

Antibody Duration After Infection From SARS-CoV-2 in the Texas Coronavirus Antibody Response Survey

Michael D. Swartz,^{1,✉} Stacia M. DeSantis,¹ Ashraf Yaseen,¹ Frances A. Brito,¹ Melissa A. Valerio-Shewmaker,² Sarah E. Messiah,³ Luis G. Leon-Novelo,¹ Harold W. Kohll,^{4,5} Cesar L. Pinzon-Gomez,¹ Tianyao Hao,¹ Shiming Zhang,¹ Yashar Talebi,¹ Joy Yoo,¹ Jessica R. Ross,¹ Michael O. Gonzalez,¹ Leqing Wu,¹ Steven H. Kelder,⁴ Mark Silberman,⁶ Samantha Tuzo,⁶ Stephen J. Pont,⁷ Jennifer A. Shuford,⁷ David Lakey,⁸ and Eric Boerwinkle¹

¹The University of Texas Health Science Center in Houston, School of Public Health in Houston, Houston, Texas, USA; ²The University of Texas Health Science Center in Houston, School of Public Health in Brownsville, Brownsville, Texas, USA; ³The University of Texas Health Science Center in Houston, School of Public Health in Dallas, Dallas, Texas, USA; ⁴The University of Texas Health Science Center in Houston, School of Public Health in Austin, Austin, Texas, USA; ⁵The University of Texas at Austin, College of Education, Department of Kinesiology and Health Education, Austin, Texas, USA; ⁶Clinical Pathology Laboratories, Austin, Texas, USA; ⁷Texas Department of State Health Services, Austin, Texas, USA; and ⁸University of Texas System, Office of Health Affairs, Austin, Texas, USA

Understanding the duration of antibodies to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus that causes COVID-19 is important to controlling the current pandemic. Participants from the Texas Coronavirus Antibody Response Survey (Texas CARES) with at least 1 nucleocapsid protein antibody test were selected for a longitudinal analysis of antibody duration. A linear mixed model was fit to data from participants ($n = 4553$) with 1 to 3 antibody tests over 11 months (1 October 2020 to 16 September 2021), and models fit showed that expected antibody response after COVID-19 infection robustly increases for 100 days postinfection, and predicts individuals may remain antibody positive from natural infection beyond 500 days depending on age, body mass index, smoking or vaping use, and disease severity (hospitalized or not; symptomatic or not).

Keywords. SARS-CoV-2; antibodies; COVID-19; antibody duration.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease 2019 (COVID-19) continue to affect large numbers of people across the globe, perpetuating the pandemic. To control the pandemic, enough of the population must be exposed to the SARS-CoV-2 virus and/or receive an effective SARS-CoV-2 vaccine and mount an immune response that confers lasting protection from acquiring COVID-19 such that the virus no longer rapidly propagates through the population [1–4]. The estimated proportion of the population with antibodies to the SARS-CoV-2 virus and its variants has gradually increased since the beginning of the pandemic, with some gaining and losing natural antibodies following infection. In the early stages of the pandemic, the source of antibodies was primarily through naturally occurring antibodies following SARS-CoV-2 infection and recovery. With the introduction of vaccines in late 2020 and massive vaccination campaigns, the prevalence of vaccine-induced antibodies has substantially increased, with

more than 50% of the population estimated to be fully vaccinated. As of December 1, 2021, it was estimated that 74.1% of the US population had immunity against severe infection for pre-omicron variants and 61.2% for the omicron variants [5].

Despite the efforts to expand vaccination, slow uptake of the vaccine requires additional understanding of the duration of antibodies acquired through infection to control the pandemic [6, 7]. Circulating antibody titers usually decrease gradually over time, and if titers reach a nonneutralizing concentration in a substantial portion of the population, control of the pandemic will be further delayed [8, 9]. Cross-sectional studies report that the presence of receptor binding domain antibodies can occur in all vaccinated individuals and up to 97% of individuals with previous infections occurring several months before the antibody tests, suggesting that these immune responses can last several months [10–12]. Furthermore, a population-based study of the longevity of SARS-CoV-2 antibody seroprevalence revealed nucleocapsid-antibody positivity based on immunoglobulin G (IgG) assays increased to 90% within the first 30 days postinfection, and show a linear decay afterwards, declining to 65% at around 300 days [11].

Other studies on the duration of the immunity from naturally occurring infection have been published [13–25]; however, the findings face several limitations that impact interpretation and use. First, recruitment of participants for many of the studies has been limited to a specific location or clinic, reducing the potential generalizability of the findings [18, 19, 23], and the resultant small sample sizes from such recruitment strategies

Received 23 December 2021; editorial decision 25 April 2022; accepted 03 May 2022; published online 6 May 2022

Correspondence: Michael D. Swartz, PhD, Department of Biostatistics and Data Science, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Pressler Street, Houston, TX 77030 (Michael.D.Swartz@uth.tmc.edu).

