



OPEN

## Seroincidence of SARS-CoV-2 infection prior to and during the rollout of vaccines in a community-based prospective cohort of U.S. adults

Denis Nash<sup>1,2,7</sup>✉, Avantika Srivastava<sup>1,2</sup>, Yanhan Shen<sup>1,2</sup>, Kate Penrose<sup>1</sup>, Sarah G. Kulkarni<sup>1</sup>, Rebecca Zimba<sup>1,2</sup>, William You<sup>1</sup>, Amanda Berry<sup>1</sup>, Chloe Mirzayi<sup>1,2</sup>, Andrew Maroko<sup>1,3</sup>, Angela M. Parcesepe<sup>1,4,5</sup>, Christian Grov<sup>1,6</sup> & McKaylee M. Robertson<sup>1</sup>

This study used repeat serologic testing to estimate infection rates and risk factors in two overlapping cohorts of SARS-CoV-2 N protein seronegative U.S. adults. One mostly unvaccinated sub-cohort was tracked from April 2020 to March 2021 (pre-vaccine/wild-type era,  $n = 3421$ ), and the other, mostly vaccinated cohort, from March 2021 to June 2022 (vaccine/variant era,  $n = 2735$ ). Vaccine uptake was 0.53% and 91.3% in the pre-vaccine and vaccine/variant cohorts, respectively. Corresponding seroconversion rates were 9.6 and 25.7 per 100 person-years. In both cohorts, sociodemographic and epidemiologic risk factors for infection were similar, though new risk factors emerged in the vaccine/variant era, such as having a child in the household. Despite higher incidence rates in the vaccine/variant cohort, vaccine boosters, masking, and social distancing were associated with substantially reduced infection risk, even through major variant surges.

Infectious disease surveillance systems, which largely rely on diagnosed case counts, emergency department visits, hospital admissions, and deaths, underestimate the true incidence of infection due to asymptomatic infections, under-ascertainment/diagnosis, and underreporting to health departments<sup>1</sup>. This has proved to be a challenge during the COVID-19 pandemic in the United States (US)<sup>2,3</sup>, where the national surveillance system relied on the reporting of positive SARS-CoV-2 real-time reverse-transcription polymerase chain reaction (RT-PCR) test results by providers and laboratories. These data have effectively been used as a proxy for the incidence and prevalence of SARS-CoV-2 infection in the US, including to inform pandemic policy decisions, often with no efforts to adjust for underestimation due to factors such as asymptomatic infections, testing and healthcare access, evolving testing practices, behaviors, and policies<sup>3-6</sup>. Studies have shown that estimates of infection based on seroprevalence far exceed the number of diagnosed and reported cases<sup>3,7-10</sup>. These issues have posed major challenges to using surveillance data to assess the true burden of infection and related risk factors and inform decision-making in an evolving pandemic<sup>1</sup>. These challenges are magnified with the end of the national emergency declaration in May 2023, as surveillance for SARS-CoV-2 further dismantles, testing behaviors change and official case counts continue to fall<sup>11</sup>.

The risk of SARS-CoV-2 infection is determined by multiple factors, including frequency of exposure, underlying medical conditions, vaccination status, virus properties and their evolution, levels of community transmission, and individual behaviors, such as the use of non-pharmaceutical interventions (NPIs) like masking and

<sup>1</sup>Institute for Implementation Science in Population Health (ISPH), City University of New York (CUNY), New York, NY, USA. <sup>2</sup>Department of Epidemiology and Biostatistics, Graduate School of Public Health and Health Policy, City University of New York (CUNY), New York, NY, USA. <sup>3</sup>Department of Environmental, Occupational, and Geospatial Health Sciences, Graduate School of Public Health and Health Policy, City University of New York (CUNY), New York, NY, USA. <sup>4</sup>Department of Maternal and Child Health, Gillings School of Public Health, University of North Carolina, Chapel Hill, NC, USA. <sup>5</sup>Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>6</sup>Department of Community Health and Social Sciences, Graduate School of Public Health and Health Policy, City University of New York (CUNY), New York, NY, USA. <sup>7</sup>CUNY Graduate School of Public Health and Health Policy, 55 W. 125th St., 6th Floor, New York, NY 10027, USA. ✉email: denis.nash@sph.cuny.edu

social distancing<sup>12–15</sup>. Vaccines, which were authorized by the FDA on December 11, 2020 and started to become more widely available in the United States in March 2021<sup>16</sup>, have dramatically reduced the risk of severe disease and death from SARS-CoV-2<sup>17</sup>, including during the Alpha (March–June 2021), Delta (June–December 2021) and Omicron (December 2021–present) variant eras<sup>18</sup>. However, with the emergence of the Delta variant<sup>19</sup> and particularly during the Omicron era, vaccine effectiveness *against infection* reduced and waned quickly after vaccination, including after a booster<sup>18,20–23</sup>. Recent population-representative, cross-sectional studies during successive Omicron surges have shown a high point prevalence of SARS-CoV-2 infection among those who were previously vaccinated and boosted<sup>1,25</sup>.

Alongside variant surges, starting in late 2021, the United States relaxed public health policies and guidelines recommending or requiring quarantine and isolation, masking, social distancing, and remote K-12 schooling in response to the increasing availability and uptake of vaccines and the availability of effective therapeutics<sup>26–30</sup>. National policies have increasingly emphasized the importance of using vaccines and boosters to reduce the burden of severe disease and healthcare system strain, with less emphasis on NPIs for preventing infection<sup>14</sup>. Once schools and childcare reopened, having school-aged children in the household may have increased risk of infection over the course of the pandemic, particularly during surges<sup>31–33</sup>. These policy choices and related public health messaging likely impacted individual-level and community-level risk factors and behavior, resulting in less utilization of NPIs such as masking, social distancing, and more<sup>34,35</sup>.

Beyond basic demographics and geography, risk factors for SARS-CoV-2 infection (vs. diagnoses), including asymptomatic infection, have not been well-characterized in the vaccine and variant eras, either via routine case-based surveillance or by cross-sectional seroprevalence studies in population-based samples<sup>1,36,37</sup>. Moreover, the degree of the potential protective effect of vaccines for preventing SARS-CoV-2 infection has not been examined in a population-based prospective cohort with systematically gathered, time-updated information on risk factors, behaviors, and infection status derived from infection-induced seroconversions.

Within a well-characterized national prospective cohort of US adults (for whom we previously characterized risk factors for seroincident SARS-CoV-2 infection in the pre-vaccine/wild-type era<sup>15</sup>), we used repeat serologic testing to assess the incidence of SARS-CoV-2 infection and risk factors among those who were N protein seronegative by March 2021 (the start of the variant era), with vaccine uptake and infection status assessed through June 2022.

## Methods

### Recruitment

We used internet-based strategies<sup>38–40</sup> to recruit a geographically and socio-demographically diverse cohort of 6740 adults into longitudinal follow-up with at-home, dried blood spot (DBS) specimen collection. To be eligible for inclusion in the cohort, individuals had to: (1) reside in the United States or a U.S. territory; (2) be  $\geq 18$  years of age; (3) provide a valid email address for follow-up; and (4) demonstrate early engagement in study activities (provision of a baseline specimen for serologic testing or completion of  $> 1$  recruitment/enrollment visit). Details of the study design and recruitment procedures<sup>40</sup> and a pre-vaccine/wild-type era serology-based incidence study in this cohort<sup>15</sup> are described elsewhere. Cohort enrollment was completed between March 28 and August 21, 2020, during which baseline demographic data collection took place. The full cohort includes participants from all 50 U.S. states, the District of Columbia, Puerto Rico, and Guam.

### Study population

The study population ( $n = 3582$ ) was divided into two *overlapping* sub-groups (henceforth cohorts, Fig. 1) which broadly correspond to those with a seronegative specimen during from April through September 2020 (Serology Period 1) with at least one follow-up serologic test (pre-vaccine/wild-type era cohort,  $n = 3421$ ), and those with a seronegative specimen during November 2020 through March 2021 (Serology Period 2) with at least one follow-up serologic test (vaccine/variant era cohort,  $n = 2735$ ) (Fig. 2)<sup>41,42</sup>. The vaccine/variant era cohort included those in the pre-vaccine/wild-type era cohort who remained seronegative on their specimen from Serology Period 2 (Fig. 2).

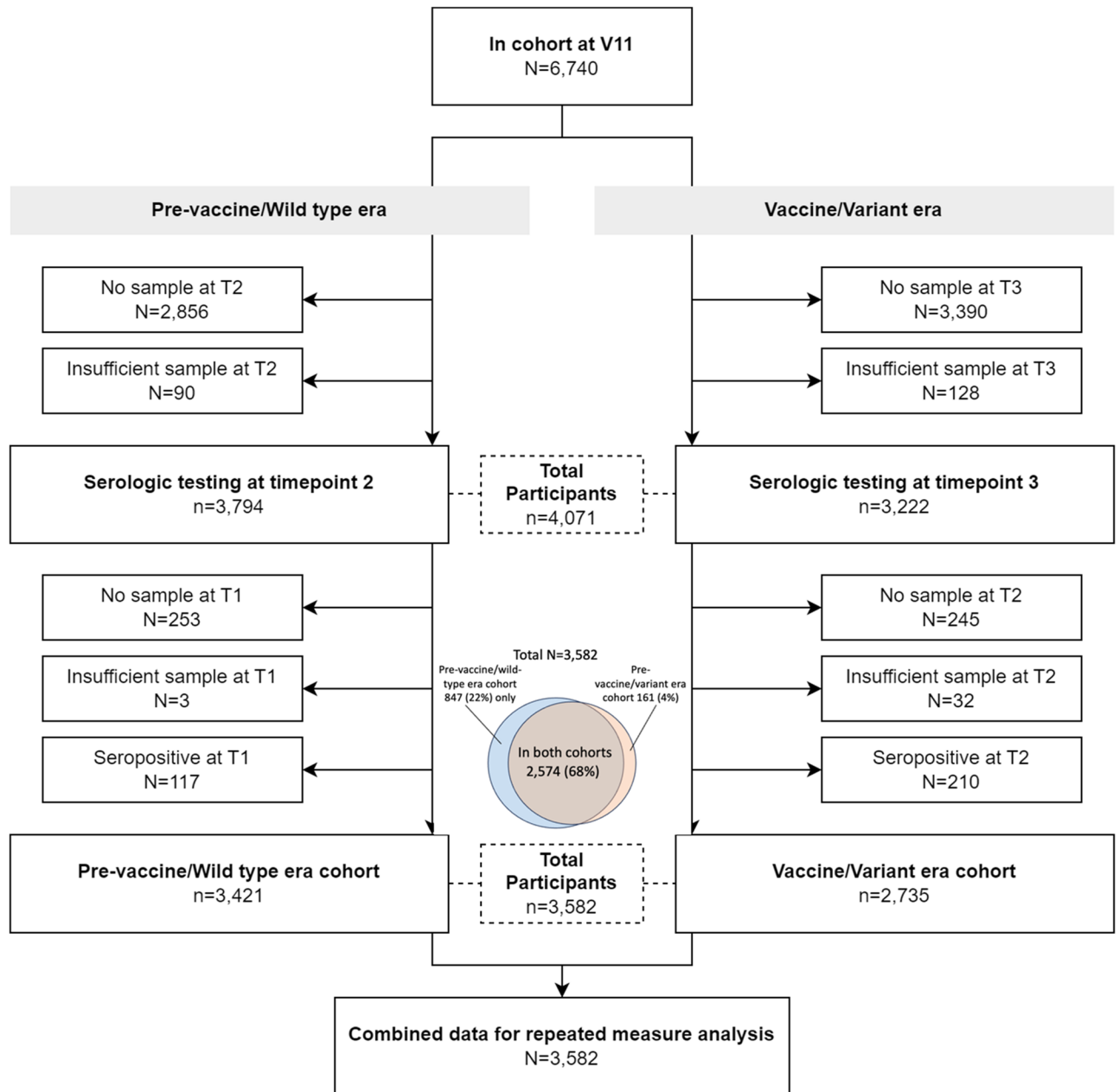
### Follow-up data collection

From 14 follow-up study encounters occurring approximately quarterly between August 2020 and July 2022, we obtained repeated measurements of epidemiologic risk factors, COVID-19 symptoms, non-study-related SARS-CoV-2 testing (PCR or rapid, at-home rapid), hospitalizations, use of NPIs, public health strategies (i.e., quarantine, isolation), and contact tracing encounters.

