

Citywide serosurveillance of the initial SARS-CoV-2 outbreak in San Francisco

Isobel Routledge*, Adrienne Epstein*, Saki Takahashi*, Owen Janson^, Jill Hakim^, Elias Duarte^, Keirstinne Turcios^, Joanna Vinden^, Kirk Sujishi, Jesus Rangel, Marcelina Coh, Lee Besana, Wai-Kit Ho, Ching-Ying Oon, Chui Mei Ong, Cassandra Yun, Kara Lynch, Alan H.B. Wu, Wesley Wu, William Karlon, Edward Thornborrow, Michael J. Peluso, Timothy J. Henrich, John E. Pak, Jessica Briggs, Bryan Greenhouse*, Isabel Rodriguez-Barraquer*

*These authors contributed equally

^These authors contributed equally

Abstract

1 Serosurveillance provides a unique opportunity to quantify the proportion of the population that has been
2 exposed to pathogens. Here, we developed and piloted *Serosurveillance for Continuous, Actionable*
3 *Epidemiologic Intelligence of Transmission* (SCALE-IT), a platform through which we systematically
4 tested remnant samples from routine blood draws in two major hospital networks in San Francisco for
5 SARS-CoV-2 antibodies during the early months of the pandemic. Importantly, SCALE-IT allows for
6 algorithmic sample selection and rich data on covariates by leveraging electronic medical record data. We
7 estimated overall seroprevalence at 4.2%, corresponding to a case ascertainment rate of only 4.9%, and
8 identified important heterogeneities by neighborhood, homelessness status, and race/ethnicity.
9 Neighborhood seroprevalence estimates from SCALE-IT were comparable to local community-based
10 surveys, while providing results encompassing the entire city that have been previously unavailable.
11 Leveraging this hybrid serosurveillance approach has strong potential for application beyond this local
12 context and for diseases other than SARS-CoV-2.

13 Introduction

14 The rapid spread of the SARS-CoV-2 virus has laid bare important gaps in routine infectious diseases
15 surveillance. Serological data, particularly when collected at high spatial and temporal resolutions, are a
16 key resource for addressing many key epidemiological questions since they directly quantify the proportion
17 of the population that has been infected by a pathogen^{1,2}. For SARS-CoV-2, serology is particularly useful

18 given the high levels of disease under-ascertainment: serologic surveillance is the gold standard for
19 estimating attack rates (the proportion of the population that has been infected) and highly complementary
20 to virologic and syndromic surveillance systems for providing vital information on where a population is
21 along the epidemic curve³. Population-based serosurveys that employ a probabilistic sampling frame are
22 considered to be the gold standard for estimating seroprevalence. However, performing large population-
23 based serosurveys can be prohibitively resource-intensive to initiate swiftly or perform repeatedly,
24 especially during an ongoing outbreak, as demonstrated by the relative sparsity of population-based vs.
25 convenience sampled serosurveys for SARS-CoV-2 that have been conducted to date³. For example, to
26 date, no population-based serosurveys have been conducted for the city of San Francisco or wider Bay
27 Area, and few have been conducted in the United States, limiting our ability to identify of risk factors for
28 infection, understand population level immunity, and determine which populations and localities may be in
29 need of targeted public health resources such as testing, contact tracing, or vaccine allocation⁴.

30 Residual blood samples from readily available sources (e.g., blood donors or remnant samples collected
31 from routine medical care visits), especially when linked to individual-level meta-data, provide a unique
32 opportunity to address these limitations and to efficiently survey a population for antibodies over an
33 extended period of time^{5,6}. Such studies were found to be useful in the 2009 H1N1 influenza pandemic⁷⁻¹³,
34 facilitating analyses on a broader spatial and temporal scale than typical cross-sectional serological surveys
35 allow. However, in most studies that use residual blood samples the source population is unknown¹⁴. This
36 presents a major limitation, as the results are difficult to interpret when it is not known whether the sampled
37 population is representative of the population of interest.

38 The San Francisco Bay Area has widely been recognized for taking an early and proactive response to
39 COVID-19. San Francisco Bay Area counties introduced a shelter-in-place order on 17 March 2020,
40 requiring residents to remain at home unless leaving the house for essential activities. Relative to many
41 other US cities, few cases were detected in San Francisco during the early months of the epidemic, a pattern
42 which continued as the pandemic progressed. However, like many other areas, a high proportion of

43 asymptomatic infections and limited access to diagnostic testing during this time makes it difficult to
44 interpret these numbers. Results from an early San Francisco seroprevalence study conducted on
45 convenience samples in late March to early April 2020 suggested that <1% of the population had been
46 infected overall¹⁶, in contrast to a seroprevalence of >6% estimated by a community study focusing on a
47 specific neighborhood, particularly among the Hispanic/Latinx population¹⁷. The lack of citywide,
48 representative seroprevalence estimates during this time period limits the ability to determine to what
49 degree these discrepancies reflect heterogenous exposure or differences in study design.

50 Here we present a blueprint and early results of the ongoing SCALE-IT study (*Serosurveillance for*
51 *Continuous, Actionable Epidemiologic Intelligence of Transmission*), leveraging residual sera samples
52 from two large hospital systems in San Francisco, California to quantify the prevalence of SARS-CoV-2
53 antibodies. Importantly, these remnant samples are linked to electronic medical records (EMRs) enabling
54 careful algorithmic selection based on demographic and clinical variables, improving their
55 representativeness to the general population. We tested over 5,000 samples collected from late March to
56 June 2020 from San Francisco residents, and calculated raw and adjusted seroprevalence estimates over
57 space, time, and socio-demographic indicators. These data provide estimates of the overall seroprevalence
58 in San Francisco during the initial phase of the local SARS-CoV-2 outbreak and highlight spatial and
59 demographic heterogeneities in transmission across the city.

60 **Methods**

61 *Data Source*

62 Residual serum samples from routine blood draws from the University of California, San Francisco (UCSF)
63 and San Francisco Department of Public Health (SFDPH) inpatient and outpatient healthcare systems were
64 sampled from March 28, 2020 onward. UCSF Medical Center is a network of 3 hospitals with
65 approximately 1.8 million outpatient visits annually¹⁹. The SFDPH hospital, Zuckerberg San Francisco
66 General Hospital (ZSFG), is a city hospital which provides trauma, medical and surgical services to a

67 heterogeneous population of largely un- or underinsured patients, including the city's homeless population,
68 and serves roughly 100,000 patients per year²⁰.