The Journal of Infectious Diseases® 2023;227:193–201

© The Author(s) 2022. Published by Oxford University Press on behalf of Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com
<https://doi.org/10.1093/infdis/jiac167>

limit their statistical power. Second, studies using alternative recruitment strategies to increase sample size, such as testing health care personnel from large institutions or using large datasets and blood banks, allowing them to increase sample size, limit their follow-up time to less than 4 months when assessing antibody levels [4, 16, 21, 26]. Third, given most have been designed as observational studies, follow-up is a concern as cases can be lost; therefore, most studies have measured antibodies within a short 1- to 4-month period [20–22], and others resort to a cross-sectional design [24, 26]. Regardless of the design, most studies have found that naturally occurring antibodies due to SARS CoV-2 infection usually last between 3 and 6 months, with at least 1 study reporting up to 11 months [21]. Furthermore, studies have identified variations in responses with weaker immune responses in asymptomatic and mild cases than in cases with more severe symptoms [23, 25]. To fully navigate the COVID-19 pandemic into the endemic phase, it is important to understand the duration and behavior of antibody levels over time in broad, diverse samples of individuals. The Texas coronavirus antibody response survey (Texas CARES; <https://sph.uth.edu/projects/texascares/>) was developed to survey the human antibody response to SARS-CoV-2 by recruiting volunteer participants across Texas. The current analysis used longitudinal data from Texas CARES to predict the duration of antibody response among individuals assessed over time and identify key predictors of individual differences in duration.

METHODS

Recruitment and design details were reported elsewhere [27]. Briefly, Texas CARES consists of a prospective convenience sample of individuals, were are longitudinally tested for SARS-CoV-2 antibody status every 3 months for a total of 3 time points from 1 October 2020 to 30 August 2021. Participants included adult retail/business employees, Kindergarten - 12th grade and university educators and students, and patients and employees from Health Resources and Services Administration designated Federally Qualified Health Centers. A consent form and survey questionnaires were administered online over all 3 time points, and participants proceeded to a convenient laboratory location for the antibody test at each time point. The current report focuses on individuals from Texas CARES who were 20 years of age and older, reported only 1 positive COVID-19 diagnosis with a date of diagnosis, and had at least 1 valid nucleocapsid (N) antibody test (N-test) result and at least 1 nonzero N-test value after their first reported COVID-19-positive diagnosis through 30 August 2021. Note that although the N-test value is nonzero, it is only considered seropositive if the index value is greater than or equal to 1.0. The Roche test used for Texas CARES is a total antibody assay for IgG, IgA, and IgM antibodies, and

does not parse the different types of antibodies; however, IgM and IgA typically decline quickly, within a few months [28], thus longer duration of antibody-positive status is more likely from IgG antibodies. The study protocols were approved by both the University of Texas Health Science Center at Houston Institutional Review Board (IRB) and deemed public health practice by the Texas Department of Health Services IRB.

Statistical Methods

The response of interest was the continuous value of the N-test, which was interpreted as positive for antibodies if the N-test index value was greater than or equal to 1.0, and negative otherwise. Predictors of interest were time (days since the self-reported positive COVID-19 diagnosis), age, race/ethnicity, gender, number of chronic conditions (categorized as none, 1, 2, or more), tobacco/vaping use (yes/no), insured (yes/no), body mass index (BMI; calculated from self-reported height and weight and categorized according to the Centers for Disease Control and Prevention [CDC] convention [29]), hospitalized during COVID-19 infection (yes/no), and symptom status (symptomatic/asymptomatic). A missing category for each variable was included in the analysis to maximize sample size. The analysis used a linear mixed model (LMM) with the continuous value of the N-test as the response, participant as a random effect to account for the correlation of the N-test values within the same individual, and a restricted cubic spline with 3 internal knots to model the time trend. Purposeful variable selection similar to that described by Hosmer et al [30] was used to select important covariates. Variables were included in the multivariable model after simple univariable models for each considered variable had a type III ANOVA F statistic *P* value less than .20. Variables were retained if the multivariable model Wald statistic *P* value was less than .05. All variables removed at the .05 level were further assessed using a likelihood ratio test (LRT) and were retained if the LRT *P* value was less than .05, or for biological relevance. After fitting the model, the predicted curves for N-test values are presented and interpreted; the antibody response was positive if the predicted N-test value was greater than or equal to 1.0, and negative otherwise.

RESULTS

Study Sample and Longitudinal Response

The analytic sample for the current study consisted of 4558 individuals who had at least 1 N-test after their first reported COVID-19-positive diagnosis, with at least 1 of the N-test values being greater than 0. The median number of days since the positive COVID-19 diagnosis was 228 days (interquartile range = 121). The median value of the N-test values was 38 units, ranging from 0 to 295. The mean age of this sample was 49.7 years (SD = 13.6 years), mostly female (70.1%) and

non-Hispanic white (71.8%), educated with a 2-year college degree or more (71.0%), employed full time (64.6%), with health insurance (89.7%), with obesity as calculated from self-reported height and weight, and reporting at least 1 chronic disease (48.8%). The majority reported no use of tobacco or vaping products (90.2%); 39.0% reported asymptomatic COVID-19 and 10.4% reported being hospitalized for COVID-19. The majority of participants in our sample were not fully vaccinated (55.0%). Full details of the demographic and clinical characteristics of the analyzed sample are provided in [Table 1](#).

The trajectory of each individual's antibody duration is shown in [Figure 1](#); each line represents an individual's trajectory in the study, and the black curve represents the model-based predicted average curve. This mixed linear model with only the restricted cubic spline representation of time since COVID-19-positive and participant-specific random effect was highly significant ($LRT = 433.0$, $df = 4$, $P < .001$). The general pattern of the N-test values over time shows a rapid increase over the first few months (90–100 days; [Figure 1](#)). Of note, in the sample, 5.5% of N-tests were negative in the first 100 days after infection, 2.8% negative between 101 and 200 days after infection, 4.5% negative between 201 and 300 days after infection, 4.3% negative between 301 and 400 days after infection, and 7.4% negative between 401 and 500 days after infection. These negative N-tests include those who were negative upon entering into our study.

