

August 2021

# Estimation of Total Immunity to SARS-CoV-2 in Texas

STACIA M. DESANTIS<sup>1</sup>, LUIS G. LEÓN-NOVELO<sup>1</sup>, MICHAEL D. SWARTZ<sup>1</sup>, ASHRAF S. YASEEN<sup>1</sup>, MELISSA A. VALERIO<sup>3</sup>, FRANCES A. BRITO<sup>1</sup>, JESSICA A. ROSS<sup>1</sup>, HAROLD W. KOHL, III<sup>4,5</sup>, SARAH E. MESSIAH<sup>2,7</sup>, STEVE H. KELDER<sup>4</sup>, LEQING WU<sup>1</sup>, SHIMING ZHANG<sup>1</sup>, KIMBERLY A. AGUILLARD<sup>1</sup>, MICHAEL O. GONZALEZ<sup>1</sup>, ONYINYE S. OMEGA-NJEMNOB<sup>4</sup>, CAMILLE J. BREAUX<sup>1</sup>, DAVID L LAKEY<sup>8</sup>, JENNIFER A. SHUFORD<sup>9</sup>, STEPHEN PONT<sup>5,9</sup>, ERIC D BOERWINKLE<sup>1</sup>.

<sup>1</sup>University of Texas Health Science Center at Houston, School of Public Health, 1200 Pressler St. Suite E805, Houston, TX, 77030, USA E-mail: [Luis.G.LeonNovelo@uth.tmc.edu](mailto:Luis.G.LeonNovelo@uth.tmc.edu)

<sup>2</sup>University of Texas Health Science Center, School of Public Health, Dallas Campus, Dallas, TX, USA

<sup>3</sup>University of Texas Health Science Center, School of Public Health, Brownsville Campus, Brownsville, TX, USA

<sup>4</sup>University of Texas Health Science Center, School of Public Health, Austin Campus, Austin, TX, USA

<sup>5</sup>University of Texas at Austin, Austin, TX, USA

<sup>6</sup>University of Texas System, Austin, TX, USA

<sup>7</sup>Center for Pediatric Population Health, UTHealth School of Public Health and Children's Health System of Texas, Dallas, TX, USA

<sup>8</sup>The University of Texas System and The University of Texas Health Science Center at Tyler

<sup>9</sup>Texas Department of State Health Services, Austin, TX, USA

**Abstract** Given the underestimate of seroprevalence in the US due to insufficient testing, accurate estimates of population immunity to SARS-CoV-2 or vaccinations do not exist. Although model-based estimates have been proposed, they require inputting unknown parameters such as viral reproduction number, longevity of immune response, and other dynamic factors. In contrast to a model-based approach for estimating population immunity, or simplistic summing of natural- and vaccine- induced immunity, the current study presents a data-driven statistical procedure for estimating the total immunity rate in a region using prospectively collected serological data along with state-level vaccination data. We present a detailed procedure so that efforts can be replicated regionally to inform policy-making decisions relevant to SARS-CoV-2. Specifically, we conducted a prospective longitudinal statewide cohort serological survey with 10,482 participants and more than 14,000 blood samples beginning on September 30, 2020. Along with Department of State Health Services vaccination data, as of July 4, 2021, the estimated percentage of those with naturally occurring antibodies to SARS-CoV-2 in Texas is 35.3% (95% CI = (33.7%, 36.9%)) and total estimated immunity is 69.1%. We conclude the seroprevalence of SARS-CoV-2 is 4 times higher than the state-confirmed COVID-19 cases (8.8%). This methodology is integral to pandemic preparedness.