### Serologic testing

Figure 2 shows the three periods of serologic testing in the cohort—from April through September 2020 (Serology Period 1), November 2020 through March 2021 (Serology Period 2), and March 2022 through June 2022 (Serology Period 3). During these periods, participants were invited to complete serologic testing using an at-home self-collected dried blood spot (DBS) specimen collection kit. DBS cards were sent from and returned to the study laboratory (Molecular Testing Laboratories [MTL], Vancouver, WA) via the U.S. Postal Service using a self-addressed, stamped envelope containing a biohazard bag.

To assess infection-induced seroconversion, all DBS specimens were tested by the study laboratory for total antibodies to the SARS-CoV-2 nucleocapsid protein (total nucleocapsid Ab) using the Bio-Rad Platelia test for IgA, IgM, and IgG (manufacturer sensitivity 98.0%, specificity 99.3%)<sup>43</sup>. Other studies have independently validated this assay and found average sensitivity and specificity of 91.7% and 98.8%, respectively<sup>44–46</sup>. This assay

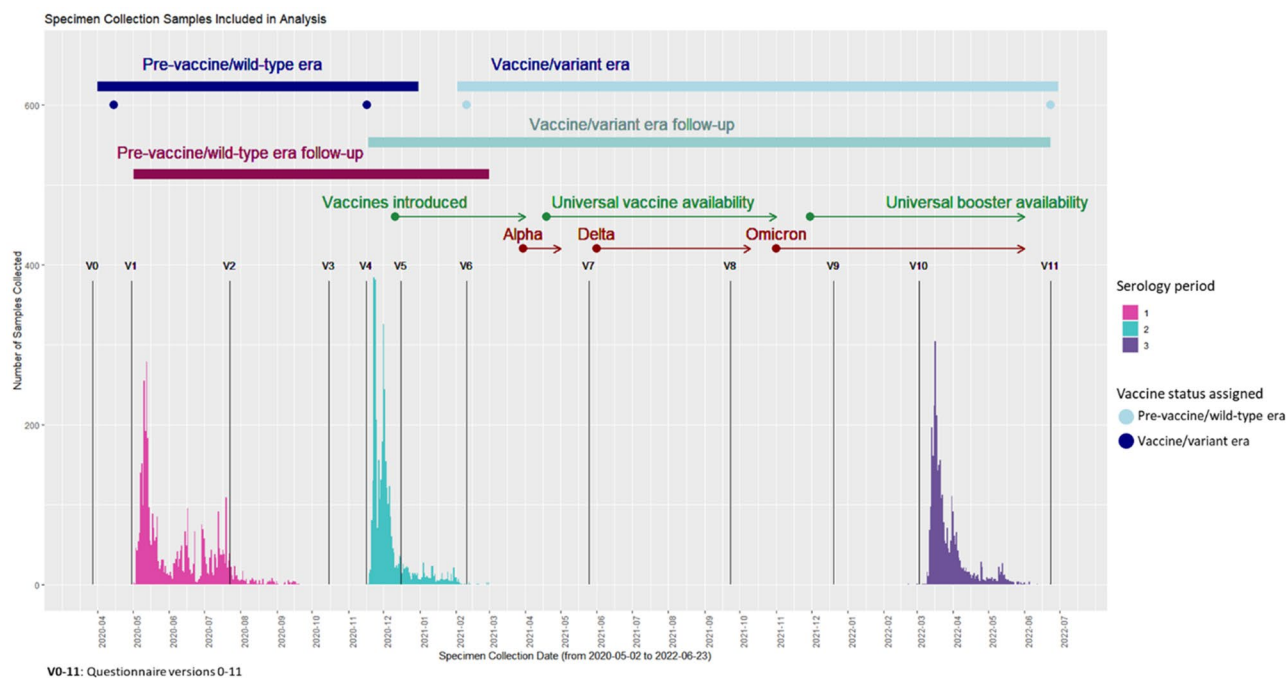


**Figure 1.** Study population and sub-cohorts.

was also validated for use with DBS by the study laboratory, which found 100% sensitivity and 100% specificity (MTL, personal communication).

**Outcome (infection-induced seroconversion)**

Participants were assessed for the outcome in each cohort if they were: (1) seronegative at the start of follow-up; and (2) had a subsequent serologic test during follow-up. Among these individuals, the outcome of infection-induced SARS-CoV-2 seroconversion in the pre-vaccine/wild-type era cohort was defined as having a negative total nucleocapsid Ab test in Serology Period 1 followed by a positive total nucleocapsid Ab test in Serology Period 2. The outcome of infection-induced SARS-CoV-2 seroconversion in the vaccine/variant era cohort was defined as a negative total nucleocapsid Ab test in Serology Period 2 followed by a positive total nucleocapsid Ab test in Serology Period 3. We estimated person-years of follow-up in each cohort using the collection dates for each specimen from that cohort’s follow-up period. Participants could contribute person-time to each cohort (pre-vaccine/wild-type or vaccine/variant era). When the specimen collection date was missing, we used the date the laboratory received the sample. For those who seroconverted, the seroconversion date in each cohort was assigned as the midpoint between the first seronegative and the subsequent seropositive specimen collection dates. If a positive SARS-CoV-2 test result (PCR or rapid test) was reported by a participant in between the specimen collection dates for serologic testing, the date of the positive test was used as the infection date.



**Figure 2.** Timing of specimen collection, vaccine rollout, and cohort follow-up.

## Exposures

### *Timing of data collection on risk factors, behaviors, and vaccination status*

All exposure measurements were derived from time-updated questionnaire data collected during each era, and only those measures taken prior to outcome measurement for a given era were used in our analyses. For the pre-vaccine/wild-type era cohort, we used exposure data from the questionnaire for study visits 1 through 4 (V1–V4; Fig. 2). For the vaccine/variant era cohort, we used data from V4–V10 questionnaires (Fig. 2).

### *Individual-level COVID-19 risk factors*

We collected time-updated information on an array of epidemiologic risk factors for SARS-CoV-2 infection reported by participants, including the following: essential worker status (those working in healthcare, emergency response, law enforcement, delivery of food/goods, transportation), household factors (household crowding defined as  $\geq 4$  people living in a single unit of a multi-unit dwelling, having a child in the household, and having a confirmed COVID-19 case in a household member before participant tested positive); spending time in public places (attending mass gatherings, indoor dining in a restaurant or bar, outdoor dining at a restaurant or bar, visiting places of worship, or visiting public parks or pools); mask use indoors (for grocery shopping, visiting non-household members, at work, and in salons or gyms); mask use outdoors; gathering in groups with 10 or more people; travel during the pandemic (air travel and public transit use); and individual-level factors that may increase the risk of infection and/or severe COVID-19 (comorbid conditions, binge drinking, regular cannabis use or un-prescribed opioid use). Binge drinking was defined as six or more drinks in one sitting during the last month, asked as part of the Alcohol Use Disorders Identification Test questions on select questionnaires<sup>47</sup>. As a measure of susceptibility to severe COVID-19, we used comorbid conditions or exposures that CDC identified as increasing the risk for COVID-19 complications, given SARS-CoV-2 infection: age  $\geq 60$  years, daily smoking, chronic lung disease, including chronic obstructive pulmonary disease, emphysema, chronic bronchitis, serious heart conditions, current asthma, type 2 diabetes, kidney disease, immunocompromised condition, or an HIV diagnosis<sup>48</sup>.

### *Risk groups*

We hypothesized that some participants may be at higher risk of SARS-CoV-2 infection in the vaccine/variant era because of membership in a group more directly affected by policy changes, including changes to guidelines and public health messaging. These groups included essential workers, those living in crowded households, and those with children in the household who might attend childcare or school. For essential worker status, household factors, and other binary variables, we assigned exposure status based on any vs. no exposure (e.g., having a confirmed COVID-19 case in the household or not) that occurred within each era.

### *Risk behaviors*

We also hypothesized that some participants may have a higher risk of SARS-CoV-2 infection in the vaccine/variant era relative to the pre-vaccine/wild-type era due to the de-implementation of policies that may change risk factors and behaviors. These risk factors/behaviors included: mask use indoors while visiting non-household members, mask use at work, social distancing with individuals the participant knows, and social distancing with

individuals the participant does not know. For time-dependent exposure variables (e.g., social distancing and masking), we assigned exposure status based on a hierarchy of exposure risk during follow-up. Specifically, participants were classified according to the highest risk strata (e.g., never masking > sometimes masking > always masking) that they reported at one or more follow-up assessments.

#### Composite risk score

We computed a composite COVID-19 risk score, as many of the above COVID-19 risk groups and behaviors are likely to be highly correlated. We applied least absolute shrinkage selection operator (LASSO) regression to select the set of risk factors that best-predicted seroconversion in the pre-vaccine/wild-type era<sup>49</sup>. The LASSO model selected household crowding, having a confirmed COVID-19 case in a household member, indoor dining in a bar/restaurant, gathering with groups of  $\geq 10$ , and no mask use indoors in salons or gyms as the most predictive of seroconversion in our cohort during the pre-vaccine/wild-type era. Scores were assigned to each participant based on their responses for each of the risk factors selected by the LASSO model. Scores were normalized between 0 and 100, with higher scores indicating more engagement in high-risk activities (details in Supplementary Statistical Appendix). The composite score was divided into tertiles for statistical analysis.

#### Vaccination status

For the pre-vaccine/wild-type cohort, vaccination status *at the start of follow-up* was assigned as unvaccinated for all participants in the cohort, since vaccines were not available during Serology Period 1 (i.e., 100% were unvaccinated at the start). Vaccination status *at the end of follow-up* was assigned at the start of Serology Period 2 (based on responses to the V4 questionnaire in Fig. 2), at which time almost the entire cohort (99%) remained unvaccinated. For the vaccine/variant era cohort, vaccination status *at the start of follow-up* was assigned based on vaccination status as of February 10, 2021, which corresponded with the first questionnaire after Serology Period 2 specimen collection (V6 in Fig. 2). Individuals in the vaccine/variant-era cohort were classified according to their vaccination/booster status *as of the end of follow-up* (June 2022), with categories as follows: un/undervaccinated (unvaccinated or did not complete primary vaccine series), completed primary vaccine series, completed primary vaccine series with one booster, and completed primary vaccine series with two or more boosters. Completing a primary vaccine series was defined as one dose for participants who indicated that they had received the Johnson & Johnson vaccine and two doses for any other COVID vaccine type specified.

