THE LANCET
Global Health

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Malani A, Shah D, Kang G, et al. Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. *Lancet Glob Health* 2020; published online Nov 13. http://dx.doi.org/10.1016/S2214-109X(20)30467-8.

Online Supplement for
Geographic variation in seroprevalence of SARS-CoV-2: Comparing slums and non-slums in Mumbai, India, during June 29-July 19, 2020

Anup Malani, Daksha Shah, Gagandeep Kang, Gayatri Nair Lobo, Jayanthi Shastri, Manoj Mohanan, Rajesh Jain, Sachee Agrawal, Sandeep Juneja, Sofia Imad, Ullas Kolthur-Seetharam

Main Tables and Figures

Table 1

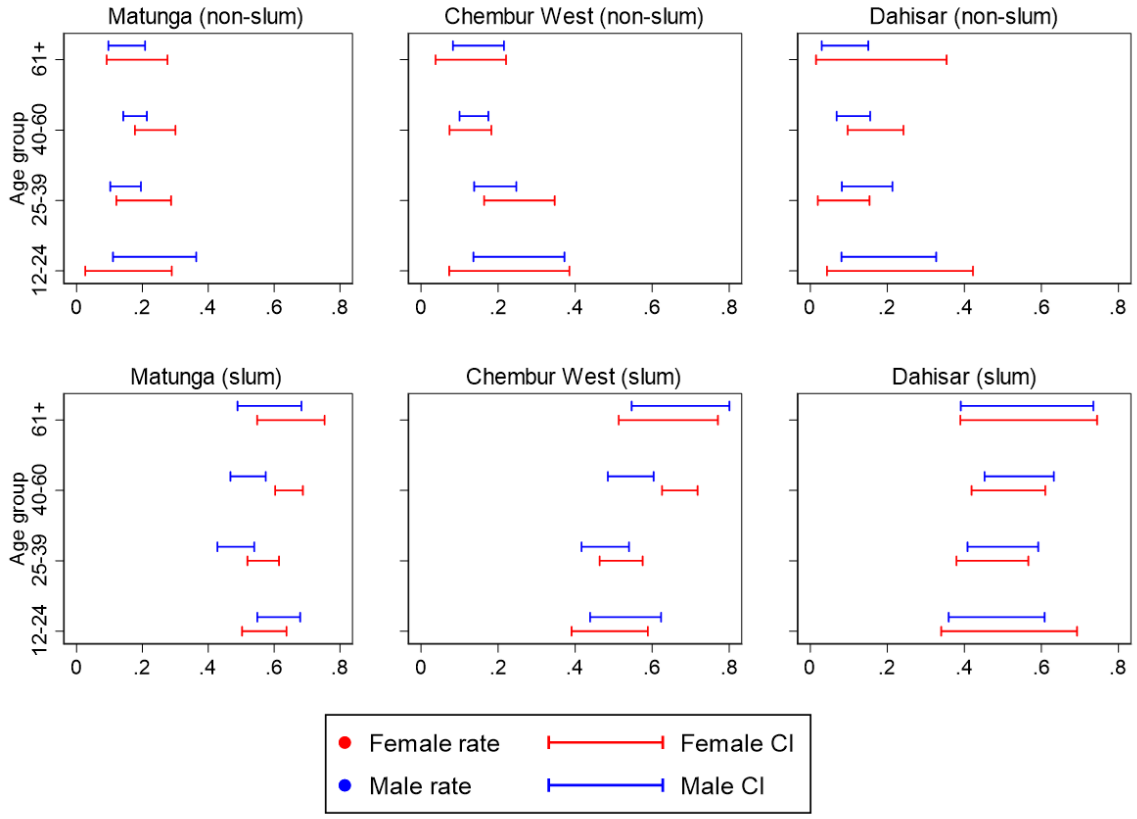
Adjusted proportion of positive tests (by location) and adjusted seroprevalence (by location and estimate of test accuracy)

	Adjusted proportion of positive tests		Sero-prevalence (sensitivity = 0.969, specificity = 0.999)[3]		Sero-prevalence (sensitivity = 0.900, specificity = 1.000)[2]	
Location	Rate	95% CI	Rate	95% CI	Rate	95% CI
Non-slums	0.160	(.148,.172)	0.162	(.149,.174)	0.171	(.158,.184)
Matunga	0.176	(.15,.202)	0.179	(.159,.199)	0.189	(.168,.21)
Chembur West	0.165	(.137,.194)	0.168	(.145,.19)	0.177	(.154,.201)
Dahisar	0.119	(.085,.153)	0.120	(.095,.145)	0.128	(.101,.154)
Slum	0.541	(.527,.555)	0.557	(.543,.572)	0.583	(.568,.599)
Matunga	0.570	(.547,.592)	0.587	(.567,.607)	0.614	(.593,.635)
Chembur West	0.551	(.524,.578)	0.567	(.544,.591)	0.594	(.569,.618)
Dahisar	0.511	(.464,.558)	0.526	(.486,.567)	0.551	(.509,.593)