**Keywords:** COVID-19; Infectious Diseases; Epidemic; Pandemic; Seroprevalence; Vaccine; Vaccination.

## 1. Introduction

It is increasingly important to estimate the percentage of individuals in the US who may be protected from the novel SARS-CoV-2 virus as a result of having circulating anti-SARS-CoV-2 antibodies. People obtain immunity through either natural infection with SARS-CoV-2 or vaccination, and total immunity is the combination of these two avenues of immunity. Usually, an estimate of total immunity is obtained using mathematical modeling and simulation, which require inputs such as duration of immunity once infected, viral reproduction rate, population mixing, and additional factors [1, 2, 3, 4, 5]. However, the contributions of these inputs are still not fully known. For example, researchers are unsure

of the duration of natural and vaccine-induced immunity, and possible T-cell cross-reactivity. Further, continual emergence of SARS-CoV-2 variants threaten progress toward immunity. [*e.g.*, 6, 7, 8]

Recent research suggests neutralizing antibodies to SARS-CoV-2 persist for at least 5 months [9, 10] or possibly longer [11], and that re-infection risk is low in the several months after initial infection [12]. Additionally, at the time of this publication, there has been great success of mass vaccination's in lowering viral transmission, *e.g.*, [13] indicate that in Israel where approximately 61% of the population are vaccinated (>80% of adults). This resultant reduction in viral spread has inspired the idea of a path to normality [NYT, 14].

Given the above most current prevailing assumptions that: 1. Reinfection with COVID-19 within a few months is rare [15]; 2. Neutralizing antibodies from natural infections typically last at least 5 months and cross-reactivity of serological tests is rare [16, 17]; and, 3. Vaccination produces a robust and reasonably long-term antibody response, make it possible to estimate regional total immunity as a combination of natural and vaccine-induced immunity [18].

The goal of this report is to demonstrate this estimation process in Texas as of July 4, 2021 using a prospectively designed serological survey. To this end, we first estimate period seroprevalence over 1-week intervals from 14,899 blood specimens collected prospectively from participants throughout Texas. We then compute a census age-adjusted seroprevalence estimate of natural infection and combine it with the Texas Department of State Health Services (DSHS) de-identified population-level vaccination data to obtain an accurate state-level estimate of total immunity. Notably, the approach is not limited to the current pandemic; it is applicable to any infectious disease.

## 2. Methods

### 2.1. Participants and Study Design

The Texas Coronavirus Antibody REsponse Survey (Texas CARES) initiative has been previously described [19, 20]. Briefly, Texas CARES is a prospective convenience sample of adult retail/business employees, K-12 and university educators and university students, those attending Health Resources and Services Administration (HRSA)- designated federally qualified health centers (FQHCs), and children 5-17 years, all of whom agreed to longitudinal monitoring of SARS-CoV-2 antibody status every three months (three time points total) from 10/1/2020-9/30/2021. A consent form and survey questionnaire were administered online at each of the three time points. More details about the study are publicly available on the Texas CARES dashboard [21].

### 2.2. Serological Assay and Vaccination Records

Antibody status was determined using the Roche Elecsys<sup>®</sup> Anti-SARS-CoV-2 (qualitative) assay detection of neutralizing antibodies against SARS-CoV-2 nucleocapsid (N) protein, hereafter referred to as "Roche N-test". The test has a sensitivity (95% confidence interval, CI) of 99.5%(97.0,100.0) and specificity of 99.82%(99.69,99.91)  $\geq$ 14 days after infection. De-identified population level daily vaccination data (2 doses for mRNA vaccines or 1 dose for Johnson and Johnson vaccine) by age group were obtained from Texas Department of State Health Services (DSHS). All protocols were approved by the UTHealth Committee for the Protection of Human Subjects and were also deemed "public health practice" by the Texas DSHS IRB.

## 2.3. Statistical Methods

The following components are estimated from the data, (1) immunity from natural infection, (2) immunity from complete vaccination, and (3) total immunity defined as immunity from either natural infection or vaccination (*e.g.*, “estimated total immunity”). Natural immunity period seroprevalence is calculated at a given time interval of Texas CARES, while vaccine-induced immunity is known (recorded by DSHS). For the current survey, a 1-week interval was deemed appropriate given the participant accrual rate into Texas CARES, and disease wave fluctuations. Within this interval, we assume serological status from prior infection and vaccine status are not independent events. This assumption is very well-supported by the data, which as of July 4th show 18.6% of those with reported prior documented COVID-19 disease are vaccinated, versus 30.7% without prior documented COVID-19. Our data and other research[22, 23] support that natural and vaccine-induced immunity are likely not independent events.

