Association between SARS-CoV-2 IgG antibody detection levels and time since viral exposure in humans

Hayden L. Smith,[1] Jacob B. Bliss,[1] Katherine R. Sittig[2]
[1]UnityPoint Health – Des Moines, 1415 Woodland Ave, Ste 140, Des Moines, IA, 50903
[2]University of Iowa – Des Moines Internal Medicine Residency Program, 1415 Woodland Ave, Ste 140, Des Moines, IA 50309

Abstract

Background: There is insufficient information on the human immunogenic response to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

Objective: This study aimed to model waning SARs-CoV-2 immunoglobulin G (IgG) antibody detection levels given time since a self-reported positive viral test.

Methods: A seroprevalence study was conducted within a United States (US) health system located in the Midwest. Participating hospital and clinic employees completing a study survey were eligible to receive a free SARS-CoV-2 IgG antibody test. A generalized linear model was fit regressing IgG detection levels on time since a known prior infection as documented in the study survey. A literature review was conducted to locate and recreate serial SARS-CoV-2 IgG antibody detection level data from external studies. These data were scored based on the study model and used to evaluate the predictive out-of-sample accuracy of the estimate.

Results: Of the 6,009 eligible employees, 2,848 completed the study survey, and 2,118 had antibody testing. Of these employees, 221 reported a prior SARS-CoV-2 infection and date they received the positive test result. These data were used to model IgG detection values given time since a prior infection. Antibody testing for these employees was taken a median of 90 (IQR: 59, 153) days since their reported infection. The study model estimated a multiplicative IgG detection decrease of 0.99 (95% CI: 0.99, 1.00) per day since the prior positive test. Five external studies were found with applicable test information and used to examine out-of-sample model accuracy. Over fifty percent of these external data were recreated and represented 131 patients and greater than 300 individual tests. The crude out-of-sample model error was 0.0 (SE: 0.1) and -0.9 (SE: 0.1) when controlling for patient clusters within studies.

Conclusions: Given the near-real-time dissemination of pandemic information, flexible modeling strategies and validation processes were explored. The presented modeling approach served to validate the possible utility of the presented IgG prediction model.

Key Words: COVID-19; SARS-CoV-2; serological testing; health personnel; healthcare providers; seroprevalence; United States; Midwestern United States

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in late 2019 and is the viral strain responsible for the ongoing pandemic (Imai 2021). Since the initial spread of coronavirus disease (COVID-19), the international scientific community has attempted to understand the immune response against SARS-CoV-2. Approximately 20% of COVID-19 cases have been documented as severe, with a significant number of infected populations asymptomatic (Wu 2020). An improved understanding of the longevity of the post-infection immune response in humans could help guide testing and interventional methods.

Evaluation of prior infections can require an examination of individual immunogenic responses utilizing new testing methods. COVID-19 has two key antigens that produce an antibody response: the internal nucleocapsid (N protein) and the external spike protein (S protein) (Imai 2021). A successful response to large-scale outbreaks requires an understanding of all individuals who have been exposed. Since most SARS-CoV-2 infections result in detectable antibody production, serology testing makes it possible to identify previous infections, regardless of the presence of physical symptoms. Data from serological surveys would improve public health understanding of the human immunogenic response to SARS-CoV-2 (Murrell 2020). The objective of this study was to model SARs-CoV-2 immunoglobulin G (IgG) detection level given time since a known exposure.

2. Methods

2.1 Study Design

A seroprevalence study was conducted between November 2020 through January 2021 within a midwestern healthcare system in the United States (US). Those eligible for the study were employees from four hospitals and associated clinics located in a single metropolitan area in Iowa. Individuals who could review study materials written in English (~98% employees) were eligible to participate in the study. These participants were asked to complete an electronic survey on job roles, health risks, and COVID-19 exposures, survey available as supplemental material in Smith 2021. Participating employees completing the survey were then eligible for a free SARS-CoV-2 antibody test. The Abbott Diagnostics Architect System SARS-CoV-2 IgG (Abbott, Abbott Park, IL, US) chemiluminescent microparticle immunoassay was used to conduct the antibody test. The immunoassay detects IgG antibodies against the SARS-CoV-2 N protein and reports an index value, representing the ratio of signal to cut-off values. This index value corresponds to the relative light units (RLU) produced by the sample compared to the RLU's produced by an assay calibrator sample. The test manufacturer recommended a 1.40 index as a cut-off value for seropositivity. Semi-quantitative test results based on the manufacturer cut-off were mailed to the study participants via self-addressed envelopes filled out by participants at the time of testing. Seroprevalence estimates for the subsample of participants with only a completed survey, survey plus an antibody test, and results transported to the study population were calculated based on methods presented in the Smith 2021 Society of Epidemiologic Research abstract. The presented study received Institutional Review Board approval (IM2020-126), which included a waiver of written consent for survey response data and written consent for antibody testing.