Note. The adjusted proportion of positive tests is estimated by, first, estimating the proportion for each demographic group and, then, calculating a weighted average that ensures each group's proportion has a weight proportional to its share of the population in a location, with weights as indicated in Table e3. Sero-prevalence and its confidence interval is calculated from the adjusted proportion of positive tests for a location using the Rogan-Gladden[4] formula. Confidence interval is estimated using a normal approximation.

Figure 1

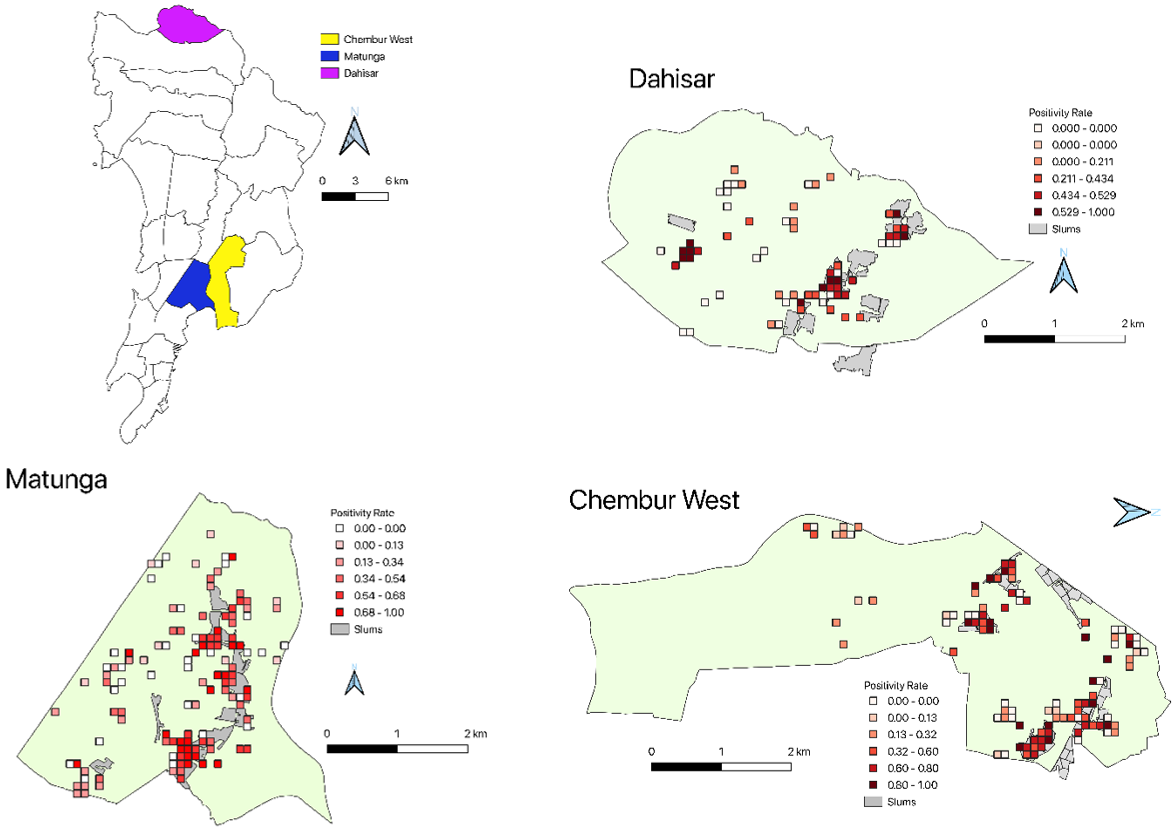
Proportion of positive tests by age group and sex for each location.



Notes. Plots present unadjusted proportion of positive tests (point) and Score 95% confidence intervals (whisker) by the age group on y-axis. Different colors represent male and females.

Figure 2

Spatial variation in crude prevalence within wards of Mumbai



Study Approvals and IRB

This study was approved by the Government of India (NITI Aayog), Government of Maharashtra, and the Municipal Corporation of Greater Mumbai. The study protocols were also approved by IRB / IEC committees at the authors' respective academic institutions:

- TIFR (TIFR/IHEC/2020-1),
- Kasturba Hospital (IRB 20/2020),
- THSTI (EC/NEW/INST/2019/275),
- Duke University (Protocol:2020-0575),
- University of Chicago (IRB20-1144).

