BMJ Open

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020- a mail-based cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051157
Article Type:	Original research
Date Submitted by the Author:	13-Mar-2021
Complete List of Authors:	Snyder, Teah; University of Massachusetts Amherst, Biostatistics and Epidemiology Ravenhurst, Johanna; University of Massachusetts Amherst, Biostatistics and Epidemiology Cramer, Estee; University of Massachusetts Amherst, Biostatistics and Epidemiology Reich, Nicholas; University of Massachusetts Amherst, Biostatistics and Epidemiology Balzer, Laura; University of Massachusetts Amherst, Biostatistics and Epidemiology Alfandari, Dominique; University of Massachusetts Amherst, Veterinary and Animal Sciences Lover, Andrew A.; University of Massachusetts Amherst, Biostatistics and Epidemiology, School of Public Health and Health Sciences
Keywords:	EPIDEMIOLOGY, Public health < INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020: a mail-based cross-sectional study

Teah Snyder,[1†] Johanna Ravenhurst,[1†] Estee Y. Cramer,[1†] Nicholas G. Reich,[1] Laura B. Balzer,[1] Dominique Alfandari,[2] and Andrew A. Lover[1*]

† Equal contributions.

[1] Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts- Amherst, Amherst MA

[2] Department of Veterinary and Animal Sciences, College of Natural Sciences, University of Massachusetts- Amherst, Amherst MA

* Correspondence to: alover@umass.edu

School of Public Health and Health Sciences
University of Massachusetts- Amherst
715 North Pleasant Street
Amherst, MA 01003-9304
phone: 413-545-7426

Abstract

Objectives: To estimate the seroprevalence of anti-SARS-CoV-2 IgG and IgM among Massachusetts residents and to better understand asymptomatic SARS-CoV-2 transmission during the summer of 2020.

Design: Mail-based cross-sectional study

Setting: Massachusetts, United States

Participants: Primary sampling group: sample of undergraduate students at the University of Massachusetts, Amherst (n = 548) and a member of their household (n = 231). Secondary sampling group: sample of graduate students, faculty, librarians and staff (n = 214) and one member of their household (n = 78). All participants were MA residents without prior COVID-19 diagnosis.

Primary and secondary outcome measures: Prevalence of SARS-CoV-2 seropositivity. Association of seroprevalence with variables including age, gender, race, geographic region, occupation, and symptoms.

Results: Approximately 27,000 persons were invited via email to assess eligibility. 1,001 households were mailed dried blood spot sample kits, 762 returned blood samples for analysis. In the primary sample group, 36 (4.6%) had IgG antibodies detected for an estimated weighed prevalence in this population of 5.3% (95% CI: 3.5 to 8.0). In the secondary sampling group, 10 (3.4%) had IgG antibodies detected for an estimated adjusted prevalence of 4.0% (95% CI: 2.2 to 7.4). No samples were IgM positive. No association was found in either group between seropositivity and self-reported work duties or customer-facing hours. In the primary sampling group, self-reported febrile illness since Feb 2020, male sex, and minority race (Black or American Indian/Alaskan Native) were associated with seropositivity. No factors except geographic regions within the state were associated with evidence of prior SARS-CoV-2 infection in the secondary sampling group.

Conclusions: This study fills a critical gap in estimating the levels of sub-clinical and asymptomatic infection. Estimates can be used to calibrate models estimating levels of population immunity over time, and these data are critical for informing public health interventions and policy.

Strengths and limitations of this study

- Our study collected serological samples in a well-defined rigorous sample frame in a contact-less (mail-based) survey in an early stage of the pandemic in an area of high SARS-CoV-2 burden.
- A range of potentially associated demographic, occupational, and behavioral factors were surveyed to contextualize seropositivity across geographic regions within the state.
- Our study sampled from populations affiliated with a large public university in Massachusetts, and may not be generalizable to the general population

Introduction

Since emergence in early 2020, the SARS-CoV-2 virus has severely impacted the entire globe. The state of Massachusetts was heavily impacted in the earliest stages of the pandemic, and a "super-spreader" event in the state in April 2020 may have seeded large case clusters throughout the country.[1] However, the trajectory of the early stages of transmission in this state, as well as across the US remain poorly understood due to changes in case definitions and limited testing of both symptomatic and asymptomatic persons during the summer of 2020.[2] To assess seroprevalence across the state, a mail-based serosurvey was implemented July-August 2020. At the time of this survey, the Massachusetts Department of Public Health had reported over 109,143 confirmed COVID-19 cases and over 8,081 deaths.[3] Seroepidemiological studies are a critical tool to explore infection dynamics, especially where asymptomatic or subclinical infections are common, as for SARS-CoV-2.[4] This study helps to fill a critical gap in estimating the levels of sub-clinical and asymptomatic infection to inform consequent levels of population-immunity.[5]

Concurrent to this study, a number of seroprevalence studies were conducted on the east coast of the United States; these studies focused on specific populations at high risk and found varying results. A survey in April 2020 in a convenience sample of 200 asymptomatic residents of Chelsea, MA found an estimated of seroprevalence of 31.5% (17.5% IgM+/IgG+, 9.0% IgM+/IgG- and 5.0% IgM-/IgG+).[6] This study used a small convenience sample and did not include any randomization.[6] A study with a larger sample of over 28,000 clinical patient samples in New York City, USA found an IgG seropositivity prevalence of 44% with over 50% of participants reporting no symptoms.[7]

Other seroprevalence surveys across the US have found generally low- to moderate prevalence in a diverse set of study populations. A study of 790 university students in Los Angeles, California conducted in April and May of 2020 estimated a prevalence of SARS-CoV-2 IgG antibody of 4.0% (95% CI: 3.0 to 5.1%).[8] During May – April of 2020, a cross-sectional study in St. Louis found IgG seropositivity to be estimated at 1.71% (95% CI 0.04% to 3.38%) in pediatric patients and 3.11% (95% CI: 0.92% to 5.32%) in adult patients. In the most comprehensive serosurvey from the spring and summer months of 2020, 16,025 clinical samples were analyzed with IgG

spike protein sero-reactivity ranging from 1.0% in the San Francisco Bay Area to 6.9% in New York City.[9] These disparate results highlight major geographic variability in the trajectory of infections, and reinforce the need for additional seroprevalence studies to more fully contextualize trends in immunity to SARS-CoV-2 targeting specific geographic regions.

Though community seroprevalence studies generally rely on serum samples collected in health facilities, the use of dried blood spot (DBS) samples is a practical and effective alternative.[10] DBS samples involve a small finger-prick sample self-collected by participants in their own homes. The use of dried blood samples for antibody assays has been validated in other work prior to the current pandemic,[10,11] and previous studies have evaluated the feasibility, validity, and acceptability of using DBS samples for SARS-CoV-2 antibody testing.[12–15] This method of sample collection facilitates efficient population-level sampling while minimizing social mixing and concurrent potential exposures.

This study estimated the prevalence of previous infection with SARS-CoV-2 in individuals who had not been diagnosed with COVID-19 and were asymptomatic with representative coverage across the entire state of Massachusetts, USA. Information from this study can provide knowledge regarding the seropositivity of this population and can be used to inform decision-making regarding community re-openings during the pandemic.

Methods

Study Design and Participants

The study population included undergraduate students, graduate students, staff, and faculty members currently affiliated with the University of Massachusetts, Amherst and their household members. On-campus classes were suspended in mid-March 2020; consequently, undergraduates had exposure to the local epidemiology within their communities from March until sampling in July-August throughout the state (primary sampling group). Conversely, graduate students, faculty, staff, librarians and their family members (secondary sampling group) almost universally reside in close proximity to Amherst, and broadly reflect transmission in the Western part of the state. UMass affiliates were eligible to participate in this study if they were above the age of 18, had been living in Massachusetts for the past eight weeks; had never received a COVID-19

diagnosis from a medical professional; and did not have a fever greater than 100.4° F at the time of survey completion. Household members were eligible for inclusion if they met all of these same criteria and were between the ages of 23 and 78 (chosen to expand sampling beyond college-age population groups).

An institutional email list was provided by university administration for recruitment. Initial emails were sent out to UMass affiliates between June 23, 2020 and June 26, 2020 for participant recruitment. The email provided information about the study and links to a screening eligibility survey, informed consent document, initial survey regarding COVID-19 risk factors, and information regarding shipping addresses. If the UMass affiliate had a household member interested in participating, a single household member was invited to complete an eligibility, consent, and initial survey forms. To increase participation rates, two reminder emails were sent to all non-respondents (day three and six after initial solicitation). All survey responses were collected and stored in REDCap.[16]

The survey was closed after a three-week enrollment period, and a subset of participants were selected to receive a test kit. To select a population representative of the broader UMass community across the entire state of Massachusetts, two sampling schemes were applied. The first consisted of all undergraduates and their household members (primary sampling group); the second sampling frame consisted of graduate students, staff, faculty members, librarians, and their household members (secondary sampling group). Within the primary sampling group, selection for biosample collection used probability proportional to population size, using the most recent census data aggregated to state-level emergency response regions due to sparse county-level populations (Figure 1).[17] For the secondary sampling group, selection for biosampling was via simple random sampling.

The full sample frame selection is shown in Figure 2. Briefly, an email invitation was sent to a total of 27,339 individuals, of which 4,124 completed the screening, informed consent, and initial survey forms. A total of 1,001 individuals were then randomly selected to receive a sampling kit. Participants were mailed all materials to safely collect and return samples,

including lancets, alcohol wipes, gauze, gloves, bandages, a bloodspot collection card, a pre-paid shipping box, and detailed printed instructions cards (including a nurse call line).

Upon mailing out the test kits, participants were also emailed a link containing a video on how to collect the DBS, along with a detailed survey form with demographics, risk factors, and any current symptoms or COVID diagnoses. No participants reported a COVID-19 diagnosis between the initial survey, and sample collection several weeks later. All shipments utilized a Biological Substance Category B (UN3373) shipping box.

Ethical approvals

This study was approved by the University of Massachusetts-Amherst Human Research Protection Office (Approval #2062; April 27, 2020).

Sample Preparation and ELISA Analysis

Upon receipt of boxes, the sample cards (Whatman® Protein Saver 903) were heat-treated (30 minutes at 56^0 C); a single blood spot per card was punched (0.25-inch diameter) and transferred to an ELISA plate. Plates were coated with 1 µg/ml of purified RBD diluted in PBS overnight at 4C and blocked with Tris-Buffered Saline with 0.1% Tween 20 (TBST) containing 5% non-fat dry milk. DBS were eluted in 500 µl of TBST overnight at 4C and 50 µl of each sample was added to the ELISA plate preloaded with 50 µl of TBST containing 2% non-fat dry milk. Samples were then assayed for SARS-CoV-2 antibodies according to published protocols.[18,19] The RBD protein was produced in-house via transfection of HEK293T cells using polyethylenimine (Plasmid was a generous gift from Pr. F. Kramer mount Sinai School of Medicine). Batches were control for purity by SDS-PAGE followed by Coomassie staining and ELISA using an anti His-tag monoclonal antibody. Optical densities were read at 405 nm, and each 96-well plate contained seven negative controls and one positive control (serum from PCR-confirmed case at 1/100 dilution). Samples were tested against IgG and positive samples were confirmed and tested with anti IgM antibodies. Optical density values were normalized to the mean optical density of negative controls daily.

Data Analysis

Sample size and power

The study was designed to assess seropositivity within the primary sampling group with sufficient precision to inform policy. With 750 persons, and an assumed 5% positivity, the 95% CI for this estimate is 3.6% to 6.9%. Within the five emergency response subregions, at 5% seropositivity, the survey is powered for a precision of 2.3% to 10.2%. The secondary sampling group (n= 250) sample size was based on logistic limitations, but was powered to a precision of 2.8% to 8.8%. All confidence intervals are binomial exact, without adjustments for study design effects or non-response.

Analysis of serology data

Finite mixture models were used to determine seropositivity cutoffs. These latent-class models estimate "breakpoints" for seropositive and seronegative subpopulations, and have been applied to a range of pathogen serosurvey data, including rubella, pertussis, and parvovirus.[20–22] From this analysis, all samples with IgG optical density ratio ≥ 2.49-fold above daily background were considered positive for SARS-CoV-2.

Adjusted estimates

All reported prevalences and prevalence ratio estimates are adjusted with non-response weights, which were estimated using inverse weighting. Briefly, logistic regression models were used to calculate propensity scores for each individual in the sample using reported gender and race categories. These were transformed to probabilities; a small number of individuals had extremely large weights due to sparse strata; these weights were truncated at 1/0.02.[23] Weights were then used for all prevalence and prevalence ratio estimations using the *survey* package in R.[24] The primary sampling group sample was self-weighting due to probability-proportional to population size sampling. Sampling weights were not used in the secondary sampling group as selection used simple random sampling.

Multivariable analyses for prevalence ratios

Prevalence ratios were estimated to assess factors associated with seropositivity, with separate Poisson models[25] for both of the two sampling groups, with robust (sandwich) errors to address clustering within households.

Bivariate analyses were performed for each factor separately. All variables with a p-value < 0.20 based on bivariate association with outcome were further evaluated for inclusion in final models. All final models were adjusted for age (continuous), race, and gender (see Table 1). Due to several very sparse categories, some were combined in final models. Specifically, all race/ethnicity categories and all geographic regions were not included in analysis of the secondary sampling group due to unstable estimates.

Model parsimony was evaluated using AIC/BIC and all tests were two-tailed, with $\alpha = 0.05$. R version 4.0.3 and SAS version 9.4 (SAS Institute, Inc, Cary, NC, USA) were used for analysis.

Patient and public involvement

All members of the university community were invited to participate, and serological testing was proved at no cost to the sampled individuals and their selected household contact.

Results

A total of 1,001 individuals were enrolled into the study; this included 752 undergraduate students, 90 graduate students, 63 faculty/librarians, and 96 staff members (Figure 2). Seventy-six percent of these (n=762) returned blood samples for analysis; 548 in the primary sampling group, and 214 in the secondary sampling group. Of the 548 participants in the primary sampling group, 230 enrolled a household member. One household member submitted a sample without the sample of the main participant, bringing the total number of undergraduate household members to 231. Of the 214 participants in the secondary sampling group, 78 enrolled a household member. Two returned samples were excluded from analysis due to unlinkable samples. A total of 1,071 samples were included in the final analyses: 762 main participants and 309 household members (Figure 2).

Demographic characteristics of both sampling groups are presented in Table 1. Race categories do not total to 100% due to non-response and multiple possible answers. Age, gender, and essential worker status were broadly similar between those invited to participate and those who completed the study (Supplemental Table 1).