69
70 We obtained daily EMRs for all patients in these networks undergoing routine blood testing, defined as
71 blood chemistries and tests for sexually transmitted infections, rubella, and lead. EMR data included
72 information on patient demographics, address, insurance provider, and diagnoses. We also obtained
73 information on all tests for respiratory infections (including SARS-CoV-2) performed on patients in the 6
74 months prior to the blood draw.

75
76 *Sampling Methodology*

77 We aimed to collect 2,000 samples monthly. We determined this sample size based on considerations of
78 both statistical power and feasibility. To estimate seroprevalence with an absolute error of 5% and at Type
79 I error of 5%, and a prior of 20% seroprevalence, a sample size of 246 individuals would need to be tested
80 each month. We determined that an overall sample size of a minimum 1230 samples per month would be
81 sufficient to allow stratification of results by five age groups (0-19, 20-39, 40-59, 60-79, 80+ years).

82
83 From the full list of residual serum samples that were available, we restricted our sampling frame to samples
84 from individuals undergoing routine blood testing. We included patients residing in San Francisco,
85 including those experiencing homelessness. We excluded individuals who were tested for SARS-CoV-2
86 during the visit when they received their blood draw (except if the test was for routine purposes, such as
87 testing prior to an elective procedure or admittance to the hospital). We restricted our sample to outpatient
88 and emergency department visits for adults; for the youngest age group, we included both inpatient and
89 outpatient visits due to small numbers of available samples. Finally, we excluded samples if a sample from
90 the same patient had been selected within the previous 30 days.

91

92 After obtaining the list of eligible samples according to the above criteria, we selected serum samples for
93 the study using a sampling algorithm aimed to ensure an adequate sample size for each of five age strata
94 and to maximize geographic representativity. After setting a daily target sample size for our overall
95 population, we divided this equally between five age bins to set a target sample size for each age bin. We
96 also set a target sample size for each zip code which was proportional to its population size. For each
97 zipcode with a larger number of eligible samples than its target size, we kept all samples from age groups
98 with sample sizes below or at their target and obtained a random sample from any age group that had an
99 eligible sample size above the target size. We intentionally over-sampled pregnant women as a healthy
100 sentinel population by aiming to obtain up to 10% of the samples from pregnant women undergoing routine
101 care, as defined by ICD-10 codes.

102

103 *Sample Processing*

104 Remnant samples were stored at +4 °C in outpatient laboratories at UCSF and ZSFG, and collected by our
105 study team twice every week. After collection, samples were centrifuged for 15 minutes at 3500 g before
106 aliquoting a working stock of 300 uL into 96 well barcoded tubes, diluting in 1:1 HEPES storage buffer,
107 and storing at +4 °C. The remainder of the sample was aliquoted into 1.4 mL barcoded tubes and stored at
108 -20 °C.

109

110 *Serologic Assays and Validation Data*

111 We used two serologic assays for this study in order to maximize assay specificity. First, we screened all
112 samples using an in-house ELISA assay, and then performed confirmatory testing on a subset of samples
113 above a threshold value using an in-house Luminex assay. The ELISA assay detected IgG to the receptor
114 binding domain (RBD) of the spike (S) protein, based on published protocols with minor modifications²¹.
115 Briefly, 1 ug of RBD was used to coat each well of 384-well high binding plates, secondary antibody was
116 diluted 1:5,000 (Southern Biotech #2048-05), and OPD was used to develop the plates. Concentration
117 values were calculated from the ELISA optical density (OD) using a plate-specific standard curve from

118 serial dilutions of a pool of positive control samples²². Samples with an ELISA concentration value above
119 0.049 were selected for confirmatory testing (see **Supplementary Text 1**).

120

121 For confirmatory testing, we used a multiplex microsphere assay (Luminex platform) to detect IgG against
122 the SARS-CoV-2 S protein, RBD, and the nucleocapsid (N) protein, based on a standardized serology
123 protocol with minor modifications²³. Briefly, plasma samples were diluted to 1:100 in blocking buffer A
124 (1xPBS, 0.05% Tween, 0.5% bovine serum albumin (BSA), 0.02% sodium azide). Antigen concentrations
125 used were as follows: S: 4 ug/mL, RBD: 2 ug/mL, and N: 3 ug/mL. As above, concentration values were
126 calculated from the Luminex median fluorescent intensity (MFI) using a plate-specific standard curve from
127 serial dilutions of a pool of positive control samples. A logistic regression model including the
128 concentration values of the three antigens for each sample was determined to have the highest cross-
129 validation accuracy for classification, and was used to establish a cutoff for positivity (see **Supplementary**
130 **Text 1**).

131

132 Serologic assays were optimized using positive and negative controls from several sources. Serum samples
133 from 127 patients with PCR confirmed SARS-CoV-2 infections (representing 266 total samples, with 1-4
134 longitudinal monthly time points per individual beginning at 3 weeks post-symptom onset) were obtained
135 from the *Long-term Impact of Infection with Novel Coronavirus* (LIINC) study
136 (<https://www.liincstudy.org/>) and used as positive controls. Importantly, participants in this cohort
137 represent a range of infection severities (ranging from asymptomatic to severe), age, sex, and ethnicity and
138 race. Serum samples from 119 individuals obtained prior to the emergence of SARS-CoV-2 were used as
139 negative controls. The overall sensitivity of our serial testing approach using positive and negative controls
140 was 94.0% (95% CrI = 89.0%, 97.2%) and specificity was 99.8% (95% CrI = 98.2%, 100.0%)
141 (**Supplementary Table 1, Supplementary Text 1**).