Study sites

The study was conducted in 6 sites: in slum and non-slum communities in each of 3 wards (Matunga, Chembur West, and Dahisar), representing the three zones (city, eastern suburbs, and western suburbs, respectively) in Mumbai.

The specific ward in each zone was selected based on logistical considerations: these are wards where we had NGO partners operating in the ward's slums and the Municipal Corporation had health officers that could accompany phlebotomists as they surveyed non-slums.

Slums are defined as communities living on land or in premises to which they do not have formal legal rights and which have been labelled as slums by the municipal corporation. Slums constitute 49% of the population (slum population 705,523; non-slum population 709,394) of the 3 wards in our sample according to the city's Mid Year Estimated Population for 2019[1].

We surveyed up to 8 of the largest slums in each ward. We chose the largest slums because there is a fixed cost to sampling in a slum, most notably setting up the relationship with the NGO(s) that operated in the slum. We chose largest by size because we have data on the area of specific slums but not the population of those slums. However, because the size of homes is similar (e.g., 10 feet x 10 feet) across slums, we assumed that population is roughly proportional to size.

Sample size calculation

The study was powered to estimate a 1.5 percentage point difference in proportion of positive tests in a two-sided test with 95% confidence in each of the six study areas. We chose a 1.5 percentage point minimum detectable effect based on budgetary and logistical considerations.

The sample-size formula we employed was $N = p(1-p)z_{\alpha/2}/(0.015)^2$, where p is our prior on prevalence.

Our estimate of p for each ward was obtained in two steps. First, we obtained a citywide estimate 0.1 for seroprevalence based on a simulation conducted by the Indian Institute of Science, Bangalore, and the Tata Institute for Fundamental Research using an agent-based model and the history of confirmed cases in the city.[2] Second, we estimated the ward-level seroprevalence by scaling the citywide estimated prevalence by the rate of confirmed cases per million in each ward relative to the rate of confirmed cases citywide (Table e1). Because we did not have data suggesting that slums had a higher seroprevalence than non-slums, we assumed seroprevalence was the same across both types of community. Indeed, a simple plot of ward-wise confirmed cases against the share of each ward's population that was in slums suggested no significant or material relationship between slum share and the rate of confirmed cases.

Our choice of $z_{\alpha/2} = 1.96$ was based on a normal approximation to the binomial, a reasonable approximation at larger sample sizes.

Our required sample sizes were 2249, 1622, and 564 participants in each of the slum and non-slum sections of Matunga, Chembur West, and Dahisar, respectively, a total of 8,870 individuals (Table e1).

Study duration

Because prevalence changes over time, we sought to estimate average prevalence over a two-week period. To balance statistical power and bias, we stopped sampling either when we hit our sample size targets or the sampling period lapsed. Slums were sampled June 29 to July 14, 2020; non-slums were sampled from July 3 to 19. Therefore, the overall duration of the study was 20 days.

Systematic sampling

Slums. Within each ward, we recruited in up to 8 of the largest slums by population to balance the fixed costs of working in each additional slum and the possibility that prevalence may vary across slums. We divided each slum into mutually exclusive, geographic polygons covering roughly 400 homes, and sampled 100 homes per polygon. Starting with the home closest to the

centroid of each polygon, we systematically sampled one person in every fourth home in one direction.

Non-slums. On maps of each ward, we drew rectangular grids such that the ward was covered with just enough cells that, if we draw 100 persons from each cell, we would meet our sample size target for non-slum areas in the ward. We started sampling at a building close to the center of the cell. At each building we sought consent from the building's residential association before entering. There were difficulties in obtaining consent from some resident associations. When allowed to enter, we recruited one household per floor. Otherwise we asked the association to request one volunteer per floor. If we did not meet our sample size target for a cell at a building, we went to an adjacent building to continue sampling. Some cells of our grid were areas with minimal residential facilities. We did not sample in these cells. We reallocated samples from such non-sampled cells uniformly to other cells.

Surveyors were given a list of 8 demographic groups (4 age bins by 2 gender bins) and asked to cycle through the list when selecting whom to survey at each home. The distribution of our final sample across these groups are a function of the population distribution across groups and of consent rates in each group.

Data collection, and testing

Each participant was administered a survey to collect socio-demographic data (age, gender, household composition), comorbidities (e.g., hypertension), and contact and travel history over the last 2 months. Phlebotomists collected 5ml of blood from each participant via venipuncture in an EDTA vacutainer.