Of the total 1,071 samples tested, 46 were positive for SARS-CoV-2 antibodies. Demographic results are stratified by IgG serostatus (Table 2); no samples showed evidence for IgM positivity. Seropositivity was low-to-moderate across the survey groups, with several important exceptions. Variation is apparent by sex, race, and across geographic regions; however, several strata have wide confidence intervals due to small sample sizes.

Of the 779 primary sampling group participants and their household members, 36 were positive for SARS-CoV-2 antibodies. This corresponds to an overall seroprevalence of 5.3% (95% CI: 3.1 – 7.5) of the population after adjustment for nonresponse and geographic location. In the secondary sampling group, of the 292 graduate students, staff, librarians, faculty members, and their household members, ten (adjusted 4.0 %, 95% CI: 1.6 - 6.5) had evidence for prior SARS-CoV-2 infection. (Table 2). Results were also further stratified by UMass affiliate vs. household member. Of the 548 undergraduate students in the primary sampling group, 27 were positive for SARS-CoV-2 IgG antibodies (population positivity rate of 5.3% (95% CI: 3.1 – 7.6%). Of the 231 household members of undergraduate participants, nine (adjusted 5.2%, 95% CI: 1.2 - 9.2) were positive for SARS-CoV-2 IgG antibodies. In the secondary sampling group, eight University affiliates (adjusted 4.3 %, 95% CI: 1.3 - 7.3) were positive for SARS-CoV-2 IgG antibodies. Of the household members in the secondary sampling group, 2 were seropositive, with a weighted seroprevalence of (3.3 %, 95% CI: 0.0 – 7.8%). (Table 2). The overall distributions of measured IgG lognormal optical density ratios by subgroups are broadly similar (Supplemental Figure 1).

After adjustments for age, gender, region, and self-reported febrile illness since February 2020, the factor with the strongest association with seropositivity in the primary sampling group was Black or AIAN Race (PR = 4.49, 95% CI = 1.57,12.9) (Table 3). This indicates that individuals who report being Black or American Indian / Alaskan Native have a prevalence 3.49 times higher than White individuals after adjustment. Additionally, after adjustments, females and those who are gender diverse were at a significantly lower risk of prior SARS-CoV-2 infection (PR = 0.5; 95% CI = 0.27, 0.92) compared to males. Those who reported a febrile illness in February were more likely to be seropositive than those who were not sick (PR = 2.42, 95% CI = 1.24, 4.75). No significant associations were found across each of the 5 regions in the primary

sampling group, however the prevalence of seropositivity was 48% higher in Region 1 compared to Region 4 (PR = 1.48, 95% CI = 0.62, 3.52).

Within the secondary sampling population, after adjustments for age, race, gender, region, household member and self-reported febrile illness since February (Table 4), participants who reported residing in Region 2, 3, 4, or 5 had greater than 4 times higher prevalence of SARS-CoV-2 antibodies as compared to those who resided in Region 1 (PR = 4.08, 95% CI = 1.09, 15.33). No other factors included in the model were significantly associated with seropositivity.

Discussion

This mail-based serosurvey of two university-affiliated populations across Massachusetts in July and August 2020 found an estimated seroprevalence of ~5% of antibodies to SARS-CoV-2. These results indicate that even with extensive morbidity and mortality across the state at the time of sampling, there had been limited exposure to SARS-CoV-2 at a population-level. This estimated seroprevalence is lower than that detected with concurrent community-based studies in other states. An estimated 14.3% of the United States population had been previously infected with SARS-CoV-2 by November 2020, as estimated in a pooled analysis of multiple seroprevalence surveys.[26]

Our estimates are substantially lower than some models of COVID-19 seroprevalence in Massachusetts. One model estimates a seroprevalence of 16.2% (no CIs provided) on Jul 27, 2020 (closest modeled date). These estimates are nearly double our measured seroprevalence with inclusion of 110,000 confirmed cases at that date (ca. 1.5%)[27]. These differences might be caused by a number of reasons, including a non-representative population by age or geographic range, or waning of antibody titers. Without CIs, we are unable to evaluate coverage outside the reported point estimate. However, alternate nowcasting estimates suggest a total statewide attack rate on July 31, 2020 of 6.9% (95% CrI: 5.5 – 8.4%) in Massachusetts[28]. Our results are closely aligned with these estimates.

Within the surveyed groups, approximately 24% of the primary sampling group and 18% of the secondary sampling group reported illness since February. Febrile illness since February was

associated with an increased prevalence of seropositivity in both sampling groups, but after multivariable analysis this association was found in only the primary sampling group. These results reinforce results from other studies: asymptomatic illness is an important contributor to observed force of infection; and important limitations of testing availability at the time of survey.

Differing antibody dynamics have been reported in other studies. A number of studies have found sustained antibody levels for over 3 months,[29,30] while others have found IgG levels can remain 6 months or more.[31–33] An additional study has reported rapid waning of routine serological markers in individuals who had lower initial antibody responses.[34]. Only 7.1% of those with high titers at baseline seroreverted to a level below the threshold for positivity within 60 days, compared to 64.9% of those with lower titers at baseline.[34] Evidence for IgM seropositivity was not detected in any of IgG positive samples, which is consistent with results from other surveys studies that included asymptomatic or subclinical populations due to rapidly waning titers.[32,35] Studies have shown that IgM levels decline more rapidly after infection than IgA and IgG levels, [30,36,37], and this is especially apparent with asymptomatic and sub-clinical infections.[32,35]

Trends in the patterns of SARS-CoV-2 antibody levels vary greatly depending on the timing of sampling and severity of disease.[31,32,35] Seroconversion times vary depending on the study, but one study found a median time-to-seroconversion for IgM of 8 days and median seroconversion for IgG of 10 days. Additionally, the SARS-CoV-2 IgG response generally begin around 10-15 days after symptom onset.[2] For this reason, repeated serial sampling, repeated serial sampling of convalescent populations should be prioritized to more fully understand the dynamics of immune response.

SARS-CoV-2 seropositivity was associated with minority race status in this survey. While the total number of non-white participants was limited, the large effect size reinforces other studies suggesting that marginalized communities have been and continue to be excessively impacted by the pandemic. Results from the primary sampling group analysis suggest that self-reported Black race is a risk factor for previous SARS-CoV-2 infection, which is consistent with findings from other studies.[38–40] A number of factors may play a role in this significant association

including the fact that Black individuals are more likely to work in frontline industries or live in areas with a higher population density.[41] No parallel associations were found in the analysis of the secondary sampling group due to limited sample size in some strata. In the secondary sampling group, the aggregation of Race categories into White race and Non-white race likely obscured meaningful associations between Race and exposure.

Results from the primary sampling group showed higher prevalence of IgG seropositivity among males. After adjusting for age, race, and region, male gender was a significant risk factor for evidence of SARS-CoV-2 infection. Other studies have similarly found that males have higher rates of infection than females for asymptomatic infections[4243] These findings may reflect differences in care seeking behavior (biased recruitment), true biological differences, or differences in health behaviors such as smoking, alcohol use, and COVID-19 prevention measures.[44] This association was not observed in the secondary sampling group.

The primary sampling group also showed increased risk of seropositivity with self-reported illness since February 2020; this association was not observed in analyses of the secondary sampling group. This finding may indicate that some of the participants in our study were not strictly asymptomatic and were simply unable to obtain a COVID-19 test due to limited availability during the beginning of the pandemic.

This study was population-based and had broad eligibility criteria but is subject to several limitations. The exclusion of persons with confirmed diagnoses and any current symptoms (due to biosafety concerns) also inherently limited capture of sub-clinical infections. As such, the estimates are likely a lower bound. However, participants who suspected they may have been previously infected with SARS-CoV-2 might be more likely to participate compared to those that were not concerned with prior infection. This is a pervasive issue in community-based studies, where characteristics of those who volunteer to participate in community-based research differ from the general population.[45]

Randomization after a three-week enrollment period helped to address this limitation, as using only the first participants to volunteer could have biased the sample to include those who were most motivated to receive their antibody test results. If participants were more motivated to

receive their results because they thought they suspected a prior exposure to SARS-CoV2, this would have inflated the observed prevalence of seropositivity in the study population.

Another limitation of the study is the self-reported response of the lack of a prior COVID-19 diagnosis and current fever. It is possible that some participants shielded their answer and submitted samples for analysis without meeting the eligibility criteria; this would have inflated our estimation of seroprevalence in asymptomatic groups, Thirdly, the limited number of non-White, and gender-diverse participants also limited some analyses. Fourthly, while multiple studies have validated DBS sampling for SARS-CoV-2,[46] waning antibodies in asymptomatic individuals could be below the limited of detection of the ELISA assay. Finally, generalizability is limited due to the recruitment of a university-affiliated population in a relatively restricted geographic area.[8]

This serosurvey estimates prevalence of prior SARS-CoV-2 infections in a university-affiliated population in Massachusetts. Risk factors for IgG seropositivity included recent self-reported febrile illness, minority race status, and male gender. This study provides estimates of seroprevalence in Massachusetts after a 'first wave' of SARS-CoV-2 infections in the spring of 2020. Repeat seroprevalence studies in this population could provide estimates in changes of seropositivity rates given subsequent waves of SARS-CoV-2 transmission. This study reinforces the critical need for targeted serosurveys in highest-risk and marginalized communities, both in Massachusetts, and nationwide.

Tables

Table 1. Demographics of study populations, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Characteristic	Primary Sampling Group (N=779)	Secondary Sampling Group (N=292)
Gender		
Female	499 (64.1%)	154 (52.7%)
Male	274 (35.2%)	136 (46.6%)
Gender Diverse	5 (0.6%)	2 (0.7%)
Missing	1 (0.1%)	0 (0.0%)
Race		
AIAN	17 (2.2%)	1 (0.3%)
Asian	78 (10.0%)	34 (11.6%)
Black	12 (1.5%)	3 (1.0%)
Hispanic	36 (4.6%)	9 (3.1%)
Multiple	37 (4.8%)	11 (3.8%)
White	545 (70.0%)	217 (74.3%)
Missing	54 (6.9%)	17 (5.8%)
Age		
Mean	29.9	41.6
Median	21	39
Range	18 – 75	21 – 75
Education		
HS/GED	102 (13.1%)	5 (1.7%)
Some College	483 (62.0%)	24 (8.2%)
BA/BS	117 (15.0%)	78 (26.7%)
More Than BA/BS	74 (9.5%)	183 (62.7%)
Missing	3 (0.4%)	2 (0.7%)
Essential Worker		
No	533 (68.4%)	224 (76.7%)
Yes	195 (25.0%)	51 (17.5%)
Missing	51 (6.6%)	17 (5.8%)
Self-reported attitude about COVID-19		
Strongest fear	135 (17.3%)	72 (24.7%)
Somewhat fearful	389 (49.9%)	122 (41.8%)
Neutral/Missing	139 (17.8%)	63 (21.6%)
Somewhat not fearful	86 (11.0%)	23 (7.9%)
Not fearful	30 (3.9%)	12 (4.1%)
Self-reported febrile illness since February	No: 534 (68.6%) Yes: 188 (24.1%) Missing: 57 (7.3%)	No: 224 (76.7%) Yes: 53 (18.2%) Missing: 15 (5.1%)
Self-reported care	No: 112 (59.6%)	No: 32 (60.4%)

Seeking (if illness since Feb)	Yes: 75 (39.9%) Missing: 1 (0.5%)	Yes: 21 (39.6%) Missing: 0 (0.0%)

*Notes: AIAN= American Indian/Alaska Native. The primary sampling group includes UMass undergraduates and household members, and the secondary sampling group includes UMass affiliated faculty, staff, and graduate students and household members.

Table 2. Weighted seropositivity by main demographic variables, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Characteristic	**Primary Sampling Group**	**Secondary Sampling Group**
Age in years, median	21 95% CI (20 - 21)	41 95% CI (38 - 44)
Prevalence of SARS-CoV-2 Antibodies, by sub-group		
Overall Population Prevalence (95% CI)	5.3% (3.5 - 8.0)	4.0% (2.2 - 7.4)
Sex %, (95% CI)		
Female	4.0% (2.4 -6.6)	4.9% (2.2 - 10.7)
Male	8.7% (5.1 - 15.0)	3.0% (1.1 - 8.6)
Gender Diverse/No Response	0.0	0.0
Race * % (95% CI)		
Primary:		
White	3.9 (2.6 - 5.9)	-
Multiple	6.3 (1.7 - 23.7)	-
Asian	6.2 (2.7 - 14.5)	-
Missing	1.9 (0.3 - 13.5)	-
Hispanic	5.4 (1.4 - 21.0)	-
Black/AIAN	21.0 (5.8 - 76.4)	-
Secondary:		
White	-	4.2 (2.2 - 7.9)
Non-White	-	1.6 (0.2 - 11.7)
Essential worker status		
Yes	4.2 (2.0 - 8.8)	7.1 (2.3 - 21.3)
No	5.8 (3.5 - 9.7)	3.6 (1.7 - 7.6)
Missing Response	3.5 (0.9 - 14.2)	0

Table 2 cont. Weighted seropositivity by main demographic variables, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Participant type		
University-affiliate	0.053 (0.035 - 0.081)	4.3 (2.1 - 8.6)
Household member	0.051 (0.024 - 0.112)	3.3 (0.8 -13.0)
State Emergency Response Region (Figure 1)		
Region 1	7.8 (3.9 -15.6)	-
Region 2	1.6 (0.2 - 11.3)	-
Region 3	3.2 (1.0 - 10.7)	-
Region 4	5.7 (3.0 - 10.8)	-
Region 5	6.2 (2.3 - 16.5)	-
Region 1	-	3.1 (1.4 - 6.5)
Regions 2/3/4/5	-	11.3 (4.1 - 31.3)

*Notes: AIAN= American Indian/Alaska Native. All proportions are adjusted for non-response.

Table 3. Multivariable associations for SARS-CoV-2 seropositivity, Primary Sampling Group, MA USA, Jul-Aug 2020.