Univariate and Multivariable Predictors of Antibody Duration

The univariate type III ANOVA F statistic *P* values for each predictor are given in [Table 2](#). Each variable in [Table 2](#) was significant at the .20 level except for insurance status. Insurance status was still maintained in the adjusted modeling as it reflects socioeconomic status. Based on the single fixed-effects model results, a multivariable model was fit considering all predictors. The Wald-test *P* values were assessed and variables were selected based on the purposeful variable selection procedure. The adjusted regression coefficients, standard errors, and Wald-statistic *P* values for each fixed effect are reported in [Table 3](#).

Gender, number of chronic conditions, insurance status, symptomatic status, and vaping/tobacco use were all not significant at the .05 level. Gender, number of chronic conditions, and insurance were determined to have biological and socioeconomic relevance a priori and were retained in the model. A likelihood ratio test comparing the full model with all predictors and the full model excluding symptomatic status and vaping/tobacco use was significant ($LRT = 29.4$, $df = 2$, $P < .001$). The likelihood ratio test assessing the spline versus the full model without the time component was also significant ($LRT = 681.3$, $df = 4$, $P < .001$). Thus, all considered predictors were included in the final model. The final adjusted model included time (days since first positive COVID-19 diagnosis), age

(categorical), race/ethnicity, gender, number of chronic conditions (categorized into 0, 1, 2+), insurance status (yes/no), BMI (categorized according to the CDC), hospitalization, symptomatic (yes/no), and vaping/tobacco use (yes/no). Because the nucleocapsid protein-related antibodies are not produced by the vaccine, we did not consider vaccination status in our modeling. Significant predictors are as follows: each age group was significantly different from the reference age group (50–64 years old); it is notable that the oldest participants (65–74 years and 75+ years) had a longer antibody duration than those 50–64 years, while those aged 40–49, 30–39, and 20–29 years had a shorter antibody duration compared to 50–64 year olds, with the largest reduction in duration seen in 30–39 year olds, followed by 20–29 year olds. Asian, non-Hispanic black, and Hispanic individuals had a longer antibody duration than non-Hispanic whites. Those self-reporting race as other had a slightly longer antibody duration compared to non-Hispanic whites. Those missing a survey response to race/ethnicity had a shorter antibody duration than non-Hispanic whites. Relative to individuals with healthy BMI, individuals categorized as with overweight or obesity had increased antibody duration. Those who were hospitalized or who did not report their status of being hospitalized did not have an increased duration of antibodies relative to those who reported not being hospitalized from COVID-19. Unsurprisingly, those who reported being symptomatic during infection exhibited an increased duration of antibodies compared to those who reported being asymptomatic or those missing a response regarding symptoms. Note that those who omitted a response for symptom status were not significantly different from those self-reporting as asymptomatic. Users of vaping or tobacco products or those not providing a response regarding vaping or tobacco use showed reduced antibody duration compared to those reporting nonuse.

Predicted Antibody Duration

The adjusted LMM was used to generate prediction curves for N-test values by various characteristics. Given the small number of individuals providing data more than 500 days postinfection ($n = 33$ individuals), prediction curves were focused within 500 days postinfection. We note that any adjusted model curves must be presented with respect to a baseline value or individual in the data set, so the below presentation uses this format. All figures include this reference individual as a black curve. The figures also include a horizontal dashed line representing the threshold for a positive N-test, and crossing the line indicates a predicted seroconversion from positive to negative.

For each panel of [Figures 2](#) and [3](#), the prediction curves for the N-test index values resulting from varying a given demographic category from the reference category are shown. Each curve can be interpreted as adjusted for the other variables in the model. The reference category is depicted as a black curve

Table 1. Demographics and Clinical Characteristics

Characteristic	Overall (n = 4552)
Days since positive COVID-19 diagnosis^a	
Mean (SD)	221.8 (110.0)
Median (Q1, Q3)	228 (158, 279)
Range	1–595
N test index values	
Mean (SD)	62.7 (64.5)
Median (Q1, Q3)	38 (10.7, 100)
Range	0–295
Age, y	
Mean (SD)	49.7 (13.7)
Median (Q1, Q3)	50.0 (40.0, 60.0)
Range	20.0–91.0
Age, categorical, y	
20–29	356 (7.8)
30–39	764 (16.8)
40–49	1133 (24.9)
50–64	1612 (35.4)
65–74	563 (12.4)
75+	124 (2.7)
Gender	
Female	3192 (70.1)
Male	1360 (29.9)
Race/ethnicity	
Non-Hispanic, white	3271 (71.9)
Non-Hispanic, black	109 (2.4)
Hispanic	944 (20.7)
Asian	138 (3.0)
Other	53 (1.2)
Missing	37 (0.8)
Education	
Some high school or less	51 (1.2)
High school graduate/General Educational Development Test (GED)	403 (9.1)
Some college, no degree	831 (18.8)
2- or 4-year college-level degree	1930 (43.6)
Advanced professional or academic degree	1213 (27.4)
Employment status	
Employed full time	2841 (64.6)
Employed part time	447 (10.2)
Not currently employed/unemployed	629 (14.3)
Other	482 (11.0)
Has health insurance/coverage	
Yes	4085 (89.7)
No	397 (8.7)
Missing	70 (1.5)
Height, inches	
Mean (SD)	66.4 (3.9)
Median (Q1, Q3)	66.0 (64.0, 69.0)
Range	48.0–92.0
Weight, pounds	
Mean (SD)	181.6 (45.3)
Median (Q1, Q3)	175.0 (149.0, 208.0)
Range	0.2–515.0
BMI, continuous	
Mean (SD)	28.8 (6.5)
Median (Q1, Q3)	27.6 (24.2, 32.2)
Range	0.0–66.6