#### Statistical analysis

Seroincidence of SARS-CoV-2 infection was calculated within each cohort and across strata of sociodemographic factors and epidemiologic risk factors, selected based on the literature and on previous published pre-vaccine/wild-type era analyses of SARS-CoV-2 incidence in this cohort<sup>15</sup>. Crude associations of each factor with SARS-CoV-2 infection were reported as rate ratios. A multivariable mixed effects Poisson model with random coefficients, using the log of total person-time as the offset and an unstructured covariance matrix, was used to estimate the rate ratio of incident SARS-CoV-2 infection stratified by vaccination status (un/undervaccinated, vaccinated, boosted once, boosted more than once) in the vaccine/variant era cohort. The entire pre-vaccine/wild-type era cohort was used as the referent group in these models. We ran a crude and multivariable overall model. To assess the association of vaccination status within different risk factor strata, we ran 12 multivariable models, one for each stratum of five different risk factor groups [essential workers, household children, household cases, social distancing with those you know, and social distancing with those you don't know]. All multivariable models were adjusted for age, sex, and the presence of comorbidities. All mixed models accounted for repeated measures among participants by including a random intercept for subject. All data were cleaned and analyzed in R and SAS.

#### Ethical approval

The study protocol was approved by the Institutional Review Board at the City University of New York (CUNY). All methods were performed in accordance with relevant guidelines and regulations, with informed consent obtained from all study participants.

## Results

### Sample characteristics

The characteristics of the study cohorts are shown in Table 1. Seventy-two percent of subjects ( $n = 2574$ ) were represented in both cohorts, 24% ( $n = 847$ ) were represented only in the pre-vaccine/wild-type era cohort and 4% ( $n = 161$ ) were represented only in the vaccine/variant era cohort (Fig. 1). Participants in each cohort were very similar on measured characteristics, except for employment status, where a slightly lower proportion was unemployed in the vaccine/variant era than in the pre-vaccine/wild-type era (6.5% vs 11.1%, respectively).

#### Vaccination status

In the pre-vaccine/wild-type era, none of the 3421 seronegative participants were vaccinated at the start of the follow-up (March 28, 2020), and only 18 participants (0.53%) had any vaccine doses as of November 17, 2020 (Table 1). In the vaccine/variant era, 282 (10.3%) of the 2735 seronegative participants had completed a primary vaccine series as of February 10, 2021 (V6 questionnaire in Fig. 1); 2497 (91.3%) were fully vaccinated by the end of follow-up, including 2246 (82.1%) boosted at least once and 723 (26.4%) boosted twice. In terms of the timing of vaccination, 87% percent of the vaccine/variant-era cohort had completed their primary series within 6 months of the start of follow-up in the vaccine/variant era (i.e., within 6 months of their seronegative specimen).

	Pre-vaccine/wild-type era cohort		Vaccine/variant era cohort	
	N	%	N	%
Total	3421	100	2735	100
No. participants included in both cohorts	2574	75.24	2574	94.11
No. participants included one cohort	847	24.76	161	5.89
Age				
Median (IQR)	42 (32–56)		42 (33–57)	
18–29	617	18.04	457	16.71
30–39	934	27.3	742	27.13
40–49	654	19.12	518	18.94
50–59	514	15.02	428	15.65
60+	702	20.52	590	21.57
Gender				
Male	1515	44.29	1157	42.3
Female	1810	52.91	1500	54.84
Non-binary/transgender	96	2.81	78	2.85
Race/ethnicity				
Non-Hispanic White	2306	67.41	1881	68.78
Hispanic	500	14.62	364	13.31
Non-Hispanic Black	261	7.63	202	7.39
Asian/PI	222	6.49	186	6.8
Other	132	3.86	102	3.73
Education at baseline				
Less than high school	37	1.08	29	1.06
High school graduate	278	8.13	206	7.53
Some college	815	23.82	640	23.4
College graduate	2291	66.97	1860	68.01
Household income				
Less than \$35,000	891	26.05	696	25.45
\$35–49,999	392	11.46	309	11.3
\$50–69,999	511	14.94	427	15.61
\$70–99,999	609	17.8	495	18.1
\$100,000+	1018	29.76	808	29.54
Employment at baseline				
Employed	2175	63.58	1850	67.64
Out of work	378	11.05	179	6.54
Homemaker	165	4.82	131	4.79
Student	192	5.61	123	4.5
Retired	511	14.94	452	16.53
Setting				
Urban	1460	42.68	1117	40.84
Suburban	901	26.34	752	27.5
Rural	1052	30.75	863	31.55
Town	8	0.23	3	0.11
Geographic region				
Northeast	963	28.15	776	28.37
Midwest	632	18.47	485	17.73
South	960	28.06	796	29.1
West	862	25.2	674	24.64
US territories	4	0.12	4	0.15
Healthcare worker				
No	3073	89.83	2465	90.13
Yes	315	9.21	245	8.96
Don't know	33	0.96	25	0.91
Essential worker				
No	2811	82.17	2272	83.07
Yes	610	17.83	463	16.93
Continued				



	Pre-vaccine/wild-type era cohort		Vaccine/variant era cohort	
	N	%	N	%
Higher risk for severe COVID (V1 or V9)				
No	2627	76.79	2045	74.77
Yes	794	23.21	690	25.23
Vaccination status at start of follow-up*				
Un/under-vaccinated	3421	100.00	2447	89.47
Primary series only	0	0.00	282	10.31
Missing	0	0.00	6	0.22
Vaccination status at end of follow-up*				
Un/under-vaccinated	3376	98.68	238	8.70
Primary series	18	0.53	251	9.18
Boosted once	0	0.00	1523	55.69
2+ boosters	0	0.00	723	26.44
Missing	27	0.79	0	0.00

**Table 1.** Characteristics of study participants in each cohort. \*Serology period was defined as time from S1–S2 and S2–S3 for pre-vaccine and vaccine eras respectively. The pre-vaccine era period was defined using v1–v4 surveys, and the post-vaccine period was defined using v6–v11 surveys.

### Seroincidence of SARS-CoV-2 infection

The seroincidence rate of SARS-CoV-2 infection in the vaccine/variant era cohort was nearly three times higher than in the pre-vaccine/wild-type era cohort. Specifically, we observed a SARS-CoV-2 infection rate of 9.61 per 100 person-years (95% CI 8.3–11.1) and 25.74 per 100 person-years (95% CI 24.2–27.3) in the pre-vaccine/wild-type era cohort and vaccine/variant era cohort, respectively (Table 2).

#### Sociodemographic factors

Table 2 also shows the SARS-CoV-2 incidence rate and univariate (crude) SARS-CoV-2 incidence rate ratios by sociodemographic factors and cohort era. Across the two cohorts, crude incidence rates were substantially higher in all sociodemographic subgroups in the vaccine/variant era cohort compared with the pre-vaccine/wild-type era cohort. Within each of the two cohorts, the SARS-CoV-2 infection rate varied substantially by sociodemographic factors, with *lower* SARS-CoV-2 infection in those aged 60 and older compared with 18–29 year-olds in both cohorts ( $IRR_{\text{pre-vaccine/wild-type}} = 0.53$  [95% CI 0.31–0.90];  $IRR_{\text{vaccine/variant}} = 0.43$  [95% CI 0.34–0.55]) and in women compared to men in the pre-vaccine/wild-type era cohort ( $IRR_{\text{pre-vaccine/wild-type}} = 0.69$  [95% CI 0.51–0.95]). Some associations appeared to be protective only in the vaccine/variant era cohort (household income above \$100,000 vs. less than \$35,000, retired vs. employed, and higher vs. lower risk of severe COVID). In each cohort, we observed *higher* seroincidence of SARS-CoV-2 infection among Hispanic ( $IRR_{\text{pre-vaccine/wild-type}} = 2.06$  [95% CI 1.41–3.01];  $IRR_{\text{vaccine/variant}} = 1.49$  [95% CI 1.24–1.80]) and non-Hispanic Black ( $IRR_{\text{pre-vaccine/wild-type}} = 1.69$  [95% CI 0.99–2.87];  $IRR_{\text{vaccine/variant}} = 1.68$  [95% CI 1.34–2.11]) participants compared with non-Hispanic White participants, essential workers compared with non-essential workers ( $IRR_{\text{pre-vaccine/wild-type}} = 1.69$  [95% CI 1.19–2.39];  $IRR_{\text{vaccine/variant}} = 1.28$  [95% CI 1.08–1.52]), and in the South compared with the Northeast ( $IRR_{\text{pre-vaccine/wild-type}} = 1.69$  [95% CI 1.10–2.60];  $IRR_{\text{vaccine/variant}} = 1.32$  [95% CI 1.10–1.58]). However, some of these associations (Hispanic vs. non-Hispanic Whites, living in the Southern United States vs. the Northeast), became less pronounced in the vaccine/variant era cohort. Individuals who were at high risk for severe COVID-19 (vs. lower risk) also had significantly lower SARS-CoV-2 infection risk in the vaccine/variant era cohort ( $IRR_{\text{vaccine/variant}} = 0.66$  [95% CI 0.56–0.79]), but not in the pre-vaccine/wild-type era cohort.

#### Vaccination status

In the vaccine/variant era cohort, the highest crude rates of SARS-CoV-2 infection were in un/under-vaccinated (51.34, 95% CI 45.02–57.63) and fully vaccinated but not boosted participants (48.94, 95% CI 42.83–55.09) (Table 2). Those who were fully vaccinated and received a booster had substantially lower rates of SARS-CoV-2 infection than the un/under-vaccinated, including those with one booster ( $IRR_{\text{vaccine/variant}} = 0.47$  [95% CI 0.39–0.57]) and two or more boosters ( $IRR_{\text{vaccine/variant}} = 0.28$  [95% CI 0.22–0.36]).

#### Epidemiologic risk factors

Table 3 shows the SARS-CoV-2 infection and univariate (crude) SARS-CoV-2 infection rate ratios by epidemiologic risk factors that were present prior to or between serologic tests for each cohort. Crude incidence rates were substantially higher in most subgroups of epidemiologic risk factors in the vaccine/variant era cohort compared with the pre-vaccine/wild-type era cohort. In both cohorts, never social distancing with people you do not know ( $IRR_{\text{pre-vaccine/wild-type}} = 3.16$  [95% CI 1.47–6.77];  $IRR_{\text{vaccine/variant}} = 2.00$  [95% CI 1.62–2.48]) and never masking or sometimes masking in a variety of settings were associated with a higher risk of SARS-CoV-2 infection. Binge drinking ( $IRR_{\text{pre-vaccine/wild-type}} = 1.47$  [95% CI 1.07–2.03];  $IRR_{\text{vaccine/variant}} = 1.45$  [95% CI 1.26–1.67]) and recent air