Characteristic	Prevalence Ratio	95% CI	p-value
Emergency Response Region			
Region 1	1.48	0.62, 3.52	0.38
Region 2	0.34	0.05, 2.45	0.28
Region 3	0.53	0.14, 1.96	0.34
Region 4	Reference		
Region 5	1.02	0.35, 2.98	0.97
Age (years)	1.04	0.96, 1.12	0.33
Gender			
Male	Reference		
Female, Gender diverse, or No response	0.50	0.27, 0.92	**0.027**
Race			
White	Reference		
Multiple	1.91	0.46, 7.98	0.38
Asian	1.66	0.66, 4.16	0.28
Missing Race	0.51	0.07, 3.71	0.51
Hispanic	1.76	0.44, 7.04	0.42
Black or AIAN	4.49	1.57, 12.9	**0.005**
Febrile illness since February			
No	Reference		
Yes	2.42	1.24, 4.75	**0.010**
Missing Response	0.33	0.04, 2.45	0.28
Other household member			
no	Reference		
yes	0.29	0.01, 7.69	0.46

Table 4. Multivariable associations for SARS-CoV-2 seropositivity, Secondary Sampling Group, Massachusetts, USA, Jul-Aug 2020.

Characteristic	Prevalence Ratio	95% CI	p-value
Age (years)	1.03	0.99, 1.07	0.17
Gender			
Male	Reference	-	-
Female or Gender Diverse	1.35	0.43, 4.31	0.61
Race			
White	Reference	-	-
All Other/Multiple/Missing	0.58	0.07, 4.86	0.62
Febrile illness since February			
No	Reference	-	-
Yes	2.56	0.68, 9.67	0.17
Missing Response	2.35	0.21, 26.73	0.49
Emergency Response Region			
Region 1	Reference	-	-
Regions 2/3/4/5	4.08	1.09, 15.33	**0.039**
Other household member			
No	Reference	-	-
Yes	0.70	0.18, 2.72	0.61

Acknowledgements

We would like to express our appreciation to all of the individuals who have enabled the completion of this study. Rob Leveille and Charlie Apicella of the UMass Mail and Distribution Services have facilitated the label printing, outgoing shipments, and incoming shipments for over 1,000 sample boxes. Sujitha Chandra Kumar, Vincent Lee, and Pratik Patel have been valuable members of the REDCap support team. Kimberly Tremblay, Jesse Mager, Katherine Dorfman, Pa Tamba Ngom, and Ryan Kurtz have graciously allowed us to utilize laboratory space for box assembly and sample processing. These individuals have worked tirelessly to support this study despite massive pandemic-related logistical challenges and tight deadlines. We are very grateful for their support.

Figures

Fig 1. Emergency Response Region sampling frames, for SARS-CoV-2 seropositivity, Massachusetts, USA, Jul-Aug 2020.

Fig 2. Participant inclusion (CONSORT) enrollment, SARS-CoV-2 serosurvey, Massachusetts, USA, Jul-Aug 2020.

Author contributions

AAL and DA led study conceptualization and design.
LB and NR informed overall aims, sampling design, and statistical analysis.
JR, TS, and EYC implemented all field implementation and associated data collection.
DA organized, performed and reported all laboratory-based testing.
JR, TS, EYC, and AAL performed data cleaning and analyses, and wrote the first draft.
All authors contributed to revisions of the final manuscript.
All authors read and approved the final manuscript.

COI statements

None of the authors report any conflicts of interest, financial or non-financial, regarding this study.

Funding

UMass Faculty Funds (A Lover; SPH-AL-001); UMass Faculty Funds (N Reich; SPH-NR-001); UMass Institute for Applied Life Science "Midigrant" (#169076; A. Lover); and D. Alfandari was supported by grants from the NIH/USPHS (RO1DE016289 and R24OD021485).

Data sharing

Fully-identified data will be deposited at: https://osf.io/437tg/.

References

1. Lemieux JE, Siddle KJ, Shaw BM, et al. Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. *Science*. Published online December 10, 2020. doi:10.1126/science.abe3261

2. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. *Nature Communications*. 2020;11(1):4507. doi:10.1038/s41467-020-18272-4

3. *Massachusetts Department of Public Health COVID-19 Dashboard*. Department of Public Health, Massachusetts; 2020. Accessed October 25, 2020. https://www.mass.gov/doc/covid-19-dashboard-july-1-2020/download

4. Bryant JE, Azman AS, Ferrari MJ, et al. Serology for SARS-CoV-2: Apprehensions, opportunities, and the path forward. *Science Immunology*. 2020;5(47). doi:10.1126/sciimmunol.abc6347

5. Gronvall G, Connell N, Kobokovich A, et al. *Developing a National Strategy for Serology (Antibody Testing) in the United States*. The Johns Hopkins Center for Health Security; 2020. https://www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/200422-national-strategy-serology.pdf

6. Naranbhai V, Chang CC, Beltran WFG, et al. High Seroprevalence of Anti-SARS-CoV-2 Antibodies in Chelsea, Massachusetts. *The Journal of Infectious Diseases*. Published online September 9, 2020:jiaa579. doi:10.1093/infdis/jiaa579

7. Reifer J, Hayum N, Heszkel B, Klagsbald I, Streva VA. SARS-CoV-2 IgG antibody responses in New York City. *Diagn Microbiol Infect Dis*. Published online July 21, 2020. doi:10.1016/j.diagmicrobio.2020.115128

8. Tilley K, Ayvazyan V, Martinez L, et al. A Cross-Sectional Study Examining the Seroprevalence of Severe Acute Respiratory Syndrome Coronavirus 2 Antibodies in a University Student Population. *Journal of Adolescent Health*. Published online October 2020:S1054139X2030522X. doi:10.1016/j.jadohealth.2020.09.001

9. 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*. Published online July 21, 2020. doi:10.1001/jamainternmed.2020.4130

10. Lim MD. Dried Blood Spots for Global Health Diagnostics and Surveillance: Opportunities and Challenges. *The American Journal of Tropical Medicine and Hygiene*. 2018;99(2):256-265. doi:10.4269/ajtmh.17-0889

11. Corran PH, Cook J, Lynch C, et al. Dried blood spots as a source of anti-malarial antibodies for epidemiological studies. *Malar J*. 2008;7(1):195. doi:10.1186/1475-2875-7-195

12. McDade TW, McNally EM, Zelikovich AS, et al. High seroprevalence for SARS-CoV-2 among household members of essential workers detected using a dried blood spot assay. Khudyakov YE, ed. *PLoS ONE*. 2020;15(8):e0237833. doi:10.1371/journal.pone.0237833

13. Karp DG, Danh K, Seftel D, Robinson P, Tsai C. *A Serological Assay to Detect SARS-CoV-2 Antibodies in at-Home Collected Finger-Prick Dried Blood Spots*. Allergy and Immunology; 2020. doi:10.1101/2020.05.29.20116004

14. Thevis M, Knoop A, Schaefer MS, et al. Can dried blood spots (DBS) contribute to conducting comprehensive SARS-CoV-2 antibody tests? *Drug Test Anal*. 2020;12(7):994-997. doi:10.1002/dta.2816

15. Valentine-Graves M, Hall E, Guest JL, et al. At-home self-collection of saliva, oropharyngeal swabs and dried blood spots for SARS-CoV-2 diagnosis and serology: Post-collection acceptability of specimen collection process and patient confidence in specimens. Kumar S, ed. *PLoS ONE*. 2020;15(8):e0236775. doi:10.1371/journal.pone.0236775

16. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*. 2019;95:103208. doi:10.1016/j.jbi.2019.103208

17. map-bt-regions-by-coalitions.pdf | Mass.gov. Accessed January 10, 2021. https://www.mass.gov/doc/map-bt-regions-by-coalitionspdf

18. Amanat F, Stadlbauer D, Strohmeier S, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Medicine*. 2020;26(7):1033-1036. doi:10.1038/s41591-020-0913-5

19. Stadlbauer D, Amanat F, Chromikova V, et al. 1 A detailed protocol for a serological assay to detect SARS-CoV-2 2 seroconversion in humans: antigen production and test setup. :13.

20. Hardelid P, Williams D, Dezateux C, et al. Analysis of rubella antibody distribution from newborn dried blood spots using finite mixture models. *Epidemiology & Infection*. 2008;136(12):1698-1706.

21. Baughman AL, Bisgard KM, Lynn F, Meade BD. Mixture model analysis for establishing a diagnostic cut-off point for pertussis antibody levels. *Stat Med*. 2006;25(17):2994-3010. doi:10.1002/sim.2442

22. Nj G. Analysis of serological surveys using mixture models: application to a survey of parvovirus B19. *Stat Med*. 1996;15(14):1567-1573. doi:10.1002/(sici)1097-0258(19960730)15:14<1567::aid-sim289>3.0.co;2-g

23. Moore KL, Neugebauer R, Laan MJ van der, Tager IB. Causal inference in epidemiological studies with strong confounding. *Statistics in Medicine*. 2012;31(13):1380-1404. doi:https://doi.org/10.1002/sim.4469

24. Lumley T. *Survey: Analysis of Complex Survey Samples*.; 2020. Accessed January 25, 2021. https://CRAN.R-project.org/package=survey

25. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio | BMC Medical Research Methodology | Full Text. Accessed January 10, 2021. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-3-21

26. Angulo FJ, Finelli L, Swerdlow DL. Estimation of US SARS-CoV-2 Infections, Symptomatic Infections, Hospitalizations, and Deaths Using Seroprevalence Surveys. *JAMA Netw Open*. 2021;4(1):e2033706. doi:10.1001/jamanetworkopen.2020.33706

27. Archive of COVID-19 cases in Massachusetts | Mass.gov. Accessed March 5, 2021. https://www.mass.gov/info-details/archive-of-covid-19-cases-in-massachusetts

28. Attack Rate 08-13 | Boni Lab. Accessed March 5, 2021. https://mol.ax/covid/attack-rate-08-13/

29. Crawford KHD, Dingens AS, Eguia R, et al. Dynamics of Neutralizing Antibody Titers in the Months After Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *The Journal of Infectious Diseases*. 2021;223(2):197-205. doi:10.1093/infdis/jiaa618

30. Isho B, Abe KT, Zuo M, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Science Immunology*. 2020;5(52). doi:10.1126/sciimmunol.abe5511

31. Figueiredo-Campos P, Blankenhaus B, Mota C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in COVID-19 patients and healthy volunteers up to six months post disease onset. *Eur J Immunol*. Published online October 21, 2020:eji.202048970. doi:10.1002/eji.202048970

32. Wang Y, Zhang L, Sang L, et al. Kinetics of viral load and antibody response in relation to COVID-19 severity. *Journal of Clinical Investigation*. 2020;130(10):5235-5244. doi:10.1172/JCI138759

33. Wu J, Liang B, Chen C, et al. SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. *medRxiv*. Published online July 24, 2020:2020.07.21.20159178. doi:10.1101/2020.07.21.20159178

34. Self WH. Decline in SARS-CoV-2 Antibodies After Mild Infection Among Frontline Health Care Personnel in a Multistate Hospital Network — 12 States, April–August 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6947a2

35. Lynch KL, Whitman JD, Lacanienta NP, et al. *Magnitude and Kinetics of Anti-SARS-CoV-2 Antibody Responses and Their Relationship to Disease Severity*. Infectious Diseases (except HIV/AIDS); 2020. doi:10.1101/2020.06.03.20121525

36. Yao X-Y, Liu W, Li Z-Y, et al. *Neutralizing and Binding Antibody Kinetics of COVID-19 Patients during Hospital and Convalescent Phases*. Infectious Diseases (except HIV/AIDS); 2020. doi:10.1101/2020.07.18.20156810

37. Beaudoin-Bussières G, Laumaea A, Anand SP, et al. Decline of Humoral Responses against SARS-CoV-2 Spike in Convalescent Individuals. *mBio*. 2020;11(5). doi:10.1128/mBio.02590-20

38. Jehi L, Ji X, Milinovich A, et al. Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing: Results From 11,672 Patients. *Chest*. 2020;158(4):1364-1375. doi:10.1016/j.chest.2020.05.580

39. Rentsch CT, Kidwai-Khan F, Tate JP, et al. Covid-19 by Race and Ethnicity: A National Cohort Study of 6 Million United States Veterans. *medRxiv*. Published online May 18, 2020. doi:10.1101/2020.05.12.20099135

40. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Network Open*. 2020;3(9):e2021892-e2021892. doi:10.1001/jamanetworkopen.2020.21892

41. Black workers face two of the most lethal preexisting conditions for coronavirus—racism and economic inequality. Economic Policy Institute. Accessed February 20, 2021. https://www.epi.org/publication/black-workers-covid/

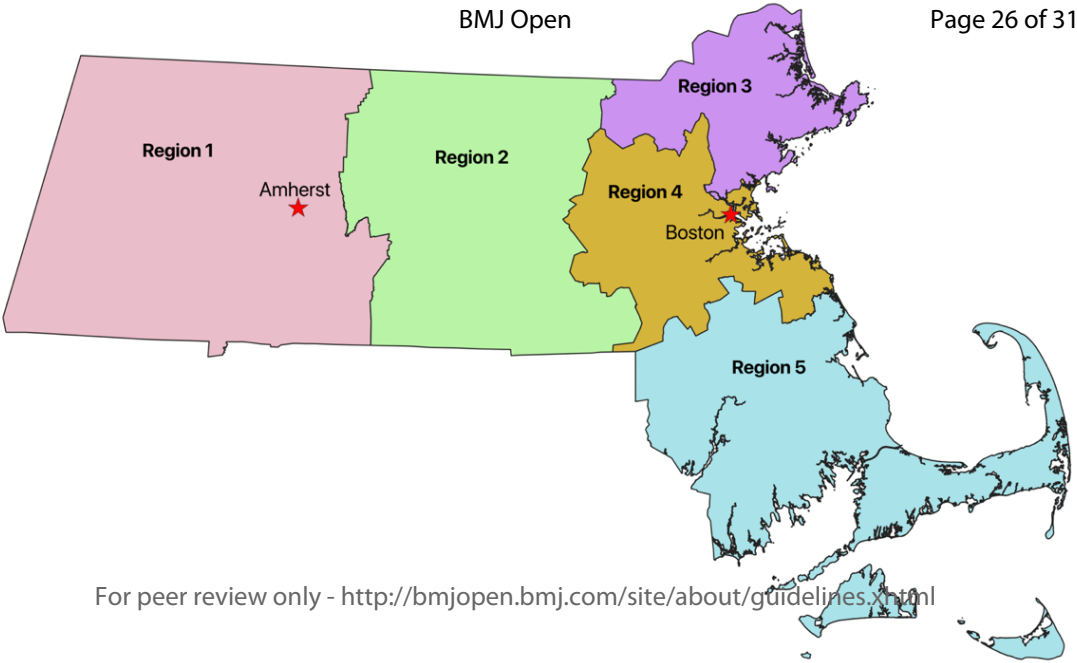
42. Al-Sadeq DW, Nasrallah GK. The incidence of the novel coronavirus SARS-CoV-2 among asymptomatic patients: A systematic review. *International Journal of Infectious Diseases*. 2020;98:372-380. doi:10.1016/j.ijid.2020.06.098