142

143

144 *Analytic Methods*

145 Raw seropositivity was determined as the proportion of all samples from unique individuals that tested
146 positive on the confirmatory assay. We then produced estimates of seroprevalence adjusted for the
147 sensitivity and specificity of the serial testing approach, incorporating potential conditional dependence of
148 the tests as described in Gardner *et al*²⁴ (see **Supplementary Text 1**). We stratified by covariates to obtain
149 seroprevalence estimates for each stratum (age, sex, insurance status, ethnicity, and neighborhood). To
150 identify neighborhoods, we geocoded sample addresses using the Google Cloud Geocoding API²⁵. Samples
151 (n=365 unique individuals) which could not be geocoded to rooftop (n=261) and/or were from homeless
152 individuals (n=157) were excluded from neighborhood level estimates of seroprevalence, however
153 estimates of seroprevalence were calculated for homeless individuals separately and provided alongside
154 neighborhood level estimates of seroprevalence. All analysis was conducted using the R statistical
155 software²⁶ and the Stan programming language²⁷. Code and data to reproduce all analyses are available at:
156 <https://github.com/EPPIcenter/scale-it>.

157

158 *Institutional Review Board (IRB) Approval*

159 This study received expedited review approval by the UCSF IRB #20-30379 (*'Serological Surveillance of*
160 *SARS-CoV-2 in Residual Serum/Plasma Samples'*). The IRB did not require patient contact or written
161 consent to use residual sera. The LIINC study (providing positive control samples) was approved by the
162 UCSF (IRB #20-30479). Pre-pandemic samples used as negative controls came from the New York Blood
163 Bank, and were de-identified and not subject to IRB review for use in this study.

164

165 **Results**

166 Between March 28 2020 and June 26 2020, we collected a total of 5,244 samples, representing 4,735
167 individual patients, from UCSF Health (n=3037 patients) and ZSFG (n=1698 patients) (**Figure 1**). By
168 design, the age distribution of sampled individuals remained consistent throughout the study period, and

169 the geographic distribution of residents matched the proportion of the San Francisco population living in
170 each zip code (Figure 2). Our sample did not achieve the target sample size for the youngest age group due
171 to the limited number of children receiving routine phlebotomy in the UCSF and ZSFG health systems
172 (Table 1). Our results were relatively representative of the San Francisco population by race and ethnicity,
173 although our sample overrepresented those who identified as Black/African American and slightly
174 underrepresented those who identified as Asian.

175
176 Overall, from 5,244 samples we identified 192/4,735 positive samples from unique patients for a raw
177 seroprevalence of 4.1%. After weighting for age group and sex to match the population structure of San
178 Francisco and correcting for test performance characteristics (overall sensitivity of 93.7% and specificity
179 of 99.6%), this corresponds to an estimated population seroprevalence of 4.2% (95% Credible Interval
180 [CrI]: 2.1%-6.3%). Based on the number of cases reported during the period covered by the study, we
181 estimate that only 4.9% of all infections were ascertained by the reporting system (95% CrI: 3.3%-9.9%)
182 (**Supplementary Text 1**). Amongst pregnant women seeking routine care (N=268), we estimated a raw
183 seroprevalence of 3.4% (9/268 seropositive), and after adjusting for test performance characteristics we
184 estimate 3.5% (95% CrI: 1.1 – 6.4%) seroprevalence amongst this group. This estimate in our sentinel
185 population group is consistent with the estimates across our overall population of samples.

186
187 We did not observe statistically significant differences in seroprevalence by age (**Figure 3A**) or hospital
188 system (**Supplementary Table 2**). We found seroprevalence to be nearly twice as high in uninsured
189 individuals (6.3%, 95% CrI: 3.1 - 9.9%) than in those with some form of insurance, [**Private/Commercial:**
190 3.4% (95% CrI: 1.6 - 4.7%); **Government:** 4.0% (95% CrI: 2.3 - 5.0%)] (**Figure 3B**). With respect to
191 race/ethnicity, seroprevalence was highest in those identifying as Hispanic (6.3%, 95% CrI: 4.4-8.3%)
192 followed by Black or African American (4.8%, 95% CrI: 2.8-7.0%), and lowest in those who identified as
193 Asian (2.3%, 95% CrI: 0.8-3.5%) (**Figure 3C**). Seroprevalence was almost twice as high in those
194 identifying as Male (5.3%, 95% CrI: 3.7%-6.6%) compared to Female (2.7%, 95% CrI: 1.1%-3.6%)

195 (Figure 3D). Although these samples were obtained over a three-month collection period, given the
196 relatively low attack rate during these initial stages of the pandemic in San Francisco, we were not able to
197 detect meaningful differences in seroprevalence over time (Supplementary Table 2).

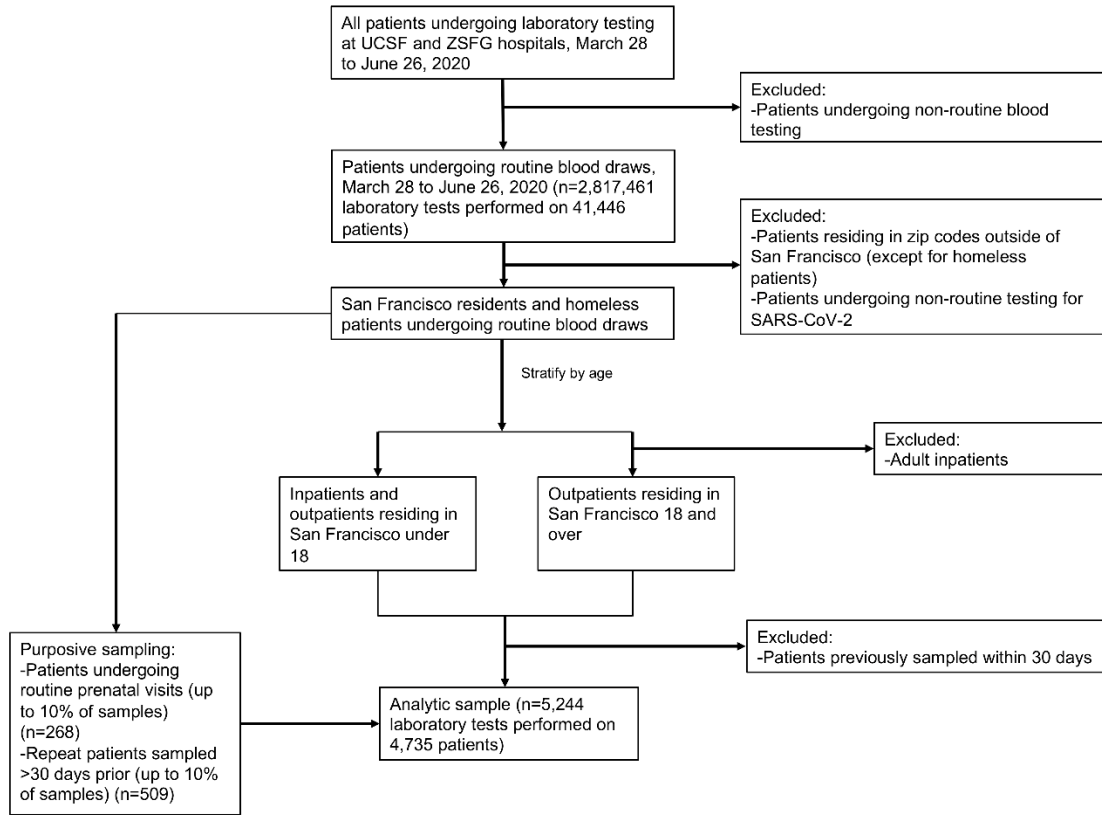
198
199 Geographically, we found seroprevalence to be highest in the Bayview neighborhood in the southeast region
200 of the city, at 8.1% (95% CrI: 4.6%, 12.3%) (Figure 4A, Supplementary Table 3). Although several other
201 neighborhoods had similarly high seroprevalences, there was much more uncertainty around these estimates
202 (Figure 4B). These findings are consistent with patterns of incidence in the city during this period of time
203 (Figure 4C). We identified 157 individuals who were homeless in our study, and amongst this group
204 seroprevalence was estimated to be 10.8% (95% CrI: 6.1%, 16.5%).