	Pre-vaccine/wild type-era cohort			Vaccine/variant-era cohort		
	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)
Total seronegative at start of follow-up	3421			2735		
Number of seroconversions during follow-up	161			815		
Overall seroincidence		9.61 (8.26, 11.14)			25.74 (24.23, 27.30)	
≥1 Self-reported PCR/rapid test during follow-up	185			686		
<b>Age</b>						
18–29	617	12.08 (8.75, 16.39)	[ref]	457	33.80 (29.71, 38.15)	[ref]
30–39	934	10.04 (7.52, 13.26)	0.83 (0.54, 1.28)	742	31.64 (28.52, 34.94)	0.94 (0.77, 1.14)
40–49	654	10.42 (7.38, 14.45)	0.86 (0.54, 1.38)	518	28.29 (24.74, 32.13)	0.84 (0.68, 1.04)
50–59	514	9.08 (5.96, 13.48)	0.75 (0.45, 1.26)	428	20.93 (17.53, 24.76)	0.62 (0.49, 0.79)
60+	702	6.44 (4.17, 9.73)	0.53 (0.31, 0.90)	590	14.63 (12.18, 17.47)	0.43 (0.34, 0.55)
<b>Gender</b>						
Male	1515	11.52 (9.39, 14.05)	[ref]	1157	27.91 (25.53, 30.43)	[ref]
Female	1810	8.00 (6.32, 10.07)	0.69 (0.51, 0.95)	1500	24.42 (22.43, 26.52)	0.87 (0.76, 1.01)
Non-binary/transgender	96	8.00 (2.59, 20.11)	0.69 (0.25, 1.89)	78	19.57 (12.48, 29.13)	0.70 (0.44, 1.11)
<b>Race/ethnicity</b>						
Non-Hispanic White	2306	7.95 (6.48, 9.69)	[ref]	1881	23.00 (21.27, 24.83)	[ref]
Hispanic	500	16.35 (11.91, 21.97)	2.06 (1.41, 3.01)	364	34.30 (29.72, 39.17)	1.49 (1.24, 1.80)
Non-Hispanic Black	261	13.42 (8.10, 21.18)	1.69 (0.99, 2.87)	202	38.74 (32.44, 45.44)	1.68 (1.34, 2.11)
Asian/PI	222	6.46 (2.86, 13.32)	0.81 (0.38, 1.75)	186	23.01 (17.65, 29.36)	1.00 (0.75, 1.34)
Other	132	14.04 (7.01, 25.49)	1.77 (0.89, 3.50)	102	27.13 (19.66, 36.08)	1.18 (0.83, 1.68)
<b>Education at baseline</b>						
Less than high school	37	12.25 (2.15, 38.96)	[ref]	29	34.69 (19.36, 53.62)	[ref]
High school graduate	278	11.32 (6.56, 18.58)	0.92 (0.21, 4.07)	206	31.76 (26.03, 38.08)	0.92 (0.49, 1.72)
Some college	815	11.42 (8.50, 15.12)	0.93 (0.23, 3.85)	640	29.74 (26.47, 33.21)	0.86 (0.47, 1.57)
College graduate	2291	8.78 (7.24, 10.60)	0.72 (0.18, 2.91)	1860	23.57 (21.81, 25.43)	0.68 (0.37, 1.23)
<b>Household income</b>						
Less than \$35,000	891	10.45 (7.78, 13.87)	[ref]	696	27.71 (24.68, 30.95)	[ref]
\$35–49,999	392	12.54 (8.36, 18.27)	1.20 (0.73, 1.97)	309	29.36 (24.66, 34.53)	1.06 (0.84, 1.34)
\$50–69,999	511	7.65 (4.79, 11.88)	0.73 (0.43, 1.25)	427	27.89 (24.01, 32.12)	1.01 (0.81, 1.24)
\$70–99,999	609	12.00 (8.65, 16.35)	1.15 (0.74, 1.78)	495	26.64 (23.09, 30.51)	0.96 (0.78, 1.18)
\$100,000+	1018	7.38 (5.34, 10.07)	0.71 (0.46, 1.09)	808	21.08 (18.55, 23.84)	0.76 (0.63, 0.92)
<b>Employment at baseline</b>						
Employed	2175	10.45 (8.72, 12.47)	[ref]	1850	28.79 (26.87, 30.78)	[ref]
Out of work	378	6.69 (3.66, 11.67)	0.64 (0.35, 1.16)	179	24.92 (19.43, 31.33)	0.87 (0.66, 1.14)
Homemaker	165	6.85 (2.55, 15.92)	0.65 (0.27, 1.60)	131	29.01 (22.01, 37.13)	1.01 (0.74, 1.37)
Student	192	13.70 (7.78, 22.65)	1.31 (0.74, 2.33)	123	29.28 (22.05, 37.66)	1.02 (0.74, 1.40)
Retired	511	7.27 (4.49, 11.44)	0.70 (0.42, 1.14)	452	12.85 (10.25, 15.97)	0.45 (0.35, 0.57)
<b>Setting</b>						
Urban	1460	8.91 (6.99, 11.27)	[ref]	1117	27.28 (24.87, 29.84)	[ref]
Suburban	901	8.80 (6.41, 11.93)	0.99 (0.66, 1.47)	752	24.67 (21.87, 27.70)	0.90 (0.76, 1.07)
Rural	1052	11.42 (8.83, 14.62)	1.28 (0.90, 1.83)	863	24.72 (22.12, 27.51)	0.91 (0.77, 1.07)
Town	8	0.00 (0.00, 60.27)	N/A	3	25.45 (1.34, 78.67)	0.93 (0.13, 6.64)
<b>Geographic region</b>						
Northeast	963	6.99 (4.96, 9.72)	[ref]	776	22.49 (19.83, 25.40)	[ref]
Midwest	632	11.06 (7.88, 15.24)	1.58 (0.98, 2.55)	485	26.33 (22.78, 30.22)	1.17 (0.95, 1.45)
South	960	11.84 (9.09, 15.25)	1.69 (1.10, 2.60)	796	29.61 (26.70, 32.70)	1.32 (1.10, 1.58)
West	862	9.20 (6.70, 12.46)	1.32 (0.83, 2.09)	674	24.65 (21.70, 27.85)	1.10 (0.90, 1.33)
US territories	4	0.00 (0.00, 80.60)	N/A	4	0.00 (0.00, 52.65)	N/A
<b>Healthcare worker</b>						
No	3073	9.29 (7.90, 10.90)	[ref]	2465	25.46 (23.88, 27.11)	[ref]
Yes	315	12.32 (7.77, 18.82)	1.33 (0.82, 2.14)	245	28.64 (23.43, 34.47)	1.13 (0.89, 1.42)
Don't know	33	13.11 (2.30, 41.09)	1.41 (0.35, 5.70)	25	25.52 (11.68, 45.99)	1.00 (0.48, 2.11)
<b>Essential worker</b>						
Continued						



	Pre-vaccine/wild type-era cohort			Vaccine/variant-era cohort		
	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)
No	2811	8.54 (7.14, 10.17)	[ref]	2272	24.59 (22.97, 26.28)	[ref]
Yes	610	14.41 (10.77, 18.97)	1.69 (1.19, 2.39)	463	31.58 (27.64, 35.80)	1.28 (1.08, 1.52)
Composite risk factors score						
Low	1141	4.13 (2.70, 6.23)	[ref]	912	8.89 (7.35, 10.72)	[ref]
Medium	1140	8.32 (6.24, 10.98)	2.01 (1.22, 3.31)	911	56.31 (52.92, 59.64)	4.00 (3.27, 4.89)
High	1140	16.42 (13.49, 19.83)	3.97 (2.51, 6.28)	912	19.78 (17.54, 22.23)	2.39 (1.93, 2.96)
Higher risk for severe COVID						
No	2627	10.04 (8.48, 11.83)	[ref]	2045	28.25 (26.44, 30.13)	[ref]
Yes	794	8.15 (5.69, 11.49)	0.81 (0.55, 1.20)	690	18.70 (16.15, 21.55)	0.66 (0.56, 0.79)
Vaccination status at end of follow-up*						
Un/under-vaccinated	3376	9.67 (8.31, 11.22)	[ref]	238	51.34 (45.02, 57.63)	[ref]
Primary series	18	0.00 (0.00, 38.58)	N/A	251	48.94 (42.83, 55.09)	0.95 (0.75, 1.22)
Boosted once	0			1523	24.16 (22.19, 26.24)	0.47 (0.39, 0.57)
2+ boosters	0			723	14.55 (12.33, 17.08)	0.28 (0.22, 0.36)
Missing	27	7.77 (0.41, 38.20)	0.80 (0.11, 5.74)	0		

**Table 2.** Crude seroincidence estimates in the pre-vaccine era and vaccine-era cohorts by sociodemographic factors and vaccination status. \*Serology period was defined as time from S1–S2 and S2–S3 for pre-vaccine and vaccine eras respectively. The pre-vaccine era period was defined using v1–v4 surveys, and the post-vaccine period was defined using v6–v11 surveys.

travel ( $IRR_{\text{pre-vaccine/wild-type}}$  1.49 [95% CI 1.04–2.13];  $IRR_{\text{vaccine/variant}}$  1.20 [95% CI 1.03–1.39]) were associated with higher risk of SARS-CoV-2 infection. A higher composite measure of risk was significantly associated with a higher risk of incident infection in both eras (Table 2).

#### Changes in epidemiologic risk factors between pre-vaccine and vaccine/variant eras

Some new associations emerged that were not present in the pre-vaccine/wild-type era cohort (Table 3). In the pre-vaccine/wild-type era cohort, people living with < 4 household members in a single unit of a multi-unit dwelling and those living with 4 or more household members in a single-family dwelling had similar incidence as those living in single-family dwellings with < 4 household members. But in the vaccine/variant era cohort, those living with < 4 household members in a single unit of a multi-unit dwelling and those living with 4 or more household members in a single-family dwelling saw their risk increase compared with those living in single-family dwellings with < 4 household members ( $IRR_{\text{vaccine/variant}}$  1.39 [95% CI 1.18–1.63] for multi-unit dwelling with < 4 household members;  $IRR_{\text{vaccine/variant}}$  1.59 [95% CI 1.32–1.92] for single-family dwelling with 4+ household members). Similarly, as shown in Table 3, having a child in the household was not associated with a higher risk of SARS-CoV-2 infection in the pre-vaccine/wild-type era cohort, but was in the vaccine/variant-era cohort ( $IRR_{\text{vaccine/variant}}$  1.41 [95% CI 1.22–1.64]). Social distancing ‘with people you don’t know’ sometimes (vs. always) was not associated with a higher risk of SARS-CoV-2 infection in the pre-vaccine/wild-type era cohort, but was significantly associated with a higher risk in the vaccine/variant era cohort ( $IRR_{\text{vaccine/variant}}$  1.25 [95% CI 1.03–1.50]). Associations for some risk factors persisted but became less pronounced in the vaccine/variant era cohort compared with the pre-vaccine/wild-type era cohort (Table 3). Specifically, having a confirmed case in the household had the highest absolute incidence rate and was the strongest risk factor in each cohort, but the strength of the association decreased in the vaccine/variant-era cohort ( $IRR_{\text{vaccine/variant}}$  8.35 [95% CI 7.22–9.66]) vs. the pre-vaccine/wild-type era cohort ( $IRR_{\text{pre-vaccine/wild-type}}$  22.34 [95% CI 14.77–33.77]). Social distancing ‘with people you know’ never (vs. always) was associated with a higher risk of SARS-CoV-2 infection in the pre-vaccine/wild-type era cohort ( $IRR_{\text{pre-vaccine/wild-type}}$  3.16 [95% CI 1.47–6.77]) than in the vaccine/variant era cohort ( $IRR_{\text{vaccine/variant}}$  2.00 [95% CI 1.62–2.48]). The risk ratio for SARS-CoV-2 infection was also lower in the vaccine/variant era cohort than the pre-vaccine/wild-type era cohort for indoor dining, visiting a place of worship, gathering indoors with 10 or more persons, and social distancing with ‘people you do not know’ never (vs. always).

Associations of mask use with incidence depended on the context. For mask use while grocery shopping, at the salon or gym, or on public transit, risk for sometimes use (vs. always) was elevated in both cohorts but was less pronounced in the vaccine/variant era cohort (Table 3). No mask use (vs. always) while indoors visiting non-household members was associated with an elevated risk of infection in both cohorts, but less so in the vaccine/variant era cohort. Mask use sometimes or never (vs. always) while indoors at work was associated with a higher risk in the vaccine/variant era, but not in the pre-vaccine/wild-type era, as was no mask use (vs. always) while at the salon or gym or while on public transit.