Table 1. Continued

Characteristic	Overall (n = 4552)
BMI, categorical	
Healthy	1263 (27.7)
Underweight	34 (0.7)
Overweight	1516 (33.3)
Obesity	1591 (35.0)
Missing	148 (3.3)
Tobacco/vaping use	
None at all	4105 (90.2)
At least some	394 (8.7)
Missing	53 (1.2)
Symptomatic/asymptomatic	
Asymptomatic	1775 (39.0)
Symptomatic	2510 (55.1)
Missing	267 (5.9)
Hospitalized	
No	3915 (86.0)
Yes	475 (10.4)
Missing	162 (3.6)
No. of different chronic diseases	
None	1953 (42.9)
1	1235 (27.1)
2 or more	985 (21.6)
Missing	379 (8.3)
Vaccination status	
Not fully vaccinated	2503 (55.0)
Fully vaccinated	1892 (41.6)
Missing	157 (3.4)

Data are No. (%) except where indicated.

^aSummary statistics of days since COVID-19–positive diagnosis is based on 5974 tests across 4552 individuals.

in each of the graphs presented, and represents the most frequently represented characteristics in our sample: 50–64 years old, non-Hispanic white, woman, not reporting any chronic conditions, healthy BMI, has insurance coverage, asymptomatic infection, no reported hospitalization, and reporting no tobacco or vaping use. For age (Figure 2A), categories 65 years and older had the highest predicted N-test curves and were not predicted to drop below 1.0 (convert to negative antibody status) within 500 days of their COVID-19 infection. Those 30–39 years old had the shortest predicted duration of antibodies and were predicted to seroconvert around 450 days postinfection. Those 20–29 years old were predicted to be close to converting to negative antibody status around 500 days postinfection. Finally, 40–49 year-olds were predicted to maintain positive antibody status at least 500 days postinfection, but all models were on a downward trajectory at 500 days.

Varying race/ethnicity (Figure 2B) predicted slightly different curves; however, none of the categories dropped below the threshold of a positive test in the duration of the time interval observed (500 days), implying that all ethnicities showed

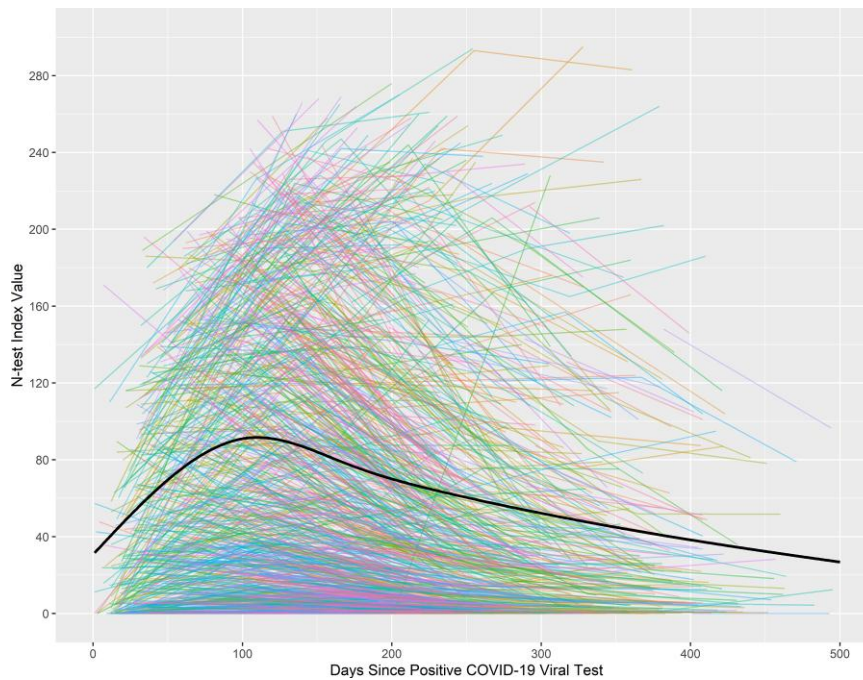


Figure 1. Plot of nucleocapsid antibody levels (N-test values) over time (unadjusted) with individual spaghetti plots. Time is measured in days since the self-reported coronavirus disease 2019 (COVID-19) infection. For each individual depicted, their infection date is day 0, and the date of their first and subsequent tests are chronologically arranged on the plot; later tests have longer days since infection.

positive antibody status lasting 500 days on average. Varying BMI status from the reference category (Figure 2C) showed underweight individuals having the shortest antibody duration, but still lasting 500 days. Antibody duration for the remaining BMI categories (healthy, overweight, obesity, and those missing BMI status) was predicted to last at least 500 days on average, but the higher curves implied longer duration than those underweight. Varying hospitalization status (Figure 3A) did not shorten predicted antibody status shorter than 500 days; however, the model predictions suggested a longer time of maintaining antibody-positive status for those who were hospitalized, compared to those who were not hospitalized.