Testing

At the Kasturba Hospital laboratory, plasma was separated and used to test for IgG antibodies via chemiluminescence (CLIA) using Abbott Diagnostics ArchitectTM N-protein based SARS-CoV-2 tests. According to the manufacturer, sensitivity was 96.77% (95% CI: 90.86 to 99.33) and specificity was 99.63% (95% CI: 99.05 to 99.90). Manufacturer validation results are consistent with [3], finding a sensitivity of 96.9% (89.5% to 99.5%) and specificity of 99.90% at ≥14 days after symptom onset. However, according to [4], sensitivity was 92.7% (95% CI: 90.2, 94.7) and specificity of 99.9% (95% CI: 99.4, 100) at ≥14 days after symptom onset. Likewise,

according to [5], sensitivity was 90% (74.4% to 96.5%) and specificity was 100% (95.4% to 100%), though time since symptom onset was not recorded. To be conservative about test accuracy, we use validation data from [4].The range from 90% (95% CI: 74.4%-96.5%)[5] to 96.9% (95% CI: 89.5%-99.5%)[3], with specificity in those studies was 100% (95% CI: 95.4%-100%)[5] and 99.90%[3], respectively.

Outcomes

Our primary outcome is the age-and-sex adjusted proportion of positive tests for each of 6 study areas. This is the outcome we used to compute our target sample size. Our secondary outcomes are (a) unadjusted proportion of positive tests by demographic group at each site, (b) the age-sex-and-population-weighted proportion of positive tests at the slum and non-slum level across all 3 wards, and (d) the adjusted prevalence accounting for the imperfect accuracy of tests.

Statistical methods

We estimated the unadjusted proportion of positive tests $\hat{p}_{ijk}$ in a demographic group i in a community j in ward k (i.e., secondary outcome (a)) as the ratio of the number of positive test results, using Abbott's recommended cutoff (1.4) for its CLIA test, and the number of participants that gave an adequate sample. We calculated this proportion for each of 8 demographic strata (i) at each of 6 locations, defined as the slums and non-slum communities (j) for each of the 3 wards (k). We provide Score confidence intervals for these proportions as Score intervals have been shown to have interval coverage probability closer to 95% than even exact intervals.[6]

We estimated the age-and-sex-weighted adjusted proportions of positive tests $\hat{p}_{jk}$ in community j in ward k (i.e., the primary outcome) as the weighted average of positive proportions in each demographic group in that location, $\hat{p}_{jk} = \sum_i p_{ijk} f_{ijk}$, where the weights (Supplement Table e3) are the fraction f_{ijk} of the population in demographic group (i) in location ($j \times k$) and $\sum_i f_{ijk} = 1$. Sampling weights were estimated from our survey, which asked how many people in each demographic group reside in a respondent's household. The confidence intervals for these weighted proportions are adjusted Wald intervals (i.e., normal approximations using an adjusted measure of proportion positive). Following Price & Bonnet (2004), these adjusted Wald intervals add $h = 2/8$ successes and $2h$ tests to each strata (i, j, k) before calculating adjusted proportions for each strata and then calculated Wald intervals using these adjusted proportions. The

denominator is 8 because there are 8 demographic strata per location. These intervals have been shown to have better coverage probabilities than Wald intervals (normal approximation with unadjusted proportions) or Score intervals.[7]

We estimated the age-sex-and-population-weighted proportion of positive tests at the slum or non-slum level across wards as the weighted average of positive proportions in each type of community, $\hat{p}_j = \sum_k p_{jk} f_{jk}$, where j indexes types of community (slum or non-slum), k indexes the three wards, and the weights f_{jk} are the fraction of slum (or non-slum) population across the 3 wards that are in ward j. Community level populations are drawn from the 2011 Census. The confidence intervals for these weighted proportions are adjusted Wald intervals. Following Price & Bonnet (2004), these adjusted Wald intervals add $h = 2/24$ successes and h failures to each strata (i, j, k) before calculating adjusted proportions for each strata and then calculated Wald intervals using these adjusted proportions. The denominator is 24 because there are 24 strata (8 demographic x 3 wards) for slums and for non-slums.

We estimate the adjusted prevalence using the Rogan-Gladen[8] correction for imperfect accuracy of tests after calculating weighted proportions. We employ weighted proportions to calculate the variance of adjusted prevalence and then employ normal approximations to estimate Wald confidence intervals for that prevalence. We do not use adjusted Wald or Score intervals because there is no empirically validated method to do so for a affine transform of the weighted proportion, the sort of transform that the Rogan-Gladen correction is.