#### Poisson models of SARS-CoV-2 seroconversion in strata of risk factors

Table 4 shows adjusted IRRs and 95% CIs from the overall multivariable model and the 12 risk-factor group-specific multivariate Poisson models stratified by vaccine status in the vaccine/variant era, with the pre-vaccine/wild-type cohort as the referent group, adjusting for age, gender, and presence of co-morbidities. Relative to the

COVID-19 risk factors	Pre-vaccine/wild type-era cohort			Vaccine/variant-era cohort		
	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)
Total seronegative at start of follow-up	3421			2735		
Number of seroconversions during follow-up	161			815		
Overall seroincidence		9.61 (8.26, 11.14)			25.74 (24.23, 27.30)	
<b>Household characteristics</b>						
Household crowding						
Single-family with < 4 members	1373	8.50 (6.55, 10.94)	[ref]	1129	20.53 (18.42, 22.81)	[ref]
Single-family with 4+ members	625	9.20 (6.26, 13.24)	1.08 (0.68, 1.71)	494	32.69 (28.80, 36.82)	1.59 (1.32, 1.92)
Multi-family with < 4 members	1160	9.18 (7.03, 11.88)	1.08 (0.74, 1.57)	917	28.49 (25.79, 31.34)	1.39 (1.18, 1.63)
Multi-family with 4+ members	171	17.58 (10.28, 28.08)	2.07 (1.15, 3.71)	128	33.67 (26.11, 42.12)	1.64 (1.21, 2.23)
Dorms/Group homes/Other congregate settings	17	11.56 (0.61, 50.66)	1.36 (0.19, 9.82)	13	18.33 (4.86, 45.50)	0.89 (0.29, 2.79)
Other	72	20.56 (9.33, 38.37)	2.42 (1.10, 5.30)	51	14.20 (7.09, 25.76)	0.69 (0.36, 1.34)
Children in the household						
No children in household	2456	9.45 (7.90, 11.26)	[ref]	2037	23.36 (21.69, 25.12)	[ref]
Children in household	965	10.05 (7.50, 13.30)	1.06 (0.75, 1.50)	698	33.03 (29.75, 36.47)	1.41 (1.22, 1.64)
Household exposures						
No confirmed case in household member <sup>^</sup>	3377	0.67 (0.57, 0.80)	[ref]	2384	1.51 (1.39, 1.65)	[ref]
Confirmed case in household member <sup>^</sup>	44	15.02 (10.30, 21.28)	22.34 (14.77, 33.77)	351	12.65 (11.29, 14.15)	8.35 (7.22, 9.66)
Social distancing						
Social distancing with people you know						
Always	1136	9.26 (7.03, 12.07)	[ref]	93	26.03 (18.39, 35.36)	[ref]
Sometimes	1787	8.94 (7.18, 11.06)	0.97 (0.68, 1.37)	839	18.14 (15.83, 20.70)	0.70 (0.47, 1.03)
Never	300	16.74 (11.22, 24.09)	1.81 (1.11, 2.94)	1731	29.27 (27.28, 31.35)	1.12 (0.77, 1.63)
NA	158	5.31 (1.71, 13.74)	0.57 (0.21, 1.59)	72	33.39 (23.35, 45.08)	1.28 (0.76, 2.18)
Social distancing with people you do not know						
Always	2615	8.85 (7.38, 10.58)	[ref]	646	19.32 (16.64, 22.31)	[ref]
Sometimes	660	10.39 (7.40, 14.35)	1.17 (0.80, 1.72)	1468	24.07 (22.08, 26.18)	1.25 (1.03, 1.50)
Never	53	27.94 (12.85, 49.52)	3.16 (1.47, 6.77)	466	38.66 (34.39, 43.10)	2.00 (1.62, 2.48)
NA	53	15.99 (5.25, 36.89)	1.81 (0.67, 4.90)	155	33.58 (26.75, 41.14)	1.74 (1.29, 2.35)
Spent time in public places						
Did not attend mass gatherings	3071	9.55 (8.13, 11.18)	[ref]	2460	24.97 (23.41, 26.61)	[ref]
Attended mass gatherings	350	10.08 (6.25, 15.70)	1.06 (0.65, 1.72)	275	32.80 (27.64, 38.40)	1.31 (1.07, 1.62)
Indoor dining/bar						
No indoor dining/bar	1666	6.52 (4.96, 8.50)	[ref]	391	20.12 (16.67, 24.06)	[ref]
Indoor dining/bar	1755	12.52 (10.42, 14.96)	1.92 (1.38, 2.67)	2344	26.73 (25.08, 28.46)	1.33 (1.07, 1.64)
Outdoor dining/bar						
No outdoor dining/bar	1551	8.75 (6.85, 11.09)	[ref]	674	24.48 (21.57, 27.64)	[ref]
Outdoor dining/bar	1870	10.28 (8.45, 12.43)	1.17 (0.86, 1.61)	2061	26.16 (24.41, 27.99)	1.07 (0.91, 1.25)
Place of worship						
Did not visit place of worship	3062	8.78 (7.42, 10.35)	[ref]	2017	23.83 (22.13, 25.61)	[ref]
Visited place of worship	359	16.81 (11.72, 23.42)	1.91 (1.28, 2.86)	718	31.38 (28.20, 34.75)	1.32 (1.14, 1.53)
Public park/pool						
Did not visit public park/pool	1087	9.48 (7.14, 12.46)	[ref]	529	23.80 (20.57, 27.34)	[ref]
Visited public park/pool	2334	9.66 (8.06, 11.53)	1.02 (0.73, 1.43)	2206	26.22 (24.53, 27.99)	1.10 (0.92, 1.31)
Gathered in groups ≥ 10						
No	2392	8.65 (7.13, 10.44)	[ref]	422	19.87 (16.57, 23.63)	[ref]
Indoors only	195	18.20 (11.25, 27.82)	2.10 (1.26, 3.52)	307	25.67 (21.29, 30.57)	1.29 (0.97, 1.71)
Outdoors only	453	9.27 (5.97, 14.02)	1.07 (0.67, 1.72)	231	23.14 (18.32, 28.74)	1.16 (0.85, 1.60)
Indoors and outdoors	353	12.11 (7.83, 18.14)	1.40 (0.88, 2.24)	1775	27.60 (25.67, 29.61)	1.39 (1.13, 1.71)
Mask use						
Mask while grocery shopping						
Always	3083	9.24 (7.85, 10.84)	[ref]	1461	19.53 (17.70, 21.50)	[ref]
Continued						

COVID-19 risk factors	Pre-vaccine/wild type-era cohort			Vaccine/variant-era cohort		
	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)	N	Seroincidence per 100 person-years (95% CI)	Rate ratio (95% CI)
Sometimes	132	20.31 (11.37, 33.16)	2.20 (1.22, 3.96)	693	29.65 (26.50, 33.00)	1.52 (1.28, 1.79)
Never	33	99.00 (93.76, 99.95)	0.00 (0.00, NA)	408	40.89 (36.35, 45.59)	2.09 (1.75, 2.50)
Did not go grocery shopping	133	9.26 (3.82, 19.72)	1.00 (0.44, 2.27)	173	29.79 (23.62, 36.76)	1.53 (1.16, 2.01)
Mask while indoors visiting non-household members						
Always	1118	7.76 (5.74, 10.39)	[ref]	154	16.89 (12.00, 23.16)	[ref]
Sometimes	1131	11.30 (8.85, 14.29)	1.46 (0.99, 2.14)	761	19.78 (17.24, 22.59)	1.17 (0.80, 1.71)
Never	463	15.12 (10.78, 20.73)	1.95 (1.24, 3.07)	1701	29.64 (27.62, 31.73)	1.75 (1.23, 2.50)
NA (did not visit indoors)	668	5.88 (3.67, 9.18)	0.76 (0.44, 1.30)	119	21.32 (15.05, 29.18)	1.26 (0.77, 2.08)
Mask while indoors at work						
Always	1372	10.53 (8.36, 13.15)	[ref]	811	21.71 (19.16, 24.48)	[ref]
Sometimes	299	12.21 (7.60, 18.87)	1.16 (0.69, 1.95)	747	30.49 (27.40, 33.75)	1.40 (1.17, 1.69)
Never	68	19.40 (8.14, 38.13)	1.84 (0.80, 4.24)	380	38.50 (33.90, 43.31)	1.77 (1.45, 2.18)
NA (did not attend indoor workplace)	1641	7.87 (6.14, 10.01)	0.75 (0.53, 1.05)	797	19.82 (17.35, 22.54)	0.91 (0.75, 1.11)
Mask while at salon/gym						
Always	1527	9.90 (7.91, 12.31)	[ref]	982	19.22 (17.02, 21.62)	[ref]
Sometimes	180	23.10 (15.00, 33.63)	2.33 (1.42, 3.82)	656	28.49 (25.29, 31.91)	1.48 (1.23, 1.79)
Never	108	13.91 (6.24, 27.20)	1.40 (0.65, 3.05)	645	35.97 (32.47, 39.62)	1.87 (1.57, 2.24)
NA (did not attend)	1566	7.37 (5.66, 9.53)	0.74 (0.53, 1.05)	452	22.48 (19.07, 26.29)	1.17 (0.94, 1.46)
Mask while on public transit						
Always	758	10.32 (7.52, 13.94)	[ref]	1123	24.08 (21.80, 26.52)	[ref]
Sometimes	32	28.82 (9.67, 58.36)	2.79 (1.00, 7.82)	109	42.93 (33.99, 52.34)	1.78 (1.33, 2.40)
Never	80	13.71 (5.16, 29.94)	1.33 (0.52, 3.37)	268	39.32 (33.75, 45.17)	1.63 (1.32, 2.02)
NA (did not use)	2510	8.98 (7.47, 10.76)	0.87 (0.60, 1.25)	1235	23.05 (20.92, 25.32)	0.96 (0.82, 1.12)
Outdoor mask use						
No mask use outdoors	707	11.04 (7.98, 15.01)	[ref]	182	32.40 (26.13, 39.36)	[ref]
Mask use outdoors	2714	9.25 (7.78, 10.96)	0.84 (0.58, 1.21)	2553	25.28 (23.73, 26.89)	0.78 (0.61, 1.00)
Movement during the pandemic						
Use of public transit						
Avoided or did not use	730	10.29 (7.44, 14.02)	[ref]	601	28.19 (24.87, 31.76)	[ref]
Used public transit	2691	9.42 (7.92, 11.16)	0.92 (0.63, 1.32)	2134	25.06 (23.38, 26.83)	0.89 (0.76, 1.05)
Recent air travel						
No	2839	8.84 (7.42, 10.50)	[ref]	2019	24.51 (22.80, 26.31)	[ref]
Yes	582	13.16 (9.63, 17.67)	1.49 (1.04, 2.13)	716	29.34 (26.23, 32.64)	1.20 (1.03, 1.39)
Alcohol/substance use						
Binge drinking						
No	2470	8.50 (7.02, 10.25)	[ref]	1878	22.64 (20.92, 24.46)	[ref]
Yes	951	12.51 (9.71, 15.95)	1.47 (1.07, 2.03)	857	32.79 (29.85, 35.86)	1.45 (1.26, 1.67)
Regular cannabis or un-prescribed opioid use						
No	2953	9.75 (8.29, 11.42)	[ref]	2146	25.26 (23.58, 27.03)	[ref]
Yes	468	8.71 (5.54, 13.32)	0.89 (0.56, 1.43)	589	27.46 (24.17, 31.01)	1.09 (0.92, 1.28)

**Table 3.** Crude seroincidence estimates in the pre-vaccine era and vaccine-era cohorts by epidemiologic risk factors. ^Person-time is in months.