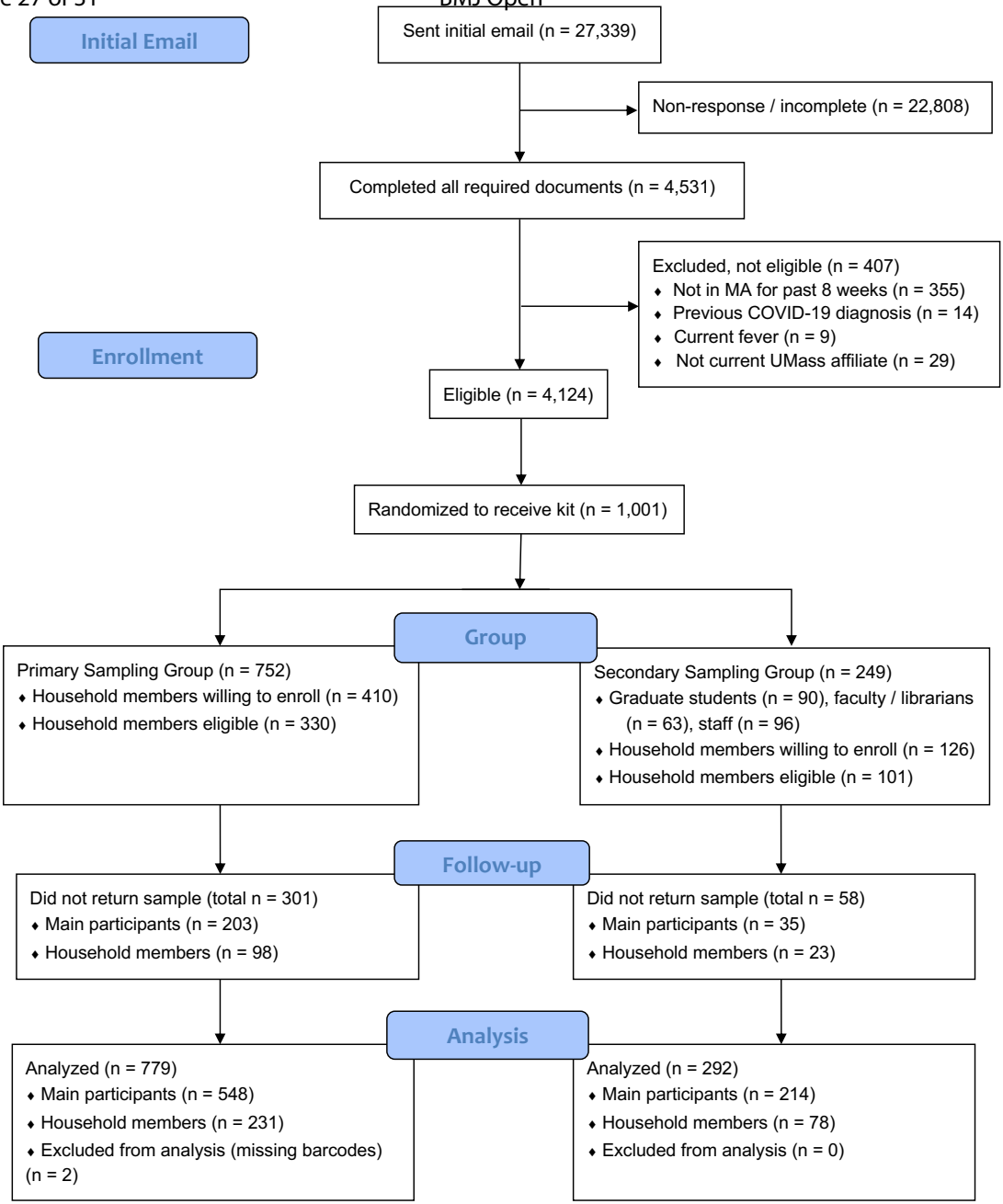
43. Barek MdA, Aziz MdA, Islam MS. Impact of age, sex, comorbidities and clinical symptoms on the severity of COVID-19 cases: A meta-analysis with 55 studies and 10014 cases. *Heliyon*. 2020;6(12):e05684. doi:10.1016/j.heliyon.2020.e05684

44. Bwire GM. Coronavirus: Why Men are More Vulnerable to Covid-19 Than Women? *SN Compr Clin Med*. Published online June 4, 2020:1-3. doi:10.1007/s42399-020-00341-w

45. Ganguli M, Lytle ME, Reynolds MD, Dodge HH. Random Versus Volunteer Selection for a Community-Based Study. *The Journals of Gerontology: Series A*. 1998;53A(1):M39-M46. doi:10.1093/gerona/53A.1.M39

46. Krishnamurthy HK, Jayaraman V, Krishna K, et al. Antibody profiling and prevalence in US patients during the SARS-CoV2 pandemic. *PLOS ONE*. 2020;15(11):e0242655. doi:10.1371/journal.pone.0242655




Initial Email
Sent initial email (n = 27,339)
Non-response / incomplete (n = 22,808)
Completed all required documents (n = 4,531)
Excluded, not eligible (n = 407)
Not in MA for past 8 weeks (n = 355)
Previous COVID-19 diagnosis (n = 14)
Current fever (n = 9)
Not current UMass affiliate (n = 29)
Enrollment
Eligible (n = 4,124)
Randomized to receive kit (n = 1,001)
Group
Primary Sampling Group (n = 752)
Household members willing to enroll (n = 410)
Household members eligible (n = 330)
Secondary Sampling Group (n = 249)
Graduate students (n = 90), faculty / librarians (n = 63), staff (n = 96)
Household members willing to enroll (n = 126)
Household members eligible (n = 101)
Follow-up
Did not return sample (total n = 301)
Main participants (n = 203)
Household members (n = 98)
Did not return sample (total n = 58)
Main participants (n = 35)
Household members (n = 23)
Analysis
Analyzed (n = 779)
Main participants (n = 548)
Household members (n = 231)
Excluded from analysis (missing barcodes) (n = 2)
Analyzed (n = 292)
Main participants (n = 214)
Household members (n = 78)
Excluded from analysis (n = 0)

Supplemental Information

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020: a mail-based cross-sectional study

* Correspondence to: alover@umass.edu

School of Public Health and Health Sciences

University of Massachusetts- Amherst

715 North Pleasant Street

Amherst, MA 01003-9304

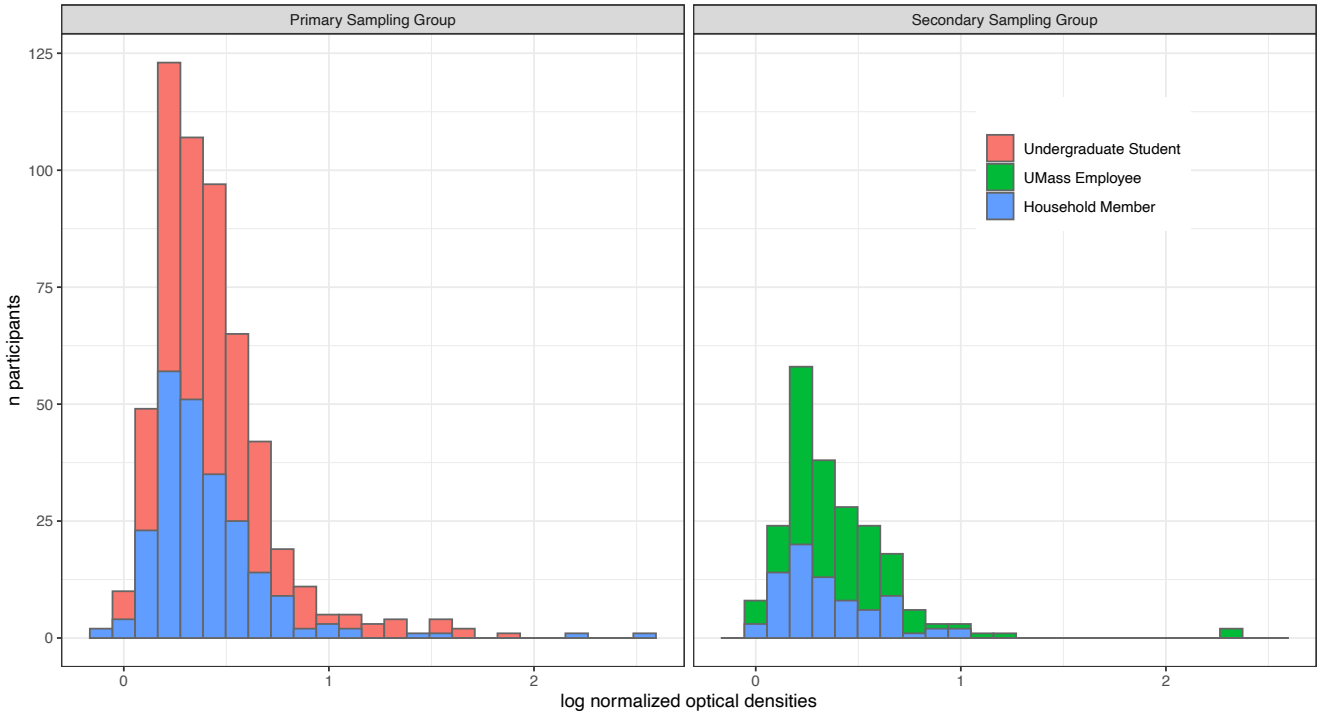
phone: 413-545-7426

Supplemental Information

Supplemental Table 1: Comparison of characteristics of participants selected for randomization and those not selected for randomization, SARS-CoV-2 serosurvey, Massachusetts, USA, Jul-Aug 2020.

	Primary Sampling Group (undergraduates)		Secondary sampling group (employees)	
	Not Randomized	Randomized	Not Randomized	Randomized
n	869	752	2253	249
Age (mean (SD))	20.64 (3.09)	20.18 (1.79)	42.05 (13.53)	40.69 (13.88)
Gender (%)				
Unknown	0 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)
Female	525 (60.4)	453 (60.2)	1262 (56.0)	147 (59.0)
Gender diverse	12 (1.4)	8 (1.1)	49 (2.2)	3 (1.2)
Male	332 (38.2)	291 (38.7)	941 (41.8)	99 (39.8)
Self-reported febrile illness since February (%)				
Unknown	1 (0.1)	2 (0.3)	6 (0.3)	0 (0.0)
No	568 (65.4)	499 (66.4)	1658 (73.6)	189 (75.9)
Not sure	75 (8.6)	57 (7.6)	154 (6.8)	16 (6.4)
Yes	225 (25.9)	194 (25.8)	435 (19.3)	44 (17.7)
Education (%)				
Unknown	4 (0.5)	3 (0.4)	6 (0.3)	1 (0.4)
BA/BS	15 (1.7)	14 (1.9)	509 (22.6)	64 (25.7)
High school / GED	90 (10.4)	108 (14.4)	46 (2.0)	2 (0.8)
More than BA/BS	2 (0.2)	0 (0.0)	1528 (67.8)	166 (66.7)
Prefer not to answer	1 (0.1)	0 (0.0)	13 (0.6)	1 (0.4)
Some college	757 (87.1)	627 (83.4)	151 (6.7)	15 (6.0)
Response to, "I am afraid of COVID-19" (%)				
Unknown	1 (0.1)	2 (0.3)	2 (0.1)	0 (0.0)
Neither agree nor disagree	199 (22.9)	141 (18.8)	399 (17.7)	51 (20.5)
Somewhat agree	390 (44.9)	374 (49.7)	1051 (46.6)	100 (40.2)
Somewhat disagree	115 (13.2)	99 (13.2)	153 (6.8)	22 (8.8)
Strongly agree	121 (13.9)	104 (13.8)	573 (25.4)	66 (26.5)
Strongly disagree	43 (4.9)	32 (4.3)	75 (3.3)	10 (4.0)

Supplemental Figure 1. Distributions of IgG log normal optical densities (OD) ratios, by subgroups, SARS-CoV-2, Massachusetts, USA, Jul-Aug 2020.



STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

Sero-MAss]	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5-6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	15 (Table 1)
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	7; 21 (Fig 2)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-9
		(*b*) Describe any methods used to examine subgroups and interactions	8
		(*c*) Explain how missing data were addressed	7-8
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	8
		(*e*) Describe any sensitivity analyses	n/a
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7; 21 (Fig 2)
		(b) Give reasons for non-participation at each stage	9
		(c) Consider use of a flow diagram	Fig. 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Fig 1.
		(b) Indicate number of participants with missing data for each variable of interest	Supp.

Outcome data	15*	Report numbers of outcome events or summary measures	16 (Table 2)
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-11
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10
Discussion			
Key results	18	Summarise key results with reference to study objectives	11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-14
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	3

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020- a mail-based cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051157.R1
Article Type:	Original research
Date Submitted by the Author:	23-Jul-2021
Complete List of Authors:	Snyder, Teah; University of Massachusetts Amherst, Biostatistics and Epidemiology Ravenhurst, Johanna; University of Massachusetts Amherst, Biostatistics and Epidemiology Cramer, Estee; University of Massachusetts Amherst, Biostatistics and Epidemiology Reich, Nicholas; University of Massachusetts Amherst, Biostatistics and Epidemiology Balzer, Laura; University of Massachusetts Amherst, Biostatistics and Epidemiology Alfandari, Dominique; University of Massachusetts Amherst, Veterinary and Animal Sciences Lover, Andrew; University of Massachusetts Amherst, Biostatistics and Epidemiology, School of Public Health and Health Sciences
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology
Keywords:	EPIDEMIOLOGY, Public health < INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020: a mail-based cross-sectional study

Teah Snyder,[1†] Johanna Ravenhurst,[1†] Estee Y. Cramer,[1†] Nicholas G. Reich,[1] Laura B. Balzer,[1] Dominique Alfandari,[2] and Andrew A. Lover[1*]

† Equal contributions.

[1] Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts- Amherst, Amherst MA

[2] Department of Veterinary and Animal Sciences, College of Natural Sciences, University of Massachusetts- Amherst, Amherst MA

* Correspondence to: alover@umass.edu

School of Public Health and Health Sciences
University of Massachusetts- Amherst
715 North Pleasant Street
Amherst, MA 01003-9304
phone: 413-545-7426

Abstract

Objectives: To estimate the seroprevalence of anti-SARS-CoV-2 IgG and IgM among Massachusetts residents and to better understand asymptomatic SARS-CoV-2 transmission during the summer of 2020.