2.2 Statistical Analysis

Quantitative antibody test results were modeled, regressing IgG index detection values on time since an employee's self-reported prior positive test using a generalized linear model with a gamma distribution and log link. The estimate was presented with a 95% confidence interval (CI).

2.3 Model Validation

A review of accessible study results searchable online (e.g., PubMed and Arxiv servers) was performed using key terms related to COVID-19 (e.g., COVID-19, SARS-CoV-2, antibodies, IgG, serological study, seroprevalence). Papers with plotted serial SARS-CoV-2 antibody test results were considered eligible for inclusion as validation data. Located figures were captured, enlarged with grid lines added, and had serial values matched and labeled. Next, study team members recorded the estimated values and days between tests and reviewed them by committee as needed to achieve consensus.

Recovered external validation data were then scored, using the first antibody test value to calculate the estimated serial value within-subject. These estimates were then subtracted from the recovered serial value to calculate the estimated errors of predictions. The mean values of these errors were naively calculated via arithmetic mean and controlled for multiple values taken from the same subject by using an empty, intercept-only, multi-level model with unstructured covariance. Lastly, a sensitivity analysis was conducted by repeating this process while leaving each external study out of the model once when calculating the mean prediction error. The output of this process, along with standard errors (SE), served to visualize the contribution of each study on the prediction error estimate.

3. Results

3.1 Study Sample

Approximately 6,009 employees were sent a study invitation, with 2,848 (47%) completing the electronic survey and 2,118 (36%) obtaining an antibody test. The seroprevalence semi-quantitative SARS-CoV-2 IgG antibody detection was 13.1% in the sample. The prevalence after inverse propensity score weighting of the sample to the characteristics of the employees that completed the study survey changed the prevalence to 13.5%. Lastly, the seroprevalence remained at 13.5% after calibrating the weighted sample to match the sex and age distributions of the study's target population.

Subject demographic information stratified by semi-quantitative antibody results is presented in Table 1. As noted, 273 of the participating employees were positive for SARS-CoV-2 IgG antibody, with 1,845 (87%) testing negative. Positivity rate did not appear to vary based on gender, full-time vs. part-time status, or whether employees provide direct patient care or care to patients undergoing aerosol-generating procedures. However, only 8% of physicians and 9% of advanced practice providers tested positive for the SARS-CoV-2 antibody, compared to 13-15% for other professional groups. There was a general trend of increasing positivity rate with the frequency of care to patients with COVID-19; only 8% of employees who never cared for patients with COVID-19 tested positive, while 20% of employees who provided daily care to COVID-19 patients were antibody positive. A relatively larger proportion (17%) of participants who had tested for COVID-19 were positive for the antibody to SARS-CoV-2. Finally, the proportion of antibody-positive employees increased with the self-reported likelihood of having had COVID-19.

3.2 Study Model

Antibody results are presented in Figure 1 and categorized by subjects' self-reported previous COVID-19 test results from the study survey. Figure 2 presents the antibody test values for subjects with a self-reported positive test result (n = 221) and date, categorized by age group and plotted against days since their self-reported test. Antibody testing for these employees was taken at a median of 90 (interquartile range: 59, 153) days since their reported infection. The fitted model revealed the antibody detection values decrease by an estimated multiple of 0.994 (95% CI: 0.992, 0.995) for every day since the date of their self-reported positive infection (Figure 2). The rate change can be examined for any number of days since infection by taking the estimate of 0.994 to the power of days of interest (e.g., an antibody test result would decrease by multiple of 0.83, thirty days after a reported positive test [i.e., 0.994**30]).

3.3 External Study Data

Five studies met inclusion criteria for modeling (Boonyaratanakornkit 2020; Murrell 2020; Taubel 2020; Buss 2021; Sakhi 2021). Descriptive information about these studies is presented in Table 2. Two of the studies were from the United Kingdom (i.e., public health center staff and patients), one from the United States (i.e., plasma donors), one from Brazil (i.e., plasma donors), and the last study from France (i.e., hemodialysis patients). Over fifty percent of these external data were recreated and represented 131 unique patients and greater than 300 individual tests.

3.4 External Model Validation

The naive mean prediction error of the study model was 0.00007 (SE: 0.08) index value units when applied to external study data. The mean prediction error from the multi-level model controlling for within and between person variability was -0.09747 (SE: 0.10) index value units. The mean and recalculated mean prediction error while removing each external study data set one time is presented in Figures 3-4. These results revealed that the error estimate had the largest mean shifts when removing the US and France studies.