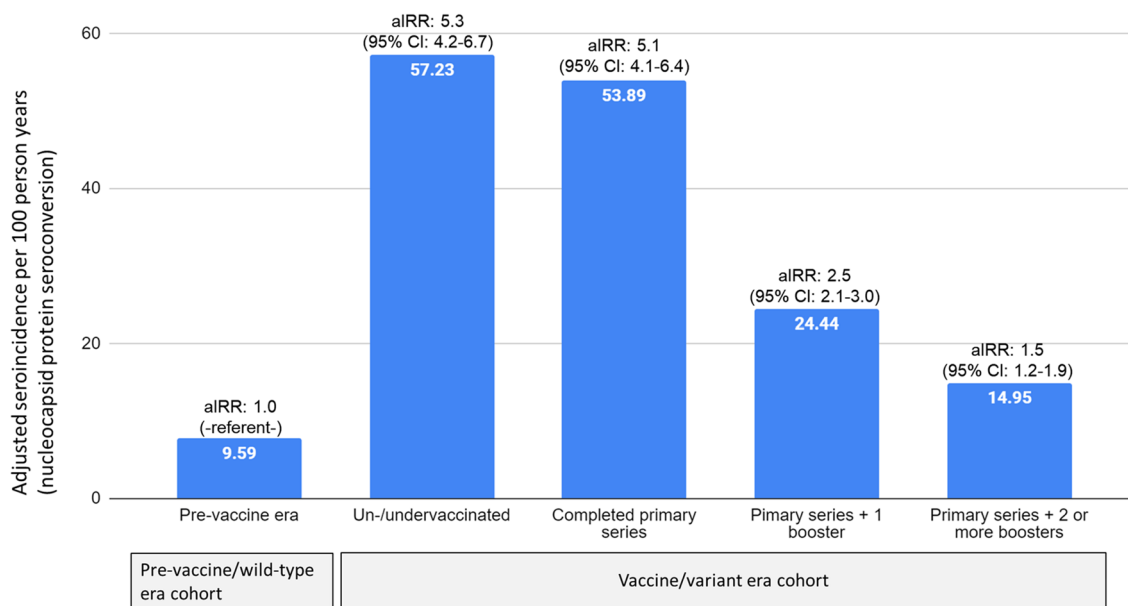
pre-vaccine/wild-type cohort, the risk of SARS-CoV-2 infection was similar between un/undervaccinated and fully vaccinated participants in the vaccine/variant era. However, the risk of infection tended to decrease with an increasing number of booster doses. Specifically, adjusted incident rate ratios (aIRR) with the pre-vaccine/wild-type cohort as the referent group were:  $aIRR_{un/undervaccinated} = 5.3$  (95% CI 4.2–6.7);  $aIRR_{primary\ series\ only} = 5.1$  (95% CI 4.1–6.4);  $aIRR_{boosted\ once} = 2.5$  (95% CI 2.1–3.0), and  $aIRR_{boosted\ twice} = 1.65$  (95% CI 1.3–2.1) (Table 4, Fig. 3). These associations were essentially unchanged within risk factor-stratified models, except for those with a confirmed case of SARS-CoV-2 in the household, where the relative change in SARS-CoV-2 infection between the pre-vaccine/wild-type cohort and the vaccine/variant-era cohort was smaller than that in other risk groups.

### Self-reported testing

Table 5 shows the number and proportion of participants who tested positive on serologic tests as part of our study as well as on self-reported PCR or rapid tests reported to have been taken by participants outside of the

	N participants	N Observations	Pre-vaccine/wild type era cohort (n = 3421)	Vaccine/variant-era cohort (n = 2735)			
				Un/undervaccinated	Primary series only	Boosted once	2+ Boosters
Total	3582	6156	3421	238	251	1523	723
Overall model (all participants)							
Crude	3582	6156	-ref-	5.30 (4.19, 6.71)	5.10 (4.05, 6.44)	2.52 (2.10, 3.02)	1.65 (1.31, 2.09)
Adjusted*	3582	6156	-ref-	5.34 (4.24, 6.73)	5.09 (4.05, 6.42)	2.51 (2.10, 3.01)	1.51 (1.20, 1.91)
Risk factor-stratified models*							
Essential worker							
No	2939	5083	-ref-	5.43 (4.14, 7.12)	5.44 (4.17, 7.10)	2.74 (2.22, 3.38)	1.79 (1.38, 2.34)
Yes	643	1073	-ref-	5.44 (3.37, 8.78)	4.17 (2.59, 6.71)	1.96 (1.36, 2.82)	1.37 (0.82, 2.29)
Household children							
No child in household	2702	4493	-ref-	4.38 (3.15, 6.08)	5.08 (3.81, 6.75)	2.49 (2.01, 3.09)	1.69 (1.29, 2.20)
Child in household	1095	1663	-ref-	6.06 (4.18, 8.79)	5.27 (3.51, 7.90)	2.61 (1.85, 3.70)	1.55 (0.91, 2.62)
Household cases							
No confirmed case in household member	3519	5761	-ref-	4.55 (3.46, 5.98)	4.27 (3.25, 5.61)	2.10 (1.71, 2.58)	1.46 (1.12, 1.91)
Confirmed case in household member	388	395	-ref-	1.54 (0.80, 2.95)	1.42 (0.75, 2.69)	1.01 (0.60, 1.73)	0.86 (0.46, 1.61)
Social distancing with people you know							
Always	1148	1229	-ref-	5.72 (2.70, 12.12)	4.33 (1.77, 10.56)	1.67 (0.79, 3.50)	1.36 (0.53, 3.49)
Sometimes/never	3139	4657	-ref-	5.15 (3.94, 6.74)	4.86 (3.73, 6.34)	2.43 (1.96, 3.02)	1.63 (1.25, 2.12)
NA	258	270	-ref-	6.01 (1.31, 27.70)	18.67 (4.81, 72.39)	5.76 (2.18, 15.23)	2.01 (0.50, 8.08)
Social distancing with people you do not know							
Always	2673	3261	-ref-	4.47 (2.94, 6.80)	4.20 (2.66, 6.64)	1.90 (1.39, 2.58)	1.48 (0.99, 2.21)
Sometimes/never	2141	2647	-ref-	4.95 (3.42, 7.17)	4.25 (2.94, 6.16)	2.20 (1.59, 3.05)	1.47 (1.01, 2.12)
NA	206	208	-ref-	1.89 (0.52, 6.92)	3.67 (1.22, 11.01)	2.01 (0.70, 5.83)	0.99 (0.29, 3.46)

**Table 4.** Incidence rate ratios (IRRs) from multivariate models comparing incidence in the pre-vaccine/wild-type era cohort to that within strata of vaccination status in the vaccine/variant era cohort. \*Adjusted for age, gender, and comorbidities.



\*Adjusted for age, gender, and presence of comorbidities

**Figure 3.** Adjusted\* incidence rates and incidence rate ratios (aIRR) for SARS-CoV-2 infection by vaccination status compared with the pre-vaccine era.

	Pre-vaccine/wild-type era cohort (n = 3421)		Vaccine/variant era cohort (n = 2735)	
	Number positive	Positivity rate (%)	Number positive	Positivity rate (%)
Positive serologic test (study)	161	4.71	815	29.80
Self-reported positive PCR/rapid test (outside study)	137	4.00	561	20.51
Ratio of self report/serology positivity rates	85%		69%	
Positive serologic test AND self-reported positive test	47	1.37	397	14.52
Positive serologic test and NO self-reported positive	114	3.33	418	15.28
Ratio of self-reported positive/positive serologic test	29%		49%	

**Table 5.** Positivity rate of serologic testing compared with self-reported PCR/rapid testing.

study for each cohort. Compared to serologic test results, participants had lower rates of test positivity on self-reported viral PCR or rapid tests. Specifically, in the pre-vaccine/wild-type era cohort, 4.0% (n = 137) of participants self-reported a positive PCR or rapid test outside of the study, compared to 4.7% (n = 161) that tested positive on serologic testing (ratio 85%). Of the pre-vaccine/wild-type era participants with positive serologic results, 29% (n = 47) had also self-reported at least one positive viral PCR or rapid test during that era. In the vaccine/variant era cohort, 21% (n = 561) self-reported a positive test, compared to 30% (n = 815) from serologic testing (ratio 69%). Among this cohort, 49% (n = 397) of participants had also self-reported at least one positive viral PCR or rapid test. Thus, the proportion of participants with an infection detected outside the serologic testing conducted by the study declined from 85% in the pre-vaccine/wild-type era cohort to 69% in the vaccine/variant era cohort (Table 5).

## Discussion

In a community-based prospective study with repeat serologic testing of SARS-CoV-2 N protein seronegative individuals, we observed a nearly threefold increase in the incidence of SARS-CoV-2 infection as measured by N protein seroconversion, coinciding with the SARS-CoV-2 vaccine/variant era (25.74 per 100 person-years) as compared with the pre-vaccine/wild-type era (9.61 per 100 person-years). This corresponds to an increase in SARS-CoV-2 infection risk from 10 to 26% of participants infected per year. The large increase in SARS-CoV-2 incidence coincided with a relaxing of guidelines (e.g., around social distancing, masking, school attendance, in-person school attendance) and with surges of increasingly transmissible, immune evasive variants: Alpha (March–June 2021)<sup>50</sup>, Delta (June–December 2021) and Omicron variant and subvariants (December 2021–present)<sup>51</sup>, all emerged as SARS-CoV-2 vaccines were being more widely taken up. Our cohort findings are consistent with widespread community transmission in the general population, particularly during the Delta and Omicron surges, including in workplaces and households with children, in the vaccine/variant era compared with the pre-vaccine/wild-type era<sup>52–56</sup>. Despite the large increase in community transmission in the vaccine/variant era, being up-to-date on vaccines (i.e., being boosted once or more) was associated with a lower risk of SARS-CoV-2 infection compared with being un/undervaccinated or only receiving the primary vaccine series. While there are likely differences in risk factors among those who were boosted compared with those who were not, the observed associations were maintained across several epidemiologic risk strata. Although being boosted was associated with a reduced incidence in the vaccine/variant era, except for the groups reporting a confirmed case in the household, the incidence was still generally 1.3–2 times higher among individuals with 2+ boosters compared with those in the pre-vaccine/wild-type era cohort (Table 4). While this highlights the potential for new variants to cause breakthrough infections even among those who are more up-to-date on vaccines, it also suggests that being up-to-date on SARS-CoV-2 vaccines can greatly reduce the risk of SARS-CoV-2 infection during major surges. In addition, many non-pharmaceutical interventions used by cohort participants (e.g., masking in many different settings, social distancing) remained associated with substantially lower SARS-CoV-2 incidence rates in the vaccine/variant-era cohort, despite large increases in absolute incidence rates (Table 3).

In multivariate models, those who only received the primary vaccine series and no booster doses had similar SARS-CoV-2 infection risk as those who were un/undervaccinated in the vaccine/variant era. In both groups, the risk was approximately 5 times higher than in the pre-vaccine/wild-type era cohort. However, receipt of booster doses beyond the primary vaccine series was associated with a lower risk of infection compared with other vaccine status groups in the vaccine/variant era cohort. In fact, the risk of infection became progressively lower as the number of vaccine booster doses increased. This could be because 87% of the vaccine/variant-era cohort was fully vaccinated by August 2021 (~6 months into cohort follow-up, Table 1), and enough time had passed such that boosters would have been needed for most participants in order to offer some protection against infection from the more immune evasive variants. There also may be differences in behaviors and other factors among these groups. Importantly, however, these associations were observed in each stratum across various risk factors, with models that adjusted for age, gender, and presence of comorbidities (Table 4).

