Hispanic (including white and African American)	1402	(47.1)	129/1402	(9.2)
<b>Household income per year</b>				
< \$50,000	655	(22.0)	98/655	(15.0)
\$50,000–\$99,999	722	(24.2)	96/722	(13.3)
≥ \$100,000	1090	(36.6)	87/1090	(8.0)
Prefer not to answer	512	(17.2)	70/512	(13.7)
<b>Had contact with anyone suspected or confirmed of COVID-19?</b>				
Yes	579	(19.4)	120/579	(20.7)
No	1850	(62.1)	191/1850	(10.3)
Don't know	546	(18.3)	40/546	(7.3)
<b>Had any SARS-CoV-2 symptoms in last 2 months?</b>				
Yes	1038	(34.8)	166/1038	(16.0)
No	1941	(65.2)	185/1941	(9.5)

**Table 1.** Unadjusted SARS-CoV-2 seroprevalence by social, demographic and survey characteristics among 2979 adults in Orange County, CA, July 10th to August 16th, 2020.

Characteristic	Prevalence ratio	(95% CI)
<b>Age</b>		
18–34 years	1.11	(0.90, 1.36)
35–54 years	1.05	(0.85, 1.29)
≥ 55 years (ref.)	1	–
<b>Gender</b>		
Female (ref.)	1	–
Male	0.85	(0.71, 1.01)
<b>Race/ethnicity</b>		
Hispanic (including multi-race)	1.47	(1.22, 1.78)
Asian (non-Hispanic, Non-African American)	1.35	(1.10, 1.67)
Other Non-Hispanic (including white and African American)	1	–
<b>Household income per year</b>		
< \$50,000	1.42	(1.14, 1.79)
\$50,000 to \$99,999	1.44	(1.15, 1.80)
≥ \$100,000 (ref.)	1	–
<b>Had contact with anyone suspected or confirmed of COVID-19?†</b>		
Yes	2.20	(1.82, 2.64)
No or Don't Know (ref.)	1	–
<b>Had any SARS-CoV-2 symptoms in last 2 months?†</b>		
Yes	1.68	(1.41, 2.00)
No (ref.)	1	–

**Table 2.** Adjusted\* SARS-CoV-2 seroprevalence ratios among adults in Orange County, CA, July 10th to August 16th, 2020. \*Estimates use population weights and adjust for age, gender, income, race/ethnicity, and sampling week. †Effects of suspected contact or symptoms are estimated in separate models.

## Discussion

We set out to provide a minimally biased estimate of SARS-CoV-2 seroprevalence among adults in Orange County, a densely-populated and diverse county in southern California. We recruited subjects outside of a clinical setting and without their knowledge that we would ultimately offer an antibody test. Results from a socioeconomically and demographically diverse population of >2900 adults using a highly specific and sensitive CoVAM microarray indicate a SARS-CoV-2 seroprevalence of approximately 12 percent. This SARS-CoV-2 point seroprevalence among adults for July to August 2020 is seven-fold greater than the cumulative incidence of diagnosed cases reported to Orange County.

Our seroprevalence estimate is greater than that reported from antibody surveys among blood donors nationwide<sup>23</sup>, but less than that reported from New York City<sup>14,24</sup>. Early community serosurveys in California ranged from 4.65% among adults in neighboring LA<sup>7</sup> to 2.8% among adults and children in Santa Clara County<sup>25</sup>, both of which suggested infection was substantially more widespread in those regions than case counts implied. Other states have similarly estimated prevalence between 2 and 4%<sup>9,26</sup>. The study by Biggs and colleagues<sup>26</sup> in Georgia employed a strong multi-stage sampling design, but we caution against direct comparisons between this study and ours given their unclear control/assessment of bias due to selective survey participation by anticipated receipt of the SARS-CoV-2 antibody test. We also hesitate to compare our study directly to others given the different course, sample (clinic vs. community) and timing of the epidemic across regions, and different sensitivity and specificity of test used.

The seroprevalence of SARS-CoV-2 in OC is 1.8 fold greater among Hispanics than among the referent group of mostly non-Hispanic whites. Previous clinic-based research in Baltimore/Washington, DC<sup>27</sup> and Orange County<sup>28</sup> also supports this finding. Greater prevalence of SARS-CoV-2 among Hispanics may arise from work in settings that do not allow for physical distancing, work under hazardous conditions due to economic necessity, and residence in relatively dense housing conditions<sup>28–31</sup>.

Strengths of our study include the recruitment of a large sample of adults without their knowledge that we would ultimately offer an antibody test, which minimizes response bias. We also recruited a diverse population in multiple languages which permits precise estimation of SARS-CoV-2 seroprevalence by age, gender, race/ethnicity, and income subgroups. The detailed survey, moreover, allowed us to assess the sensitivity of main results to biased participation in blood testing among symptomatic individuals. The CoVAM microarray technology is a study strength in that it quantifies antibody responses to a wide range of SARS-CoV-2 proteins. The specificity and sensitivity of CoVAM compares favorably to other antibody test kits used to estimate seroprevalence of past infection in the US. In addition, the lack of reliance of CoVAM on any one specific SARS-CoV-2 protein (e.g., the “N” protein) minimizes the risk of a false negative test resulting from inter-individual variability in antibody profiles across antigens. CoVAM also requires only a few microliters of blood, which is practical for increasing compliance during large surveillance programs.