4. Discussion

4.1 Summary of Findings

This study utilized an adaptive approach for model validation. There was a relatively large initial study sample, with nearly 3,000 health system employees completing the questionnaire and over 2,000 participating in the blood sample collection. The rate of positive tests for the SARS-CoV-2 IgG antibody was about 13.5% in the study population. This positivity rate was a weighted percentage based on employee demographics. There were 273 individuals who had a test value above the positive threshold set by Abbott, and there were 221 individuals with a reported previous coronavirus test with a known date. When looking at how frequently health care workers cared for known COVID-19 positive patients, overall positive test values were highest in those who reported "daily" interactions with these patients compared to less frequent interactions with COVID-19 patients. Lastly, while observing the fitted model created from the study data, antibody detection decreased by an estimated multiple of 0.994 for every day since the self-reported date of a previous positive test result. From a public health perspective, this is a valuable model to identify individuals that had a prior COVID-19 infection and understand the waning of IgG values.

The SARS-CoV-2 IgG anti-nucleocapsid antibodies appear to decrease over time (Boonyaratanakornkit 2020; Murrell 2020; Taubel 2020). Through combining data

collected from other studies of similar design, we developed a more comprehensive understanding of the waning process for SARS-CoV-2 antibodies. When comparing the presented model results to the five external studies, the model nominally overpredicted the waning process. The long-term immunological response to COVID-19 is unknown given it is a novel disease. Based on the model constructed from around 200 individuals, their antibody levels decrease from the day they receive a positive test. Given the plotted data, it is estimated that antibody production only remains high enough to result in a semi-quantitative positive antibody test result for around seven months. There are still a significant number of cases that do not follow this general trend. It is still unknown what causes a variance in immune response, whether age, gender, infection severity, or other health problems. Another potential latent variable may be secondary exposures to the virus after a documented positive test. Due to the small sample size, we could not investigate and extrapolate our data to make any definite conclusions about what causes a variance in immune responses. Knowing that antibody production is not a long-term response is essential from a public health perspective. It highlights the importance of utilizing all precautions necessary to protect oneself from possible COVID-19 exposures and helps in understanding the possible utility of vaccines.

4.2 Limitations

The study sample included individuals who voluntarily completed the self-reported survey and received an antibody test. The data utilized from the survey of the 221 individuals included a documented past positive test result allowing for the calculation of time since prior known infection. These individuals most likely had a known exposure or were symptomatic of COVID-19, so the generalizability of study results may be limited to comparable individuals. In addition, the antibody test in this study was conducted before the establishment of viral variants such as the delta variant in the US. The external data recovery methods were limited by the number of values that could be discerned from published figures, though the process was systematically completed and included confirmation by committee when necessary. Multi-level modeling of the mean prediction errors adjusted for values clustered in subjects but did not address for possible heterogeneity between studies – which was post hoc examined in the sensitivity analyses. Finally, all antibody results used in this study were based on one type of antibody test. Of note, at the time of this study, there were no known other studies with visualized results based on a different antibody test that had to be excluded from analyses.

4.3 Future Considerations

If additional external data became available, the waning of IgG values could be analyzed based on patient characteristics. This analysis would allow for the investigation of potential differences in antibody waning between subgroups of patients. Additionally, this same process could be replicated for examining the S protein in the vaccine. Researchers could then examine waning IgG antibody levels related to receiving the COVID-19 vaccine and provide information toward implementing a possible booster shot to attempt to maintain a measurable immune response. However, these pursuits would need to be coupled with individual's infection and symptom data as well.

5. Conclusion

Given the near-real-time dissemination of pandemic information, a flexible modeling strategy was examined. The presented approach served to validate the possible utility of an

IgG detection prediction model. The novel process provided insight into the immunological response longevity and could be emulated to address similar research topics or inquiries.

Acknowledgments

The authors would like to acknowledge Lisa A. Veach, Megan J. Romine, Maura J. Prescher, Addison L. Bratvold, Kristen A. Beal, and Angel C. Mueller for their contributions to the design and implementation of this study. The authors would also like to thank Della Guzman and the Wellness Services staff (Alex Aller, Maria Di Gioia, Susan Hawks, Alex Leon, Nikkol Johnson, Lindsey Petty, Sara Plucar, Patty Teirnan, Marcie Troyer, Jennifer Umphfleet) for the collection and assistance with the blood samples used in antibody testing.

References

Boonyaratanakornkit J, Morishima C, Selke S, et al. Clinical, laboratory, and temporal predictors of neutralizing antibodies to SARS-CoV-2 after COVID-19. medRxiv 2020.10.06.20207472; doi: https://doi.org/10.1101/2020.10.06.20207472.