Table 2. Type III ANOVA F Statistic P Values of Univariable Linear Mixed Models with the N-Test Value as Response and Participant Specific Random Effect for Each Predictor of Interest

Variable	P Value
Days since positive COVID-19 viral test	<.001
Age, categorized	<.001
Race/ethnicity	<.001
Gender	.011
Number of chronic conditions, 0, 1, 2, or more	<.001
Health insurance/coverage, yes/no	.629
BMI category	<.001
Hospitalized, yes/no	<.001
Tobacco or vaping substance use, yes/no	.012
Symptomatic, yes/no	.019

Finally, varying symptomatic status (Figure 3B) followed a similar yet less pronounced effect on antibody duration as hospitalization.

DISCUSSION

The large sample size, demographic and clinical diversity, and geographic distribution of Texas CARES offers a unique opportunity to examine the long-term presence of IgG antibodies against SARS-CoV-2. A purposefully selected linear mixed-effects model shows that expected IgG antibody response after COVID-19 infection robustly increases for 100 days postinfection, and that duration of antibody response is reasonably long lasting. Those reporting prior COVID-19 infection are predicted to possibly maintain antibodies from natural infection beyond 500 days. The longitudinal modeling indicates the initially robust IgG response begins to decline after 90–100 days and continues to decline over the remaining follow-up period of the study (maximum follow-up 500 days); with rate of decline differing by age, body mass index, smoking or vaping use, and disease severity (hospitalized or not; symptomatic or not). Although the immune responses vary for different pathogens and between individuals, these findings demonstrate that the kinetics of antibodies against SARS-CoV-2 follow a relatively usual pattern of high levels peaking shortly after symptom onset, followed by a gradual decrease in the following months [31].

Table 3. Fixed Effect Coefficients and Wald Statistic P Values From the Multivariable Model

Variable	Regression Coefficient	Standard Error	P Value
Age group, y			
20–29	–19.47	3.49	<.001
30–39	–24.69	2.60	<.001
40–49	–14.59	2.27	<.001
50–64 (Ref) ^a
65–74	8.43	2.86	.003
75+	21.76	5.42	<.001
Race/ethnicity			
Non-Hispanic, white (Ref)
Non-Hispanic, black	9.53	5.61	.089
Hispanic	15.00	2.21	<.001
Asian	15.81	5.05	.002
Other	2.15	8.00	.788
Missing	–6.96	9.63	.470
Gender			
Female (Ref)
Male	0.97	1.94	.617
No. of chronic conditions			
None (Ref)
One	1.27	2.16	.556
Two or more	0.64	2.52	.799
Missing	–0.73	3.41	.830
Health insurance/coverage			
Yes (Ref)
No	–0.89	3.08	.772
Missing	–0.45	8.61	.958
BMI category			
Underweight	–4.67	10.01	.641
Healthy (Ref)
Overweight	10.44	2.27	<.001
Obesity	23.83	2.36	<.001
Missing	18.50	5.65	.001
Hospitalized			
No (Ref)
Yes	23.70	2.87	<.001
Missing	11.26	4.57	.014
Symptomatic/asymptomatic			
Asymptomatic (Ref)
Symptomatic	4.11	1.81	.024
Missing	–0.60	4.09	.883
Vaping/tobacco use			
None at all (Ref)
At least some	–7.04	3.10	.023
Missing	–5.36	10.18	.598

^aReference (Ref) age group selected as the largest group.

Consistent with other findings, participants reporting hospitalization with more severe infection were found to have a more robust response than the nonhospitalized participants [18, 23, 24]. The data and models show a more robust response in the first 100 days postinfection among those who have been hospitalized versus nonhospitalized participants; with the

response remaining higher in the follow-up period compared to nonhospitalized participants. While evidence of these differences in response has been previously reported, the size and scope of the current survey provides substantial evidence given the longitudinal design [18, 20, 23, 24, 26]. This has important implications for risk of reinfection in younger people whose employment, education, or social interactions may have increased vulnerability to exposure. Further investigation of reinfections observed in Texas CARES may be beneficial to understanding how antibody levels and other risk factors correlate with risk of reinfection in this population. Finally, those reporting comorbidities do not appear to have any diminished antibody response compared to those with no history of comorbidities [32].

As noted previously, the SARS-CoV-2 IgG antibody status has been described in the recent literature; however, most studies have represented a single setting (clinical, system, or organization) [23], with small sample sizes ($n < 200$) [13–15, 18, 20–23, 33] and with short follow-up periods (commonly 3–6 months) [18–20, 23, 24]. In most of these above-referenced studies, convenience samples (such as health care personnel or business employees within a specific industry) have been followed [18, 19, 23, 26]. As such, people are generally healthy and derived from a single institution or city, with results not necessarily generalizable to the broader population, due to lack of representation.

The current study provides robust evidence of the duration of antibody status in a large sample. First, the community-based sample of 57 198 adults, from which the infected sample for the current study was pulled, includes substantial numbers of demographically diverse individuals located in both urban and rural areas across the state of Texas with long-term (9–12 month) follow-up periods. Second, the extent of the baseline survey with 2 follow-up periods, including a demographic and behavioral survey with vaccine and natural antibody testing in Texas CARES, allows us to better determine and identify the influence of behavioral and clinical demographics in natural antibody responses to SARS-CoV-2. The deeply granular look at duration of antibody status, that is, the sample size and covariate information collected, enables fitting of fully adjusted linear mixed-effects models with restricted cubic splines (to model the nonlinearity of antibody response over time). To this end, we are able to report with high power and accuracy, fully covariate-adjusted demographic effects on antibody duration in a very large, diverse sample.