Design: Mail-based cross-sectional study

Setting: Massachusetts, United States

Participants: Primary sampling group: sample of undergraduate students at the University of Massachusetts, Amherst (n = 548) and a member of their household (n = 231). Secondary sampling group: sample of graduate students, faculty, librarians and staff (n = 214) and one member of their household (n = 78). All participants were residents of Massachusetts without prior COVID-19 diagnosis.

Primary and secondary outcome measures: Prevalence of SARS-CoV-2 seropositivity. Association of seroprevalence with variables including age, gender, race, geographic region, occupation, and symptoms.

Results: Approximately 27,000 persons were invited via email to assess eligibility. 1,001 households were mailed dried blood spot sample kits, 762 returned blood samples for analysis. In the primary sample group, 36 (4.6%) had IgG antibodies detected for an estimated weighted prevalence in this population of 5.3% (95% CI: 3.5 to 8.0). In the secondary sampling group, 10 (3.4%) had IgG antibodies detected for an estimated adjusted prevalence of 4.0% (95% CI: 2.2 to 7.4). No samples were IgM positive. No association was found in either group between seropositivity and self-reported work duties or customer-facing hours. In the primary sampling group, self-reported febrile illness since Feb 2020, male sex, and minority race (Black or American Indian/Alaskan Native) were associated with seropositivity. No factors except geographic regions within the state were associated with evidence of prior SARS-CoV-2 infection in the secondary sampling group.

Conclusions: This study fills a critical gap in estimating the levels of sub-clinical and asymptomatic infection. Estimates can be used to calibrate models estimating levels of population immunity over time, and these data are critical for informing public health interventions and policy.

Strengths and limitations of this study

- Our study collected serological samples in a well-defined rigorous sample frame in a contact-less (mail-based) survey in an early stage of the pandemic in an area of high SARS-CoV-2 burden.
- A range of potentially associated demographic, occupational, and behavioral factors were surveyed to contextualize seropositivity across geographic regions within the state.
- Our study sampled from populations affiliated with a large public university in Massachusetts, and may not be generalizable to the general population.

Introduction

Since emergence in early 2020, the SARS-CoV-2 virus has severely impacted the entire globe. The state of Massachusetts was heavily impacted in the earliest stages of the pandemic, and a "super-spreader" event in the state in April 2020 may have seeded large case clusters throughout the country.[1] However, the trajectory of the early stages of transmission in the state, as well as across the US remain poorly understood due to changes in case definitions and limited testing of both symptomatic and asymptomatic persons during the summer of 2020.[2] To assess seroprevalence across the state, a mail-based serosurvey was implemented July-August 2020. At the time of this survey, the Massachusetts Department of Public Health had reported over 109,143 confirmed COVID-19 cases and over 8,081 deaths.[3] Sero-epidemiological studies are a critical tool to explore infection dynamics, especially where asymptomatic or subclinical infections are common, as for SARS-CoV-2.[4] This study helps to fill a critical gap in estimating the levels of sub-clinical and asymptomatic infection to inform consequent levels of population-immunity.[5]

Concurrent to this study, a number of seroprevalence studies were conducted on the east coast of the United States; these studies focused on specific populations at high risk and found varying results. A survey in April 2020 in a convenience sample of 200 asymptomatic residents of Chelsea, MA found an estimated of seroprevalence of 31.5% (17.5% IgM+/IgG+, 9.0% IgM+/IgG- and 5.0% IgM-/IgG+).[6] This study used a small convenience sample and did not include any randomization.[6] A study with a larger sample of over 28,000 clinical patient samples in New York City, USA found an IgG seropositivity prevalence of 44% with over 50% of participants reporting no symptoms.[7]

Other seroprevalence surveys across the US have found generally low- to moderate prevalence in a diverse set of study populations. A study of 790 university students in Los Angeles, California conducted in April and May of 2020 estimated a prevalence of SARS-CoV-2 IgG antibody of 4.0% (95% CI: 3.0 to 5.1%).[8] During May – April of 2020, a cross-sectional study in St. Louis found IgG seropositivity to be estimated at 1.71% (95% CI 0.04% to 3.38%) in pediatric patients and 3.11% (95% CI: 0.92% to 5.32%) in adult patients. In the most comprehensive serosurvey from the spring and summer months of 2020, 16,025 clinical samples were analyzed with IgG

spike protein sero-reactivity ranging from 1.0% in the San Francisco Bay Area to 6.9% in New York City.[9] These disparate results highlight major geographic variability in the trajectory of infections, and reinforce the need for additional seroprevalence studies to more fully contextualize trends in immunity to SARS-CoV-2 targeting specific geographic regions.

Though community seroprevalence studies generally rely on serum samples collected in health facilities, the use of dried blood spot (DBS) samples is a practical and effective alternative.[10] DBS samples involve a small finger-prick sample self-collected by participants in their own homes. The use of dried blood samples for antibody assays has been validated in other work prior to the current pandemic,[10,11] and previous studies have evaluated the feasibility, validity, and acceptability of using DBS samples for SARS-CoV-2 antibody testing.[12–15] This method of sample collection facilitates efficient population-level sampling while minimizing social mixing and concurrent potential exposures.

This study estimated the prevalence of previous infection with SARS-CoV-2 in individuals who had not been diagnosed with COVID-19 and were asymptomatic with representative coverage across the entire state of Massachusetts, USA. Information from this study can provide knowledge regarding the seropositivity of this population and can be used to inform decision-making regarding community re-openings during the pandemic.

Methods

Study Design and Participants

The study population included undergraduate students, graduate students, staff, and faculty members currently affiliated with the University of Massachusetts, Amherst (UMass) and their household members. On-campus classes were suspended in mid-March 2020; consequently, undergraduates had exposure to the local epidemiology within their communities from March until sampling in July-August throughout the state (primary sampling group). Conversely, graduate students, faculty, staff, librarians and their family members (secondary sampling group) almost universally reside in close proximity to Amherst, and broadly reflect transmission in the Western part of the state. UMass affiliates were eligible to participate in this study if they were above the age of 18, had been living in Massachusetts for the past eight weeks; had never

received a COVID-19 diagnosis from a medical professional; and did not have a fever greater than 100.4° F at the time of survey completion. Household members were eligible for inclusion if they met all of these same criteria and were between the ages of 23 and 78 (chosen to expand sampling beyond college-age population groups). Both UMass affiliates and their household members had to complete online consent forms in order to participate in the study. Upon meeting eligibility criteria, participants were directed to a consent form which they reviewed prior to providing their first and last name, the date, and an electronic signature.

An institutional email list was provided by university administration for recruitment. Initial emails were sent out to UMass affiliates between June 23, 2020 and June 26, 2020 for participant recruitment. The email provided information about the study and links to a screening eligibility survey, informed consent document, initial survey regarding COVID-19 risk factors, and information regarding shipping addresses. If the UMass affiliate had a household member interested in participating, a single household member was invited to complete the eligibility, consent, and initial survey forms. The household member was invited to participate prior to analyzing samples from the main participant. To increase participation rates, two reminder emails were sent to all non-respondents (day three and six after initial solicitation). All survey responses were collected and stored in REDCap.[16]

The survey was closed after a three-week enrollment period, and a subset of participants were selected to receive a test kit. To select a population representative of the broader UMass community across the entire state of Massachusetts, two sampling schemes were applied. The first consisted of all undergraduates and their household members (primary sampling group); the second sampling frame consisted of graduate students, staff, faculty members, librarians, and their household members (secondary sampling group). Within the primary sampling group, selection for biosample collection used probability proportional to population size, using the most recent census data aggregated to state-level emergency response regions due to sparse county-level populations (Figure 1).[17] For the secondary sampling group, selection for biosampling was via simple random sampling.

The full sample frame selection is shown in Figure 2. Briefly, an email invitation was sent to a total of 27,339 individuals, of which 4,124 completed the screening, informed consent, and initial survey forms. A total of 1,001 individuals were then randomly selected to receive a sampling kit. Participants were mailed all materials to safely collect and return samples, including lancets, alcohol wipes, gauze, gloves, bandages, a bloodspot collection card, a pre-paid shipping box, and detailed printed instruction cards (including a nurse call line).

Upon mailing out the test kits, participants were also emailed a link containing a video on how to collect the DBS, along with a detailed survey form with demographics, risk factors, and any current symptoms or COVID diagnoses. No participants reported a COVID-19 diagnosis between the initial survey and sample collection several weeks later. All shipments utilized a Biological Substance Category B (UN3373) shipping box.

Ethical approvals

This study was approved by the University of Massachusetts-Amherst Human Research Protection Office (Approval #2062; April 27, 2020).

Sample Preparation and ELISA Analysis

Upon receipt of boxes, the sample cards (Whatman® Protein Saver 903) were heat-treated (30 minutes at 56° C); a single blood spot per card was punched (0.25-inch diameter) and transferred to an ELISA plate. Plates were coated with 1 µg/ml of purified RBD diluted in PBS overnight at 4° C and blocked with tris-buffered saline with 0.1% Tween 20 (TBST) containing 5% non-fat dry milk. DBS were eluted in 500 µl of TBST overnight at 4° C and 50 µl of each sample was added to the ELISA plate preloaded with 50 µl of TBST containing 2% non-fat dry milk. Samples were then assayed for SARS-CoV-2 antibodies according to published protocols.[18,19] The RBD protein was produced in-house via transfection of HEK293T cells using polyethylenimine (plasmid was a generous gift from Pr. F. Kramer Mount Sinai School of Medicine). Batches were controlled for purity by SDS-PAGE followed by Coomassie staining and ELISA using an anti His-tag monoclonal antibody. Optical densities were read at 405 nm, and each 96-well plate contained seven negative controls and one positive control (serum from PCR-confirmed case at 1/100 dilution). Samples were tested against IgG and positive samples

were confirmed and tested with anti-IgM antibodies. Optical density values were normalized to the mean optical density of negative controls daily.

Data Analysis

Sample size and power

The study was designed to assess seropositivity within the primary sampling group with sufficient precision to inform policy. With 750 persons, and an assumed 5% positivity, the 95% CI for this estimate is 3.6% to 6.9%. Within the five emergency response subregions, at 5% seropositivity, the survey is powered for a precision of 2.3% to 10.2%. The secondary sampling group (n= 250) sample size was based on logistic limitations, but was powered to a precision of 2.8% to 8.8%. All confidence intervals are binomial exact, without adjustments for study design effects or non-response.

Analysis of serology data

Finite mixture models were used to determine seropositivity cutoffs. These latent-class models estimate breakpoints for seropositive and seronegative subpopulations, and have been applied to a range of pathogen serosurvey data, including rubella, pertussis, and parvovirus.[20–22] From this analysis, all samples with an IgG optical density ratio ≥ 2.49-fold above daily background were considered positive for SARS-CoV-2 (Distributions shown in Supplemental Figure 1; and sensitivity analysis with alternative cutpoints can be found in Supplemental Table 2).

Adjusted estimates

All reported prevalences and prevalence ratio estimates are adjusted with non-response weights, which were estimated using inverse weighting. Briefly, logistic regression models were used to calculate propensity scores for each individual in the sample using reported gender and race categories. These were transformed to probabilities; a small number of individuals had extremely large weights due to sparse strata; these weights were truncated at 1/0.02.[23] Weights were then applied to all prevalence and prevalence ratio estimations using the *survey* package in R.[24] The primary sampling group sample was self-weighting due to probability proportional to population size sampling. Sampling weights were not used in the secondary sampling group as selection used simple random sampling.

Multivariable analyses for prevalence ratios

Prevalence ratios were estimated to assess factors associated with seropositivity, with indivikdual Poisson models[25] for both of the two sampling groups, with robust (sandwich) errors to address clustering within households.

Bivariate analyses were performed for each factor separately. All variables with a p-value < 0.20 based on bivariate association with outcome were further evaluated for inclusion in final models. All final models were adjusted for age (continuous), race, and gender (see Table 1). Due to several very sparse categories, some were combined in the final models. Specifically, all race/ethnicity categories and all geographic regions were not included in analysis of the secondary sampling group due to unstable estimates.

Model parsimony was evaluated using AIC/BIC and all tests were two-tailed, with $\alpha = 0.05$. R version 4.0.3 and SAS version 9.4 (SAS Institute, Inc, Cary, NC, USA) were used for analysis.

Patient and public involvement

All members of the university community were invited to participate, and serological testing was proved at no cost to either the sampled individuals or to their selected household contact.

Results

A total of 1,001 individuals were enrolled into the study; this included 752 undergraduate students, 90 graduate students, 63 faculty/librarians, and 96 staff members (Figure 2). Seventy-six percent of these (n=762) returned blood samples for analysis; 548 in the primary sampling group, and 214 in the secondary sampling group. Of the 548 participants in the primary sampling group, 230 enrolled a household member. One household member submitted a sample without the sample of the main participant, bringing the total number of undergraduate household members to 231. Of the 214 participants in the secondary sampling group, 78 enrolled a household member. Two returned samples were excluded from analysis due to unlinkable samples. A total of 1,071 samples were included in the final analyses: 762 main participants and 309 household members (Figure 2).

Demographic characteristics of both sampling groups are presented in Table 1. Race categories do not total to 100% due to non-response and multiple possible answers. Age, gender, and essential worker status were broadly similar between those invited to participate and those who completed the study (Supplemental Table 1).

Of the total 1,071 samples tested, 46 were positive for SARS-CoV-2 antibodies. Demographic results are stratified by IgG serostatus (Table 2); no samples showed evidence for IgM positivity. Seropositivity was low-to-moderate across the survey groups, with several important exceptions. Variation is apparent by sex, race, and across geographic regions; however, several strata have wide confidence intervals due to small sample sizes.

Of the 779 primary sampling group participants and their household members, 36 were positive for SARS-CoV-2 antibodies. This corresponds to an overall seroprevalence of 5.3% (95% CI: 3.1 – 7.5) of the population after adjustment for nonresponse and geographic location. In the secondary sampling group, of the 292 graduate students, staff, librarians, faculty members, and their household members, ten (adjusted 4.0 %, 95% CI: 1.6 - 6.5) had evidence for prior SARS-CoV-2 infection. (Table 2). Results were also further stratified by UMass affiliate vs. household member. Of the 548 undergraduate students in the primary sampling group, 27 were positive for SARS-CoV-2 IgG antibodies (population positivity rate of 5.3% (95% CI: 3.1 – 7.6%). Of the 231 household members of undergraduate participants, nine (adjusted 5.2%, 95% CI: 1.2 - 9.2) were positive for SARS-CoV-2 IgG antibodies. In the secondary sampling group, eight University affiliates (adjusted 4.3 %, 95% CI: 1.3 - 7.3) were positive for SARS-CoV-2 IgG antibodies. Of the household members in the secondary sampling group, 2 were seropositive, with a weighted seroprevalence of (3.3 %, 95% CI: 0.0 – 7.8%). (Table 2). The overall distributions of measured IgG lognormal optical density ratios by subgroups are broadly similar (Supplemental Figure 1).