Buss LF, Prete CA Jr, Abrahim CMM, et al. Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. Science 2021;371(6526):288-292. doi: 10.1126/science.abe9728.

Imai K, Kitagawa Y, Tabata S, et al. Antibody response patterns in COVID-19 patients with different levels of disease severity in Japan. J Med Virol 2021;93(5):3211-3218. doi: 10.1002/jmv.26899.

Murrell, I, Forde, D, Tyson, L, et al. A longitudinal comparison of spike and nucleocapsid SARS-CoV-2 antibody responses in a tertiary hospital's laboratory workers with validation of DBS specimen analysis. medRxiv 2020.10.29.20219931; doi: https://doi.org/10.1101/2020.10.29.20219931.

Sakhi H, Dahmane D, Attias P, et al. Kinetics of Anti-SARS-CoV-2 IgG Antibodies in Hemodialysis Patients Six Months after Infection. J Am Soc Nephrol 2021;32(5):1033-1036. doi: 10.1681/ASN.2020111618.

Smith HL Prescher MJ, Sittig KR, Veach LA. Severe acute respiratory coronavirus virus 2 (SARS-CoV-2) vaccine acceptance in employees in an integrated health system in the Midwest. ASHE 2021; 1(1):e14, 1-4. doi:10.1017/ash.2021.166.

Smith HL. Analytic weighting scheme for recovering target sample and population characteristics in a SARS-CoV-2 sero-prevalence study. 54th Annual Society for Epidemiologic Research (SER) Meeting, virtual conference, June 22-25, 2021.

Taubel, J, Cole, ST, Christopher S, et al. Why the SARS-CoV-2 antibody test results may be misleading: insights from a longitudinal analysis of COVID-19. medRxiv 2020.12.08.20245894; doi: https://doi.org/10.1101/2020.12.08.20245894.

Wu Z, McGoogan JM. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention. JAMA 2020;323(13):1239-1242. doi: 10.1001/jama.2020.2648.

Table 1. Participating employee characteristics by SARS-CoV-2 Antibody test status in a Midwestern health system, n=2118

Characteristic	Positive Sero-Antibody Test[a]	
	Positive n=273 (13%)	**Negative n=1845 (87%)**
Age[b]		
21 – 30	75 (18%)	353 (82%)
31 – 40	62 (10%)	580 (90%)
41 – 50	63 (13%)	412 (87%)
51 – 60	57 (15%)	323 (85%)
61 or older	14 (8%)	170 (92%)
Gender[c]		
Female	239 (13%)	1620 (87%)
Male	33 (13%)	221 (87%)
Work Status		
Full time	231 (13%)	1561 (87%)
Part time	42 (13%)	284 (87%)
Primary Work Setting		
Inpatient/ED	137 (12%)	999 (84%)
Outpatient	92 (14%)	586 (86%)
Non-clinical	44 (14%)	187 (86%)
Occupation Group		
Physicians	12 (8%)	144 (92%)
PA/NP	7 (9%)	71 (91%)
Nurses, nursing assistant	118 (13%)	794 (87%)
Allied health	59 (14%)	367 (86%)
Support roles	30 (13%)	195 (87%)
Other nonpatient care	47 (15%)	274 (85%)
Provided Direct Patient Care (Yes)	203 (13%)	1360 (87%)
Patient Population[d]		
Adult	114 (15%)	662 (85%)
Pediatric	19 (7%)	238 (93%)
Both	70 (13%)	460 (87%)
On average, how often do you provide direct patient care to patients known or suspected to have COVID-19?	60 (20%)	239 (80%)
Daily	63 (13%)	410 (87%)
Multiple times per week	14 (12%)	105 (88%)
Once per week	28 (12%)	215 (88%)
Multiple times per month	19 (10%)	167 (90%)
Less than once per month	11 (8%)	130 (92%)
Never		
Provide Care to Patients Undergoing AGPs [d]	73 (13%)	489 (87%)
How likely do you think it is that you've had COVID-19?	9 (9%)	92 (91%)
Extremely unlikely	22 (4%)	587 (96%)
Unlikely	57 (7%)	811 (93%)
Equally likely and unlikely	30 (11%)	241 (89%)
Likely	155 (58%)	114 (42%)
Extremely likely		

Previously Tested for COVID-19 (Yes)	212 (17%)	1014 (83%)
If yes, was test positive?	150 (67%)	71 (32%)
How do you think you got COVID-19?		
Unsure	43 (64%)	24 (36%)
Work exposure from patient	33 (72%)	13 (28%)
Home exposure	32 (63%)	19 (37%)
Community exposure	23 (77%)	7 (23%)
Work exposure from coworker	18 (69%)	8 (31%)
Prefer not to answer	1 (100%)	0 (0%)

AGP = Aerosol-Generating Procedure

[a]See manuscript test for information about a positive antibody test.