### 2.3.1. Calculating Total Immunity to SARS-CoV-2 in Texas

We describe the methods used to compute the total immunity in Texas over time. Let  $H$  denote the total number of census age groups,  $h = 1, \dots, H$ . Assume we have serological and vaccination data for  $T$  weeks. Let  $t$  index the time window (week), where  $t = 1, \dots, T$ . Now define:

- $\nu_{ht}$ : Vaccination proportion in age group  $h$  at week  $t$  (provided by state records). Since the vaccination status is cumulative,  $\nu_{h1} \leq \nu_{h2} \leq \dots \leq \nu_{hT}$  for  $h = 1, \dots, H$ , and  $\nu_{ht}$  is known with certainty.
- $\eta_{ht}$ : Natural immunity proportion in age group  $h$  at week  $t$ . This is unknown but is estimated cross-sectionally using the Roche N-test from Texas CARES.
- $w_h$ : Proportion of the Texas population in age group  $h$ . Thus,  $w_h$  is also known, and  $\sum_h w_h = 1$ .

The immunity rate in time window,  $t$ . (defined as having received the vaccine or testing positive for antibodies in the time window) in age group  $h$  is

$$\begin{aligned}
 \iota_{ht} &= P[\text{natural immunity or vaccine immunity}] \\
 &= P[\text{natural immunity}] + P[\text{vaccine immunity \& no natural immunity}] \\
 &= P[\text{natural immunity}] + P[\text{no natural immunity} \mid \text{vaccine immunity}]P[\text{vaccine immunity}] \\
 &= \eta_{ht} + \kappa_h \nu_{ht}
 \end{aligned} \tag{2.1}$$

In (2.1) above, for brevity, we omit the text “in group age  $h$  at week  $t$ ”. The definition of

$$\kappa_h := P[\text{no natural immunity in group } h \text{ at week } t \mid \text{vaccine immunity}].$$

This implicitly assumes the probability is equal across all weeks,  $t = 1, \dots, T$ . Also notice that the proportion of the population with both natural and vaccine induced immunity is easily estimated as  $(1 - \kappa_h)\nu_{ht}$ , so a mathematically equivalent expression to (2.1) is

$$\begin{aligned}
 \iota_{ht} &= P[\text{natural immunity or vaccine immunity}] \\
 &= P[\text{natural immunity}] + P[\text{vaccine immunity}] - P[\text{natural \& vaccine immunity}] \\
 &= \eta_{ht} + \nu_{ht} - (1 - \kappa_h)\nu_{ht}
 \end{aligned} \tag{2.2}$$

which may appear more intuitive: the total immunity rate is equal to the sum of the natural and vaccine induced immunity rates, minus their overlap. The population seroprevalence at week  $t$  is,

$$SPR_t = \sum_h w_h \eta_{ht} \quad (2.3)$$

and the immunity rate (natural or vaccine induced) at week  $t$  in the population is,

$$IR_t = \sum_h w_h \iota_{ht} = \sum_h w_h (\eta_{ht} + \kappa_h \nu_{ht}) \quad (2.4)$$

where  $\iota_{ht}$  is given in (2.1). The population proportion with both natural and vaccine-induced immunity is,

$$\sum_h w_h (1 - \kappa_h) \nu_{ht}.$$

In order to estimate  $SPR_t$  and  $IR_t$  we must estimate  $\eta_{ht}$  and  $\kappa_h$ . We show these steps in the following subsection. We also note that had we assumed independence between natural and vaccine induced immunity,  $\kappa_h = 1 - \eta_{ht}$ , as expected.

