8-2020

Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York

Eli S. Rosenberg
University at Albany, State University of New York

James M. Tesoriero
University at Albany, State University of New York

Elizabeth M. Rosenthal
University at Albany, State University of New York

Rakkoo Chung
University at Albany, State University of New York

Meredith A. Barranco
University at Albany, State University of New York

See next page for additional authors

Authors

Eli S. Rosenberg, James M. Tesoriero, Elizabeth M. Rosenthal, Rakkoo Chung, Meredith A. Barranco, Linda M. Styer, Monica M. Parker, Shu-Yin John Leung, Johanne E. Morne, Danielle Greene, David R. Holtgrave, Dina Hoefer, Jessica Kumar, Tomoko Udo, Brad Hutton, and Howard A. Zucker

Annals of Epidemiology 48 (2020) 23–29

Contents lists available at ScienceDirect

Annals of Epidemiology

Original article

Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York

Eli S. Rosenberg, PhD [a,*], James M. Tesoriero, PhD [b], Elizabeth M. Rosenthal, MPH [a], Rakkoo Chung, PhD [b], Meredith A. Barranco, MPH [a], Linda M. Styer, PhD [c], Monica M. Parker, PhD [c], Shu-Yin John Leung, MA [b], Johanne E. Morne, MS [b], Danielle Greene, DrPH [b], David R. Holtgrave, PhD [a], Dina Hoefer, PhD [b], Jessica Kumar, DO [b], Tomoko Udo, PhD [a], Brad Hutton, MPH [b], Howard A. Zucker, MD [b]

[a] University at Albany School of Public Health, State University of New York, Rensselaer
[b] New York State Department of Health, Albany, NY
[c] Wadsworth Center, New York State Department of Health, Albany, NY

ARTICLE INFO

Article history:
Received 3 June 2020
Accepted 10 June 2020
Available online 17 June 2020

Keywords:
Coronavirus
Infectious diseases
Epidemiology
Surveillance
Epidemics
Seroepidemiologic studies
Seroprevalence

ABSTRACT

Purpose: New York State (NYS) is an epicenter of the SARS-CoV-2 pandemic in the United States. Reliable estimates of cumulative incidence in the population are critical to tracking the extent of transmission and informing policies.
Methods: We conducted a statewide seroprevalence study in a 15,101 patron convenience sample at 99 grocery stores in 26 counties throughout NYS. SARS-CoV-2 cumulative incidence was estimated from antibody reactivity by first poststratification weighting and then adjusting by antibody test characteristics. The percent diagnosed was estimated by dividing the number of diagnoses by the number of estimated infection-experienced adults.
Results: Based on 1887 of 15,101 (12.5%) reactive results, estimated cumulative incidence through March 29 was 14.0% (95% confidence interval [CI]: 13.3%–14.7%), corresponding to 2,139,300 (95% CI: 2,035,800–2,242,800) infection-experienced adults. Cumulative incidence was highest in New York City 22.7% (95% CI: 21.5%–24.0%) and higher among Hispanic/Latino (29.2%), non-Hispanic black/African American (20.2%), and non-Hispanic Asian (12.4%) than non-Hispanic white adults (8.1%, $P < .0001$). An estimated 8.9% (95% CI: 8.4%–9.3%) of infections in NYS were diagnosed, with diagnosis highest among adults aged 55 years or older (11.3%, 95% CI: 10.4%–12.2%).
Conclusions: From the largest U.S. serosurvey to date, we estimated >2 million adult New York residents were infected through late March, with substantial disparities, although cumulative incidence remained less than herd immunity thresholds. Monitoring, testing, and contact tracing remain essential public health strategies.

Introduction

The first cases of COVID-19 were identified in New York State (NYS) in early March, 2020, and since then NYS, particularly the metropolitan New York City (NYC) area, has become one of the most-impacted communities in the United States [1,2]. As of June 2, 2020, over 370,000 laboratory-confirmed diagnoses have been made, accounting for approximately 25% of diagnoses in the United States [2,3]. As with most infections, laboratory-confirmed diagnoses undercount the true population-level burden of infections; with SARS-CoV-2, the virus that causes COVID-19, key factors that contribute to underdiagnosis include absent or mild symptoms and access to testing [4]. Thus, although NYS has tested more residents for COVID-19 than any other state (over 2,229,000 persons tested through June 2, 2020), it is likely that laboratory-confirmed cases represent a relatively small portion of the total number of persons with a history of infection in NYS [3].

* Corresponding author. Department of Epidemiology and Biostatistics, University at Albany, School of Public Health, 1 University Place, Room 123, Rensselaer, NY 12144. Tel.: +1-518-486-9667; fax: +1 518-402-0380.
E-mail address: erosenberg2@albany.edu (E.S. Rosenberg).

https://doi.org/10.1016/j.annepidem.2020.06.004
1047-2797/

Estimates of COVID-19 cumulative incidence (i.e., prevalence of previous or current infection) can inform the extent of epidemic spread and the number of persons still susceptible and progress toward herd immunity, which are critical for parameterizing simulation models and informing policies, including those for altering societal restrictions [5]. Furthermore, such data provide needed denominators for understanding the extent of diagnosis, rates of hospitalization, morbidity, and mortality, and geographic differences.

Antibody testing for SARS-CoV-2 has emerged as an important tool for understanding infection history. Although a several-week window period for development of IgG antibodies and evidence that not all persons with infection develop an antibody response limit their utility for diagnostics, and their interpretation for short- and long-term immunity remain uncertain, as with other infections, antibody prevalence serostudies with validated assays can assess population-level cumulative incidence in the recent past [6–11].

Antibody serostudies for SARS-CoV-2 are being conducted in other countries and in the United States are occurring on the national and county levels, but none have been conducted at the state level, and only one population-based serostudy has been peer-reviewed [12–15]. The current array of recommendations against individual movement and business operation during the pandemic complicates study specimen collection. A recent RNA survey in Iceland and serosurveys in two California counties conducted sampling at centralized testing sites, which offer ease of execution particularly in small geographies, with potentially large self-selection biases [13,15,16]. Alternative approaches include random at-home mail-in testing and community-intercept studies in high-traffic locations that remain open [14].

To provide a statewide picture of COVID-19 infection through late-March and diagnoses by early-April 2020, during April 19–28, 2020, the NYS Department of Health (NYSDOH) conducted a community-based serostudy throughout NYS. Cumulative incidence among non-institutionalized adults, by geographic and demographic features, was estimated from weighted reactivity rates that were adjusted for validated test characteristics. Combining these findings with cumulative diagnoses enabled estimation of the percent of infections diagnosed.

Methods

Field study

The NYSDOH conducted a convenience sample of over 15,000 New Yorkers attending 99 grocery stores across 26 counties, which contain 87.3% of the state's population, located in all regions of NYS (Fig. 1). Grocery stores were chosen as the testing venue because they were classified an essential business to remain open and, due to the necessity of grocery shopping, they attract a heterogeneous clientele [17]. Store locations were chosen to increase sample coverage of the racial and ethnic diversity of the statewide population.