[b]9 subjects under 21 years of age not listed

[c]3 employees reported another category for gender.

[d]Subsample of employees reporting direct patient care.

Table 2. External study data sources used to validate a model of change in immunoglobin G antibody levels given time since a reported prior COVID-19 infection.

Study (Year)	Location (population)	Eligible Values	Recovered Values	Recovered Clusters[a]
Taubel (2020)	UK Public Health (staff/visitors)	65	65	15
Murrell (2020)	UK Public Health (staff)	31	31	9
Boonyaratanakornkit (2020)	US Public (plasma donors)	82	30	15
Buss (2021)	Brazil (plasma donors)	208	146	73
Sakhi (2021)	France (hemodialysis patients)	224	55	18
Totals		**610**	**327**	**131**

UK: United Kingdom; US: United States.
[a]Number of unique subjects, which could be smaller than the number of recovered values since in some studies subjects had greater than two serial antibody tests.

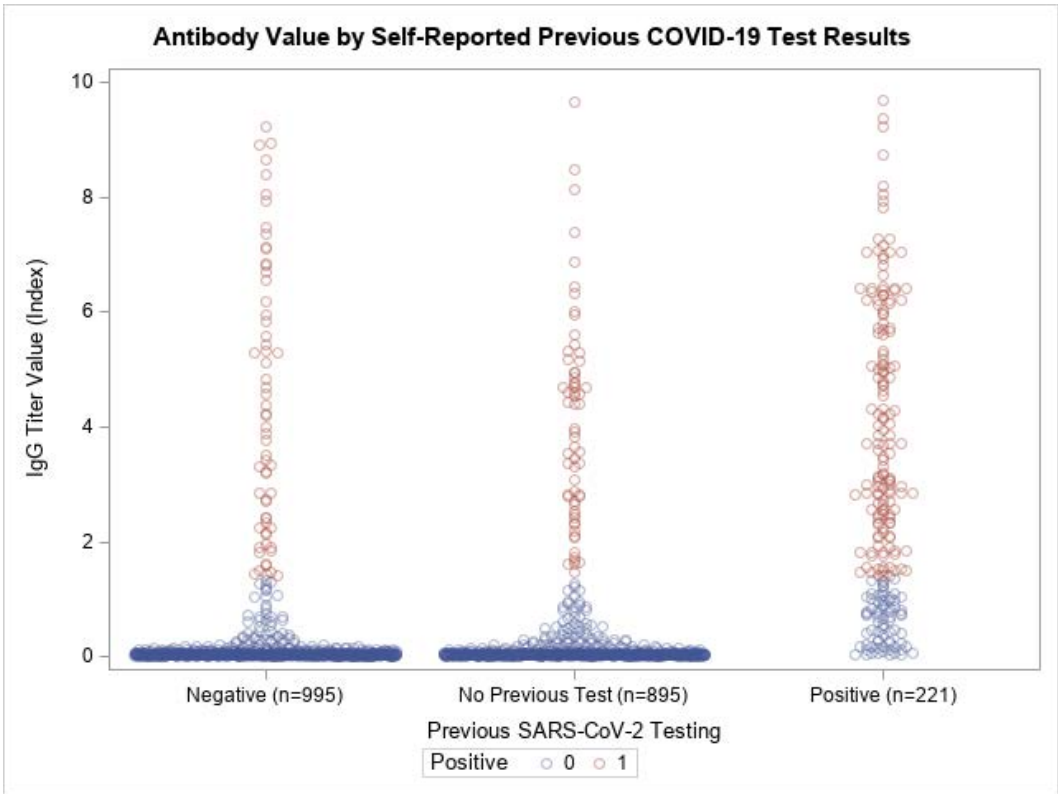

Figure 1: Plotted SARS-CoV-2 antibody detection value against self-reported previous COVID-19 test results for study participants from an integrated health system. Red empty dots represent positive antibody test results given the test manufacturers recommended threshold (i.e., 1.4 index value).