### 2.3.2. Estimation of Parameters for calculation of Total Immunity

We estimate  $\kappa_h$  and  $\eta_{ht}$  using the Roche N-test results. First,  $\kappa_h$  is estimated using the information from all  $T = 26$  study weeks since January 1, 2021, as the following sample proportion,

$$\tilde{\kappa}_h = \frac{\# \text{ of participants without natural immunity \& vaccinated in age group } h \text{ at any week}}{\# \text{ of vaccinated participants in age group } h} \quad (2.5)$$

and  $\eta_{ht}$  is initially estimated as,

$$\hat{\eta}_{ht} = \frac{\# \text{ participants in group } h \text{ tested at week } t \text{ with natural immunity}}{\# \text{ participants in group } h \text{ tested at week } t}. \quad (2.6)$$

Once we have  $\hat{\eta}_{i1}, \dots, \hat{\eta}_{iT}$  (and the denominators in the 2 equations above) we compute the isotonic version (across index  $t$ ) of  $\hat{\eta}_{ht}$ ,  $\tilde{\eta}_{ht}$  such that that  $\tilde{\eta}_{h1} \leq \tilde{\eta}_{h2} \leq \dots \leq \tilde{\eta}_{hT}$  for  $h = 1, \dots, H$ . See Supplementary Materials Subsection S.1 for details of this calculation. The isotonic estimate of  $\eta_{ht}$  is appropriate here because it reflects the fact that seroprevalence should not decrease over a short time interval (even though its raw estimate  $\hat{\eta}_{ht}$  can decrease due to expected sampling error in a small window,  $t$ ). Once these estimates are obtained, we compute the estimates of  $SPR_t$  and  $IR_t$ , called  $\widetilde{SPR}_t$  and  $\widetilde{IR}_t$ , by substituting the values of  $\tilde{\eta}_{ht}$  and  $\tilde{\kappa}_h$  into equations (2.3) and (2.4). Construction of a 95% confidence interval for  $\widetilde{SPR}_t$  is based on that for a proportion from a stratified design in which the outcome variable is binary [e.g., 24, 25] (details provided in Supplementary Materials S.2).

### 2.3.3. Algorithm to Estimate the Total Immunity Curve from Jan 1, 2021 to July 4, 2021

Recalling,  $H$  is the total number of age groups and  $T$  the total number of weeks. The algorithm is:

1. For,  $h = 1, \dots, H$ 
  - a) Using the Roche N-test, compute  $\tilde{\kappa}_h$  in (2.5).
  - b) Obtain the (cumulative) state vaccination rate for week  $t$  by age group, denoted  $\nu_{ht}$ , from the Texas Dept of State Health Services database. Since they are cumulative,  $\nu_{h1} \leq \nu_{h2} \leq \dots \leq \nu_{hT}$  for  $h = 1, \dots, H$ .

- c) For  $t = 1, \dots, T$ , compute the preliminary estimated N-test positive rate in the study at week  $t$ ,  $\hat{\eta}_{ht}$  in (2.6). Next, compute the isotonic version of  $\hat{\eta}_{ht}$ ,  $\tilde{\eta}_{ht}$ , such that  $\tilde{\eta}_{h1} \leq \tilde{\eta}_{h2} \leq \dots \leq \tilde{\eta}_{hT}$
2. Estimate the age-adjusted seroprevalence rate at week  $t$ ,

$$\widetilde{SPR}_t = \sum_h w_h \times \tilde{\eta}_{ht}$$

and then the total immunity rate at week  $t$ ,

$$\widetilde{IR}_t = \sum_h w_h \times [\tilde{\eta}_{ht} + \tilde{\kappa}_h \nu_{ht}].$$

3. Plot  $t$  v.s.  $\widetilde{SPR}_t$  and  $t$  v.s.  $\widetilde{IR}_t$

### 3. Results

Demographics of the full sample, and adults 18 years and over, respectively are shown in Tables 1 and 2. The mean (standard deviation) age of all participants was 45.9 years (16.1) and most participants were in the 50-54 year age group (30.9%). Most were female (69.4%), White (88.9%), and from urban locations (92.0%). Most adults reported having some college education or an advanced or professional degree, and were employed full time.