The Rogan-Gladen estimator has two weakneses. First, it requires knowledge of sensitivity and specificity. In our case, however, we have several estimates of this for the CLIA test we use. Second, when positive rates are very high or low they can produce estimates above or below one. However, our positive rates are neither close to 1 or 0 and so the estimator does not produce estimates that are "out of bounds." Even when estimates are out of bounds, however, the implication is that estimates of sensitivity and specificity must be wrong, not that the Rogan-Gladen formula, which come from Bayes Theorem, is wrong[9]. In our case, we have no indication that our estimates of sensitivity and specificity are wrong because our estimates are in bounds

One can use Bayesian methods, as recommended by Lewis and Torgerson (2012)[9], in lieu of the Rogan-Gladen formula. But unless one has priors on infection, Bayesian methods with uninformative priors produce largely the same results as Maximum Likelihood Estimators (MLE). We do not have strong priors on prevalence. The big advantage of Rogan-Gladen over MLE is

that it is much simpler to implement. Moreover, as sample sizes get large, the MLE and Rogan-Gladen formula should converge. In our sample, the community with the smallest sample size still has an N=564.

The estimated sensitivity of CLIA tests range from 90% (95% CI: 74.4% to 96.5%)[5] to 96.9% (95% CI: 89.5% to 99.5%)[3], while specificity in those studies was 100% (95% CI: 95.4% to 100%)[5] and 99.90%[3], respectively. We present estimates of prevalence assuming both low[5] and high[3] estimates of sensitivity reported and associated estimates of specificity.

All analyses are conducted in Excel 2016 or in Stata Version 16 (StataCorp).

Results

Table e2 provides data on the breakdown of our sample by age group, gender, and location, where location is defined both by ward and whether the community is a slum or not. Table e4 provides additional demographic detail by aggregating, without weights, across slums and across non-slums. Our sample has fewer females in non-slums. It has few members in the youngest age group (12-24 years old) and fewer in the oldest age group (61 and older) in slums.

Although the age distribution in our sample is consistent with that in the general population (Table e3), which is younger in slums, the distribution across gender is skewed more towards males in our sample.

We use the age distribution of the slum and non-slum population in Table e3, gathered from surveys of the demographic composition of sample households, as weights to convert positive test rates at the demographic group and location level to positive test rates at the location level. Although the data from the non-slums of Dahisar have a similar age-and-sex distribution as those of non-slums in other sample wards, the household size in non-slums of Dahisar is much larger. Using the non-slum distribution of other wards to reweight Dahisar data does not materially affect our estimates of adjusted proportions or adjusted seroprevalence.

Table e5 reports the results of a binomial regression (generalized linear model with binomial family) of proportion of positive test on indicators for different demographic groups (indicators for age groups 12-24, 25-39, 40-60 and over 61 and for females). Observations are at the individual level. This regression is run separately on a sample from non-slums and then one from slums. We report coefficients as risk differences. The table permits examination of whether

unadjusted proportions across demographic groups within each type of community are significantly different from one another.

We examine whether age-and-sex-weighting causes significant changes to unadjusted positive proportions in Table e6, which reports unweighted proportions and age-and-sex-weighted proportions for each of our 6 sites, along with Wald adjusted confidence intervals.

We examine the sensitivity of our estimates of positive test rates to the cutoff of IgG titer used to label test results as positive or negative. The manufacturer recommends a cutoff of 1.4 for its Architect CLIA test for N-antibodies to SARS-CoV-2. Figure e1 shows that the cumulative distribution of IgG titers in slums stochastically dominates that in non-slums, implying that average titer is higher in slums. Non-slums have a much higher concentration (roughly 75% of the population) than slums (roughly 25%) of IgG concentrations near zero. While there are no jumps in the data around the 1.4 cutoff, changes in the cutoff impact estimates of the positive rate in slums much more than non-slums (Table e5).

Supplementary tables and figures

Table e1

Estimated prevalence for sample-size calculations

Ward Name	Ward No	Number infected by June 2	Population (Census 2011, 100,000's)	Confirmed infections per person (2 June 2020)	Scaled prevalence est. (based on model)	Samples to get 1.5% half width
Matunga	FN	2726	5.29	0.52%	15.61%	2,249
Chembur West	MW	1446	4.12	0.35%	10.63%	1,622
Dahisar	RN	486	4.31	0.11%	3.42%	576
Total Mumbai		41068	124.4	0.33%	10.00%	

Note. Total Mumbai prevalence is based on simulation of an agent-based model for spread of infection. The simulation used as inputs confirmed cases in Mumbai until June 2.

Table e2

Demographic profile of sample, by location.