Testing occurred over 6 distinct days from April 19, 2020 through April 28, 2020. Each store had a team of 6–8 staff responsible for recruiting participants, collecting specimens, recording data, and managing specimen transport to Wadsworth Center Laboratory (Albany, NY) for analysis. Eligible subjects were adults aged 18 years or older, New York residents irrespective of county, recruited through a recruitment flyer posted at stores and by systematically approaching each patron as they entered the store. To minimize selection bias, community testing site locations were not announced ahead of time and were changed frequently (i.e., 99 venues in 6 days). Most locations were used only once, and no individual site was used more than twice. Testing was halted at locations that became publicized on social media.

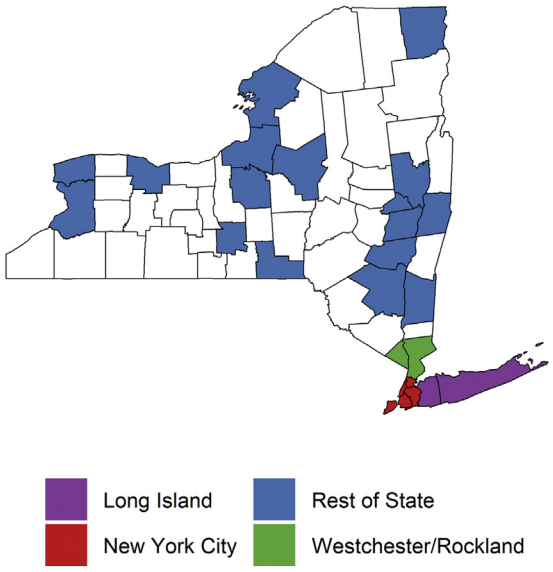

Fig. 1. New York State counties included in the New York State Department of Health Serological Testing Survey[1]. [1]Sampled counties—Long Island: Nassau, Suffolk; New York City: Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island; Westchester, Rockland Counties; Rest of State: Albany, Broome, Clinton, Dutchess, Erie, Greene, Jefferson, Monroe, Niagara, Oneida, Onondaga, Oswego, Rensselaer, Saratoga, Schenectady, Tompkins, and Ulster.

Patrons were given information about the testing, and if interested, completed written informed consent. Procedures included a brief demographic questionnaire and dried blood spot (DBS) collection by trained personnel. Approximately 13% of participants initially had missing demographic data. Staff attempted to capture these data through >2500 follow-up phone calls, reaching all but approximately 75 participants, who were subsequently excluded from analyses. Test results were delivered to participants by text message if nonreactive and by phone if indeterminant or reactive.

Testing approach

Blood was collected by fingerstick onto custom 903 filter paper cards labeled with a specimen ID. Cards were dried for 3–4 hours at ambient temperature and transported to the Wadsworth Center. A fully saturated greater than or equal to 3-mm-diameter DBS was required. A total of 525 DBS cards from eligible individuals were rejected; 433 with insufficient or improperly collected blood, 92 with no specimen ID. Acceptable DBS cards were processed for testing.

SARS-CoV IgG testing was conducted using a microsphere immunoassay developed and validated for DBS by the NYSDOH Wadsworth Center. Briefly, nucleocapsid (N) antigen-coupled magnetic beads were incubated with blood eluted from a 3-mm DBS punch. Phycoerythrin-labeled goat anti-human IgG secondary antibody was used to detect microsphere-bound IgG antibodies, and the median fluorescence intensity (MFI) was determined using a FlexMap 3D (Luminex Corp., Austin, TX). The mean MFI of 90–100 negative DBS was used to set cutoffs; results greater than the mean MFI plus 6 SDs were reported as reactive; results less than the mean MFI + 3 SDs were nonreactive, and results between the mean MFI +3 to +6 were indeterminate. Serosurvey testing was initiated with SARS-CoV IgG v1, which used SARS-CoV-1N antigen (Wadsworth Center, Albany, NY) and was completed using SARS-CoV IgG

Table 1
Reactivity and test characteristic–adjusted cumulative incidence of COVID-19, overall and by demographic factors and region

Stratum	% Of adults in New York	Reactivity		Test characteristic–adjusted estimated cumulative incidence			*P*-value
		Unweighted # reactive/Total sample	Weighted percent	% (95% CI)	Infection-experienced adults* (95% CI)	% Of infection-experienced adults	
Overall	100.0	1887/15,101	12.5	14.0 (13.3–14.7)	2,139,300 (2,035,800–2,242,800)	100.0	
Sex							.03
Male	47.6	918/6635	13.2	14.8 (13.8–15.8)	1,076,500 (1,001,900–1,151,100)	50.3	
Female	52.4	969/8466	11.9	13.3 (12.4–14.2)	1,062,200 (990,500–1,133,800)	49.7	
Race and ethnicity							<.0001
Hispanic or Latino	17.4	757/2735	25.8	29.2 (27.2–31.2)	775,800 (722,700–829,000)	36.6	
Non-Hispanic white	58.0	623/9545	7.3	8.1 (7.4–8.7)	715,400 (657,100–773,700)	33.7	
Non-Hispanic black/African American	13.9	388/1913	18.0	20.2 (18.1–22.3)	428,000 (382,700–473,400)	20.2	
Non-Hispanic Asian	8.6	75/629	11.1	12.4 (9.4–15.4)	161,700 (122,600–200,800)	7.6	
Multiracial/other	2.1	44/279	10.7	11.9 (6.4–17.5)	38,800 (20,800–56,800)	1.8	
Age group (y)							.0002
18–34	30.7	377/3151	13.0	14.6 (13.1–16.1)	682,600 (612,000–753,200)	31.8	
35–44	15.9	334/2628	13.7	15.3 (13.7–17.0)	371,800 (331,700–411,900)	17.3	
45–54	17.4	479/3345	14.3	16.0 (14.6–17.5)	424,700 (386,400–463,100)	19.8	
55+	36.1	697/5977	10.9	12.1 (11.2–13.1)	667,800 (615,600–719,900)	31.1	
Region							<.0001
New York City †	43.3	1319/5946	20.2	22.7 (21.5–24.0)	1,504,400 (1,421,300–1,587,500)	70.1	
Westchester/Rockland Counties	6.4	134/980	14.4	16.1 (13.2–19.0)	156,500 (128,400–184,600)	7.3	
Long Island ‡	14.4	241/2074	11.9	13.2 (11.4–15.1)	291,800 (250,600–332,900)	13.6	
Rest of NYS §	35.9	193/6101	3.4	3.6 (3.0–4.1)	194,600 (162,600–226,600)	9.1	

* Stratified estimates may not exactly sum to total because of rounding and differences between the weighting scheme and noninstitutionalized population totals.
† Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
‡ Nassau and Suffolk Counties.
§ Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, and Yates counties.

v2, which used SARS-CoV-2N protein (Sino Biological, Wayne, PA) after validation studies confirmed comparable performance.