We applied the method with  $H = 10$  age groups, 0-15, 16-17, 18-29, 30-39, 40-49, 50-64, 65-74, 75-79, 80-84 and 85+ years. The census age-adjusted Texas COVID-19 seroprevalence using the Roche N-protein test over time (*i.e.*,  $t$  v.s.  $\widetilde{SPR}_t$ ) along with the 95% confidence band is shown in Figure 1. The vertical line on the graph delineates the time of first vaccine availability. The surges in seroprevalence correspond well to the known waves of SARS-CoV-2 in Texas [26].

We note that  $\tilde{\kappa}_h \approx 0.83$  in all age groups; thus, the proportion of study participants who reported having had both COVID-19, and being fully vaccinated were roughly  $1 - \tilde{\kappa}_h \approx 17\%$ . This indicates a violation of independence of natural infection and vaccination, which was expected. As these people must not be counted twice in the total immunity estimate, they are subtracted appropriately in each time period (week) per Equation (2.2).

The estimated age-adjusted total period immunity in Texas, defined as immunity from either natural infection or full vaccination (solid line) over time (*i.e.*,  $t$  v.s.  $\widetilde{IR}_t$ ) is shown in Figure 2. As of July 4, 2021, total immunity is estimated at 69% of the Texas population, with approximately 35.3% (95% CI = (33.7, 36.9)) resulting from natural infection. To our knowledge, this is the most robust and accurate non-model based estimate of total immunity to date in the state of Texas.

We do not include a confidence interval for total immunity since the proportion vaccinated is a known (fixed) population quantity rather than an estimate, and thus does not lend itself to an estimate of variability. While the seroprevalence is not known or fixed, the large sample of 14,899 blood specimens would result in a very small range for the 95% confidence interval, if one were to be produced for the total immunity line.

### 4. Discussion

Using the methods proposed, the estimated proportion of the Texas population with antibodies against the SARS-CoV-2 virus, either from natural infection or induced by the vaccine, is nearly 70% as of

July 4, 2021. This means 70% of the population benefit from some degree of protection from reinfection from SARS-CoV-2 and acquiring COVID-19. There are several challenges to further practical or applied interpretation of these data. First, we do not know the relative degree of protection from antibodies from a natural infection compared to antibodies from the vaccine. The titer of antibodies from a full vaccine regimen is higher than a typical natural infection [27], but the diverse epitopes of a natural infection may offer advantages over antibodies targeting only spike protein. In addition, the SARS-CoV-2 mutates producing new strains that will likely influence the degree of protection of circulating antibodies [28]. Second, there is limited data on how long antibodies to the vaccine and to natural infection last [*e.g.*, 29, 30, 31, 32, 33]. Though early findings about the duration of natural and vaccine-generated immunity are promising, it is reasonable to expect the proportion of people with detectable antibodies will decline over time. And finally, protection from an infectious agent is complex, and the concept of seroprevalence and protection does not take into account cell-mediated immunity and physical barriers, such as masks.

In contrast to model-based approaches, the current research will allow researchers and health departments to calculate regional estimates of total immunity in the least biased manner.