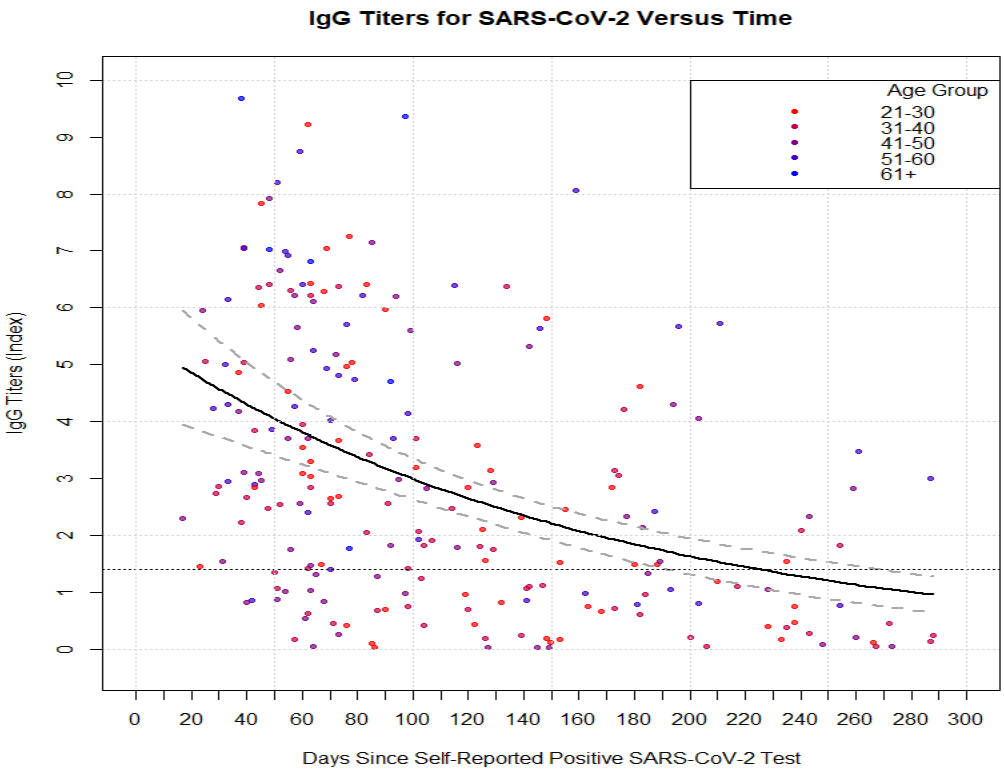

Figure 2: Plotting of SARS-CoV-2 antibody test results against time since a self-reported positive test for study participants from an integrated health system in a serosurvey conducted in December 2020-January 2021, n=221. Visualized is the model's estimated fit with 95% confidence interval. The model did not control for age category. The dashed horizontal line represents the manufacturer's suggested positive value for the antibody test (i.e., 1.4 index value).