We recognize the limitations of the current study. First, cohort studies are nonrandom and self-selection biases may be seen; however, the larger cohort size allows us to control for covariates. Loss to follow-up does occur in cohort studies, but loss in Texas CARES is partially mitigated by the large baseline sample size ($n = 4553$ in the current study for time 1, $n = 1074$ for time 2, and $n = 486$ for time 3) with continued outreach efforts being made to increase retention over time to allow us to more

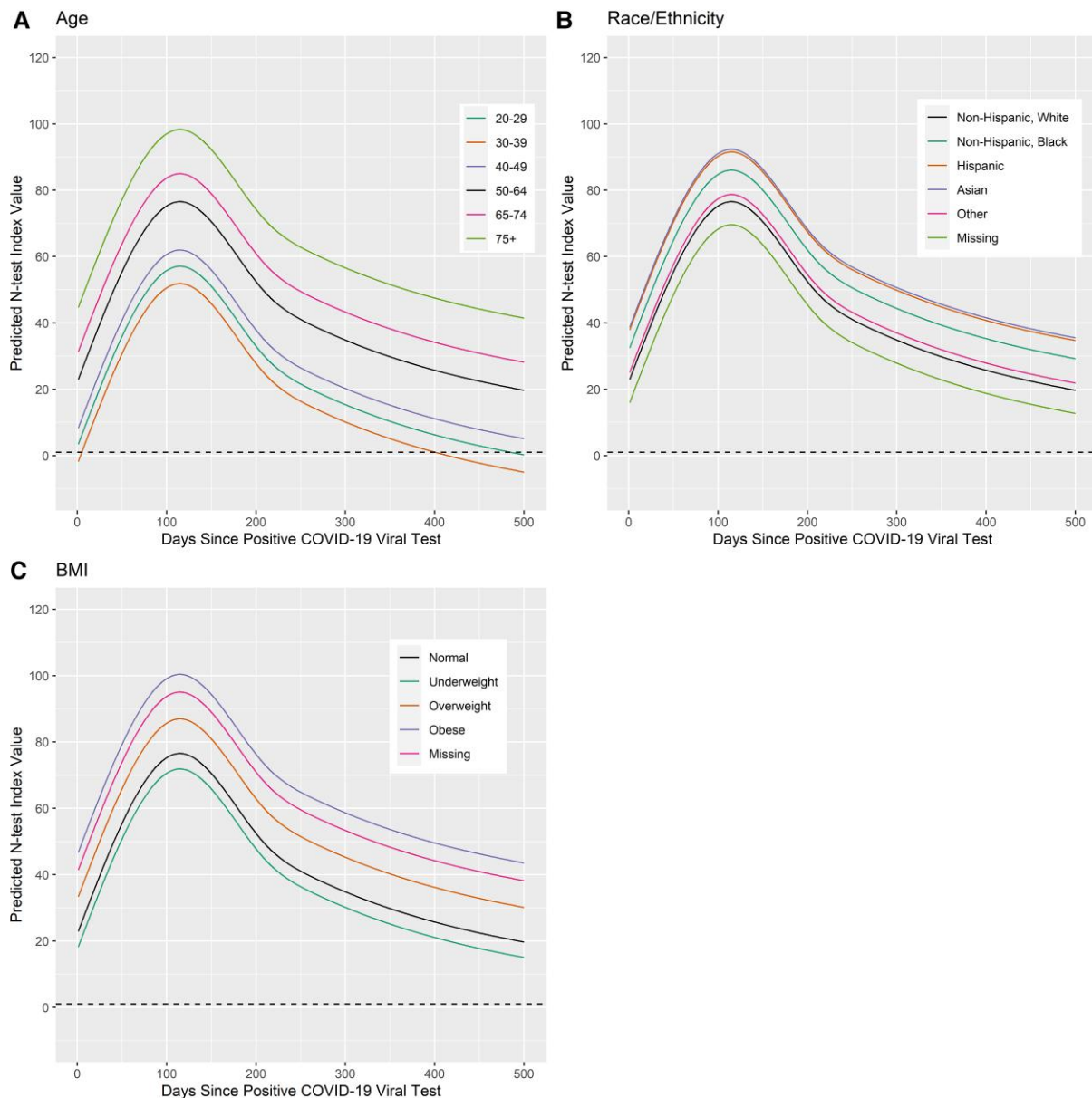


Figure 2. Predicting nucleocapsid antibody status varying demographics in the reference model. The dashed line at 1 represents when the N-test reports a positive antibody response. N-test value ≥ 1 is a positive antibody response. The solid black line represents the reference prediction curve, that is for a 50 to 64-year-old non-Hispanic white woman without chronic conditions, with healthy BMI, who has insurance coverage, was asymptomatic in their infection, who was not hospitalized, and did not use tobacco or vaping. The effect of varying a covariate across its domain is shown: (A) varying age, (B) varying race/ethnicity, and (C) category of BMI status. Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019; N-test, nucleocapsid antibody test.

accurately estimate response over time. Finally, COVID-19 infection dates were self-reported so attendant limitations in self-report must be recognized. It is possible that people incorrectly reported infection or dates of infection. However, as the vast majority of self-reported COVID-19 infections were confirmed by positive N-tests (93.9%), it appears the self-reported laboratory-based diagnoses and dates are generally accurate. The full scope of naturally induced immunity cannot be established in the current study as its current focus is only the antibodies against the nucleocapsid proteins. Specifically, these

antibody tests do not detect receptor binding domain antibodies, shown to be longer lasting in some studies [12], and although not many participants had access to them, the effect of monoclonal antibody treatments for COVID-19 on these antibodies is not well known. In that vein, the predicted durations are the average N-test index value based on characteristics reported in the data and the robustness of individual immunity results by clinical standards will vary.