Limitations may occur in observational serological surveys; *e.g.*, sample demographics may not be fully representative of the state, which is true for some variables in the current survey. Sampling variability or selection biases may operate within small time windows of a serological survey, and can result in inaccuracies in seroprevalence estimates. It will therefore generally be necessary to smooth estimates using a chosen time window dependent on factors such as the magnitude of the wave of infection and participant accrual rate. Fortunately, we observe that the application of an isotonic restriction to reflect the assumption that seroprevalence should not decrease in a reasonably small time window mostly overcomes the issue of daily or weekly sampling variability. Further, it is necessary to estimate the percentage of people who have both had natural COVID-19 infection and are fully vaccinated in a given time window in order to subtract that proportion from the overall sum. Finally, it is important to age-adjust estimated serological and vaccination rates to the state census so they are commensurate with population demographics. This is especially important since vaccination was rolled out by age group, with older adults first priority in January-March 2021.

To our knowledge, this is the first fully data-driven estimation of total immunity to SARS-CoV-2 in the state of Texas, which is the second largest state in the US with a population of 29.2 million. The method proposed and applied here can be applied to any state or geographic area using vaccine counts, and an estimate of seroprevalence. As the pandemic unfolds and new variants are introduced, immunity will require further investigation and re-estimation.

Table 1: Texas CARES participants' demographics.

	N=10,482	
<b>Age (years)</b>		
Mean (SD)	45.9	(16.1)
Missing	0	
<b>Age (categorical)</b>		
0-15	402	(3.8%)
16-17	124	(1.2%)
18-29	1257	(12.0%)
30-39	1802	(17.2%)
40-49	2331	(22.2%)
50-64	3244	(30.9%)

*Total Immunity to SARS-CoV-2 in Texas*

7

65-74	1100	(10.5%)
75-79	170	(1.6%)
80-84	39	(0.4%)
85+	13	(0.1%)
Missing	0	
<b>Gender</b>		
Female	7262	(69.4%)
Male	3208	(30.6%)
None of these describe me	1	(0.0%)
Missing	11	
<b>Race</b>		
American Indian or Alaskan Native	63	(0.6%)
Asian	470	(4.7%)
Black	382	(3.8%)
Hawaiian or Other Pacific Islander	18	(0.2%)
Multi-racial	192	(1.9%)
White	8974	(88.9%)
Missing	383	
<b>Ethnicity</b>		
Hispanic	2711	(26.8%)
Non-Hispanic	7409	(73.2%)
Missing	362	
<b>BMI (categorical)</b>		
Underweight	126	(1.3%)
Normal	3250	(32.6%)
Overweight	3131	(31.4%)
Obese	3458	(34.7%)
Missing	517	
<b>Geographic Location</b>		
Rural	824	(8.0%)
Urban	9488	(92.0%)
Missing	170	

Table 2: Texas CARES participants' demographics for participants 18 and older.

Adults $\geq$ 18 years	N=9,956
<b>Education</b>	
Some high school or less	121 (1.3%)
High school graduate/GED	802 (8.4%)
Some college, no degree	1505 (15.8%)
Two or four year college level degree	3808 (39.9%)
Advanced professional or academic degree	3301 (34.6%)
Missing	419