Our study showed substantial increases in SARS-CoV-2 incidence rates in the vaccine/variant era cohort compared to the pre-vaccine/wild-type era cohort within every sociodemographic subgroup and epidemiologic risk factor that we examined. Importantly, many factors that appeared protective against SARS-CoV-2 in the pre-vaccine/wild-type era cohort (e.g., NPIs such as masking and social distancing) remained protective in the vaccine/variant era cohort, despite the major increases in community transmission that occurred. This suggests that NPIs play an important role in limiting SARS-CoV-2 transmission even during major surges of new variants,



and in protecting those most vulnerable to SARS-CoV-2 infection. However, it should be noted that while the incidence rates in those engaging in protective behaviors in the vaccine/variant era cohort were lower than those who didn't engage in protective behaviors, absolute incidence rates were still very high in most instances in the vaccine/variant era cohort compared with the pre-vaccine/wild-type era cohort, highlighting that protective behaviors can reduce *but not eliminate* risk when community transmission rates are very high. For example, there was a lower infection risk associated with wearing masks while at work in the vaccine/variant era cohort, but the absolute incidence rate among those wearing masks at work was also much higher than in the pre-vaccine/wild-type era cohort. This could be due to higher levels of exposure inside the workplace, more individuals returning to in-person work, or higher levels of exposure from other sources (such as at home or on public transportation) or other locations that had since opened up after the pre-vaccine/wild-type era (e.g., gyms, indoor dining). We also noted that some new epidemiologic risk factors emerged in the vaccine/variant cohort (e.g., having a child in the household), likely reflecting the higher incidence of infection in the general population, including new sources of exposure such as children attending in-person school or daycare. Lastly, looking at the composite risk score, there was a clear dose–response with infection risk in the pre-vaccine/wild-type era cohort; but in the vaccine/variant era, those in the highest level of the composite risk score had lower infection risk than that of the medium risk score. This could be because the risk associated with having a case in the household, by far the strongest risk factor in both eras (Table 3), cannot increase as much in comparison with that of other risk factors when moving from the pre-vaccine/wild-type era to the vaccine/variant era (i.e., a ceiling effect).

Our study aligns with recent cross-sectional, population-representative surveys that were conducted during Omicron variant surges in some ways and not others. For example, similar to our study, recent NYC-based and national surveys conducted during major surges found high absolute point prevalence of SARS-CoV-2 infection but substantially lower relative point prevalence estimates among older (vs. younger) adults, those with comorbidities (vs. those without), and higher relative point prevalence estimates among those in households with school-aged children (vs. those without)<sup>25,57</sup>. However, in contrast to our study, these surveys found that vaccinated and boosted respondents had similar point prevalence estimates of SARS-CoV-2 infection to unvaccinated respondents. Reasons for this discrepancy could be that the surveys captured self-reported infection (positive point of care test, home test, or symptoms plus close contact) during the two weeks prior to the survey, while our study examined SARS-CoV-2 infection prospectively using serologic testing and over a longer time frame. In our cohort, compared with the seropositivity rate, the positivity rate on self-reported PCR/rapid tests over the same time period was lower in both the pre-vaccine/wild-type (85%) and vaccine/variant era cohorts (69%; Table 5). The reasons for the lower ratio in the vaccine/variant era cohort are not clear, but may be due to the fact that this was a highly vaccinated cohort, and fully vaccinated individuals with breakthrough infections may be less likely to be symptomatic, have a lower viral load, and/or experience a shorter duration of infection/illness<sup>58–60</sup>. As such, these participants may not have recognized signs of SARS-CoV-2 infection and/or were less likely to feel the need to test for SARS-CoV-2.

As the COVID-19 pandemic evolves, it remains important to monitor the incidence of SARS-CoV-2 infections. It may become more difficult, however, to identify cases through routine provider/laboratory reporting of PCR or rapid antigen tests. Individuals are increasingly less likely to be required to test in certain scenarios, may choose not to test, or may exclusively use at-home tests that are not captured in routine surveillance<sup>1</sup>. Thus, using serologic testing in cohort studies is a useful strategy to characterize SARS-CoV-2 incidence and risk factors. Strengths of our study include its prospective nature, with time-updated exposure measurement prior to outcome ascertainment. We also used repeat serologic testing to examine SARS-CoV-2 nucleocapsid seroconversion as the outcome measure of incident infection. Our design, which compared incidence in the two cohorts, leveraged the use of individuals as their own controls, since 75% of the overall sample was represented in both cohorts (Fig. 1), helping to reduce confounding when comparing incidence in the two eras. Finally, comparing incidence rates within models specific to different strata of risk factors also helps to limit confounding of the association of vaccination status with incidence by risk behaviors.

Our study also has limitations worth noting. The observed cumulative incidence in our cohort may be lower than the true cumulative incidence in our cohort because of the imperfect nature of serologic testing and the potential waning of SARS-CoV-2 antibodies<sup>61</sup>, particularly for milder infections<sup>62,63</sup>. Studies of SARS-CoV-2 antibody persistence have suggested the waning of antibodies to both nucleocapsid and spike proteins<sup>64,65</sup>. Boosted individuals who have an infection after vaccination may experience a more rapid waning of nucleocapsid antibodies<sup>66</sup>. Our study required total nucleocapsid seronegativity for inclusion. Because of the timing of specimen collection relative to infection in our cohort (median of 476 days in the vaccine/variant era cohort and 191 days in the pre-vaccine/wild type era cohort), this could mean that we have underestimated the true cumulative incidence due to waning. Additionally, immunocompromised status for study participants was not collected, and fully vaccinated status could therefore not be accurately assigned using three doses among this subgroup. For these participants, the third primary series dose may have been misidentified as a booster dose or skipped entirely, resulting in a fully vaccinated status. We did, however, adjust for the presence of comorbidities.

Crude associations between SARS-CoV-2 risk factors and incidence are subject to confounding. For example, behaviors between risk groups likely differ, with interpretation for some associations further hampered by small sample sizes in some exposure strata. Some risk behaviors may have been underreported (e.g., due to social desirability), which would bias observed associations toward the null. While our study was prospective, because we used a midpoint method to infer the timing of infection between a negative and positive serologic test, it is possible that some measured exposures, including vaccination, did not temporally precede infection. Also, some infections may have occurred in between vaccine doses. Finally, while our study was able to examine the role of behavioral risk factors over time, because of the timing of serologic testing, we could not distinguish any variant-specific effects (e.g., wild-type vs. Alpha, Delta vs. Omicron).



## Conclusion

Increases in the incidence of infection and newly emerging risk factors in the vaccine/variant era likely resulted from multiple co-occurring factors related to policy changes, individual- and community-level behavior changes (due to the availability of vaccines and relaxation of restrictions), and changing virus properties (i.e., more transmissible, immune evasive variants). While SARS-CoV-2 incidence increased markedly in most groups in the vaccine/variant cohort, being up to date on vaccines and the use of NPIs (masking, distancing) was associated with a greatly reduced risk of SARS-CoV-2 infection during major surges, making them relevant strategies to mitigate the impact of future SARS-CoV-2 surges, including those due to new variants that may evade existing vaccine-induced and hybrid immunity.

## Data availability

A limited, de-identified dataset of the CHASING COVID Cohort study data is publicly available through the Zenodo repository, at <https://zenodo.org/records/7305435>.

Received: 9 October 2023; Accepted: 29 December 2023

Published online: 05 January 2024

## References

- Qasmieh, S., Robertson, M. & Nash, D. 'Boosting' surveillance for a more impactful public health response during protracted and evolving infection disease threats: Insights from the COVID-19 pandemic. *Health Secur.* <https://doi.org/10.1089/hs.2023.0046> (2023).
- Alwan, N. A. Surveillance is underestimating the burden of the COVID-19 pandemic. *Lancet* **396**(10252), e24 (2020).
- Qasmieh, S. A. *et al.* The prevalence of SARS-CoV-2 infection and other public health outcomes during the BA.2/BA.2.12.1 surge, New York City, April–May 2022. *Commun. Med.* <https://doi.org/10.1101/2022.05.25.22275603v3.full-text> (2022).
- Reese, H. *et al.* Estimated incidence of coronavirus disease 2019 (COVID-19) illness and hospitalization—United States, February–September 2020. *Clin. Infect. Dis.* **72**(12), e1010–e1017 (2021).
- Wu, S. L. *et al.* Substantial underestimation of SARS-CoV-2 infection in the United States. *Nat. Commun.* **11**(1), 4507 (2020).
- Al-Sadeq, D. W. & Nasrallah, G. K. The incidence of the novel coronavirus SARS-CoV-2 among asymptomatic patients: A systematic review. *Int. J. Infect. Dis.* **98**, 372–380 (2020).
- Havers, F. P. *et al.* Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23–May 12, 2020. *JAMA Intern. Med.* **180**, 1576. <https://doi.org/10.1001/jamainternmed.2020.4130> (2020).
- Wiegand, R. E. *et al.* Estimated SARS-CoV-2 antibody seroprevalence trends and relationship to reported case prevalence from a repeated, cross-sectional study in the 50 states and the District of Columbia, United States—October 25, 2020–February 26, 2022. *Lancet Reg. Health Am.* **18**, 100403 (2023).
- O'Brien, S. C. *et al.* SARS-CoV-2 seroprevalence compared with confirmed COVID-19 cases among children, Colorado, USA, May–July 2021. *Emerg. Infect. Dis.* **29**(5), 929–936 (2023).
- Bergeri, I. *et al.* Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: A systematic review and meta-analysis of standardized population-based studies. *PLoS Med.* **19**(11), e1004107 (2022).
- Jetelina, K. End of PHE: A shift in data. In *Your Local Epidemiologist*. [https://yourlocalepidemiologist.substack.com/p/end-of-phe-a-shift-in-data?utm\\_source=substack&utm\\_medium=email](https://yourlocalepidemiologist.substack.com/p/end-of-phe-a-shift-in-data?utm_source=substack&utm_medium=email) (Accessed 20 May 2023) (2023).
- Honein, M. A. *et al.* Summary of guidance for public health strategies to address high levels of community transmission of SARS-CoV-2 and related deaths, December 2020. *Morb. Mortal. Wkly. Rep.* **69**(49), 1860–1867 (2020).
- Accorsi, E. K. *et al.* Association between 3 doses of mRNA COVID-19 vaccine and symptomatic infection caused by the SARS-CoV-2 omicron and delta variants. *JAMA* **327**(7), 639–651 (2022).
- National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases. *Science Brief: Indicators for Monitoring COVID-19 Community Levels and Making Public Health Recommendations*. COVID-19 Science Briefs (Centers for Disease Control and Prevention, 2022).
- Nash, D. *et al.* SARS-CoV-2 incidence and risk factors in a national, community-based prospective cohort of US adults. *Clin. Infect. Dis.* **76**, e375. <https://doi.org/10.1093/cid/ciac423> (2022).
- Assistant Secretary for Public Affairs (ASPA). *COVID-19 Vaccines*. *Hhs.gov*. <https://www.hhs.gov/coronavirus/covid-19-vaccines/index.html> (Accessed 15 May 2023) (2020).
- Rahmani, K. *et al.* The effectiveness of COVID-19 vaccines in reducing the incidence, hospitalization, and mortality from COVID-19: A systematic review and meta-analysis. *Front. Public Health* **10**, 873596 (2022).
- Abu-Raddad, L. J. *et al.* Effect of mRNA vaccine boosters against SARS-CoV-2 omicron infection in Qatar. *N. Engl. J. Med.* **386**(19), 1804–1816 (2022).
- Brown, C. M. *et al.* Outbreak of SARS-CoV-2 infections, including COVID-19 vaccine breakthrough infections, associated with large public gatherings—Barnstable County, Massachusetts, July 2021. *Morb. Mortal. Wkly. Rep.* **70**(31), 1059–1062. <https://doi.org/10.15585/mmwr.mm7031e2> (2021).
- Feikin, D. R. *et al.* Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: Results of a systematic review and meta-regression. *Lancet* **399**(10328), 924–944 (2022).
- Andrews, N. *et al.* Covid-19 vaccine effectiveness against the omicron (B.1.1.529) variant. *N. Engl. J. Med.* **386**(16), 1532–1546 (2022).
- Gram, M. A. *et al.* Vaccine effectiveness against SARS-CoV-2 infection or COVID-19 hospitalization with the alpha, delta, or omicron SARS-CoV-2 variant: A nationwide Danish cohort study. *PLoS Med.* **19**(9), e1003992 (2022).
- Huiberts, A. J. *et al.* Vaccine effectiveness of primary and booster COVID-19 vaccinations against SARS-CoV-2 infection in the Netherlands from 12 July 2021 to 6 June 2022: A prospective cohort study. *BioRxiv*. <https://doi.org/10.1101/2023.01.09.23284335v1> (2023).
- Qasmieh, S. A. *et al.* The prevalence of SARS-CoV-2 infection and long COVID in US adults during the BA.5 surge, June–July 2022. *BioRxiv* **376**, 20200283. <https://doi.org/10.1101/2022.09.04.22279588> (2022).
- Qasmieh, S. A. *et al.* The prevalence of SARS-CoV-2 infection and long COVID in US adults during the BA.4/BA.5 surge, June–July 2022. *Prev. Med.* **169**, 107461 (2023).
- Smith, M., Mervosh, S. & Bosman, J. What does the new mask guidance really mean? *The New York Times*. <https://www.nytimes.com/2021/05/13/us/cdc-mask-guidelines-vaccinated.html> (Accessed 26 October 2022) (2021).
- People vaccinated against Covid-19 can go without masks indoors and outdoors, CDC says. *CNN*. <https://www.cnn.com/2021/05/13/health/cdc-mask-guidance-vaccinated/index.html> (Accessed 26 October 2022) (2021).
- CDC. *Isolation and Precautions for People with COVID-19* (Centers for Disease Control and Prevention, 2022).