205
206 As validation of the representativity of our approach using curated remnant samples, we compared results
207 from this study to two contemporaneous community-based serosurveys conducted in specific
208 neighborhoods of San Francisco. First, we compared these results to a cross-sectional serosurvey carried
209 out in a census tract within the Mission District (census tract 022901, zip code 94110) between April 25
210 and April 28, 2020¹⁷. Chamie *et al* tested 2,545 census tract residents for SARS-CoV-2 antibodies and
211 estimated seroprevalence to be 3.1% (95% CI: 2.5-3.9%). This is consistent with our findings of 3.8%
212 seroprevalence (95% CrI: 1.8-6.3%) between April and June 2020 in the broader Mission District
213 neighborhood. Second, we compared our results to a cross-sectional serosurvey carried out in two census
214 tracts in San Francisco's 10th District between May 30 and June 2, 2020 ([https://unitedinhealth.org/sf-](https://unitedinhealth.org/sf-district-10)
215 [district-10](https://unitedinhealth.org/sf-district-10)), located in the Bayview neighborhood. Among the nearly 1,600 individuals tested for
216 antibodies, seroprevalence was estimated at 5.6% in Latinx participants (n=320), 2.3% in Black participants
217 (N= 397) and 0.4% in white participants (n=231). The relatively high seroprevalence we detected in the
218 Bayview neighborhood through our study is comparable to the results of this community-based study, and
219 the disparities by race/ethnicity were similar in direction, though different in magnitude, to those identified
220 through our remnant sample study as well. It is worth noting that the community studies available for

221 comparison also rely upon convenience sampling as participation in the studies was voluntary, and therefore
222 may contain inherent selection biases themselves.

223

224



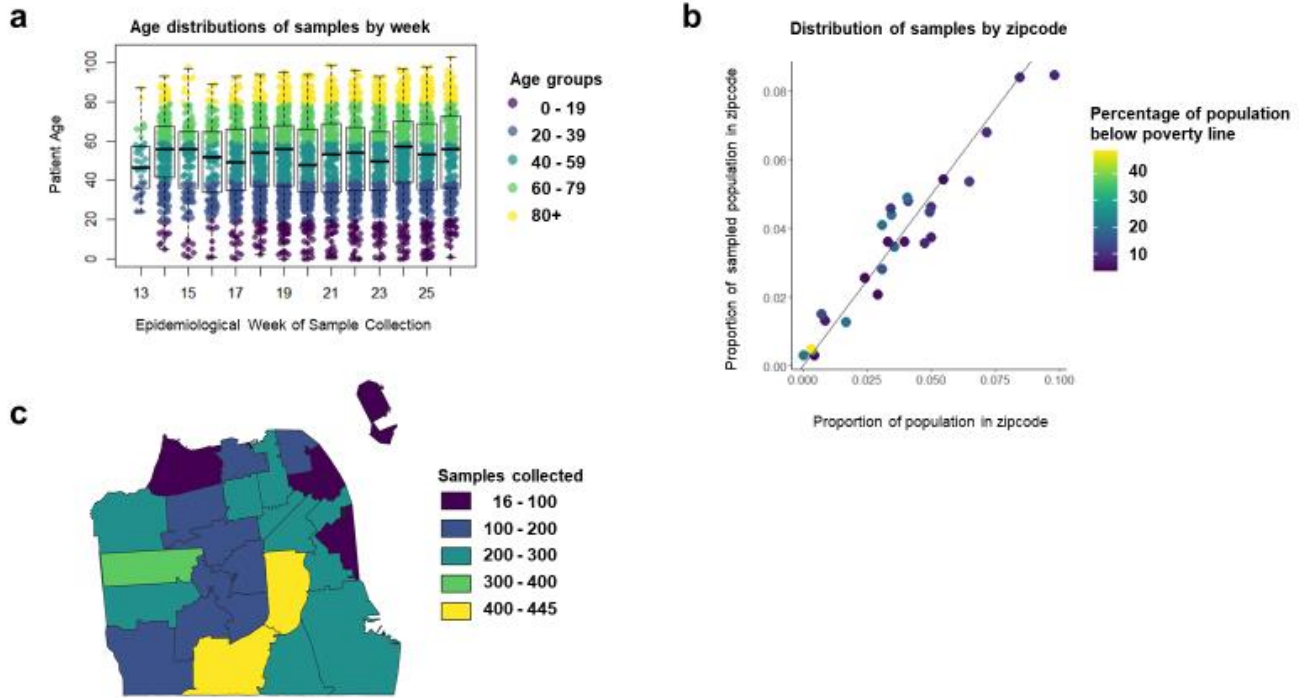
225

226 **Figure 1: Flow diagram of sampling algorithm**

Table 1. Socio-demographic characteristics of patients sampled in SCALE IT and of the San Francisco population (2019).