We did not randomly sample the entire adult population in OC owing to the logistical constraint of having neither a full roster of residents nor complete demographic information on individuals in the commercially obtained database. In addition, over a third of respondents who consented did not provide a blood test. Although this circumstance *per se* does not introduce bias<sup>32–34</sup>, we cannot rule out the possibility of differential nonresponse with respect to past SARS-CoV-2 infection. We also cannot exclude the possibility that willingness to participate was conditionally dependent on COVID-19 antibody status after accounting for demographic characteristics. Our analyses, however, permit evaluation of the potential role of bias, suggests minimal bias in estimates, and makes explicit our assumptions, thereby permitting replication (Supplemental Material, Bias Analysis).

Whereas we retrieved blood samples over a span of six weeks, relatively few samples for any particular week precluded calculations over time of SARS-CoV-2 seroprevalence precisely when OC reported an increase in reported cases. In addition, CoVAM detects past infection by quantifying antibodies that respond to a broad set of SARS-CoV-2 proteins, but it may fail to detect a small subset of new infections that have not yet elicited a sufficient antibody response. We also did not examine children < 18 years of age. Finally, findings pertain to OC only—an area thought to be affected earlier by SARS-CoV-2 than many other US regions, but not as affected as New York City—and should not be used to generalize to other places or times.

The implications of our study are three-fold. First, the widespread seroprevalence of SARS-CoV-2 in OC warrants continued public health measures of physical distancing, proper and consistent use of face masks, ventilation, and hand hygiene. As has been reported elsewhere, SARS-CoV-2 infections are underreported; our seroprevalence survey suggests that infections are underreported by a factor of 7.6. Second, in addition to contact tracing, authorities may want to consider active surveillance of novel infections. This active surveillance would involve both a population-based component (e.g., 800–1000 tests per week in a representative sample of county or city residents) as well as a targeted component on higher-risk groups or places (e.g., specific census tracts, nursing home residents, or laborers in high-density settings). Such surveillance, unlike clinic- or hospital-based strategies, would provide a less-biased estimate of the rate of new SARS-CoV-2 infections. Third, our estimates of infection fatality rates for adults 18–64 years and those 65 years and above is several-fold lower than that using county reports but lies within the range of other fatality risk studies from other US regions<sup>35</sup>. This updated estimate should inform the broader policy debate in the US regarding the relative benefits and limitations of various SARS-CoV-2 mitigation strategies.