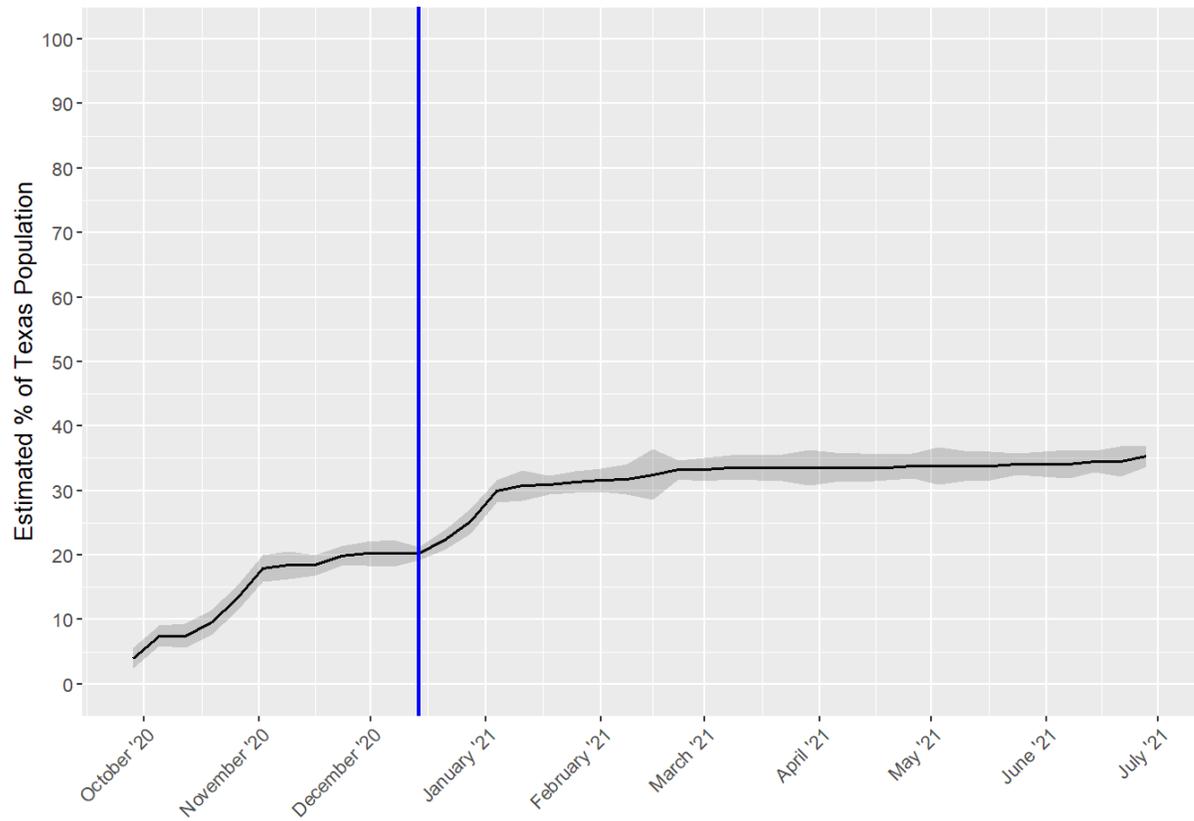
**Employment Status**

Employed-full time	6326 (67.1%)
Employed-part time	902 (9.6%)
Not currently employed/Unemployed	1277 (13.5%)
Other	920 (9.8%)
Missing	531