After adjustments for age, gender, region, and self-reported febrile illness since February 2020, the strongest association with seropositivity in the primary sampling group was Black or AIAN Race (PR = 4.49, 95% CI = 1.57,12.9) (Table 3). This indicates that individuals who reported being Black or American Indian/Alaskan Native had a prevalence 3.49 times higher than White

individuals after adjustment. Additionally, after adjustments, females and those who are gender diverse were at a significantly lower risk of prior SARS-CoV-2 infection (PR = 0.5; 95% CI = 0.27, 0.92) compared to males. Those who reported a febrile illness in February were more likely to be seropositive than those who did not report any febrile illness in this time period (PR = 2.42, 95% CI = 1.24, 4.75). No significant associations were found across each of the 5 geographic regions in the primary sampling group, however the prevalence of seropositivity was 48% higher in Region 1 compared to Region 4 (PR = 1.48, 95% CI = 0.62, 3.52).

Within the secondary sampling population, after adjustments for age, race, gender, region, household member and self-reported febrile illness since February (Table 4), participants who reported residing in either Region 2, 3, 4, or 5 had greater than 4 times higher prevalence of SARS-CoV-2 antibodies as compared to those who resided in Region 1 (PR = 4.08, 95% CI = 1.09, 15.33). No other factors included in the model were significantly associated with seropositivity.

Discussion

This mail-based serosurvey of two university-affiliated populations across Massachusetts in July and August 2020 found an estimated seroprevalence of ~5% of antibodies to SARS-CoV-2. These results indicate that even with extensive morbidity and mortality across the state at the time of sampling, there had been limited exposure to SARS-CoV-2 at a population-level. This estimated seroprevalence is lower than that detected with concurrent community-based studies in other states. An estimated 14.3% of the United States population had been previously infected with SARS-CoV-2 by November 2020, as estimated in a pooled analysis of multiple seroprevalence surveys.[26]

Our estimates are substantially lower than some models of COVID-19 seroprevalence in Massachusetts. One model estimates a seroprevalence of 16.2% (no CIs provided) on Jul 27, 2020 (the closest modeled date to these surveys). These estimates are nearly double our measured seroprevalence with inclusion of 110,000 confirmed cases at that date (ca. 1.5%).[27] These differences may be due to several factors, including a non-representative population by age or geographic range, or waning of antibody titers. Without CIs, we are unable to evaluate

coverage outside the reported point estimate. However, alternate nowcasting estimates suggest a total statewide attack rate on July 31, 2020 of 6.9% (95% CrI: 5.5 – 8.4%) in Massachusetts,[28] which are closely aligned with our estimates.

The primary sampling group also showed increased risk of seropositivity with self-reported illness since February 2020; this association was not observed in analyses of the secondary sampling group in multivariable analysis. Within the surveyed groups, approximately 24% of the primary sampling group and 18% of the secondary sampling group reported illness since February. This finding may indicate that some participants in our study may not have been strictly asymptomatic, and were simply unable to obtain a COVID-19 test due to limited availability during the beginning of the pandemic. This finding reinforces results suggesting asymptomatic illness is an important contributor to the observed pandemic trajectories.

Differing antibody dynamics have been reported in other studies. A number of studies have found sustained antibody levels for over 3 months,[29,30] while others studies suggest IgG levels can remain 6 months or more.[31–33] An additional study has reported rapid waning of routine serological markers in individuals who had lower initial antibody responses.[34]. Only 7.1% of those with high titers at baseline seroreverted to a level below the threshold for positivity within 60 days, compared to 64.9% of those with lower titers at baseline.[34] Evidence for IgM seropositivity was not detected in any of IgG positive samples, which is consistent with results from other surveys studies that included asymptomatic or subclinical populations due to rapidly waning titers.[32,35] Studies have shown that IgM levels decline more rapidly after infection than IgA and IgG levels, [30,36,37] and this is especially apparent with asymptomatic and sub-clinical infections.[32,35]

Trends in SARS-CoV-2 antibody levels are complex, and vary greatly depending on the measured antibody, timing of sampling, and severity of disease.[31,32,35] Seroconversion times vary depending on the study, but one study found a median time-to-seroconversion for IgM of 8 days and median seroconversion for IgG of 10 days. Additionally, the SARS-CoV-2 IgG response generally begin around 10-15 days after symptom onset.[2] For this reason, repeated serial

sampling of convalescent populations should be prioritized to more fully understand the dynamics of immune response.

SARS-CoV-2 seropositivity was associated with minority race status in this survey. While the total number of non-white participants was limited, the large effect size reinforces other studies suggesting that marginalized communities have been and continue to be excessively impacted by the pandemic. Results from the primary sampling group analysis suggest that self-reported Black race is a risk factor for previous SARS-CoV-2 infection, which is consistent with findings from other studies.[38–40] A number of factors may contribute to this significant association including data suggesting that Black individuals are more likely to work in frontline industries or live in areas with a higher population density in many settings.[41] No parallel associations were found in the analysis of the secondary sampling group due to limited sample size in some strata. In the secondary sampling group, the aggregation of Race categories into White race and Non-white race likely obscured meaningful associations between Race and seropositivity.

Results from the primary sampling group showed increased prevalence of IgG seropositivity among males. After adjusting for age, race, and region, male gender remained a statistically significant risk factor for evidence of prior SARS-CoV-2 infection. Other studies have similarly found that males have higher rates of infection than females for asymptomatic infections[4243] These findings may reflect differences in care seeking behavior (recruitment biases), true biological differences, or differences in health behaviors such as smoking, alcohol use, and COVID-19 prevention measures.[44] This association was not observed in the secondary sampling group.

This study was population-based and had broad eligibility criteria but is subject to several limitations. The exclusion of persons with confirmed diagnoses and any current symptoms (due to biosafety concerns) also inherently limited capture of sub-clinical infections. As such, the estimates are likely a lower bound. However, participants who suspected they may have been previously infected with SARS-CoV-2 might be more likely to participate compared to those that were not concerned with prior infection. This is a pervasive issue in community-based studies, where characteristics of those who volunteer to participate in community-based research differ

from the general population.[45] Randomization after a three-week enrollment period helped to address this limitation, as using only the first participants to volunteer could have biased the sample to include those who were most motivated to receive their antibody test results. If participants were more motivated to receive their results because they thought they suspected a prior exposure to SARS-CoV2, this would have inflated the observed prevalence of seropositivity in the study population.

Another limitation of the study is the self-reported response of the lack of a prior COVID-19 diagnosis and current fever. It is possible that some participants shielded their answer and submitted samples for analysis without meeting the eligibility criteria; this would have inflated our estimation of seroprevalence in asymptomatic groups, The limited number of non-white, and gender-diverse participants also limited some analyses and restricted out ability to assess any differences in prevalence across all racial groups. Next, while multiple studies have validated DBS sampling for SARS-CoV-2,[46] waning antibodies in asymptomatic individuals could be below the limited of detection of the ELISA assay. Additionally, the cutoff was determined in this study to be 2.49 for seropositivity. When using a single cut-point for a continuous variable, there is the possibility of outcome misclassification, however we attempted to reduce misclassification through the use of a finite mixture model. Finally, generalizability is limited due to the recruitment of a university-affiliated population in a relatively restricted geographic area.[8] The population in our study primarily included young adults, those of working age, and their household members. Individuals in older age groups or who reside in institutional settings would not have been recruited for our study. This population may not be representative of the broader US population and may be healthier, include few non-Whites, has higher education levels, and may have differing socio-political attitudes with consequent health impacts. Additionally, samples were collected during the summer months of 2020 in Massachusetts during which the state was in phase 3 of the reopening plan. The state government implemented a strict "lockdown" in March 2020 and progression to each reopening phase required a reduction in COVID-19 cases and hospitalizations. Massachusetts also had a mandatory mask order in all public spaces beginning May 1, 2020. Other states in New England followed similar timelines, however the implementation timelines and effectiveness differed widely across the United States.

It is probable that the government policy measures on a state-wide level influenced seroprevalence, with stricter guidelines resulting in lower viral exposure.

In conclusion, this serosurvey estimates prevalence of prior SARS-CoV-2 infections in a university-affiliated population in Massachusetts, with adjusted prevalences of 5.3% in the primary sampling group and 4.0% within the secondary sampling group. Risk factors for IgG seropositivity included self-reported recent febrile illness, minority race status, and male gender. This study reinforces the critical need for targeted serosurveys in highest-risk and marginalized communities, both in Massachusetts, and nationwide.

This study provides important estimates of seroprevalence in Massachusetts after the "first wave" of SARS-CoV-2 infections in the spring of 2020. These are critical to benchmark modeling studies, and to more comprehensively understand the dynamics of population-level sero-status throughout the continuing pandemic, especially as vaccines become widely available.

Tables

Table 1. Demographics of study populations, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Characteristic	Primary Sampling Group (N=779)	Secondary Sampling Group (N=292)
Gender		
Female	499 (64.1%)	154 (52.7%)
Male	274 (35.2%)	136 (46.6%)
Gender Diverse	5 (0.6%)	2 (0.7%)
Missing	1 (0.1%)	0 (0.0%)
Race		
AIAN	17 (2.2%)	1 (0.3%)
Asian	78 (10.0%)	34 (11.6%)
Black	12 (1.5%)	3 (1.0%)
Hispanic	36 (4.6%)	9 (3.1%)
Multiple	37 (4.8%)	11 (3.8%)
White	545 (70.0%)	217 (74.3%)
Missing	54 (6.9%)	17 (5.8%)
Age		
Mean	29.9	41.6
Median	21	39
Range	18 – 75	21 – 75
Education		
HS/GED	102 (13.1%)	5 (1.7%)
Some College	483 (62.0%)	24 (8.2%)
BA/BS	117 (15.0%)	78 (26.7%)
More Than BA/BS	74 (9.5%)	183 (62.7%)
Missing	3 (0.4%)	2 (0.7%)
Essential Worker		
No	533 (68.4%)	224 (76.7%)
Yes	195 (25.0%)	51 (17.5%)
Missing	51 (6.6%)	17 (5.8%)
Self-reported attitude about COVID-19		
Strongest fear	135 (17.3%)	72 (24.7%)
Somewhat fearful	389 (49.9%)	122 (41.8%)
Neutral/Missing	139 (17.8%)	63 (21.6%)
Somewhat not fearful	86 (11.0%)	23 (7.9%)
Not fearful	30 (3.9%)	12 (4.1%)
Self-reported febrile illness since February	No: 534 (68.6%) Yes: 188 (24.1%) Missing: 57 (7.3%)	No: 224 (76.7%) Yes: 53 (18.2%) Missing: 15 (5.1%)

Self-reported care seeking (if reporting illness since Feb)	No: 112 (59.6%) Yes: 75 (39.9%) Missing: 1 (0.5%)	No: 32 (60.4%) Yes: 21 (39.6%) Missing: 0 (0.0%)

*Notes: AIAN= American Indian/Alaska Native. The primary sampling group includes UMass undergraduates and household members, and the secondary sampling group includes UMass affiliated faculty, staff, and graduate students and household members.

Table 2. Weighted seropositivity by main demographic variables, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Characteristic	**Primary Sampling Group**	**Secondary Sampling Group**
Age in years, median	21 95% CI (20 - 21)	41 95% CI (38 - 44)
Prevalence of SARS-CoV-2 antibodies, by sub-group		
Overall Population Prevalence (95% CI)	5.3% (3.5 - 8.0)	4.0% (2.2 - 7.4)
Sex %, (95% CI)		
Female	4.0% (2.4 -6.6)	4.9% (2.2 - 10.7)
Male	8.7% (5.1 - 15.0)	3.0% (1.1 - 8.6)
Gender Diverse/No Response	0.0	0.0
Race * % (95% CI)		
Primary:		
White	3.9 (2.6 - 5.9)	-
Multiple	6.3 (1.7 - 23.7)	-
Asian	6.2 (2.7 - 14.5)	-
Missing	1.9 (0.3 - 13.5)	-
Hispanic	5.4 (1.4 - 21.0)	-
Black/AIAN	21.0 (5.8 - 76.4)	-
Secondary:		
White	-	4.2 (2.2 - 7.9)
Non-White	-	1.6 (0.2 - 11.7)
Essential worker status		
Yes	4.2 (2.0 - 8.8)	7.1 (2.3 - 21.3)
No	5.8 (3.5 - 9.7)	3.6 (1.7 - 7.6)
Missing Response	3.5 (0.9 - 14.2)	0

Table 2 cont. Weighted seropositivity by main demographic variables, SARS-CoV-2 serology surveys in university-affiliated populations, Massachusetts Jul-Aug, 2020.

Participant type		
University-affiliate	5.3 (3.5 – 8.1)	4.3 (2.1 - 8.6)
Household member	5.1 (2.4 – 11.2)	3.3 (0.8 -13.0)
State Emergency Response Region (see Figure 1)		
Region 1	7.8 (3.9 -15.6)	-
Region 2	1.6 (0.2 - 11.3)	-
Region 3	3.2 (1.0 - 10.7)	-
Region 4	5.7 (3.0 - 10.8)	-
Region 5	6.2 (2.3 - 16.5)	-
Region 1	-	3.1 (1.4 - 6.5)
Regions 2/3/4/5	-	11.3 (4.1 - 31.3)

*Notes: AIAN= American Indian/Alaska Native. All prevalences are adjusted for non-response.

Table 3. Multivariable associations for SARS-CoV-2 seropositivity, Primary Sampling Group, Massachusetts USA, Jul-Aug 2020.