		Non-slum			Slum		
Ward	Age bins	Female	Male	Total	Female	Male	Total
Matunga	12-24	21	38	59	203	211	414
	25-39	84	217	301	407	304	711
	40-60	180	429	609	492	330	822
	61+	61	153	214	79	95	174
	Total	346	837	1183	1181	940	2121
Chembur West	12-24	22	47	69	94	109	203
	25-39	82	193	275	304	249	553
	40-60	136	316	452	392	266	658
	61+	42	103	145	49	48	97
	Total	282	659	941	839	672	1511
Dahisar	12-24	13	35	48	27	58	85
	25-39	53	104	157	106	110	216
	40-60	96	192	288	101	114	215
	61+	12	73	85	26	28	54
	Total	174	404	578	260	310	570

Note. This table presents the number of sample members in each age and gender bin, by location.

Table e3

Age distribution among the population in each community, by sex, ward and slum status

		Non-slums			Slums		
Ward	Age group	No of hhd members	Male (share)	Female (share)	No of hhd members	Male (share)	Female (share)
Matunga	All	3079	0.551	0.449	8156	0.506	0.494
	12-24	540	0.100	0.075	2697	0.177	0.154
	25-39	812	0.143	0.121	2930	0.177	0.182
	40-60	1173	0.211	0.170	2083	0.123	0.132
	61+	554	0.097	0.082	446	0.029	0.026
Chembur West	All	3175	0.523	0.477	5848	0.503	0.497
	12-24	633	0.110	0.089	1569	0.143	0.125
	25-39	871	0.141	0.133	1973	0.171	0.166
	40-60	1160	0.190	0.175	1920	0.158	0.170
	61+	511	0.081	0.080	386	0.031	0.035
Dahisar	All	4928	0.523	0.477	2480	0.546	0.454
	12-24	1041	0.108	0.103	834	0.204	0.133
	25-39	1367	0.143	0.134	707	0.147	0.138
	40-60	1776	0.190	0.170	674	0.140	0.131
	61+	744	0.081	0.070	265	0.055	0.052

Note. These age distributions are calculated from surveys of sample households. The surveys asked the number of individuals in age group by gender in the sample member's household. The survey enumerates high household size in non-slums of Dahisar. The fractions sum to one within a location. Calculations drop observations where sample member does not answer question about household composition.

Table e4

Demographic features of the sample, unweighted and by residence in slum

	Count	Mean	(Exact 95% Conf. Int.)	
Non-slum	2702			
Female	802	0.297	0.280	0.314
Ages 12-24	176	0.065	0.056	0.075
Ages 25-39	733	0.271	0.255	0.288
Ages 40-60	1349	0.499	0.480	0.518
Ages 61+	444	0.164	0.151	0.179
Slum	4202			
Female	2280	0.543	0.527	0.558
Ages 12-24	702	0.167	0.156	0.179
Ages 25-39	1480	0.352	0.338	0.367
Ages 40-60	1695	0.403	0.388	0.418
Ages 61+	325	0.077	0.069	0.086

Note. Table presents the fraction of slum and non-slum population that are female and in each age group.

Table e5

Correlation between the proportion of positive tests and demographics

Community:	Non-slum	Slum
Dependent variable:	Positive test	Positive test
Age 12-24	0.183	0.530
	(0.131 - 0.235)	(0.463 - 0.597)
Age 25-39	0.158	0.480
	(0.154 - 0.162)	(0.459 - 0.502)
Age 40-60	0.150	0.563
	(0.142 - 0.157)	(0.519 - 0.608)
Age > 60	0.122	0.599
	(0.115 - 0.129)	(0.569 - 0.628)
Female	0.0182	0.0604
	(0.00467 - 0.0317)	(0.0227 - 0.0980)
Observations	2,702	4,202
Dep. var. mean	0.155	0.564
Dep. var. SD	0.362	0.496

Note. Coefficients are from a binomial regression (generalized linear model with binomial family) of CLIA test result with manufacturer recommended cutoff of 1.4 on indicators for age and sex. The regression in the first column includes only samples from non-slums. That in the second column includes only samples from slums. Each sample is weighted equally.