Received: 27 October 2020; Accepted: 21 January 2021

Published online: 04 February 2021



## References

1. Thacker, S. B. Historical Development. In *Principles and Practice of Public Health Surveillance* (eds Teutsch, S. M. & Churchill, R. E.) 1–16 (Oxford University Press, New York, 2000).
2. Groseclose, S. L. & Buckeridge, D. L. Public health surveillance systems: recent advances in their use and evaluation. *Annu. Rev. Public Health*. **38**, 57–79 (2017).
3. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**(5), 533–534 (2020).
4. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**(8), 1200–1204 (2020).
5. Pan, X. *et al.* Asymptomatic cases in a family cluster with SARS-CoV-2 infection. *Lancet Infect. Dis.* **20**(4), 410–411 (2020).
6. Fauci, A. S., Lane, H. C. & Redfield, R. R. Covid-19—navigating the Uncharted. *N. Engl. J. Med.* **382**(13), 1268–1269 (2020).
7. 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.
8. Stringhini, S. *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet* **396**(10247), 313–319 (2020).
9. Menachemi, N. *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.* **69**(29), 960–964 (2020).
10. Vogel G. First antibody surveys draw fire for quality, bias. *Science (New York, NY)*. 2020;368(6489):350–351.
11. American Community Survey 2018. In: US Census Bureau, 2020.
12. COVID-19 Case Counts and Testing Figures, Orange County, CA. 2020. <https://ocovid19.ochealthinfo.com/coronavirus-in-oc>. Accessed 8/16/20.
13. 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.
14. Rosenberg, E. S. *et al.* Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. *Ann. Epidemiol.* **48**(23–29), e24 (2020).
15. de Assis RR, Jain A, Nakajima R, *et al.* Analysis of SARS-CoV-2 Antibodies in COVID-19 Convalescent Plasma using a Coronavirus Antigen Microarray. *bioRxiv*. 2020.
16. Jain, A. *et al.* Evaluation of quantum dot immunofluorescence and a digital CMOS imaging system as an alternative to conventional organic fluorescence dyes and laser scanning for quantifying protein microarrays. *Proteomics* **16**(8), 1271–1279 (2016).
17. US Food and Drug Administration. In Vitro Diagnostics EUAs. In. Food and Drug Administration 2020. <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/vitro-diagnostics-euas> (Accessed on 9/15/20).
18. Khan S, Jain A, Taghavian O, *et al.* Use of an Influenza Antigen Microarray to Measure the Breadth of Serum Antibodies Across Virus Subtypes. *J Vis Exp*. 2019(149).
19. Nakajima R, Supnet M, Jasinskas A, *et al.* Protein Microarray Analysis of the Specificity and Cross-Reactivity of Influenza Virus Hemagglutinin-Specific Antibodies. *mSphere*. 2018;3, 6.
20. Lumley TS. *Complex Surveys: A Guide to Analysis Using R*. Wiley; 2010.
21. Fotini T, Evangelia V, Michail V. *Weighting of responses in the Consumer Survey: Alternative approaches – Effects on variance and tracking performance of the Consumer Confidence Indicator*. Foundation for Economic & Industrial Research 2013.
22. Goldstein, J. R. & Lee, R. D. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc. Natl. Acad. Sci. USA* **117**(36), 22035–22041 (2020).
23. Vassallo RR, Bracvo MD, Dumont LJ, Hazegh K, H. K. Seroprevalence of Antibodies to SARS-CoV-2 in US Blood Donors. *Medrxiv*. 2020.
24. Stadlbauer D, Tan J, Jiang K, *et al.* Seroconversion of a city: Longitudinal monitoring of SARS-CoV-2 seroprevalence in New York City. 2020.
25. Bendavid E, Mulaney B, Sood N, *et al.* COVID-19 Antibody Seroprevalence in Santa Clara County, California. *Medrxiv*. 2020.
26. Biggs HM, Harris JB, Breakwell L, *et al.* Estimated community seroprevalence of SARS-CoV-2 antibodies—two Georgia counties, April 28–May 3, 2020. *Morbidity and Mortality Weekly Report*, **69**(29), p965. 2020;69(29):965.
27. Martinez, D. A. *et al.* SARS-CoV-2 positivity rate for latinos in the Baltimore–Washington, DC Region. *JAMA* **324**(4), 392–395 (2020).
28. Chow DS, Soun J, Gavis-Bloom J, *et al.* The disproportionate rise in COVID-19 cases among Hispanic/Latinx in disadvantaged communities of Orange County, California: A socioeconomic case-series. *medrxiv*. 2020.
29. Page KR, Venkataramani M, Beyrer C, Polk S. Undocumented U.S. Immigrants and Covid-19. *The New England journal of medicine*. 2020;382(21):e62.
30. Yancy CW. COVID-19 and African Americans. *Jama*. 2020.
31. Saucedo JA, Neilands TB, Lightfoot M, Saberi P. Findings from a probability-based survey of U.S. households about prevention measures based on race, ethnicity, and age in response to SARS-CoV-2. *J Infect Dis*. 2020.
32. Halbesleben, J. R. & Whitman, M. V. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. *Health Serv. Res.* **48**(3), 913–930 (2013).
33. Davern, M. Nonresponse rates are a problematic indicator of nonresponse bias in survey research. *Health Serv. Res.* **48**(3), 905–912 (2013).
34. Groves, R. M. & Peytcheva, E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public opinion quarterly*. **72**(2), 167–189 (2008).
35. Ionnadis J. The infection fatality rate of COVID-19 inferred from seroprevalence data. *Medrxiv*. 2020.

## Acknowledgements

We are grateful to Dr. Neeraj Sood at the University of Southern California Price School of Public Policy for advice regarding study design, and for providing us the survey instrument used in their Los Angeles, CA seroprevalence study in conjunction with Dr. Paul Simon at the Los Angeles County Department of Public Health.

## Author contributions

T.B. and B.B.A. procured the funding for this study. T.B., D.P., V.V., E.D., B.A., M.Z., and B.B.A. contributed to the study design; D.P., S.B., V.V., B.A., B.B.A. collected the data; S.B., D.P., T.B., S.K., conducted the data analysis; T.B., D.P., S.B., S.K., A.N., B.B.A. led the interpretation of results; T.B. conducted the literature search; T.B. drafted the manuscript; D.P., S.B., V.V., S.K., B.A., M.Z., and B.B.A. wrote several sections of the manuscript. T.B., D.P., and S.B. have accessed and verified the underlying data. All Authors reviewed the manuscript for accuracy and approve the final version for submission.

## Funding

Orange County Healthcare Agency.

## Competing interests

Dr. Saahir Khan reports grants from University of California Office of the President and from National Institutes of Health during the conduct of the study. In addition, Dr. Khan has a patent “Devices For Detecting Antibodies To Coronavirus Antigens And Methods for Using Them,” Provisional Patent No. 62/993,610, filed March 23, 2020, licensed to Nanommune Inc. Saahir Khan is a consultant for and owns shares in Nanommune Inc. which is commercializing the technology used in this study. All other Authors (Drs. Bruckner, Parker, Bartell, Vieira, Drum, Noymer, Albala, Boden-Albala, Zahn) declare no potential conflict of interest.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-82662-x>.

**Correspondence** and requests for materials should be addressed to T.A.B. or B.B.-A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021