29. Nash, D. The CDC got lost trying to follow the science. *Barrons*. <https://www.barrons.com/articles/cdc-guidance-covid-isolation-pandemic-51641847943> (Accessed 10 March 2022) (2022).
30. Stolberg, S. G., Anthes, E., Mervosh, S. & Taylor, K. The C.D.C. issues new school guidance, with emphasis on full reopening. *The New York Times*. <https://www.nytimes.com/2021/07/09/health/cdc-schools-reopening-guidelines.html> (Accessed 26 October 2022) (2021).
31. Qasmieh, S. A. *et al.* The prevalence of SARS-CoV-2 infection and other public health outcomes during the BA.2/BA.2.12.1 surge, New York City, April–May 2022. *Commun. Med.* **3**(1), 92 (2023).
32. Hatoun, J., Correa, E. T., Donahue, S. M. A. & Vernacchio, L. Social distancing for COVID-19 and diagnoses of other infectious diseases in children. *Pediatrics* **146**(4), e2020006460 (2020).
33. Ferguson, N. *et al.* *Report 9: Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand* (Imperial College London, 2020).
34. IHME. *Institute for Health Metrics and Evaluation*. <https://covid19.healthdata.org/united-states-of-america?view=mask-use&tab=trend> (2022).
35. KFF COVID-19 Vaccine Monitor. *KFF*. <https://www.kff.org/coronavirus-covid-19/dashboard/kff-covid-19-vaccine-monitor-dashboard/> (2022).
36. CDC. *COVID Data Tracker* (Centers for Disease Control and Prevention, 2020).
37. Clarke, K. E. N. *et al.* Seroprevalence of infection-induced SARS-CoV-2 antibodies—United States, September 2021–February 2022. *Morb. Mortal. Wkly. Rep.* **71**(17), 606–608 (2022).
38. Nash, D. *et al.* A web-based study of HIV prevention in the era of pre-exposure prophylaxis among vulnerable HIV-negative gay and bisexual men, transmen, and transwomen who have sex with men: Protocol for an observational cohort study. *JMIR Res. Protoc.* **8**(9), e13715. <https://doi.org/10.2196/13715> (2019).
39. Grov, C., Westmoreland, D., Rendina, H. J. & Nash, D. Seeing is believing? Unique capabilities of internet-only studies as a tool for implementation research on HIV prevention for men who have sex with men: A review of studies and methodological considerations. *J. Acquir. Immune Defic. Syndr.* **82**(Suppl 3), S253–S260 (2019).
40. Robertson, M. M. *et al.* Cohort profile: A national, community-based prospective cohort study of SARS-CoV-2 pandemic outcomes in the USA—The CHASING COVID cohort study. *BMJ Open* **11**(9), e048778 (2021).
41. Sun, J. *et al.* Association between immune dysfunction and COVID-19 breakthrough infection after SARS-CoV-2 vaccination in the US. *JAMA Intern. Med.* **182**(2), 153–162 (2022).
42. CDC. *CDC Museum COVID-19 Timeline* (Centers for Disease Control and Prevention, 2023).
43. Center for Devices, Radiological Health. *EUA Authorized Serology Test Performance*. <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance> (2021).
44. Tré-Hardy, M. *et al.* Analytical and clinical validation of an ELISA for specific SARS-CoV-2 IgG, IgA, and IgM antibodies. *J. Med. Virol.* **93**(2), 803–811 (2021).
45. Traub, M.-A. *et al.* Comparison of eight commercial, high-throughput, automated or ELISA assays detecting SARS-CoV-2 IgG or total antibody. *J. Clin. Virol.* **132**, 104613 (2020).
46. Plaga, A. *et al.* Evaluation of the clinical performance of 7 serological assays for SARS-CoV-2 for use in clinical laboratories. *J. Appl. Lab. Med.* **6**(4), 998–1004 (2021).
47. Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D. & Bradley, K. A. The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking. Ambulatory care quality improvement project (ACQUIP). Alcohol use disorders identification test. *Arch. Intern. Med.* **158**(16), 1789–1795 (1998).
48. Robertson, M. M. *et al.* Racial/ethnic disparities in exposure, disease susceptibility, and clinical outcomes during COVID-19 pandemic in national cohort of adults, United States. *Emerg. Infect. Dis.* **28**(11), 2171–2180 (2022).
49. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **58**(1), 267–288 (1996).
50. Paul, P. *et al.* Genomic surveillance for SARS-CoV-2 variants circulating in the United States, December 2020–May 2021. *Morb. Mortal. Wkly. Rep.* **70**(23), 846–850 (2021).
51. *CoVariants: Per Country*. <https://covariants.org/per-country?region=United+States> (2023).
52. Singanayagam, A. *et al.* Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: A prospective, longitudinal, cohort study. *Lancet Infect. Dis.* **22**(2), 183–195 (2022).
53. Holmdahl, I., Kahn, R., Slifka, K. J., Dooling, K. & Slayton, R. B. Modeling the impact of vaccination strategies for nursing homes in the context of increased severe acute respiratory syndrome coronavirus 2 community transmission and variants. *Clin. Infect. Dis.* **75**(1), e880–e883 (2022).
54. Chen, F., Tian, Y., Zhang, L. & Shi, Y. The role of children in household transmission of COVID-19: A systematic review and meta-analysis. *Int. J. Infect. Dis.* **122**, 266–275 (2022).
55. Head, J. R., Andrejko, K. L. & Remais, J. V. Model-based assessment of SARS-CoV-2 delta variant transmission dynamics within partially vaccinated K-12 school populations. *MedRxiv* **8**, e644. <https://doi.org/10.1101/2021.08.20.21262389> (2021).
56. Koirala, A. *et al.* Understanding SARS-CoV-2 delta and omicron variant transmission and vaccine impact in schools and child-care settings in Australia: A population-based study. *Lancet Reg. Health West. Pac.* **1**, 1 (2023).
57. Qasmieh, S. A., Robertson, M. M., Teasdale, C. A., Kulkarni, S. G. & Nash, D. Estimating the period prevalence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection during the omicron (BA.1) surge in New York City (NYC), 1 January to 16 March 2022. *Clin. Infect. Dis.* **76**(3), e499–e502 (2023).
58. Ronchini, C. *et al.* Lower probability and shorter duration of infections after COVID-19 vaccine correlate with anti-SARS-CoV-2 circulating IgGs. *PLoS ONE* **17**(1), e0263014 (2022).
59. Qassim, S. H. *et al.* Effects of BA.1/BA.2 subvariant, vaccination, and prior infection on infectiousness of SARS-CoV-2 omicron infections. *J. Travel Med.* **29**, 68. <https://doi.org/10.1093/jtm/taac068> (2022).
60. Abu-Raddad, L. J. *et al.* Relative infectiousness of SARS-CoV-2 vaccine breakthrough infections, reinfections, and primary infections. *Nat. Commun.* **13**(1), 532 (2022).
61. Sullivan, P. S. *et al.* SARS-CoV-2 cumulative incidence, United States, August–December 2020. *Clin. Infect. Dis.* **74**, 1141. <https://doi.org/10.1093/cid/ciab626> (2021).
62. Liu, W. *et al.* Predictors of nonseroconversion after SARS-CoV-2 infection. *Emerg. Infect. Dis.* **27**(9), 2454–2458 (2021).
63. Van Elslande, J. *et al.* Longitudinal follow-up of IgG anti-nucleocapsid antibodies in SARS-CoV-2 infected patients up to eight months after infection. *J. Clin. Virol.* **136**, 104765 (2021).
64. Bajema, K. L. *et al.* Estimated SARS-CoV-2 seroprevalence in the US as of September 2020. *JAMA Intern. Med.* **181**, 450. <https://doi.org/10.1001/jamainternmed.2020.7976> (2020).
65. Korte, W. *et al.* SARS-CoV-2 IgG and IgA antibody response is gender dependent; and IgG antibodies rapidly decline early on. *J. Infect.* **82**(1), e11–e14 (2021).
66. Dhakal, S. *et al.* Reconsideration of antinucleocapsid IgG antibody as a marker of SARS-CoV-2 infection postvaccination for mild COVID-19 patients. *Open Forum Infect. Dis.* **10**(1), 677 (2023).

## Acknowledgements

The authors wish to thank the participants of the CHASING COVID Cohort Study. The authors are grateful to you for your contributions to the advancement of science around the SARS-CoV-2 pandemic. They are grateful to MTL Labs for serologic testing of our cohort's specimens.

## Author contributions

D.N., M.M.R., A.S. conceptualized the study. A.S. performed statistical analyses. D.N. and M.S.R. wrote the first draft of the paper. All authors contributed to interpreting the data and revising of the manuscript. S.G.K., W.Y., M.M.R. A.B., S.K., and D.N. contributed to data collection, cleaning and management. D.N., M.M.R., S.G.K., and C.G. contributed to obtaining funding for the research.

## Funding

Funding for this project is provided by The National Institute of Allergy and Infectious Diseases (NIAID), award number 3UH3AI133675-04S1 (MPIs: D Nash and C Grov), the CUNY Institute for Implementation Science in Population Health ([www.cunyisph.org](http://www.cunyisph.org)) and the COVID-19 Grant Program of the CUNY Graduate School of Public Health and Health Policy; Pfizer, Inc (PI: Nash), and the National Institute of Mental Health (NIMH), award number 1RF1MH132360 (MPIs: D Nash and A Parcesepe).

## Competing interests

Dr. Nash received consulting fees from Gilead Sciences, Inc. and Abbvie, Inc. All other authors have no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-51029-9>.

**Correspondence** and requests for materials should be addressed to D.N.

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

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



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

© The Author(s) 2024