Table e6

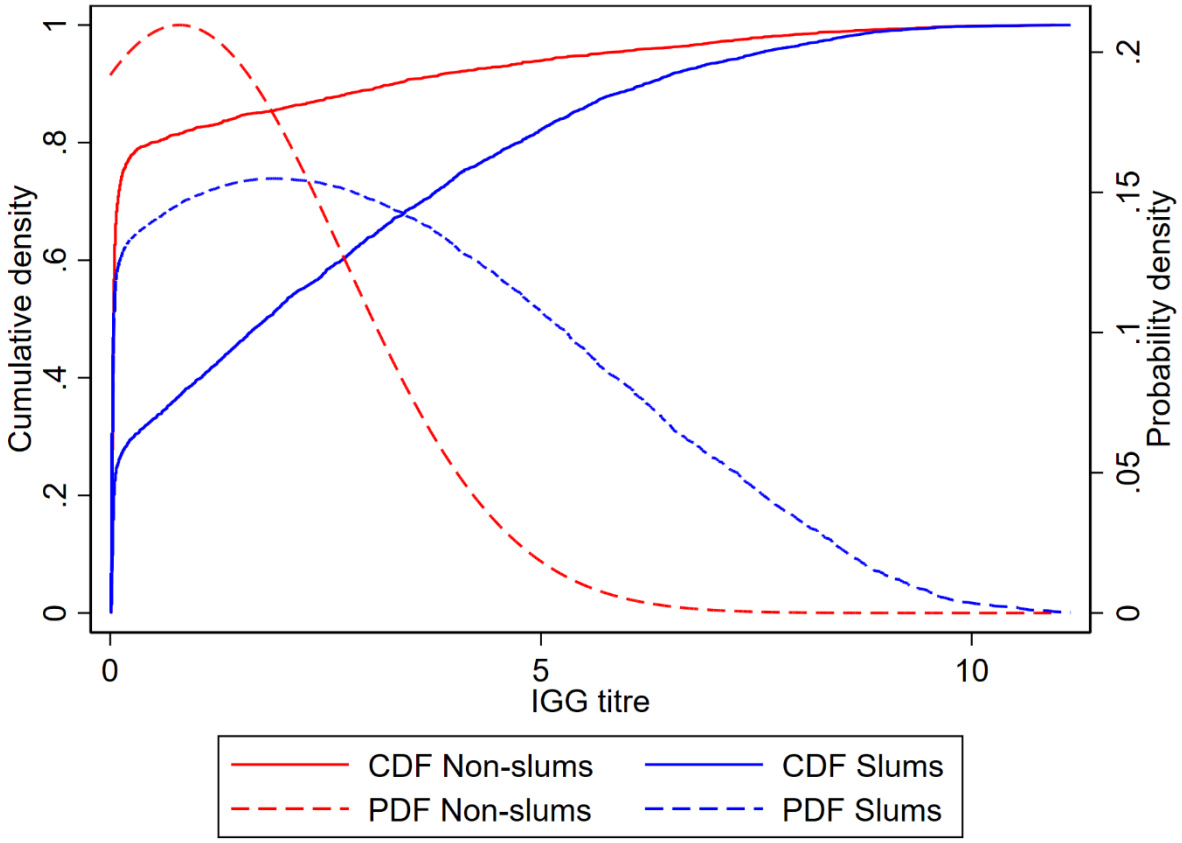
Unadjusted and adjusted proportion of positive tests (by location)

		Unadjusted proportion of positive tests		Adjusted proportion of positive tests	
Location		Rate	95% CI	Rate	95% CI
Non-slums					
	Matunga	0.169	(.148,.191)	0.176	(.15,.202)
	Chembur West	0.166	(.142,.19)	0.165	(.137,.194)
	Dahisar	0.116	(.09,.143)	0.119	(.085,.153)
Slum					
	Matunga	0.582	(.561,.603)	0.570	(.547,.592)
	Chembur West	0.572	(.547,.597)	0.551	(.524,.578)
	Dahisar	0.523	(.482,.563)	0.511	(.464,.558)

Note. The unadjusted proportion of positive tests is a simple average across 8 demographic groups. The adjusted proportion of positive tests is estimated by, first, estimating the proportion for each demographic group and, then, calculating a weighted average that ensures each group's proportion has a weight proportional to its share of the population in a location, with weights as indicated in Table e3. Confidence interval is estimated using a normal approximation.