	UCSF (n=3,037)	ZSFG (n=1,698)	Total sampled individuals (n=4,735)	SF Population (ACS 2019)
Sex				
Female	1,733 (57.1%)	758 (44.6%)	2,491 (52.6%)	49.3%
Male	1,302 (42.9%)	929 (54.7%)	2,231 (47.1%)	50.8%
Unknown	2 (0.1%)	11 (0.6%)	13 (0.3%)	N/A
Age				
0-19	246 (8.1%)	35 (2.1%)	281 (5.9%)	15.0%
20-39	836 (27.5%)	425 (25.0%)	1,261 (26.6%)	38.0%
40-59	731 (24.1%)	591 (34.8%)	1,322 (27.9%)	25.3%
60-79	834 (27.5%)	556 (32.7%)	1,390 (29.4%)	17.3%
80+	390 (12.8%)	91 (5.4%)	481 (10.2%)	4.3%
Race/Ethnicity				
American Indian or Alaska Native	3 (0.1%)	9 (0.5%)	12 (0.3%)	0.3%
Asian	783 (25.8%)	423 (24.9%)	1,206 (25.5%)	34.6%
Black or African American	283 (9.3%)	308 (18.1%)	591 (12.5%)	5.2%
Other	214 (7.0%)	73 (4.3%)	287 (6.1%)	4.5%
Other Pacific Islander	28 (0.9%)	17 (1.0%)	45 (1.0%)	0.4%
White	1,317 (43.4%)	358 (21.1%)	1,675 (35.4%)	39.8%
Unknown or Declined	43 (1.4%)	18 (1.1%)	61 (1.3%)	N/A
Hispanic*	366 (12.1%)	492 (29.0%)	858 (18.1%)	15.2%
Insurance Type				
Uninsured	119 (3.9%)	150 (8.8%)	269 (5.7%)	N/A
Government	1,462 (48.1%)	1,475 (86.9%)	2,937 (62.0%)	N/A
Private or Employer	1,351 (44.5%)	70 (4.1%)	1,421 (30.0%)	N/A
Unknown	105 (3.5%)	3 (0.2%)	108 (2.3%)	N/A

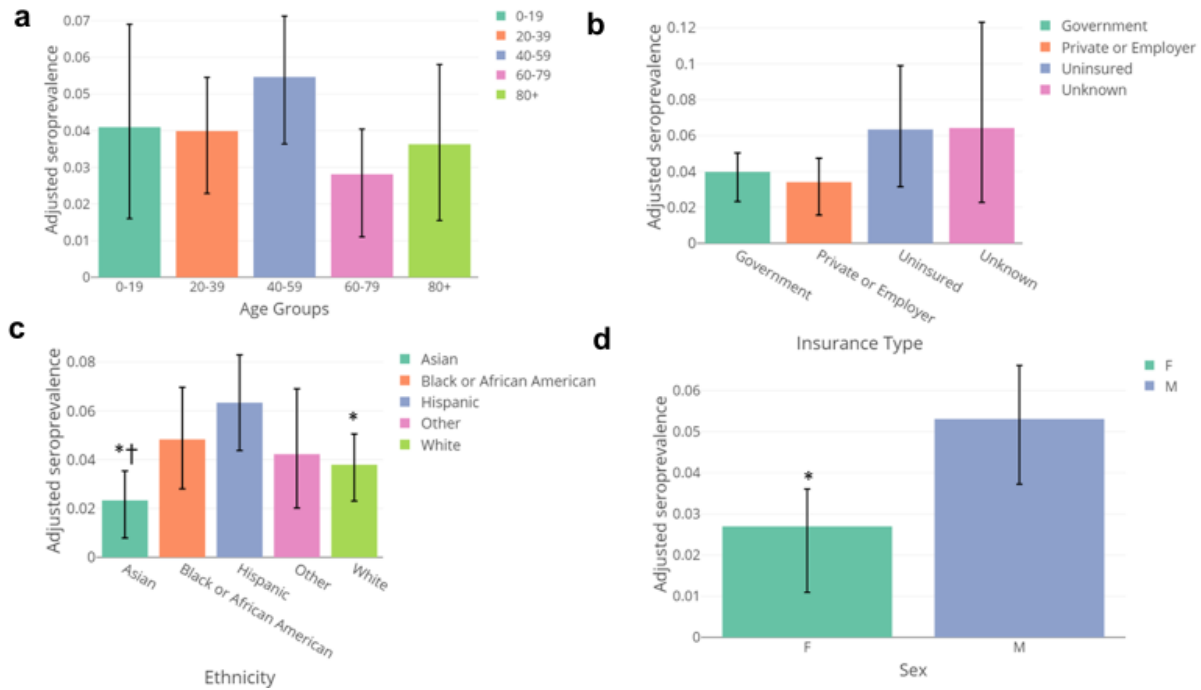
*Hispanic includes respondents of any race. Other categories are non-Hispanic.