Assay validation studies are described in Supplementary Tables 1–5 and Supplementary Figure 1. Specificity was estimated as 99.75%, based on two studies of DBS collected before December 2019 that found 99.5% (95% confidence interval [CI]: 98.5%–100%) and 100% (95% CI: 96.1%–100%) specificity. Serum collected from individuals diagnosed with non–COVID-19 respiratory and non-respiratory agents were tested to assess cross-reactivity; only 1 of 85 samples was reactive. Of 232 SARS-CoV-2 PCR-positive DBS collected, a median of 35 days after the symptom onset, 204 (87.9%, 95% CI: [83.7%–92.1%]) were reactive, informing sensitivity and thus incorporating both test performance and the proportion of infected persons who never develop IgG [6,7].

Analysis

We estimated SARS-Cov-2 cumulative incidence from observed antibody reactivity using two sequential steps: (1) post-stratification weighting to standardize to the NYS population and (2) adjustment by estimated antibody test characteristics.

Using the National Center for Health Statistics bridged-race file, weights were assigned to each participant based on their membership in each of 160 strata of sex, race, and ethnicity (Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic Asian, and non-Hispanic other), age (18–34, 35–44, 45–54, ≥55 years), and residential region (NYC, Westchester/Rockland, Long Island, Rest of State [ROS]) [18]. Poststratification weights were defined as the proportion each stratum is represented in the state's population divided by the analogous proportion in the sample [19,20]. Next, we computed weighted frequencies for the percent reactive statewide, with one-way stratifications by sex, race and ethnicity, age group, and region and two-way stratifications within levels of region, including 95% CIs, with differences assessed using Rao-Scott χ^2 tests [21]. Indeterminate results were assumed nonreactive, and statistical procedures were two-sided at $\alpha = 0.05$.

In the second step, weighted reactivity estimates ($p_{reactive}$) and their 95% CI bounds were corrected for test sensitivity and specificity, based on validation data, to yield cumulative incidence, per Bayes' Rule as applied to the diagnostic 2 × 2 table: $cumulative\ incidence = \frac{p_{reactive}+specificity-1}{sensitivity+specificity-1}$ [13,22]. Primary analyses used the sensitivity and specificity point estimates from the validation studies, with sensitivity analyses at the extremes of test characteristics' 95% CI ([96.1% specificity, 92.1% sensitivity], [100% specificity, 83.7% sensitivity]). Test characteristic–adjusted cumulative incidence values were multiplied by the one- and two-way noninstitutionalized adult populations (e.g., excluding settings such as prisons and nursing homes) from the American Community Survey 2014–2018 Public Use Microdata Sample file [23]. This yielded the estimated total 'infection-experienced' adults with SARS-CoV-2 within each stratum. With a study midpoint of April 23, and literature estimates of mean 4 days from infection to the symptom onset and mean 21 days from the onset to IgG detection, results represent cumulative incidence through approximately March 29 [6,8,24].

In NYS, diagnostic testing for SARS-CoV-2 is mandatorily reported electronically to the NYSDOH. Using cumulative diagnoses reported and the total numbers of infection-experienced adults, we estimated the percent of infections diagnosed overall and by region, sex, and age. For primary analyses, we accumulated diagnoses through April 9, based on the March 29 final infection date, 4 days to the symptom onset, and mean 7 days from the onset to diagnosis. Supplemental upper-bound estimates used the last plausible diagnosis date of May 8, based on the April 28 final study day, 4 days being earliest time from the onset to IgG detection and allowing PCR detection up to 14 days after the onset [8].

Results

Across NYS, a total of 15,626 adult residents with complete data were tested, of whom 15,101 (96.6%) had suitable specimens, of which 1887 (12.5%) were reactive and 340 (2.3%) indeterminate. After weighting, 12.5% were estimated reactive and after further adjustment for test characteristics, estimated cumulative incidence was 14.0% (95% CI: 13.3%–14.7%), corresponding to 2,139,300 (95% CI: 2,035,800–2,242,800) infection-experienced adults in NYS through approximately March 29 (Table 1). In sensitivity analyses at the extremes of test characteristics, the cumulative incidence ranged from 9.8% (95% CI: [9.1%–10.5%]) to 15.0% (95% CI: [14.3%–15.7%]), representing a total of 1,494,700 (95% CI: [1,391,800–1,597,600]) to 2,286,600 (95% CI: [2,178,200–2,395,100]) adults in NYS (Supplementary Tables 6 and 7).

The cumulative incidence was higher among males (14.8%, 95% CI: [13.8%–15.8%]) than females (13.3%, 95% CI: [12.4%–14.2%], $P = .03$), with males comprising 50.3% of adult infections. This differed significantly by race and ethnicity, with Hispanic/Latino (29.2%, 95% CI: 27.2%–31.2%), non-Hispanic black/African American (20.2% [95% CI, 18.1%–22.3%]), and non-Hispanic Asian (12.4%, 95% CI: [9.4%–15.4%]) adults having higher cumulative incidence than non-Hispanic white adults (8.1%, 95% CI: [7.4%–8.7%], $P < .0001$). Given these differences, Hispanics comprised the plurality (36.6%) of infection-experienced adults. Significant differences were also observed by age ($P = .0002$), ranging from highest levels among persons 45–54 years old (16.0%, 95% CI: [14.6%–17.5%]) to lowest among persons aged 55 years or older (12.1% [95% CI: 11.2%–13.1%]).

We observed regional heterogeneity in cumulative incidence, ranging from 22.7% (95% CI: 21.5%–24.0%) in NYC residents, to 16.1% (95% CI: 13.2%–19.0%) and 13.2% (11.4%–15.1%) in the respective metropolitan areas of Westchester/Rockland Counties and Long Island, to 3.6% (95% CI: [3.0–4.1]) in ROS ($P < .0001$). Demographic patterns were heterogenous by region (Table 2). Males had significantly higher cumulative incidence in all regions outside of, but not within, NYC. The patterns of racial disparity observed statewide were similar and statistically significant within NYC, Westchester/Rockland, and Long Island, but not in ROS. In each of the former 3 regions, Hispanic/Latino persons represented greater than 37% of infection-experienced adults, whereas in the latter non-Hispanic whites comprised a majority of infection-experienced adults (79.4%).

An estimated 8.9% (95% CI: 8.4%–9.3%) of infections in NYS were diagnosed as of April 9, 2020 (Table 3). Males (9.4%, 95% CI: 8.8%–10.1%) had higher diagnosis levels than females (8.2%, 95% CI: 7.7%–8.8%). Those aged 55 years or older were most likely to be diagnosed (11.3%, 95% CI: 10.4%–12.2%). Diagnosis rates in NYC (7.1%, 95% CI: 6.7%–7.5%) and ROS (7.5%, 95% CI: 6.4%–8.9%) were about half those observed in the other regions. Considering the May 8 upper bound for diagnoses, a maximum of 15.7% (95% CI: 15.0%–16.5%) of overall infections could have been diagnosed, with similar patterns observed across levels of each factor (Supplementary Table 8).