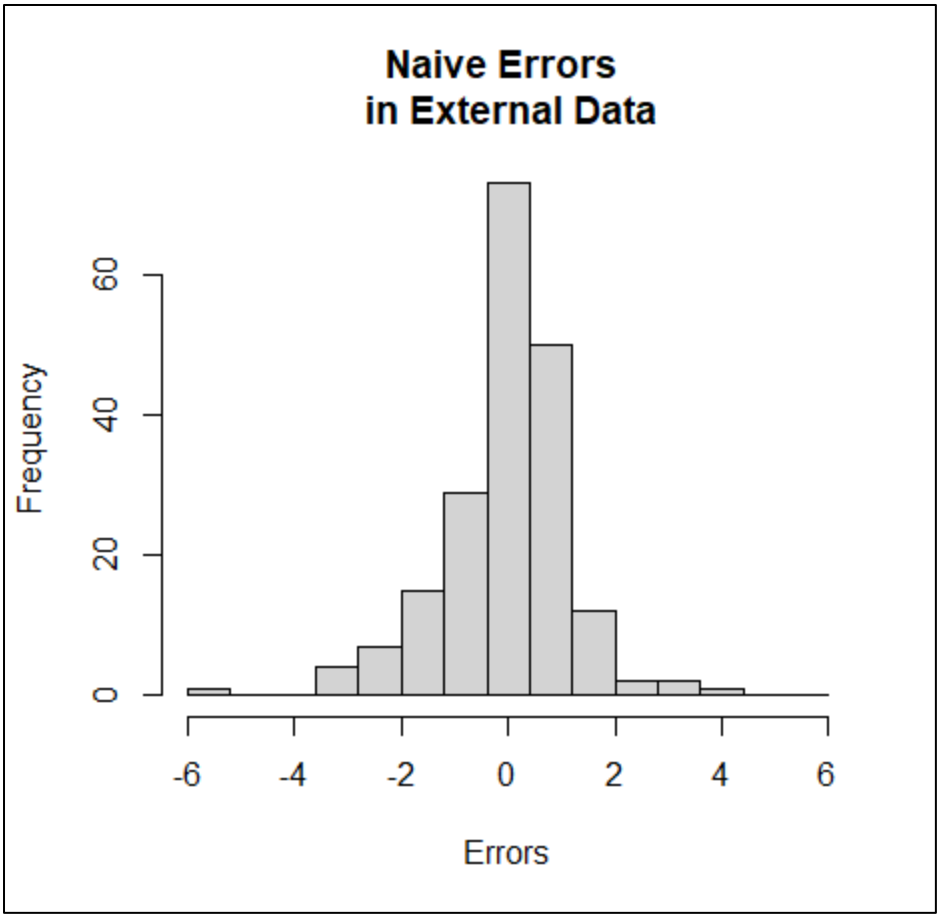

Figure 3: Histogram of naïve model errors from scored external study data on the estimated wanning of immunoglobin G values given time between antibody tests for COVID-19. The value at the approximate "-6" marker was removed from the validation dataset, as it appeared to be a subject that tested negative and then subsequently tested positive on their serial test.

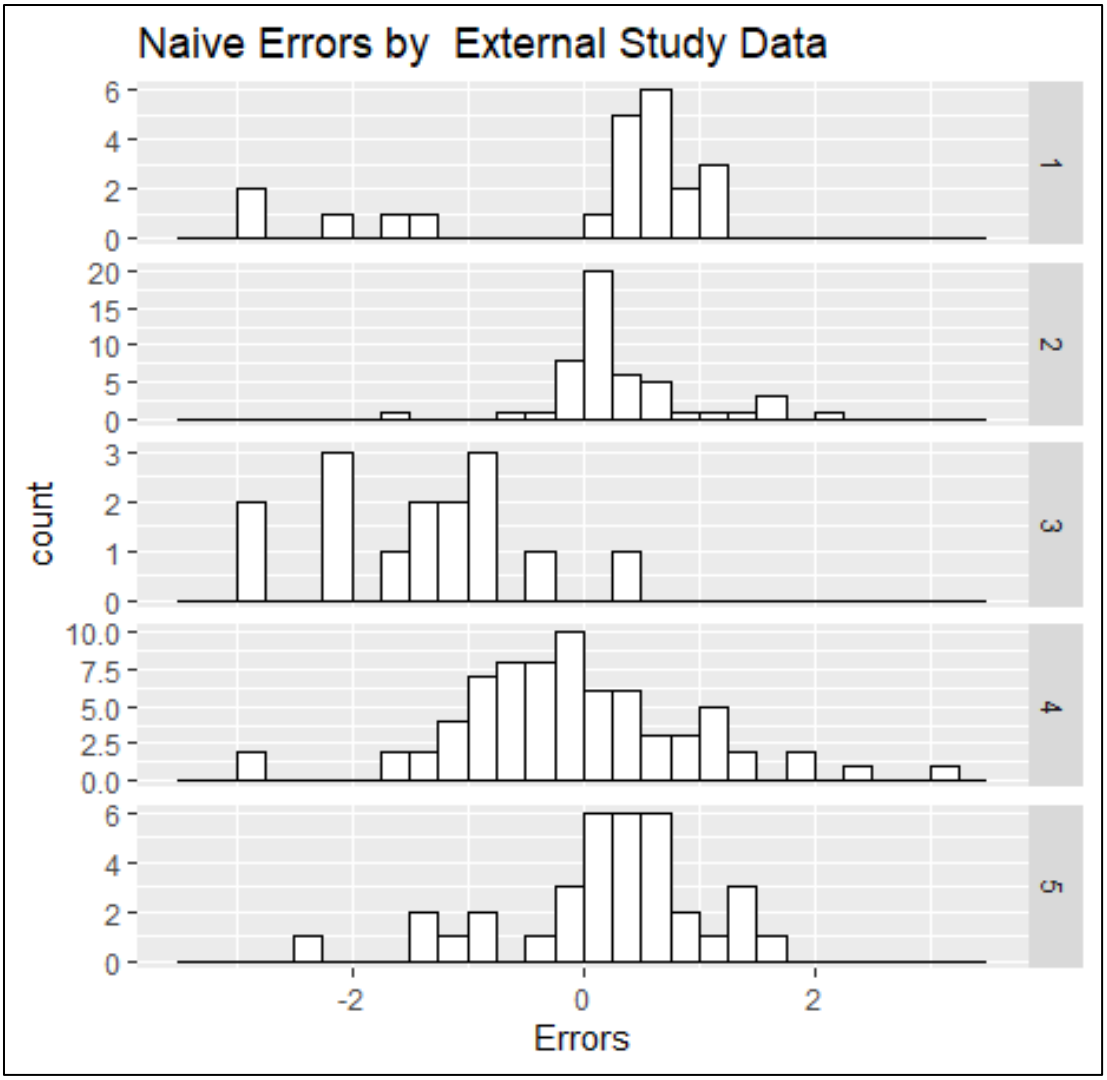

Figure 4: Histograms of model errors in each external study dataset used to validate wanning of immunoglobin G values given time estimate. 1= Taubel (2020); 2 = Murrell (2020); 3 = Boonyaratanakornkit (2020); 4 = Buss (2021); and 5 = Sakhi (2021).