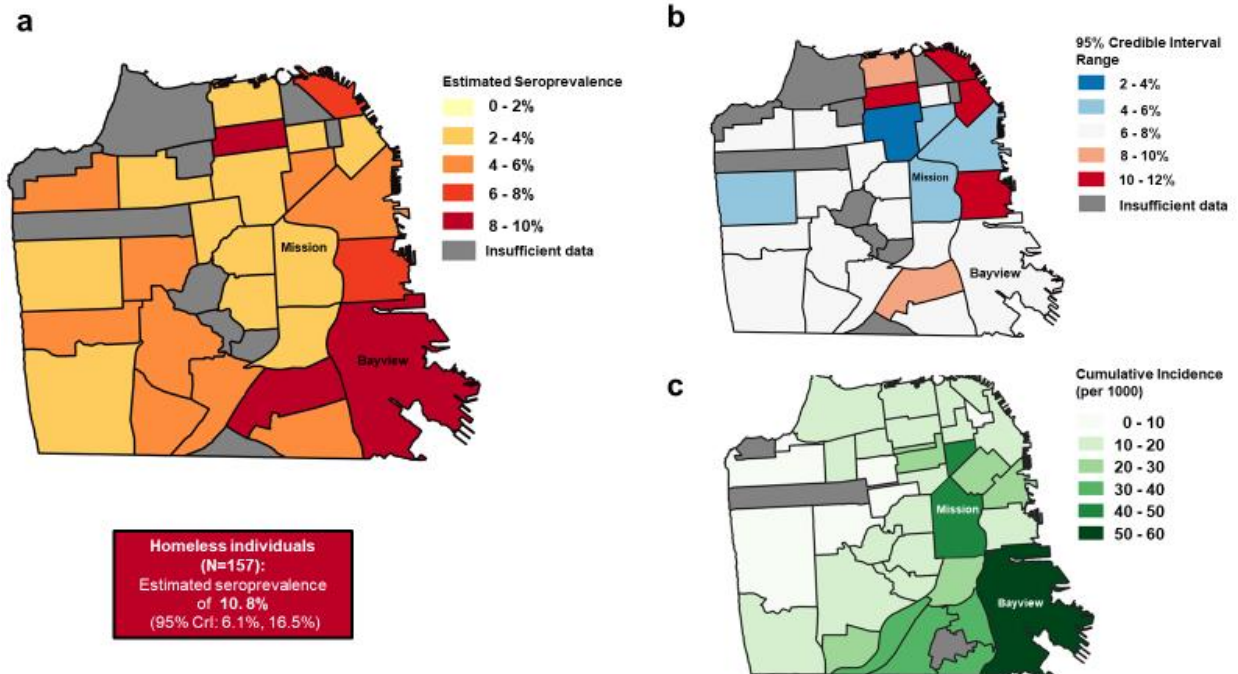
230

231 **Figure 2:** Distributions of SCALE-IT samples by A) epidemiological week and age group, B) zip code and

232 percentage below the poverty line, and C) map of counts of samples collected by zip code.



233
 234 **Figure 3:** Stratified seroprevalence by A) age, B) insurance type, C) ethnicity (groups with $N < 50$ were
 235 excluded from plot) and D) sex. Estimates are adjusted for test performance, and error bars show 95%
 236 credible intervals. For C), stars (*) indicate the ethnic groups where the 2.5% and 97.5% quantiles of
 237 (Figure 3 continued) the differences in posterior estimates for seroprevalence between samples from
 238 Hispanic patients and that group did not cross zero. Crosses (†) indicate the ethnic groups where the
 239 2.5% and 97.5% quantiles of the differences in posterior estimates for seroprevalence between samples
 240 from Black or African American patients and that group did not cross zero. For D) a star (*) indicates
 241 that the 2.5% and 97.5% quantiles of the differences in posterior estimates for seroprevalence between
 242 Males and Females did not cross zero.



243
 244 **Figure 4:** Multipanel map showing A) seroprevalence by neighborhood, adjusted for test performance.
 245 Box shows adjusted seroprevalence in individuals experiencing homelessness. B) range of 95% Credible
 246 interval of estimates, C) cumulative incidence by planning neighborhood from March - June 2020, using
 247 data from SFDPH ([https://data.sfgov.org/COVID-19/COVID-19-Cases-by-Geography-and-Date/d2ef-](https://data.sfgov.org/COVID-19/COVID-19-Cases-by-Geography-and-Date/d2ef-idww)
 248 [idww](https://data.sfgov.org/COVID-19/COVID-19-Cases-by-Geography-and-Date/d2ef-idww)). For A) and B), estimates for neighborhoods with under 50 samples from unique individuals are
 249 not plotted and shown in grey.

250 **Discussion**

251 In this study, we developed and piloted a scalable and systematic pipeline using remnant samples from two
252 major hospital networks in San Francisco to select, collect, and test specimens for SARS-CoV-2 antibodies
253 (SCALE-IT). Through this effort, we estimated seroprevalence during the early months of the epidemic to
254 be relatively low throughout San Francisco (4.2%), but still representing more than 20 times the number of
255 infections identified by PCR-confirmed cases at that time. This may be due to the limited availability of
256 PCR testing during the beginning of the pandemic and the lack of testing of asymptomatic individuals. We
257 also identified important disparities in seroprevalence at the neighborhood level, with highest
258 seroprevalence in the Bayview neighborhood in the southeast region of the city, as well as
259 disproportionately higher seroprevalence in individuals experiencing homelessness and those identifying
260 as Hispanic, Black/African American, or male. Leveraging this hybrid serosurveillance approach has
261 potential for broad application beyond this local context and for diseases other than SARS-CoV-2.

262
263 The heterogeneities in seroprevalence we observed by race/ethnicity and socio-economic status -- here
264 obtained from EMR data on health insurance status and whether individuals were housed -- echo patterns
265 which have been highlighted over the course of the pandemic at national and global levels^{29,30}. Specific to
266 San Francisco, our results provide estimates of SARS-CoV-2 cumulative exposure at a granular spatial
267 resolution with a scope covering the entire city; despite low overall seroprevalence, we identified specific
268 neighborhoods with disproportionately higher seroprevalence. Interestingly, we also found seroprevalence
269 to be approximately twice as high in those identifying as male compared to female. Potential explanations
270 for this difference include differential pathogen exposure by sex, which is supported by findings of other
271 studies in San Francisco, finding PCR positivity rates of 1.2% (20/1658) in women and 3.3% (63/1908)
272 in men, with an odds ratio of 2.71 (1.64-4.69) for PCR positivity in males, and also that the majority (74%,)
273 of those who tested positive by PCR or were seropositive for SARS-CoV-2 were frontline workers and
274 unable to shelter-in-place¹⁷, it has been found that males and females mount different immune responses
275 and infection severity³¹, which could affect assay sensitivity, however we believe this is unlikely to explain

276 the large difference we see in our estimates as we do not see sex-based differences in the sensitivity of our
277 assay on the positive controls used in the study, which represent a range of disease severities.