Discussion

From the largest U.S. SARS-CoV-2 serosurvey to date, we estimated that over 2 million adult NYS residents were infected through the end of March. Our findings estimate the extent of transmission of and community experience with SARS-CoV-2, particularly in the NYC metropolitan region. Despite large numbers of persons acquiring SARS-CoV-2, this represents only 14.0% of adult residents, suggesting that, even in this COVID-19 epicenter, the epidemic is substantially less than the estimated ~70% U.S. herd immunity threshold [25]. Against this remaining epidemic potential, ongoing vigilance through rigorous and extensive epidemic monitoring, testing, and contact tracing is a necessary component for predicting, preventing, and/or mitigating a second epidemic wave, consistent with state and federal guidance for reopening [5,26]. This vigilance is needed even in the rest of NYS outside the metropolitan region, which are in the first phases of reopening in NYS, and where lowest cumulative incidence suggests the highest proportion susceptible.

Our finding of higher cumulative incidence in the regions of the NYC metropolitan area, particularly NYC, is consistent with the known distribution of diagnoses. Furthermore, in these regions of high urbanicity, significant racial/ethnic disparities in infection history were found, with minority communities experiencing disproportionate risk. The drivers of greater COVID-19 risk and disparities in urban areas continue to be studied, but may relate to population density and the mechanisms by which transportation, employment, housing, and other socioeconomic or environmental factors shape opportunities for transmission [27–29]. A recent NYS study on a random sample of COVID-19 hospitalizations showed limited racial/ethnic differences in clinical outcomes, suggesting that observed differences in mortality by race and ethnicity may be in large part driven by different infection histories in the community [3,30–32]. Research is needed to understand the drivers of increased COVID-19 risk experienced by minority communities, followed by actions to improve health equity.

The finding that over 8.9% of adults were diagnosed reveals both the opportunities for further expansion of diagnostic testing in NYS, yet in the context of far higher diagnosis and testing levels than other U.S. settings suggests substantial progress to date [1,13]. Compared to all persons with infection history, there was a higher representation of males and those older than 55 years among diagnosed persons. Given the lower reactivity rates observed among this age group, our results expand observations from previous studies that older adults may be more likely to exhibit symptoms or illness or be more likely to seek care [30,33–35].

Although not an aim of this analysis, we note that in conjunction with 12,822 publicly reported COVID-19 deaths for NYS through April 17 (reflecting median 19 days-post-infection to death), our findings suggest an infection fatality ratio of 0.6%. This estimate is in line with estimates of 0.5%–1.0% observed in other countries; however, additional analyses are needed to more precisely estimate the infection fatality ratio in NYS [36,37].

Strengths of our study include a large sample, which contained 0.1% of the adult NYS population, and a systematic sampling approach in one of the only open public venues in the state, where a necessary commodity is purchased. Although a convenience sample, survey weights adjusted for biased demographic/geographic representation, noting that the general agreement of unweighted and weighted results suggests demographic representativeness of the study sample, and we further adjusted results for assay performance, under varied scenarios. Our study may nevertheless be limited by residual nonrepresentativeness of the underlying population. This includes potential undersampling of persons from vulnerable groups who might be less likely to go grocery shopping. For this to impact our findings, those remaining home would need to have differential antibody prevalence compared with their age/sex/racial-ethnic/regional group peers. If persons staying at home had lower prevalence because of self-isolation, our study's cumulative incidence would be a slight overestimate. Furthermore, our sample did not include those who have died from COVID-19 or those who reside in long-term care facilities, which have been differentially impacted, causing a slight underestimate, nor those in the hospital or at home due to COVID-19 illness, some of whom

Table 2
Reactivity and test characteristic–adjusted cumulative incidence of COVID-19, demographic factors within the region

Stratum	% Of adults in region	Reactivity		Test characteristic–adjusted estimated cumulative incidence			*P*-value
		Unweighted # reactive/Total sample[*]	Weighted percent	% (95% CI)	Infection-experienced adults[†] (95% CI)	% Of infection-experienced adults in region	
New York City [‡]							
Sex							.26
Male	46.7	629/2727	20.9	23.5 (21.6–25.4)	726,300 (668,500–784,100)	48.3	
Female	53.3	690/3219	19.6	22.1 (20.4–23.7)	778,000 (718,300–837,600)	51.7	
Race and ethnicity							<.0001
Hispanic or Latino	27.4	624/2103	29.2	33.0 (30.6–35.4)	599,900 (556,800–643,000)	39.8	
Non-Hispanic white	33.8	264/1758	14.8	16.6 (14.6–18.5)	371,300 (327,800–414,800)	24.6	
Non-Hispanic black/African American	21.7	329/1392	22.4	25.2 (22.5–27.9)	361,700 (322,900–400,500)	24.0	
Non-Hispanic Asian	14.5	68/509	13.0	14.5 (11.0–18.0)	139,000 (105,400–172,700)	9.2	
Multiracial/other	2.5	34/184	18.2	20.4 (13.7–27.2)	34,400 (23,100–45,700)	2.3	
Age group (y)							.04
18–34	34.0	252/1257	19.3	21.8 (19.2–24.4)	490,200 (432,200–548,300)	32.5	
35–44	17.4	243/1144	20.8	23.4 (20.6–26.2)	270,400 (238,100–302,700)	18.0	
45–54	16.3	334/1328	23.5	26.5 (23.8–29.2)	286,700 (257,700–315,700)	19.0	
55+	32.2	490/2217	19.1	21.5 (19.6–23.5)	459,000 (417,700–500,300)	30.5	
Westchester/Rockland							
Sex							.049
Male	47.6	72/450	17.1	19.2 (14.6–23.8)	88,700 (67,400–109,900)	56.7	
Female	52.4	62/530	11.9	13.3 (9.8–16.9)	67,800 (49,600–85,900)	43.3	
Race and ethnicity							.0008
Hispanic or Latino	21.0	37/141	25.3	28.6 (20.2–37.0)	58,300 (41,200–75,400)	37.8	
Non-Hispanic white	57.9	62/654	9.7	10.8 (8.0–13.5)	60,400 (45,200–75,600)	39.1	
Non-Hispanic black/African American	13.3	32/152	20.1	22.7 (15.0–30.3)	29,200 (19,400–39,000)	18.9	
Non-Hispanic Asian	6.3	**	7.7	8.5 (0–25.1)	5200 (0–15,300)	3.4	
Multiracial/other	1.6	**	7.1	7.8 (0–19.4)	1200 (0–2900)	0.8	
Age group (y)							.78
18–34	26.8	29/184	16.1	18 (11.8–24.3)	46,900 (30,600–63,200)	29.9	
35–44	16.3	22/156	15.2	17 (9.3–24.8)	27,000 (14,700–39,200)	17.2	
45–54	18.9	33/247	14.5	16.2 (10.7–21.8)	29,800 (19,600–40,000)	19.0	
55+	37.9	50/393	12.9	14.4 (10.0–18.9)	53,100 (36,700–69,400)	33.9	
Long Island [§]							
Sex							.15
Male	48.3	122/913	13.1	14.7 (11.7–17.6)	156,100 (125,100–187,100)	53.5	
Female	51.7	119/1161	10.7	11.9 (9.5–14.3)	135,700 (108,700–162,600)	46.5	
Race and ethnicity							<.0001
Hispanic or Latino	15.9	89/301	28.3	32.0 (26.1–37.9)	112,100 (91,400–132,900)	38.4	
Non-Hispanic white	67.3	126/1599	7.9	8.7 (7.2–10.3)	129,500 (106,800–152,200)	44.4	
Non-Hispanic black/African American	8.8	16/111	14.1	15.8 (6.6–25.0)	30,800 (12,900–48,700)	10.6	
Non-Hispanic Asian	6.3	**	7.7	8.4 (0–18.7)	11,800 (0–26,200)	4.0	
Multiracial/other	1.7	**	18.4	20.7 (3.9–37.5)	7600 (1400–13,800)	2.6	
Age group (y)							.73
18–34	26.3	45/429	12.6	14.1 (9.2–19.0)	81,800 (53,500–110,000)	28.0	
35–44	15.5	40/317	12.4	13.8 (9.6–18.1)	47,300 (32,800–61,800)	16.2	
45–54	19.5	61/468	12.8	14.3 (10.5–18.0)	61,200 (45,100–77,300)	20.9	
55+	38.7	95/860	10.7	12.0 (9.5–14.4)	102,200 (81,500–122,900)	34.9	
Rest of New York State [‖]							
Sex							.04
Male	48.5	95/2545	3.9	4.2 (3.2–5.1)	111,200 (85,800–136,600)	57.3	
Female	51.5	98/3556	2.8	2.9 (2.3–3.6)	83,100 (63,800–102,300)	42.7	
Race and ethnicity							.90
Hispanic or Latino	5.2	**	4.4	4.7 (0.8–8.7)	13,600 (2300–24,900)	7.0	
Non-Hispanic white	83.6	171/5534	3.2	3.4 (2.8–3.9)	154,300 (128,800–179,800)	79.4	
Non-Hispanic black/African American	6.6	11/258	4.0	4.3 (1.4–7.1)	15,500 (5100–25,800)	8.0	
Non-Hispanic Asian	2.7	**	4.0	4.3 (0–9.4)	6300 (0–13,900)	3.2	
Multiracial/other	1.9	**	4.2	4.5 (0–13.6)	4700 (0–14,300)	2.4	
Age group (y)							.04
18–34	29.0	51/1281	4.1	4.3 (3.0–5.7)	69,100 (47,500–90,700)	35.2	
35–44	14.2	29/1011	3.5	3.8 (2.1–5.4)	29,100 (16,300–41,900)	14.8	
45–54	17.5	51/1302	4.3	4.6 (3.2–6.0)	43,900 (30,500–57,400)	22.4	
55+	39.3	62/2507	2.5	2.5 (1.8–3.2)	54,000 (39,100–68,900)	27.5	