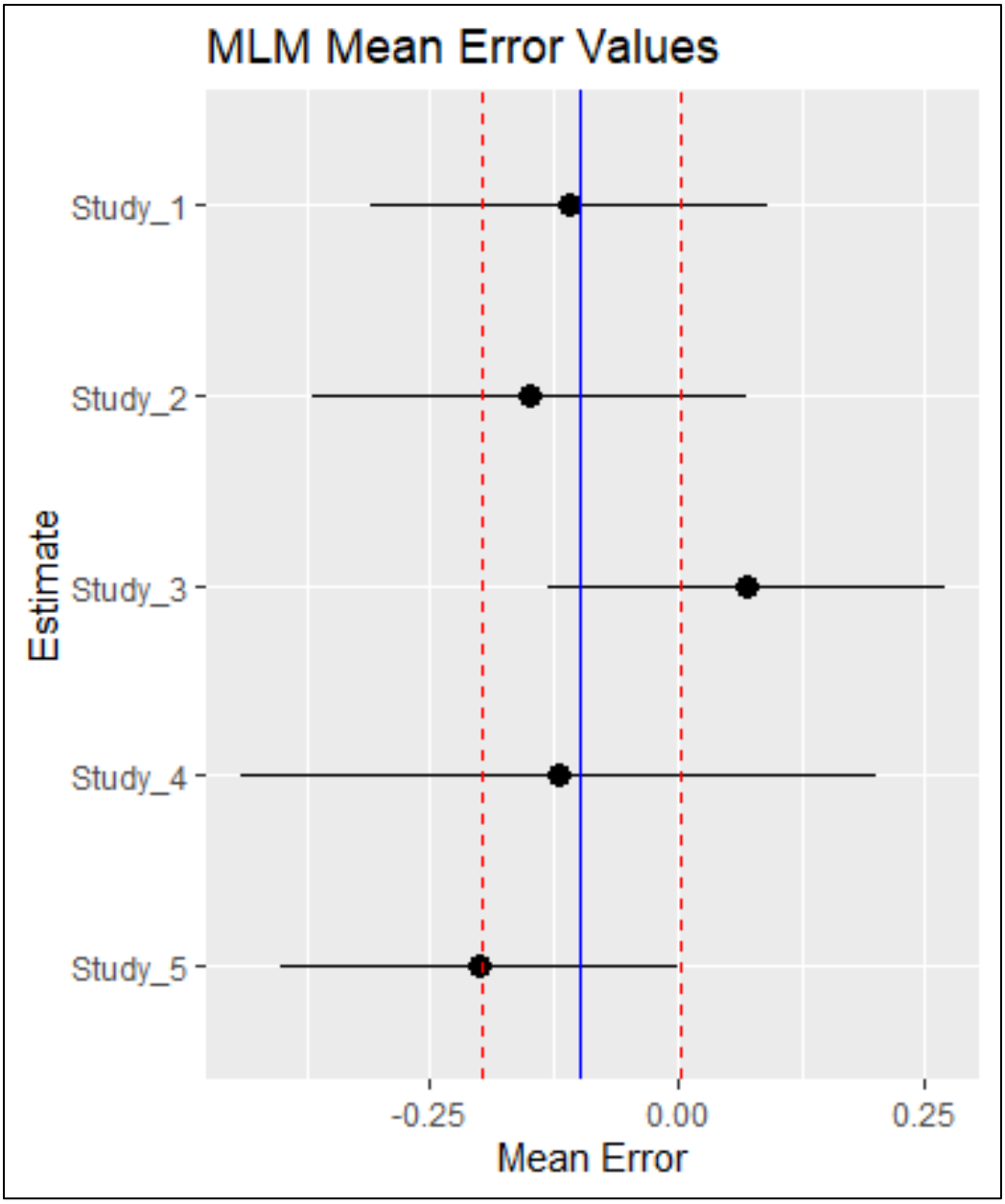

Figure 5: Histograms of mean prediction errors based on multi-level model (MLM) when leaving each study out of the sensitivity analysis one time. The blue line represents the mean prediction error when no studies are withheld, and the red dashed lines represent +/- 2 standard errors on this estimate. 1= Taubel (2020); 2 = Murrell (2020); 3 = Boonyaratanakornkit (2020); 4 = Buss (2021); and 5 = Sakhi (2021).