Findings from the ongoing Texas CARES survey will continue to elucidate understanding of the length of seropositivity

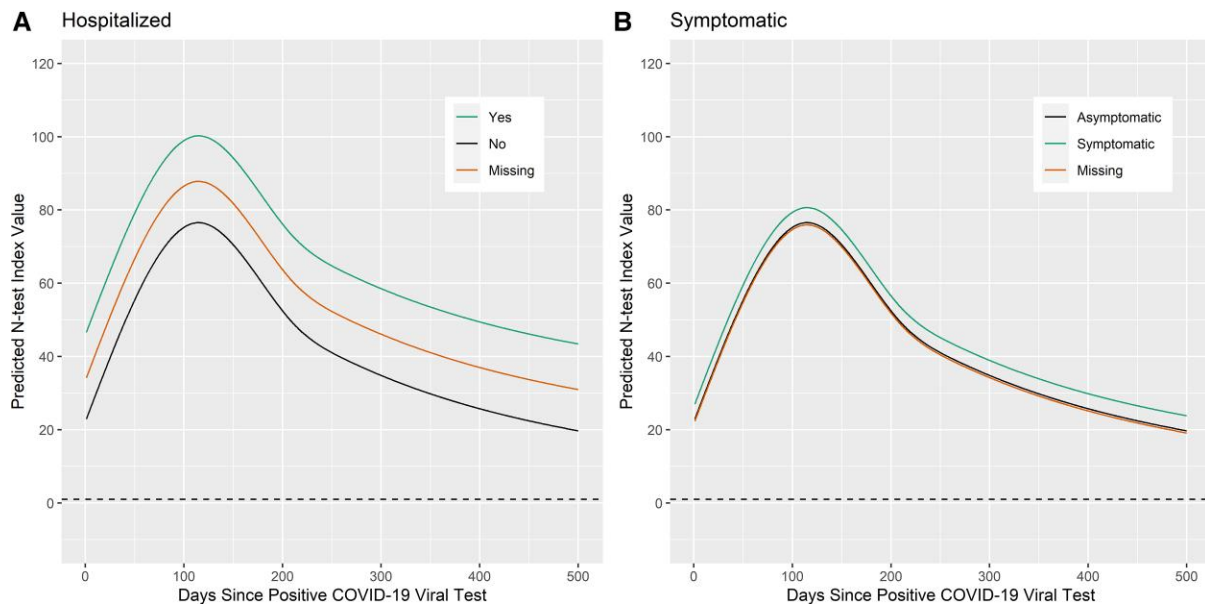


Figure 3. Predicting nucleocapsid antibody status by hospitalization and symptomatic status. The dashed line at 1 represents when the N-test reports a positive antibody response. N-test value ≥ 1 is a positive antibody response. The solid black line represents the reference prediction curve, that is for a 50 to 64-year-old non-Hispanic white woman without chronic conditions, with healthy BMI, who has insurance coverage, was asymptomatic in their infection, was not hospitalized, and did not use tobacco or vaping. The effect of varying a covariate across its domain is shown: (A) hospitalization status and (B) symptomatic/asymptomatic status. Abbreviations: BMI, body mass index; COVID-19, coronavirus disease 2019; N-test, nucleocapsid antibody test.

from natural as well as vaccine-induced human response. The knowledge gained by the state will allow Texas to target vaccine messaging to those most at risk and time responses to needs across populations.

As mitigation behaviors, such as mask use and social distancing, become more relaxed, and more variants emerge, estimating and predicting the duration of antibodies from natural infection becomes more important. Duration of the immunity produced by vaccines has been studied widely since initial clinical trials to measure their safety and efficacy. Antibody titers continue to be assessed on a population level to ascertain the need for vaccine boosters and determine the most at-risk populations to inform recommendations for distribution and programming. Given the proportion of people hesitant or unwilling to get vaccinated, understanding the duration of antibodies following a natural infection will play a key role in navigating the SARS-CoV-2 pandemic. Future research should consider longitudinal designs, collection of self-reported behaviors, and adjustments for behavioral and clinical characteristics to better understand natural and vaccine-induced response to SARS-CoV-2 across populations over time.

Notes

Financial support. This work was supported by the Texas Department of State Health Services (grant number HHS000866600001 for Texas CARES). Funding to pay the Open

Access publication charges for this article was provided by Texas DSHS (grant number HHS000866600001) for Texas CARES.

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

- Liu H, Zhang J, Cai J, et al. Investigating vaccine-induced immunity and its effect in mitigating SARS-CoV-2 epidemics in China. *BMC Med* **2022**; 20:37.
- Kadkhoda K. Herd immunity to COVID-19: alluring and elusive. *Am J Clin Pathol* **2021**; 155:471–2.
- Dyer O. Covid-19: delta infections threaten herd immunity vaccine strategy. *BMJ* **2021**; 374:n1933.
- Fine P, Eames K, Heymann DL. Herd immunity": a rough guide. *Clin Infect Dis* **2011**; 52:911–6.
- Klaassen F, Chitwood MH, Cohen T, et al. Population immunity to pre-omicron and omicron SARS-CoV-2 variants in US states and counties through December 1, 2021. medRxiv, doi: [10.1101/2021.12.23.21268272](https://doi.org/10.1101/2021.12.23.21268272), 1 March **2022**, preprint: not peer reviewed.
- Tram KH, Saeed S, Bradley C, et al. Deliberation, dissent, and distrust: understanding distinct drivers of coronavirus disease 2019 vaccine hesitancy in the United States. *Clin Infect Dis* **2021**; 74:1429–41.