[*] Unweighted results with numerator < 10 are suppressed to protect participant confidentiality, indicated by **.
[†] Stratified estimates may not exactly sum to total because of rounding and differences between the weighting scheme and noninstitutionalized population totals.
[‡] Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
[§] Nassau and Suffolk Counties.
[‖] Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, and Yates counties.

Table 3
Estimated percentage of SARS-CoV-2 infections diagnosed*

Stratum	Estimated infection-experienced adults	Diagnosed adults through April 9, 2020		% Of diagnosed adults
	Adults (95% CI)	Diagnoses	% Diagnosed (95% CI)	
Overall	2,139,300 (2,035,800–2,242,800)	189,383	8.9 (8.4–9.3)	100.0
Sex				
Male	1,076,500 (1,001,900–1,151,100)	101,030	9.4 (8.8–10.1)	53.7
Female	1,062,200 (990,500–1,133,800)	87,196	8.2 (7.7–8.8)	46.3
Unknown		1157		
Age (y)				
18–34	682,600 (612,000–753,200)	41,335	6.1 (5.5–6.8)	22.4
35–44	371,800 (331,700–411,900)	32,845	8.8 (8.0–9.9)	17.8
45–54	424,700 (386,400–463,100)	35,307	8.3 (7.6–9.1)	19.1
55+	667,800 (615,600–719,900)	75,124	11.3 (10.4–12.2)	40.7
Missing/invalid		491		
Region				
New York City †	1,504,400 (1,421,300–1,587,500)	106,401	7.1 (6.7–7.5)	56.2
Westchester/Rockland counties	156,500 (128,400–184,600)	23,557	15.1 (12.8–18.3)	12.4
Long Island ‡	291,800 (250,600–332,900)	44,907	15.4 (13.5–17.9)	23.7
Rest of State §	194,600 (162,600–226,600)	14,518	7.5 (6.4–8.9)	7.7

* Complete statewide data on the case race and ethnicity are not currently available.
† Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
‡ Nassau and Suffolk Counties.
§ Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, and Yates counties.

would be expected to have detectable antibodies [38,39]. Such actively symptomatic persons would be expected to be a small portion of the cumulative infection burden since the outbreak's commencement, and given most would have been infected after March 29, their exclusion also likely causes observed values to be overestimated.

Although data are limited on the potential for self-selection to alter our results, a recent Icelandic study found comparable prevalence when participants were tested after online self-registration versus random invitation [16]. This finding, in conjunction with our systematic community intercept approach, suggests that this bias may be small, outside of outright nonresponse. We note that although every effort was made to ensure unbiased sampling through a DOH staff-led recruitment process, patron-initiated requests for testing were honored, and in some sites, accounted for a significant percentage of total tests performed. It is possible that customers who seek out testing may be more likely to have been exposed to SARS-CoV-2. If true, our estimate of cumulative incidence would be overestimated. Another source of potential recruitment bias comes from patron refusal to be tested, either on initial request or after agreeing to participate. Although not systematically collected, nightly report outs by testing leads indicated that most persons approached agreed to be tested and that few persons left after agreeing to be tested, regardless of wait time, supporting low nonresponse. Results presented may differ from publicly discussed preliminary estimates, given both our inclusion of more participants and analytic adjustments for test characteristics. Timeframes used for cumulative infections and diagnoses are approximate, being based on the evolving SARS-CoV-2 immunological and testing literature, with the 10-day sampling period during a linear growth phase of the epidemic.

The findings of this study suggest extensive SARS-CoV-2 transmission in NYS and highlight the remaining opportunities for prevention and diagnosis. As the epidemic grows in other regions of the country, this study offers a potential model for other jurisdictions to monitor their epidemic. Estimates of cumulative incidence can be combined with diagnostic totals, or other epidemic markers such as mortality, to provide a holistic epidemic view during a time of unprecedented pandemic and to best craft high-impact approaches to prevention, containment, treatment, and mitigation.