Characteristic	Prevalence Ratio	95% CI	p-value
Emergency Response Region			
Region 1	1.48	0.62, 3.52	0.38
Region 2	0.34	0.05, 2.45	0.28
Region 3	0.53	0.14, 1.96	0.34
Region 4	Reference	-	-
Region 5	1.02	0.35, 2.98	0.97
Age (years)	1.04	0.96, 1.12	0.33
Gender			
Male	Reference	-	-
Female, Gender diverse, or No response	0.50	0.27, 0.92	**0.027**
Race			
White	Reference	-	-
Multiple	1.91	0.46, 7.98	0.38
Asian	1.66	0.66, 4.16	0.28
Missing Race	0.51	0.07, 3.71	0.51
Hispanic	1.76	0.44, 7.04	0.42
Black or AIAN	4.49	1.57, 12.9	**0.005**
Febrile illness since February			
No	Reference	-	-
Yes	2.42	1.24, 4.75	**0.010**
Missing Response	0.33	0.04, 2.45	0.28
Other household member			
no	Reference	-	-
yes	0.29	0.01, 7.69	0.46

Table 4. Multivariable associations for SARS-CoV-2 seropositivity, Secondary Sampling Group, Massachusetts, USA, Jul-Aug 2020.

Characteristic	Prevalence Ratio	95% CI	p-value
Age (years)	1.03	0.99, 1.07	0.17
Gender			
Male	Reference	-	-
Female or Gender Diverse	1.35	0.43, 4.31	0.61
Race			
White	Reference	-	-
All Other/Multiple/Missing	0.58	0.07, 4.86	0.62
Febrile illness since February			
No	Reference	-	-
Yes	2.56	0.68, 9.67	0.17
Missing Response	2.35	0.21, 26.73	0.49
Emergency Response Region			
Region 1	Reference	-	-
Regions 2/3/4/5	4.08	1.09, 15.33	**0.039**
Other household member			
No	Reference	-	-
Yes	0.70	0.18, 2.72	0.61

Acknowledgements

We would like to express our appreciation to all of the individuals who have enabled the completion of this study. Rob Leveille and Charlie Apicella of the UMass Mail and Distribution Services have facilitated the label printing, outgoing shipments, and incoming shipments for over 1,000 sample boxes. Sujitha Chandra Kumar, Vincent Lee, and Pratik Patel have been valuable members of the REDCap support team. Kimberly Tremblay, Jesse Mager, Katherine Dorfman, Pa Tamba Ngom, and Ryan Kurtz have graciously allowed us to utilize laboratory space for box assembly and sample processing. These individuals have worked tirelessly to support this study despite massive pandemic-related logistical challenges and tight deadlines. We are very grateful for their support.

Figures

Fig 1. Study sampling frames for SARS-CoV-2 seropositivity surveys, Massachusetts, USA, Jul-Aug 2020. (Note: State-level Emergency Response Regions shown in orange; anonymized participant locations shown as red markers).

Fig 2. Participant enrollment diagram, (CONSORT), SARS-CoV-2 serosurvey, Massachusetts, USA, Jul-Aug 2020.

Author contributions

AAL and DA led study conceptualization and design.
LB and NR informed overall aims, sampling design, and statistical analysis.
JR, TS, and EYC implemented all field implementation and associated data collection.
DA organized, performed and reported all laboratory-based testing.
JR, TS, EYC, and AAL performed data cleaning and analyses, and wrote the first draft.
All authors contributed to revisions of the final manuscript.
All authors read and approved the final manuscript.

COI statements

None of the authors report any conflicts of interest, financial or non-financial, regarding this study.

Funding

UMass Faculty Funds (A Lover; SPH-AL-001); UMass Faculty Funds (N Reich; SPH-NR-001); UMass Institute for Applied Life Science "Midigrant" (#169076; A. Lover); and D. Alfandari was supported by grants from the NIH/USPHS (RO1DE016289 and R24OD021485).

Data sharing

Fully-identified data will be deposited at: https://osf.io/437tg/.

References

1. Lemieux JE, Siddle KJ, Shaw BM, et al. Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. *Science*. Published online December 10, 2020. doi:10.1126/science.abe3261

2. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. *Nature Communications*. 2020;11(1):4507. doi:10.1038/s41467-020-18272-4

3. *Massachusetts Department of Public Health COVID-19 Dashboard*. Department of Public Health, Massachusetts; 2020. Accessed October 25, 2020. https://www.mass.gov/doc/covid-19-dashboard-july-1-2020/download

4. Bryant JE, Azman AS, Ferrari MJ, et al. Serology for SARS-CoV-2: Apprehensions, opportunities, and the path forward. *Science Immunology*. 2020;5(47). doi:10.1126/sciimmunol.abc6347

5. Gronvall G, Connell N, Kobokovich A, et al. *Developing a National Strategy for Serology (Antibody Testing) in the United States*. The Johns Hopkins Center for Health Security; 2020. https://www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/200422-national-strategy-serology.pdf

6. Naranbhai V, Chang CC, Beltran WFG, et al. High Seroprevalence of Anti-SARS-CoV-2 Antibodies in Chelsea, Massachusetts. *The Journal of Infectious Diseases*. Published online September 9, 2020:jiaa579. doi:10.1093/infdis/jiaa579

7. Reifer J, Hayum N, Heszkel B, Klagsbald I, Streva VA. SARS-CoV-2 IgG antibody responses in New York City. *Diagn Microbiol Infect Dis*. Published online July 21, 2020. doi:10.1016/j.diagmicrobio.2020.115128

8. Tilley K, Ayvazyan V, Martinez L, et al. A Cross-Sectional Study Examining the Seroprevalence of Severe Acute Respiratory Syndrome Coronavirus 2 Antibodies in a University Student Population. *Journal of Adolescent Health*. Published online October 2020:S1054139X2030522X. doi:10.1016/j.jadohealth.2020.09.001

9. 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*. Published online July 21, 2020. doi:10.1001/jamainternmed.2020.4130

10. Lim MD. Dried Blood Spots for Global Health Diagnostics and Surveillance: Opportunities and Challenges. *The American Journal of Tropical Medicine and Hygiene*. 2018;99(2):256-265. doi:10.4269/ajtmh.17-0889

11. Corran PH, Cook J, Lynch C, et al. Dried blood spots as a source of anti-malarial antibodies for epidemiological studies. *Malar J*. 2008;7(1):195. doi:10.1186/1475-2875-7-195

12. McDade TW, McNally EM, Zelikovich AS, et al. High seroprevalence for SARS-CoV-2 among household members of essential workers detected using a dried blood spot assay. Khudyakov YE, ed. *PLoS ONE*. 2020;15(8):e0237833. doi:10.1371/journal.pone.0237833

13. Karp DG, Danh K, Seftel D, Robinson P, Tsai C. *A Serological Assay to Detect SARS-CoV-2 Antibodies in at-Home Collected Finger-Prick Dried Blood Spots*. Allergy and Immunology; 2020. doi:10.1101/2020.05.29.20116004

14. Thevis M, Knoop A, Schaefer MS, et al. Can dried blood spots (DBS) contribute to conducting comprehensive SARS-CoV-2 antibody tests? *Drug Test Anal*. 2020;12(7):994-997. doi:10.1002/dta.2816

15. Valentine-Graves M, Hall E, Guest JL, et al. At-home self-collection of saliva, oropharyngeal swabs and dried blood spots for SARS-CoV-2 diagnosis and serology: Post-collection acceptability of specimen collection process and patient confidence in specimens. Kumar S, ed. *PLoS ONE*. 2020;15(8):e0236775. doi:10.1371/journal.pone.0236775

16. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*. 2019;95:103208. doi:10.1016/j.jbi.2019.103208

17. map-bt-regions-by-coalitions.pdf | Mass.gov. Accessed January 10, 2021. https://www.mass.gov/doc/map-bt-regions-by-coalitionspdf

18. Amanat F, Stadlbauer D, Strohmeier S, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Medicine*. 2020;26(7):1033-1036. doi:10.1038/s41591-020-0913-5

19. Stadlbauer D, Amanat F, Chromikova V, et al. 1 A detailed protocol for a serological assay to detect SARS-CoV-2 2 seroconversion in humans: antigen production and test setup. :13.

20. Hardelid P, Williams D, Dezateux C, et al. Analysis of rubella antibody distribution from newborn dried blood spots using finite mixture models. *Epidemiology & Infection*. 2008;136(12):1698-1706.

21. Baughman AL, Bisgard KM, Lynn F, Meade BD. Mixture model analysis for establishing a diagnostic cut-off point for pertussis antibody levels. *Stat Med*. 2006;25(17):2994-3010. doi:10.1002/sim.2442

22. Nj G. Analysis of serological surveys using mixture models: application to a survey of parvovirus B19. *Stat Med*. 1996;15(14):1567-1573. doi:10.1002/(sici)1097-0258(19960730)15:14<1567::aid-sim289>3.0.co;2-g

23. Moore KL, Neugebauer R, Laan MJ van der, Tager IB. Causal inference in epidemiological studies with strong confounding. *Statistics in Medicine*. 2012;31(13):1380-1404. doi:https://doi.org/10.1002/sim.4469

24. Lumley T. *Survey: Analysis of Complex Survey Samples*.; 2020. Accessed January 25, 2021. https://CRAN.R-project.org/package=survey

25. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio | BMC Medical Research Methodology | Full Text. Accessed January 10, 2021. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-3-21

26. Angulo FJ, Finelli L, Swerdlow DL. Estimation of US SARS-CoV-2 Infections, Symptomatic Infections, Hospitalizations, and Deaths Using Seroprevalence Surveys. *JAMA Netw Open*. 2021;4(1):e2033706. doi:10.1001/jamanetworkopen.2020.33706

27. Archive of COVID-19 cases in Massachusetts | Mass.gov. Accessed March 5, 2021. https://www.mass.gov/info-details/archive-of-covid-19-cases-in-massachusetts

28. Attack Rate 08-13 | Boni Lab. Accessed March 5, 2021. https://mol.ax/covid/attack-rate-08-13/

29. Crawford KHD, Dingens AS, Eguia R, et al. Dynamics of Neutralizing Antibody Titers in the Months After Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *The Journal of Infectious Diseases*. 2021;223(2):197-205. doi:10.1093/infdis/jiaa618

30. Isho B, Abe KT, Zuo M, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Science Immunology*. 2020;5(52). doi:10.1126/sciimmunol.abe5511

31. Figueiredo-Campos P, Blankenhaus B, Mota C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in COVID-19 patients and healthy volunteers up to six months post disease onset. *Eur J Immunol*. Published online October 21, 2020:eji.202048970. doi:10.1002/eji.202048970

32. Wang Y, Zhang L, Sang L, et al. Kinetics of viral load and antibody response in relation to COVID-19 severity. *Journal of Clinical Investigation*. 2020;130(10):5235-5244. doi:10.1172/JCI138759

33. Wu J, Liang B, Chen C, et al. SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. *medRxiv*. Published online July 24, 2020:2020.07.21.20159178. doi:10.1101/2020.07.21.20159178

34. Self WH. Decline in SARS-CoV-2 Antibodies After Mild Infection Among Frontline Health Care Personnel in a Multistate Hospital Network — 12 States, April–August 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6947a2

35. Lynch KL, Whitman JD, Lacanienta NP, et al. *Magnitude and Kinetics of Anti-SARS-CoV-2 Antibody Responses and Their Relationship to Disease Severity*. Infectious Diseases (except HIV/AIDS); 2020. doi:10.1101/2020.06.03.20121525

36. Yao X-Y, Liu W, Li Z-Y, et al. *Neutralizing and Binding Antibody Kinetics of COVID-19 Patients during Hospital and Convalescent Phases*. Infectious Diseases (except HIV/AIDS); 2020. doi:10.1101/2020.07.18.20156810

37. Beaudoin-Bussières G, Laumaea A, Anand SP, et al. Decline of Humoral Responses against SARS-CoV-2 Spike in Convalescent Individuals. *mBio*. 2020;11(5). doi:10.1128/mBio.02590-20

38. Jehi L, Ji X, Milinovich A, et al. Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing: Results From 11,672 Patients. *Chest*. 2020;158(4):1364-1375. doi:10.1016/j.chest.2020.05.580

39. Rentsch CT, Kidwai-Khan F, Tate JP, et al. Covid-19 by Race and Ethnicity: A National Cohort Study of 6 Million United States Veterans. *medRxiv*. Published online May 18, 2020. doi:10.1101/2020.05.12.20099135

40. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Network Open*. 2020;3(9):e2021892-e2021892. doi:10.1001/jamanetworkopen.2020.21892

41. Black workers face two of the most lethal preexisting conditions for coronavirus—racism and economic inequality. Economic Policy Institute. Accessed February 20, 2021. https://www.epi.org/publication/black-workers-covid/

42. Al-Sadeq DW, Nasrallah GK. The incidence of the novel coronavirus SARS-CoV-2 among asymptomatic patients: A systematic review. *International Journal of Infectious Diseases*. 2020;98:372-380. doi:10.1016/j.ijid.2020.06.098

43. Barek MdA, Aziz MdA, Islam MS. Impact of age, sex, comorbidities and clinical symptoms on the severity of COVID-19 cases: A meta-analysis with 55 studies and 10014 cases. *Heliyon*. 2020;6(12):e05684. doi:10.1016/j.heliyon.2020.e05684

44. Bwire GM. Coronavirus: Why Men are More Vulnerable to Covid-19 Than Women? *SN Compr Clin Med*. Published online June 4, 2020:1-3. doi:10.1007/s42399-020-00341-w

45. Ganguli M, Lytle ME, Reynolds MD, Dodge HH. Random Versus Volunteer Selection for a Community-Based Study. *The Journals of Gerontology: Series A*. 1998;53A(1):M39-M46. doi:10.1093/gerona/53A.1.M39

46. Krishnamurthy HK, Jayaraman V, Krishna K, et al. Antibody profiling and prevalence in US patients during the SARS-CoV2 pandemic. *PLOS ONE*. 2020;15(11):e0242655. doi:10.1371/journal.pone.0242655