278

279 While a key strength of our approach was leveraging residual sera from two large health system networks
280 and using data from EMRs to algorithmically select samples for inclusion, there are limitations to this type
281 of surveillance that require consideration. Most obviously, patient samples may not be fully representative
282 of the underlying population. This may be particularly true during “shelter-in-place” periods, when
283 behavioral changes may affect the availability and characteristics of the patient population. These issues
284 can ideally be mitigated by careful sample selection, as done here by focusing on a subset of outpatients,
285 with the possibility of further refinement by inclusion of additional selection criteria (e.g., by restricting or
286 weighting sampling to consider specific visit types or underlying conditions). Representativity of the
287 serosurveillance system could also be enhanced by including a broader network of local health systems.
288 We also recognize that the generalizability of our findings may differ by age groups, and is likely to be
289 lower in children who were under-represented in our sample set despite the stratified sampling framework.
290 Additional study designs, such as school-based serosurveys, could be leveraged to augment these data to
291 prospectively assess seroprevalence in specific age-groups, possibly by using non-invasive, saliva-based
292 antibody testing³². Despite including over 5,000 samples, our study was not powered to detect differences
293 between covariates or by time in a multiple regression framework, in part due to San Francisco’s success
294 in maintaining low transmission and thus low seroprevalence during this time period. Lastly, while we
295 validated our estimates against results from a couple of available community based studies, further
296 validation would be ideal to assess validity of results and findings.

297

298 In this pilot study, we developed and implemented a SARS-CoV-2 serosurveillance system to detect
299 population-level pathogen exposure in near-real time, and demonstrated how data collected through this
300 platform were comparable to results from more resource intensive community-based serological studies
301 and incidence data. The appeal of this hybrid approach is that it achieves many of the strengths of

302 population-based surveys and provides rich data, while leveraging existing infrastructure to allow for much
303 greater efficiencies often seen in convenience sampling approaches. Using EMR data, we were able to
304 develop a stratified sampling frame, ensuring improved representativeness of the results in contrast to
305 serosurveys performed using convenience samples without these key pieces of information¹⁴. At the same
306 time, we used these data to identify important spatial and demographic heterogeneities in seroprevalence
307 within our study site; serosurveys performed on residual samples are often limited to coarser levels of meta-
308 data on the sampled population³³. The relative ease with which SCALE-IT can be implemented means that
309 it can be deployed over a broad geographic scale, continuously over time, and dynamically adjusted to
310 address specific surveillance needs.

311
312 We envision multiple lines of work for future directions. First, the samples that we have selected, collected,
313 and processed in this work could serve as a valuable biorepository for future applications. The ability to
314 link rich EMR data to a large bank of well-curated serum samples opens up opportunities for additional
315 analysis including longitudinal studies of patients. Second, as serosurveillance efforts will be fundamental
316 to monitor SARS-CoV-2 transmission rates and evaluate the impact of control interventions (both NPIs and
317 pharmaceuticals) over the coming months and years, future work could leverage these and prospective
318 serological data to parametrize mechanistic models and to study the effects of control strategies on infection
319 rate. Third, as discussed by others^{1,2}, our local SCALE-IT platform could easily be expanded to contribute
320 to a ‘Global Immunological Observatory’ to perform serosurveillance for other pathogens beyond the
321 SARS-CoV-2 virus. Data generated by such an observatory could be used to address specific public health
322 gaps including serosurveillance for seasonal pathogens such as influenza or emerging infections. Lastly,
323 the insights gained from developing this platform could serve as a blueprint for adoption by other health
324 systems in various contexts.

325

326 **Acknowledgements**

327

328 We acknowledge the significant contribution to this work made by the following persons and organizations:
329 Dr. Kim Rhoads, Dr. Diane Havlir and the Unidos en Salud United in Health partnership, the Office of
330 Community Engagement at the UCSF Helen Diller Family Comprehensive Cancer Center, and the District
331 10 community partners and participants at the Rafiki Coalition for Health and Wellness, J & J Community
332 Resource Center, The Samoan Community Development Center, and the Young Community Developers,
333 for providing information from community-based testing and response efforts in the Bayview
334 neighborhood. We also acknowledge Jennifer Creasman, Dalia Martinez, and Susan Sudduth at the UCSF
335 Clinical & Translational Science Institute (CTSI) and Janet Nguyen at ZSFG for their valuable assistance
336 in accessing the EMR databases. We also acknowledge the clinical research, laboratory, and epidemiology
337 teams for collecting valuable samples and data from the LIINC cohort.

338

339 **Author Contributions**

340 IR, AE, ST, BG, JB, and IRB conceived of the study. IR and AE managed sample selection activities with
341 support from JV. Plasma specimens were collected by KS, JR, MC, LB, WKH, CYO, CMO, CY, KL, AW,
342 and WK. OJ, JH, ED, KT, and JV performed antibody assays with proteins provided by JP and WW. MP
343 and TH and provided and analyzed serum from positive controls. IR and ST performed data analyses with
344 support from AE. The manuscript and figures were prepared by IR, AE, and ST, with additional input from
345 BG and IRB. All authors contributed to interpretation of the results and edited the manuscripts. All authors
346 read and approved the final manuscript.

347

348 **Role of the Funding Source & Declaration of Interests**

349

350 Sources of support included funding from the Schmidt Science Fellows, in partnership with the Rhodes
351 Trust (ST), Chan Zuckerberg Biohub Investigator program (BG), the ZSFG Department of Medicine and
352 Division of HIV, ID, and Global Medicine, the MIDAS Coordination Center (MIDASNI2020-
353 5) by a grant from the National Institute of General Medical Science (3U24GM132013-02S2), and the
354 National Institutes of Health/National Institute of Allergies and Infectious Diseases (NIH/NIAID
355 3R01AI141003-03S1).

356

357 The funders had no role in study design, data collection and analysis, decision to publish, or preparation of
358 the manuscript. None of the authors have conflicts of interest to disclose.