Acknowledgments

The co-authors wish to acknowledge the following essential contributors to this work. Peter Cichetti, Antibody Sampling and Testing Team Co-Lead, Lyla Hunt and Patrick McKeage, Antibody Sampling and Testing Team Co-Coordinators, Thomas Sullivan, Logistical Support Lead for the Antibody Sampling and Testing Team. Data Leadership Team members Jason Ganns, Melissa Kamal, Alison Pingelski, Mary McCormick, Ann Lowenfels and Rebecca Hoen. Michelle Cummings for data management. Amy Kelly for literature review contributions. All members of the New York State Antibody Sampling and Testing Team, including the Call Center team who placed thousands of phone calls to deliver test results and to collect demographic data. Office of Quality and Patient Safety team members James Kirkwood and Meng Wu. Eric Hall at Emory University for map assistance. We thank Dr. William Lee for assistance with assay development and validation, Jean Rock and the Wadsworth Center COVID-19 serology team especially Rachel Bievenue, Seth Blumerman, Theresa Hattenrath, Jim Long, Kate Mastraccio, Erica Miller, Katie Nemeth, and Alyssa Sossei, and numerous members of Wadsworth Center's Newborn Screening Program especially Beth Vogel, Michele Caggana and Rhonda Hamel. Special thanks to Adrienne Mazeau for her leadership and guidance. Special recognition and thanks to DHSES (Division of Homeland Security and Emergency Services) for their tremendous support to the NYSDOH.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions: ESR contributed to methodology, writing—original draft, review, and editing, and formal analysis. JMT contributed to supervision, investigation, writing—original draft, review, and editing. EMR, RC, MAB, and S-YJL contributed to formal analysis. LMS and MMP contributed to investigation, resources, and validation. JEM and DG contributed to project administration, supervision, investigation, and resources. DRH, JK, and TU contributed to writing—review and editing. DH contributed to data curation

and writing—review and editing. BH and HAZ contributed to project administration, supervision, investigation, and resources.

References

[1] Rosenberg ES, Dufort EM, Blog DS, Hall EW, Hoefer D, Backenson BP, et al. COVID-19 testing, epidemic features, hospital outcomes, and household prevalence, New York State-March 2020. Clin Infect Dis 2020. https://doi.org/10.1093/cid/ciaa549.

[2] JHU. COVID-19 United States Cases 2020. https://coronavirus.jhu.edu/us-map. [Accessed 3 June 2020].

[3] Health NYSDo. NYSDOH COVID-19 Tracker 2020. https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-DailyTracker?%3Aembed=yes&%3Atoolbar=no&%3Atabs=n. [Accessed 3 June 2020].

[4] Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 2020;368(6490):489–93.

[5] Amid Ongoing COVID-19 Pandemic. Governor Cuomo outlines blueprint to Un-PAUSE New York [press release]. 2020. https://www.governor.ny.gov/news/amid-ongoing-covid-19-pandemic-governor-cuomo-outlines-blueprint-un-pause-new-york. [Accessed 3 June 2020].

[6] Xiang F, Wang X, He X, Peng Z, Yang B, Zhang J, et al. Antibody detection and dynamic characteristics in patients with COVID-19. Clin Infect Dis 2020. https://doi.org/10.1093/cid/ciaa461.

[7] Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis 2020. https://doi.org/10.1093/cid/ciaa344.

[8] Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2. JAMA 2020. https://doi.org/10.1001/jama.2020.8259.

[9] Kirkcaldy RD, King BA, Brooks JT. COVID-19 and postinfection immunity: limited evidence, many remaining questions. JAMA 2020. https://doi.org/10.1001/jama.2020.7869.

[10] Hofmeister MG, Rosenthal EM, Barker LK, Rosenberg ES, Barranco MA, Hall EW, et al. Estimating prevalence of hepatitis C virus infection in the United States, 2013-2016. Hepatology 2019;69(3):1020–31.

[11] Angulo FJ, Finelli L, Swerdlow DL. reopening society and the need for real-time assessment of COVID-19 at the community level. JAMA 2020. https://doi.org/10.1001/jama.2020.7872.

[12] Xu X, Sun J, Nie S, Li H, Kong Y, Liang M, et al. Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. Nat Med 2020. https://doi.org/10.1038/s41591-020-0949-6.

[13] Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. medRxiv 2020. https://doi.org/10.1101/2020.04.14.20062463.

[14] NIH. NIH begins study to quantify undetected cases of coronavirus infection. 2020. https://www.nih.gov/news-events/news-releases/nih-begins-study-quantify-undetected-cases-coronavirus-infection. [Accessed 3 June 2020].

[15] Sood N, Simon P, Ebner P, Eichner D, Reynolds J, Bendavid E, et al. Seroprevalence of SARS-CoV-2–Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. JAMA 2020. https://doi.org/10.1001/jama.2020.8279.

[16] Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. N Engl J Med 2020. https://doi.org/10.1056/NEJMoa2006100.

[17] Executive Order [Cuomo] No. 202.6 [9 NYCRR 8.202.6]: New York State on PAUSE. 2020. https://www.governor.ny.gov/news/governor-cuomo-issues-guidance-essential-services-under-new-york-state-pause-executive-order. [Accessed 3 June 2020].

[18] NCHS. U.S. Census. Populations With Bridged Race Categories. 2020. https://www.cdc.gov/nchs/nvss/bridged_race.htm. [Accessed 3 June 2020].

[19] Holt D, Smith T. Post Stratification. J R Stat Soc 1979;142(1):33–46.

[20] Korn EL, Graubard BI. Analysis of health surveys. New York: John Wiley & Sons; 1999.

[21] Rao JNK, Scott AJ. On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. Ann Statist 1984;12(1):46–60.

[22] Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.

[23] Bureau UC. Public use microdata sample (PUMS) documentation. 2020. https://www.census.gov/programs-surveys/acs/technical-documentation/pums.html. [Accessed 3 June 2020].

[24] Lee YL, Liao CH, Liu PY, Cheng CY, Chung MY, Liu CE, et al. Dynamics of anti-SARS-Cov-2 IgM and IgG antibodies among COVID-19 patients. J Infect 2020. https://doi.org/10.1016/j.jinf.2020.04.019.

[25] Kwok KO, Lai F, Wei WI, Wong SYS, Tang JWT. Herd immunity - estimating the level required to halt the COVID-19 epidemics in affected countries. J Infect 2020. https://doi.org/10.1016/j.jinf.2020.03.027.

[26] House TW. Opening Up America Again. 2020. https://www.whitehouse.gov/openingamerica/. [Accessed 3 June 2020].

[27] Rocklöv J, Sjödin H. High population densities catalyse the spread of COVID-19. J Travel Med 2020. https://doi.org/10.1093/jtm/taaa038.

[28] Chowkwanyun M, Reed AL. Racial health disparities and Covid-19 — Caution and Context. N Engl J Med 2020. https://doi.org/10.1056/NEJMp2012910.

[29] Millett GA, Jones AT, Benkeser D, Baral S, Mercer L, Beyrer C, et al. Assessing differential impacts of COVID-19 on black communities. Ann Epidemiol 2020. https://doi.org/10.1016/j.annepidem.2020.05.003.

[30] Rosenberg ES, Dufort EM, Udo T, Wilberschied LA, Kumar J, Tesoriero J, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York State. JAMA 2020. https://doi.org/10.1001/jama.2020.8630.

[31] Azar KMJ, Shen Z, Romanelli RJ, Lockhart SH, Smits K, Robinson S, et al. Disparities in outcomes among COVID-19 patients in a large health care system In California. Health Aff 2020. https://doi.org/10.1377/hlthaff.2020.00598.10.1377/hlthaff.2020.00598.