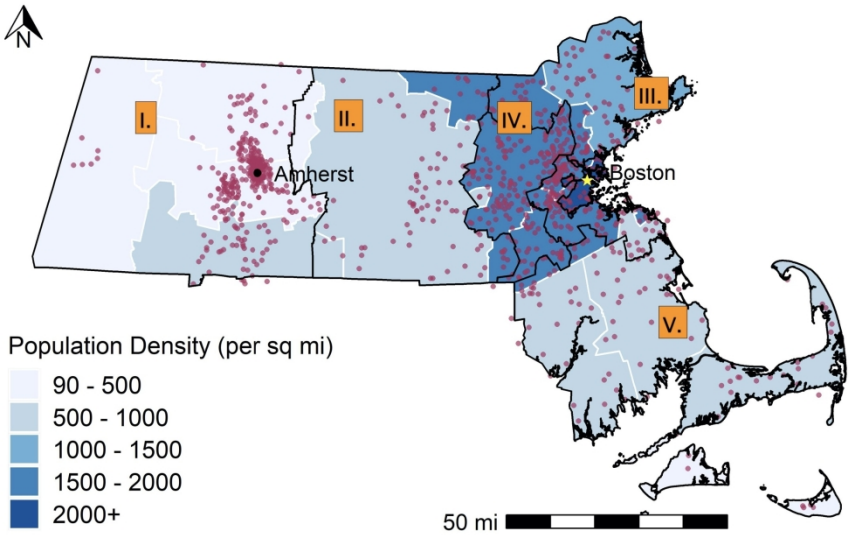
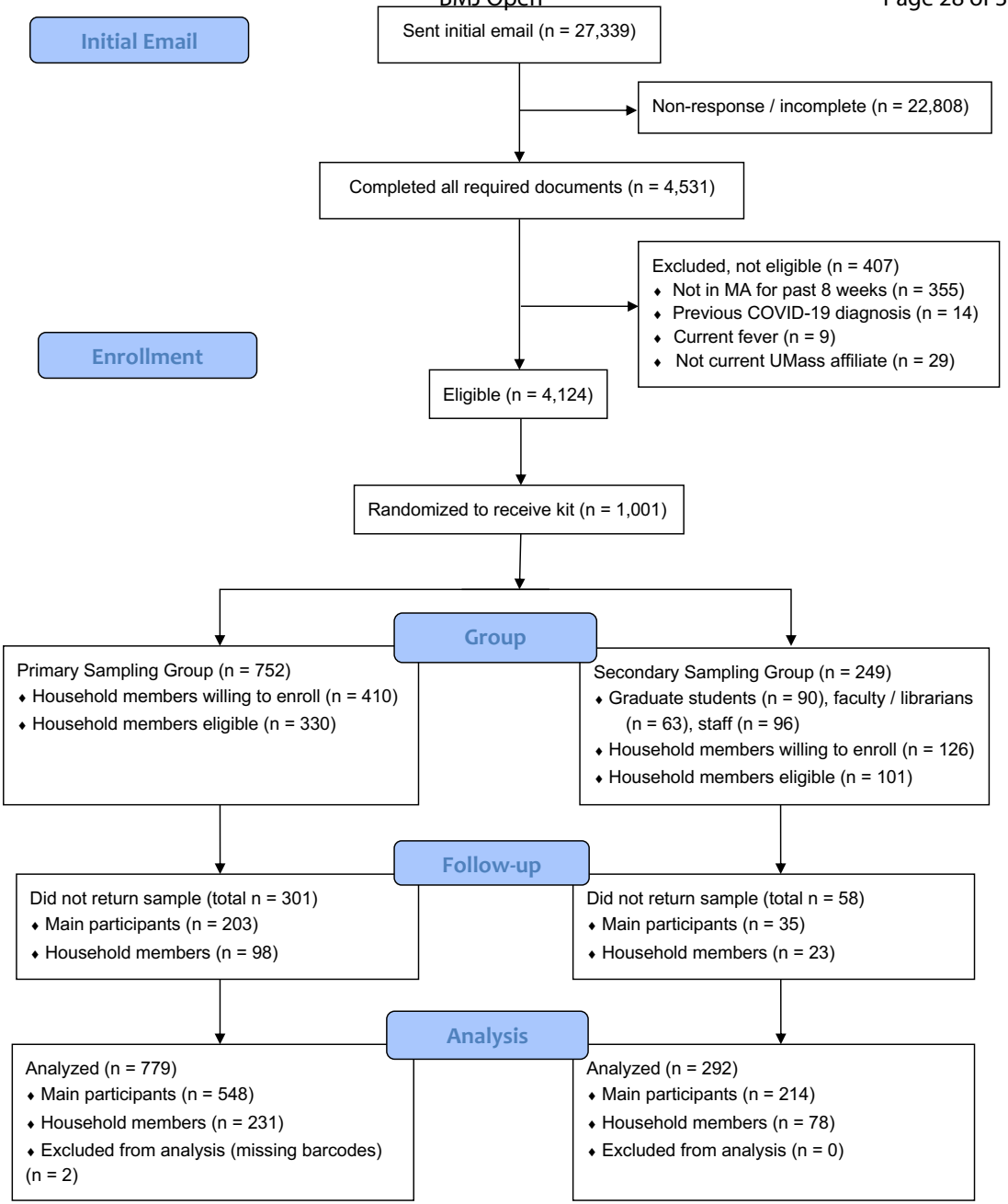


Fig 1. Study sampling frames for SARS-CoV-2 seropositivity surveys, Massachusetts, USA, Jul-Aug 2020. (Note: State-level Emergency Response Regions shown in orange; anonymized participant locations shown as red markers).

447x254mm (144 x 144 DPI)


Initial Email
Sent initial email (n = 27,339)
Non-response / incomplete (n = 22,808)
Completed all required documents (n = 4,531)
Excluded, not eligible (n = 407)
Not in MA for past 8 weeks (n = 355)
Previous COVID-19 diagnosis (n = 14)
Current fever (n = 9)
Not current UMass affiliate (n = 29)
Enrollment
Eligible (n = 4,124)
Randomized to receive kit (n = 1,001)
Group
Primary Sampling Group (n = 752)
Household members willing to enroll (n = 410)
Household members eligible (n = 330)
Secondary Sampling Group (n = 249)
Graduate students (n = 90), faculty / librarians (n = 63), staff (n = 96)
Household members willing to enroll (n = 126)
Household members eligible (n = 101)
Follow-up
Did not return sample (total n = 301)
Main participants (n = 203)
Household members (n = 98)
Did not return sample (total n = 58)
Main participants (n = 35)
Household members (n = 23)
Analysis
Analyzed (n = 779)
Main participants (n = 548)
Household members (n = 231)
Excluded from analysis (missing barcodes) (n = 2)
Analyzed (n = 292)
Main participants (n = 214)
Household members (n = 78)
Excluded from analysis (n = 0)

Supplemental Information

Serological surveys to estimate cumulative incidence of SARS-CoV-2 infection in adults (Sero-MAss study), Massachusetts, July-August 2020: a mail-based cross-sectional study

* Correspondence to: alover@umass.edu

School of Public Health and Health Sciences
University of Massachusetts- Amherst
715 North Pleasant Street
Amherst, MA 01003-9304
phone: 413-545-7426

Supplemental Information

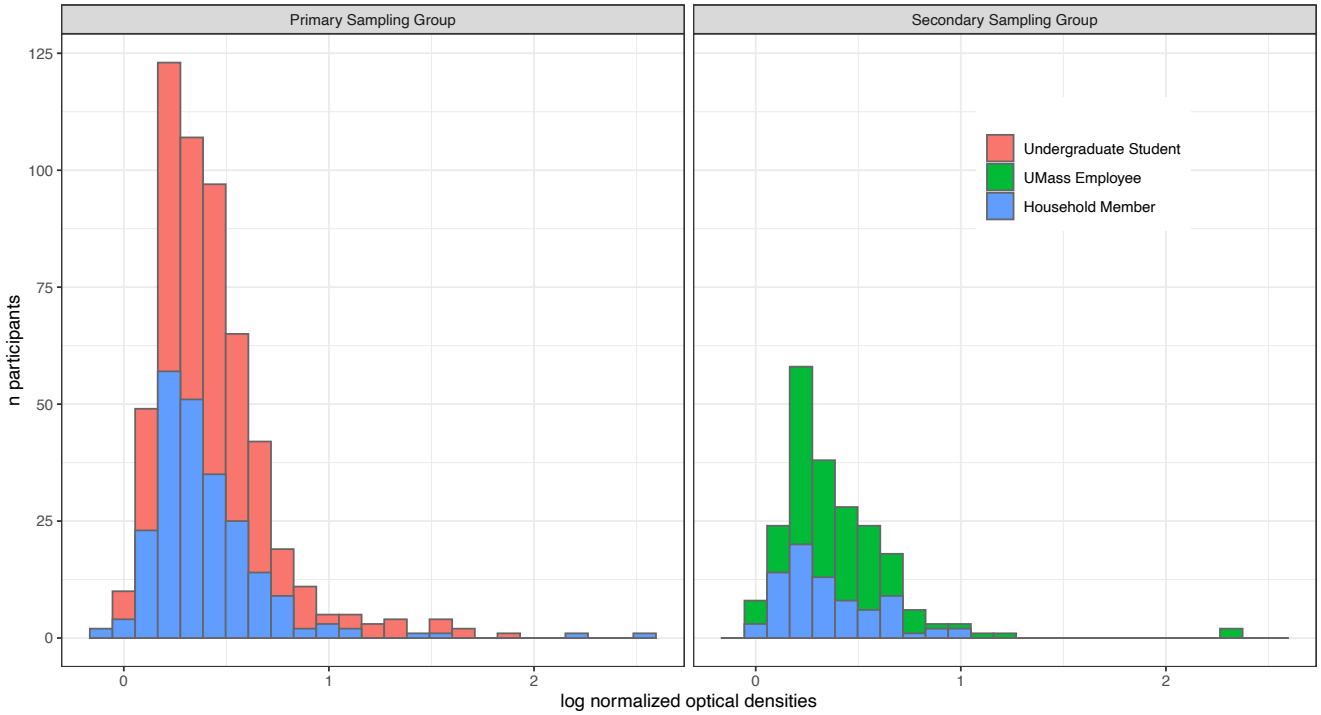
Supplemental Table 1: Comparison of characteristics of participants selected for randomization and those not selected for randomization, SARS-CoV-2 serosurvey, Massachusetts, USA, Jul-Aug 2020.

	Primary Sampling Group (undergraduates)		**Secondary sampling group (employees)**	
	Not Randomized	Randomized	Not Randomized	Randomized
n	869	752	2253	249
Age (mean (SD))	20.64 (3.09)	20.18 (1.79)	42.05 (13.53)	40.69 (13.88)
Gender (%)				
Unknown	0 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)
Female	525 (60.4)	453 (60.2)	1262 (56.0)	147 (59.0)
Gender diverse	12 (1.4)	8 (1.1)	49 (2.2)	3 (1.2)
Male	332 (38.2)	291 (38.7)	941 (41.8)	99 (39.8)
Self-reported febrile illness since February (%)				
Unknown	1 (0.1)	2 (0.3)	6 (0.3)	0 (0.0)
No	568 (65.4)	499 (66.4)	1658 (73.6)	189 (75.9)
Not sure	75 (8.6)	57 (7.6)	154 (6.8)	16 (6.4)
Yes	225 (25.9)	194 (25.8)	435 (19.3)	44 (17.7)
Education (%)				
Unknown	4 (0.5)	3 (0.4)	6 (0.3)	1 (0.4)
BA/BS	15 (1.7)	14 (1.9)	509 (22.6)	64 (25.7)
High school / GED	90 (10.4)	108 (14.4)	46 (2.0)	2 (0.8)
More than BA/BS	2 (0.2)	0 (0.0)	1528 (67.8)	166 (66.7)
Prefer not to answer	1 (0.1)	0 (0.0)	13 (0.6)	1 (0.4)
Some college	757 (87.1)	627 (83.4)	151 (6.7)	15 (6.0)
Response to, "I am afraid of COVID-19" (%)				
Unknown	1 (0.1)	2 (0.3)	2 (0.1)	0 (0.0)
Neither agree nor disagree	199 (22.9)	141 (18.8)	399 (17.7)	51 (20.5)
Somewhat agree	390 (44.9)	374 (49.7)	1051 (46.6)	100 (40.2)
Somewhat disagree	115 (13.2)	99 (13.2)	153 (6.8)	22 (8.8)
Strongly agree	121 (13.9)	104 (13.8)	573 (25.4)	66 (26.5)
Strongly disagree	43 (4.9)	32 (4.3)	75 (3.3)	10 (4.0)

Supplemental Table 2: Comparison of estimated seropositivity by sampling groups using alternative cutpoints for seropositivity, SARS-CoV-2 serosurvey, Massachusetts, USA, Jul-Aug 2020.

Sampling Group	Cutoff Value	Unadjusted Prevalence (%)	Adjusted Prevalence (%) (95% confidence interval)
Primary sampling group (undergraduates)			
	2.49	4.6	5.3 (3.5 – 8.0)
	2.0	10.9	11.6 (9.1 – 15.0)
	2.1	8.7	9.2 (6.9 – 12.3)
	2.2	7.1	7.6 (5.5 – 10.6)
	2.3	5.9	6.5 (4.5 – 9.3)
	2.4	5.0	5.6 (3.8 – 8.4)
	2.5	4.5	5.1 (3.3 – 7.9)
	2.6	4.0	4.5 (2.8 – 7.1)
	2.7	3.7	4.3 (2.6 – 6.9)
	2.8	3.3	3.9 (2.3 – 6.5)
	2.9	2.8	3.5 (2.0 – 6.2)
	3.0	2.4	3.3 (1.8 – 5.9)
Secondary sampling group (employees)			
	2.49	3.4	4.0 (2.2 – 7.4)
	2.0	7.5	8.2 (5.3 – 12.7)
	2.1	5.5	6.0 (3.7 – 9.8)
	2.2	4.8	5.4 (3.2 – 9.2)
	2.3	4.8	5.4 (3.2 – 9.2)
	2.4	3.8	4.5 (2.5 – 8.0)
	2.5	3.4	4.0 (2.2 – 7.4)
	2.6	2.4	3.0 (1.5 – 6.2)
	2.7	2.1	2.6 (1.2 – 5.7)
	2.8	2.1	2.6 (1.2 – 5.7)
	2.9	1.4	1.7 (0.7 – 4.6)
	3.0	1.0	1.3 (0.4 – 4.0)

Supplemental Figure 1. Distributions of IgG lognormal optical density (OD) ratios, by subgroups, SARS-CoV-2, Massachusetts, USA, Jul-Aug 2020.



STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

Sero-MAss]	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5-6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	15 (Table 1)
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	7; 21 (Fig 2)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-9
		(*b*) Describe any methods used to examine subgroups and interactions	8
		(*c*) Explain how missing data were addressed	7-8
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	8
		(*e*) Describe any sensitivity analyses	n/a
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7; 21 (Fig 2)
		(b) Give reasons for non-participation at each stage	9
		(c) Consider use of a flow diagram	Fig. 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Fig 1.
		(b) Indicate number of participants with missing data for each variable of interest	Supp.

Outcome data	15*	Report numbers of outcome events or summary measures	16 (Table 2)
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-11
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10
Discussion			
Key results	18	Summarise key results with reference to study objectives	11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-14
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	3

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.