359

360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385

References

1. Metcalf, C. J. E. *et al.* Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *The Lancet* **388**, 728–730 (2016).
2. Mina, M. J. *et al.* A Global Immunological Observatory to meet a time of pandemics. *eLife* **9**, e58989 (2020).
3. Arora, R. K. *et al.* SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. *Lancet Infect. Dis.* **0**, (2020).
4. Bubar, K. M. *et al.* Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *medRxiv* 2020.09.08.20190629 (2020) doi:10.1101/2020.09.08.20190629.
5. Metcalf, C. J. E., Mina, M. J., Winter, A. K. & Grenfell, B. T. Opportunities and challenges of a World Serum Bank – Authors’ reply. *The Lancet* **389**, 252 (2017).
6. Clapham, H. *et al.* Seroepidemiologic Study Designs for Determining SARS-COV-2 Transmission and Immunity - Volume 26, Number 9—September 2020 - Emerging Infectious Diseases journal - CDC. doi:10.3201/eid2609.201840.
7. Bandaranayake, D. *et al.* Risk Factors and Immunity in a Nationally Representative Population following the 2009 Influenza A(H1N1) Pandemic. *PLoS ONE* **5**, (2010).
8. Gilbert, G. L. *et al.* Influenza A (H1N1) 2009 Antibodies in Residents of New South Wales, Australia, after the First Pandemic Wave in the 2009 Southern Hemisphere Winter. *PLoS ONE* **5**, (2010).
9. Dowse, G. K. *et al.* Incidence of pandemic (H1N1) 2009 influenza infection in children and pregnant women during the 2009 influenza season in Western Australia - a seroprevalence study. *Med. J. Aust.* **194**, 68–72 (2011).
10. Reed, C., Katz, J. M., Hancock, K., Balish, A. & Fry, A. M. Prevalence of Seropositivity to Pandemic Influenza A/H1N1 Virus in the United States following the 2009 Pandemic. *PLOS ONE* **7**, e48187 (2012).

- 386 11. Waalen, K. *et al.* High prevalence of antibodies to the 2009 pandemic influenza A(H1N1) virus in the
387 Norwegian population following a major epidemic and a large vaccination campaign in autumn 2009.
388 *Eurosurveillance* **15**, 19633 (2010).
- 389 12. Hoschler, K. *et al.* Seroprevalence of Influenza A(H1N1)pdm09 Virus Antibody, England, 2010 and
390 2011. *Emerg. Infect. Dis.* **18**, 1894–1897 (2012).
- 391 13. Mak, G. C. *et al.* Sero-immunity and serologic response to pandemic influenza A (H1N1) 2009 virus
392 in Hong Kong. *J. Med. Virol.* **82**, 1809–1815 (2010).
- 393 14. Havers, F. P. *et al.* Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United
394 States, March 23–May 12, 2020. *JAMA internal medicine* **180**, 1576–1586 (2020).
- 395 15. Johns Hopkins Coronavirus Resource Center. COVID-19 United States Cases by County. [Online]
396 Available at: <https://coronavirus.jhu.edu/us-map>.
- 397 16. Ng, D. L. *et al.* SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood.
398 *Nat. Commun.* **11**, 4698 (2020).
- 399 17. Chamie, G. *et al.* Community Transmission of Severe Acute Respiratory Syndrome Coronavirus 2
400 Disproportionately Affects the Latinx Population During Shelter-in-Place in San Francisco. *Clin.*
401 *Infect. Dis.* (2020) doi:10.1093/cid/ciaa1234.
- 402 18. Appa, A. *et al.* Universal PCR and antibody testing demonstrate little to no transmission of SARS-
403 CoV-2 in a rural community. *Open Forum Infect. Dis.* (2020) doi:10.1093/ofid/ofaa531.
- 404 19. UCSF Health. Annual Reports.[online] Available at: [https://www.ucsfhealth.org/about/annual-](https://www.ucsfhealth.org/about/annual-reports)
405 reports. [Accessed 11 January 2021].
- 406 20. Zuckerberg San Francisco General Hospital and Trauma Center. About UCSF At ZSFG. [online]
407 Available at: <https://zsfg.ucsf.edu/about-ucsf-zsfg> [Accessed 11 January 2021].
- 408 21. Roy, V. *et al.* SARS-CoV-2-specific ELISA development. *J. Immunol. Methods* **484**, 112832 (2020).
- 409 22. EPPIcenter/flexfit: Flexible format standard curve fitting and data processing (R package).
410 <https://github.com/EPPIcenter/flexfit>.
- 411 23. Wu, L. *et al.* Optimisation and standardisation of a multiplex immunoassay of diverse Plasmodium

412 falciparum antigens to assess changes in malaria transmission using sero-epidemiology. *Wellcome*
413 *Open Res.* **4**, 26 (2020).

414 24. Gardner, I. A., Stryhn, H., Lind, P. & Collins, M. T. Conditional dependence between tests affects the
415 diagnosis and surveillance of animal diseases. *Prev. Vet. Med.* **45**, 107–122 (2000).

416 25. Google Geocoding API [online] available at:
417 <https://developers.google.com/maps/documentation/geocoding/overview>. [Accessed 11 January 2021].

418 26. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical
419 Computing, Vienna, Austria. (2017). <https://www.R-project.org/>.

420 27. Stan Development Team. 2020. Stan Modeling Language Users Guide and Reference Manual (2020).
421 <https://mc-stan.org>

422 28. Long, Q.-X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* **26**,
423 845–848 (2020).

424 29. Johns Hopkins Coronavirus Resource Center. COVID-19 Racial Data Transparency [online].
425 Available at: <https://coronavirus.jhu.edu/data/racial-data-transparency>. [Accessed 11 January 2021].

426 30. Centers for Disease Control and Prevention. COVID-19 Hospitalization and Death by Race/Ethnicity
427 (COVID-19). [online]. Available at: [https://www.cdc.gov/coronavirus/2019-ncov/covid-](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html)
428 [data/investigations-discovery/hospitalization-death-by-race-ethnicity.html](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html) [Accessed 11 January
429 2021].

430 31. Takahashi, T. *et al.* Sex differences in immune responses that underlie COVID-19 disease outcomes.
431 *Nature* **588**, 315–320 (2020).

432 32. Cooch, P. *et al.* Supervised self-collected SARS-CoV-2 testing in indoor summer camps to inform
433 school reopening. *medRxiv* 2020.10.21.20214338 (2020) doi:10.1101/2020.10.21.20214338.

434 33. Anand, S. *et al.* Prevalence of SARS-CoV-2 antibodies in a large nationwide sample of patients on
435 dialysis in the USA: a cross-sectional study. *The Lancet* **396**, 1335–1344 (2020).