[32] Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among black patients and white patients with Covid-19. N Engl J Med 2020. https://doi.org/10.1056/NEJMsa2011686.

[33] CDC. Coronavirus Disease 2019 (COVID-19): Cases in U.S. 2020. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html. [Accessed 3 June 2020].

[34] Livingston E, Bucher K. Coronavirus Disease 2019 (COVID-19) in Italy. JAMA 2020. https://doi.org/10.1001/jama.2020.4344.

[35] Goyal P, Choi JJ, Pinheiro LC, Schenck EJ, Chen R, Jabri A, et al. Clinical Characteristics of Covid-19 in New York City. N Engl J Med 2020. https://doi.org/10.1056/NEJMc2010419.

[36] Roques L, Klein EK, Papaïx J, Sar A, Soubeyrand S. Using early data to estimate the actual infection fatality ratio from COVID-19 in France. Biology 2020;9(5):97. https://doi.org/10.3390/biology9050097.

[37] Nishiura H, Kobayashi T, Yang Y, Hayashi K, Miyama T, Kinoshita R, et al. The rate of underascertainment of novel Coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights. J Clin Med 2020;9(2):419. https://doi.org/10.3390/jcm9020419.

[38] McMichael TM, Clark S, Pogosjans S, Kay M, Lewis J, Baer A, et al. COVID-19 in a long-term care facility - King County, Washington, February 27-March 9, 2020. MMWR Morb Mortal Wkly Rep 2020;69(12):339–42.

[39] Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al. Asymptomatic and Presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, March 2020. MMWR Morb Mortal Wkly Rep 2020;69(13):377–81.

Appendix

Validation studies

Part 1. Assay specificity

Table 1
Cutoff assay using normal human dry blood spots

Assay	Number tested	Number > mean + 3 SDs	Number > mean + 6 SDs	Specificity	95% CI
SARS-CoV-1N	196	5	1	195/196 = 99.5%	98.5%–100%
SARS-CoV-2N	92	2	0	92/92 = 100%	96.1%–100%

Normal DBS collected before December 2019 were tested to establish background levels of reactivity to the SARS-CoV-1N and SARS-CoV-2N proteins. Specificity for primary analyses was based on the mean of observed reactivity to each protein (99.5% and 100%, respectively), which was 99.75%.

Table 2
Serum collected from patients with antibodies to the indicated nonrespiratory agents tested for reactivity to the SARS-CoV-1N protein

Analyte	Number tested	Number > mean + 3 SDs	Number >mean + 6 SDs
Antinuclear antibodies	3	0	0
Chikungunya virus	3	0	0
Dengue virus	3	0	0
Enterovirus	2	0	0
Epstein Barr Virus	3	0	0
Hepatitis C virus	3	0	0
Herpes Simplex virus	3	0	0
HIV	3	0	0
Measles	6	0	0
Mumps	6	0	0
Parvovirus B19	4	0	0
Rheumatoid factor	3	0	0
Rubella	3	0	0
Syphilis	3	1	1
Varicella Zoster virus	3	0	0
West Nile virus	3	0	0
Zika virus	3	0	0
TOTAL	57	1	1

Table 3
Serum collected from patients diagnosed with the indicated non–COVID-19 respiratory agents tested for reactivity to the SARS-CoV-1N and SARS-CoV-2N protein

Analyte	SARS-CoV-1N			SARS-CoV-2N		
	Number tested	Number > mean + 3 SDs	Number > mean + 6 SDs	Number tested	Number > mean + 3 SDs	Number > mean + 6 SDs
Human coronavirus NL63	6	0	0	4	0	0
Human coronavirus OC43	2	0	0	2	0	0
Human parainfluenza virus	1	0	0	1	0	0
Influenza	6	0	0	15	1	0
Metapneumovirus	6	0	0	6	0	0
Rhinovirus	7	0	0	6	0	0
Total	28	0	0	34	1	0

Part 2. Assay sensitivity

Table 4
SARS-CoV-2N assay results for a set of 232 individuals with self-reported positive PCR results, with median (IQR) days from the symptom onset to DBS collection

Result	#	Median days after the symptom onset	IQR
Nonreactive	21	33	28–37
Indeterminate	7	30	28–34
Reactive	204	35	30–40
Total	232		

Based on 204 of 232 reactive, estimate sensitivity was 87.9% (exact 95% CI: [83.7–92.1%]).

Part 3 – SARS-CoV-1N versus SARS-CoV-2N assay comparison

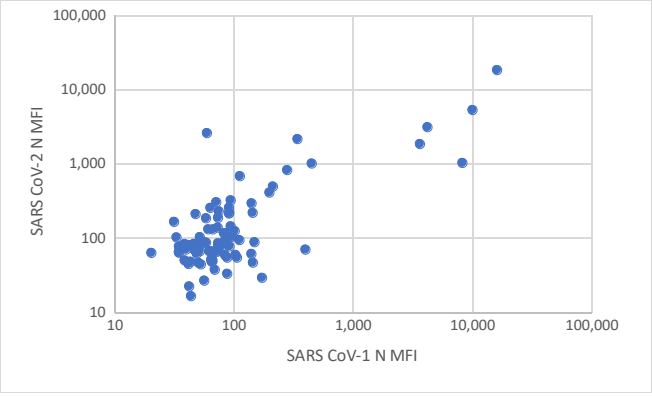

Fig. 1. Comparison of MFI values for DBS samples tested by the SARS-CoV-1N versus SARS-CoV-2N assay.

Table 5
Comparison of qualitative results of the same samples tested by the SARS-CoV-1N versus SARS-CoV-2N assay

Result		SARS-CoV-2N			
		R	IND	NR	Total
SARS-CoV-1N	R	17	0	0	17
	IND	0	0	0	0
	NR	5	6	145	156
	TOTAL	22	6	145	173

Concordance	162/173	94%	Results match exactly (NR-NR, IND-IND, R-R)
Discordance	5/173	3%	Reactive = Nonreactive, Nonreactive = Reactive

Kappa of 0.855 (95% CI: 0.731–0.979), as measured on the 2 × 2 table excluding IND results, indicates strong agreement above chance.