Figure e1

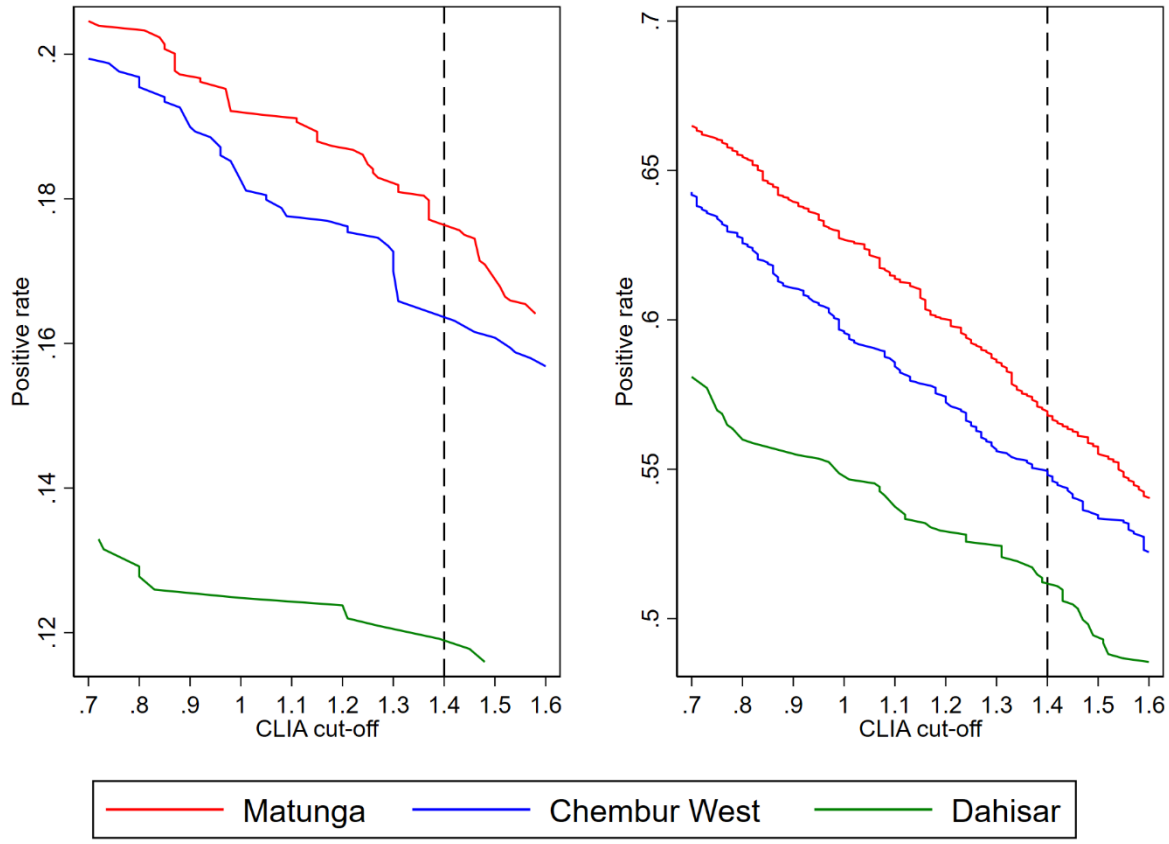
Empirical cumulative and probability density functions for IGG scores, separately for slum and non-slum communities and weighted to reflect age and gender distribution in the population.



Note. Within slums and within non-slums in each of 3 wards, the positive rates are first weighted to ensure positive rates reflect population weighting of age and genders. When aggregating across wards to calculate slum and non-slum positive rates, we weight each ward equally.

Figure e2

Relationship between weighted positive test rate and CLIA cut-off value, by ward for non-slums and slums



Note. CLIA cutoff recommended by Abbott, the manufacturer, is 1.4. Range for CLIA cutoff is the range analyzed by Bryan et al. (2020).[3] Within slums and within non-slums in each of 3 wards, the positive rates are weighted to ensure positive rate reflect population weighting of age and genders.

References

1. Municipal Corporation of Greater Mumbai. Vital Statistics Report. 2020. https://portal.mcgm.gov.in/irj/portal/anonymous/qlvitalstatsreport.
2. Harsha P, Juneja S, Patil P, et al. COVID-19 Epidemic Study II: Phased Emergence From the Lockdown in Mumbai. arXiv; 2020.
3. Bryan A, Pepper G, Wener MH, et al. Performance Characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and Seroprevalence in Boise, Idaho. *Journal of Clinical Microbiology* 2020: JCM.00941-20.
4. Public Health England, Oxford University Hospitals. Evaluation of sensitivity and specificity of four commercially available SARS-CoV-2 antibody immunoassays. 2020.
5. U.S. Food and Drug Administration. Serology Test Evaluation Report for "Architect i1000 SARS-CoV-2 IgG" from Abbott. 2020.
6. Agresti A, Coull BA. Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician* 1998; **52**(2): 119-26.
7. Price RM, Bonett DG. An improved confidence interval for a linear function of binomial proportions. *Computational Statistics & Data Analysis* 2004; **45**(3): 449-56.
8. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *American journal of epidemiology* 1978; **107**(1): 71-6.
9. Lewis FI, Torgerson PR. A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic. *Emerging Themes in Epidemiology* 2012; **9**(1): 9.