7. Hildreth JEK, Alcendor DJ. Targeting COVID-19 vaccine hesitancy in minority populations in the US: implications for herd immunity. *Vaccines (Basel)* **2021**; 9:489.
8. Terpos E, Stellas D, Rosati M, et al. SARS-CoV-2 antibody kinetics eight months from COVID-19 onset: Persistence of spike antibodies but loss of neutralizing antibodies in 24% of convalescent plasma donors. *Eur J Intern Med* **2021**; 89:87–96.
9. Grandjean L, Saso A, Torres Ortiz A, et al. Long-term persistence of spike protein antibody and predictive modeling of antibody dynamics after infection with severe acute respiratory syndrome coronavirus 2. *Clin Infect Dis* **2022**; 74:1220–9.
10. Sood N, Pernet O, Lam CN, et al. Seroprevalence of antibodies specific to receptor binding domain of SARS-CoV-2 and vaccination coverage among adults in Los Angeles county, April 2021: the LA pandemic surveillance cohort study. *JAMA Netw Open* **2022**; 5:e2144258.
11. Alfego D, Sullivan A, Poirier B, et al. A population-based analysis of the longevity of SARS-CoV-2 antibody seropositivity in the United States. *EClinicalMedicine* **2021**; 36:100902.
12. Alejo JL, Mitchell J, Chang A, et al. Prevalence and durability of SARS-CoV-2 antibodies among unvaccinated US adults by history of COVID-19. *JAMA* **2022**; 327:1085–7.
13. Rodda LB, Netland J, Shehata L, et al. Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell* **2021**; 184:169–83.e17.
14. Long QX, Tang XJ, Shi QL, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* **2020**; 26:1200–4.
15. Gaebler C, Wang Z, Lorenzi JCC, et al. Evolution of antibody immunity to SARS-CoV-2. *Nature* **2021**; 591:639–44.
16. Gudbjartsson DF, Norddahl GL, Melsted P, et al. Humoral immune response to SARS-CoV-2 in Iceland. *N Engl J Med* **2020**; 383:1724–34.
17. Dan JM, Mateus J, Kato Y, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **2021**; 371:eabf4063.
18. Wang K, Long QX, Deng HJ, et al. Longitudinal dynamics of the neutralizing antibody response to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. *Clin Infect Dis* **2021**; 73:e531–9.
19. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat Microbiol* **2020**; 5:1598–607.
20. Efrati S, Catalogna M, Abu Hamed R, et al. Early and long term antibody kinetics of asymptomatic and mild disease COVID-19 patients. *Sci Rep* **2021**; 11:13780.
21. Ferrari D, Di Resta C, Tomaiuolo R, et al. Long-term antibody persistence and exceptional vaccination response on previously SARS-CoV-2 infected subjects. *Vaccine* **2021**; 39:4256–60.
22. Marchi S, Viviani S, Remarque EJ, et al. Characterization of antibody response in asymptomatic and symptomatic SARS-CoV-2 infection. *PLoS One* **2021**; 16:e0253977.
23. Crawford KHD, Dingens AS, Eguia R, et al. Dynamics of neutralizing antibody titers in the months after severe acute respiratory syndrome coronavirus 2 infection. *J Infect Dis* **2021**; 223:197–205.
24. Tian X, Liu L, Jiang W, Zhang H, Liu W, Li J. Potent and persistent antibody response in COVID-19 recovered patients. *Front Immunol* **2021**; 12:659041.
25. Wang H, Yuan Y, Xiao M, et al. Dynamics of the SARS-CoV-2 antibody response up to 10 months after infection. *Cell Mol Immunol* **2021**; 18:1832–4.
26. Scozzari G, Costa C, Migliore E, et al. Prevalence, persistence, and factors associated with SARS-CoV-2 IgG seropositivity in a large cohort of healthcare workers in a tertiary care university hospital in northern Italy. *Viruses* **2021**; 13:1064.
27. Valerio-Shewmaker MA, DeSantis S, Swartz M, et al. Strategies to estimate prevalence of SARS-CoV-2 antibodies in a Texas vulnerable population: results from phase I of the Texas coronavirus antibody response survey. *Front Public Health* **2021**; 9:753487.
28. Qu J, Wu C, Li X, et al. Profile of immunoglobulin G and IgM antibodies against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin Infect Dis* **2020**; 71:2255–8.
29. Centers for Disease Control and Prevention. About adult BMI. https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html. Accessed 4 October **2021**.
30. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. 3rd edn. Hoboken, NJ: John Wiley and Sons, **2013**.
31. Figueiredo-Campos P, Blankenhaus B, Mota C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in COVID-19 patients and healthy volunteers up to 6 months post disease onset. *Eur J Immunol* **2020**; 50:2025–40.
32. Lynch KL, Whitman JD, Lacanienta NP, et al. Magnitude and kinetics of anti-severe acute respiratory syndrome coronavirus 2 antibody responses and their relationship to disease severity. *Clin Infect Dis* **2021**; 72:301–8.
33. Wajnberg A, Amanat F, Firpo A, et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **2020**; 370:1227–30.