Sensitivity analyses for cumulative incidence and percent diagnosed

Table 6
Reactivity and test characteristic–adjusted cumulative incidence of COVID-19, overall and by demographic factors and region: sensitivity 92.1% and specificity 96.1%

Stratum	Reactivity		Test characteristic–adjusted estimated cumulative incidence				
	Unweighted # reactive/Total sample	Weighted percent	%	(95% CI)	Infection-experienced adults*	(95% CI)	% Of infection-experienced adults
Overall	1887/15,101	12.5	9.8	(9.1–10.5)	1,494,700	(1,391,800–1,597,600)	100.0
Sex							
Male	918/6635	13.2	10.6	(9.6–11.6)	769,200	(695,000–843,400)	51.5%
Female	969/8466	11.9	9.1	(8.2–10.0)	724,900	(653,700–796,200)	48.5%
Race/ethnicity							
Hispanic or Latino	757/2735	25.8	24.9	(22.9–26.9)	661,300	(608,400–714,100)	24.9
NH-white	623/9545	7.3	3.9	(3.2–4.5)	344,500	(286,500–402,500)	3.9
NH-black/African American	388/1913	18.0	15.9	(13.8–18.1)	337,800	(292,700–382,900)	15.9
NH-Asian	75/629	11.1	8.2	(5.2–11.1)	106,600	(67,800–145,500)	8.2
NH-Multiracial/other	44/279	10.7	7.7	(2.2–13.2)	25,100	(7300–43,000)	7.7
Age group (y)							
18–34	377/3151	13.0	10.4	(8.9–11.9)	484,900	(414,700–555,100)	10.4
35–44	334/2628	13.7	11.1	(9.4–12.7)	269,100	(229,200–308,900)	11.1
45–54	479/3345	14.3	11.8	(10.3–13.2)	312,500	(274,300–350,600)	11.8
55+	697/5977	10.9	7.9	(7.0–8.9)	435,900	(384,000–487,800)	7.9
Region							
New York City †	1319/5946	20.2	18.5	(17.2–19.7)	1,221,700	(1,139,100–1,304,300)	**
Westchester/Rockland Counties	134/980	14.4	11.9	(9.0–14.8)	115,400	(87,500–143,400)	**
Long Island ‡	241/2074	11.9	9.0	(7.2–10.9)	198,800	(157,900–239,700)	**
Rest of NYS §, ‖	193/6101	3.4	**	**	**	**	**

* Stratified estimates may not exactly sum to total because of rounding and differences between the weighting scheme and noninstitutionalized population totals.
† Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
‡ Nassau and Suffolk Counties.
§ Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, and Yates counties.
‖ Estimated values for Rest of NYS are negative because of antibody prevalence <1—specificity (observed results may be dominated by false positives). True cumulative incidence is greater than 0, as evidence by diagnosed cases, rendering study-based cumulative incidence inestimable under these test characteristics. Values that cannot be estimated are indicated with ** in the table.

Table 7
Reactivity and test characteristic–adjusted cumulative incidence of COVID-19, overall and by demographic factors and region: sensitivity 83.7% and specificity 100%

Stratum	Reactivity		Test characteristic–adjusted estimated cumulative incidence		
	Unweighted # reactive/Total sample	Weighted percent	% (95% CI)	Infection-experienced adults * (95% CI)	% Of infection-experienced adults
Overall	1887/15,101	12.5	15.0 (14.3–15.7)	2,286,600 (2,178,200–2,395,100)	100.0
Sex					
Male	918/6635	13.2	15.8 (14.7–16.9)	1,149,500 (1,071,300–1,227,600)	50.3
Female	969/8466	11.9	14.2 (13.3–15.2)	1,136,600 (1,061,500–1,211,700)	49.7
Race/ethnicity					
Hispanic or Latino	757/2735	25.8	30.9 (28.8–33.0)	820,700 (765,000–876,400)	36.2
NH-white	623/9545	7.3	8.8 (8.1–9.4)	775,900 (714,800–837,000)	34.2
NH-black/African American	388/1913	18.0	21.5 (19.2–23.7)	454,700 (407,200–502,200)	20.1
NH-Asian	75/629	11.1	13.3 (10.1–16.4)	173,300 (132,300–214,300)	7.6
NH-Multiracial/other	44/279	10.7	12.8 (7.0–18.6)	41,600 (22,800–60,500)	1.8
Age group (y)					
18–34	377/3151	13.0	15.6 (14.0–17.2)	729,100 (655,100–803,100)	31.8
35–44	334/2628	13.7	16.3 (14.6–18.1)	396,700 (354,700–438,700)	17.3
45–54	479/3345	14.3	17.1 (15.6–18.6)	452,800 (412,700–493,000)	19.7
55+	697/5977	10.9	13.0 (12.0–14.0)	716,000 (661,300–770,600)	31.2
Region					
New York City †	1319/5946	20.2	24.1 (22.8–25.4)	1,595,700 (1,508,700–1,682,800)	69.5
Westchester/Rockland Counties	134/980	14.4	17.2 (14.2–20.2)	166,800 (137,400–196,300)	7.3
Long Island ‡	241/2074	11.9	14.2 (12.2–16.1)	312,200 (269,100–355,400)	13.6
Rest of NYS §	193/6101	3.4	4.0 (3.4–4.6)	220,200 (186,700–253,700)	9.6

* Stratified estimates may not exactly sum to total because of rounding and differences between the weighting scheme and noninstitutionalized population totals.
† Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
‡ Nassau and Suffolk Counties.
§ Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, and Yates counties.

Table 8
Estimated percentage of SARS-COV-2 infections diagnosed by the last plausible diagnosis date

	Estimated infection-experienced adults	Diagnosed adults through May 8, 2020		
	Adults 95% CI	Diagnoses	% diagnosed (95% CI)	% of diagnosed adults
Overall	2,139,300 (203,5800–2,242,800)	335,770	15.7 (15.0–16.5)	100.0
Sex				
Male	1,076,500 (1,001,900–1,151,100)	172,238	16.0 (15.0–17.2)	51.6
Female	1,062,200 (990,500–1,133,800)	161,799	15.2 (14.3–16.3)	48.4
Unknown		1733		
Age (y)				
18–34	682,600 (612,000–753,200)	71,427	10.5 (9.5–11.7)	21.9
35–44	371,800 (331,700–411,900)	54,792	14.7 (13.3–16.5)	16.8
45–54	424,700 (386,400–463,100)	61,054	14.4 (13.2–15.8)	18.8
55+	667,800 (615,600–719,900)	138,312	20.7 (19.2–22.5)	42.5
Missing/invalid		851		
Region				
New York City [*]	1,504,400 (1,421,300–1,587,500)	187,623	12.5 (11.8–13.2)	55.9
Westchester/Rockland Counties	156,500 (128,400–184,600)	39,967	25.5 (21.7–31.1)	11.9
Long Island [†]	291,800 (250,600–332,900)	75,100	25.7 (22.6–30.0)	22.4
Rest of State [‡]	194,600 (162,600–226,600)	33,080	17.0 (14.6–20.3)	9.9

[*] Boroughs of Bronx, Brooklyn, Manhattan, Queens, Staten Island.
[†] Nassau and Suffolk Counties.
[‡] Albany, Allegany, Broome, Cattaraugus, Cayuga, Chautauqua, Chemung, Chenango, Clinton, Columbia, Cortland, Delaware, Dutchess, Erie, Essex, Franklin, Fulton, Genesee, Greene, Hamilton, Herkimer, Jefferson, Lewis, Livingston, Madison, Monroe, Montgomery, Niagara, Oneida, Onondaga, Ontario, Orange, Orleans, Oswego, Otsego, Putnam, Rensselaer, St. Lawrence, Saratoga, Schenectady, Schoharie, Schuyler, Seneca, Steuben, Sullivan, Tioga, Tompkins, Ulster, Warren, Washington, Wayne, Wyoming, Yates counties.