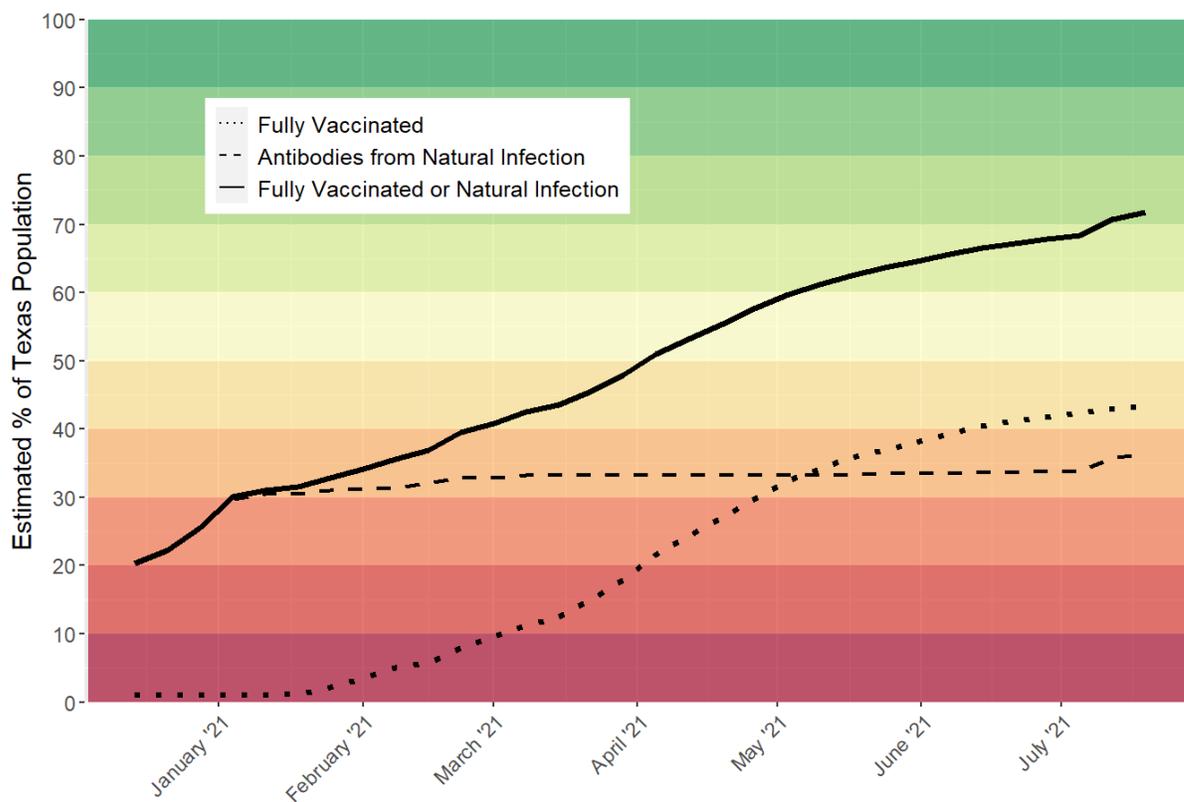
**Employment Industry**

Accommodation and Food Services	170 (2.4%)
Administrative and Support and Waste Management	83 (1.2%)
Agriculture, Forestry, Fishing & Hunting	45 (0.6%)
Arts, Entertainment and Recreation	106 (1.5%)
Central Administrative Office Activity	222 (3.1%)
Construction	146 (2.1%)
Educational Services	2109 (29.7%)
Finance and Insurance	210 (3.0%)
Health Care and Social Assistance	2381 (33.6%)
Information	157 (2.2%)
Management of Companies and Enterprises	98 (1.4%)
Manufacturing	139 (2.0%)
Mining	11 (0.2%)
Other	59 (0.8%)
Professional, Scientific and Technical Services	621 (8.8%)
Real Estate and Rental and Leasing	126 (1.8%)
Retail Trade	197 (2.8%)
Transportation and Warehousing	133 (1.9%)
Utilities	54 (0.8%)
Wholesale Trade	25 (0.4%)
Missing	2864

---



**Figure 1.** Weekly Natural Immunity Texas Cares Roche N-Test. Horizontal axis labels denote the first day of the month.



**Figure 2.** Estimated total immunity in Texas (i.e., weekly percentage of fully vaccinated or naturally occurring antibodies). Horizontal axis labels denote the first day of the month. The estimate as of July 4, 2021 is 69.08%.

## References

- [1] Tom Britton, Frank Ball, and Pieter Trapman. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science*, 369(6505):846–849, 2020. [1](#)
- [2] Sunpeng Wang, Yang Pan, Quanyi Wang, Hongyu Miao, Ashley N Brown, and Libin Rong. Modeling the viral dynamics of SARS-CoV-2 infection. *Mathematical biosciences*, 328:108438, 2020. [1](#)
- [3] Hongyu Miao, Qianmiao Gao, Han Feng, Chengxue Zhong, Pengwei Zhu, Liang Wu, Michael D Swartz, Xi Luo, Stacia M DeSantis, Dejian Lai, et al. Mathematical modeling of business re-opening when facing SARS-CoV-2 pandemic: Protection, cost, and risk. *Frontiers in Applied Mathematics and Statistics*, 6:35, 2020. [1](#)
- [4] Martin Richard Bicher, Claire Rippinger, Günter Richard Schneckenreither, Nadine Weibrecht, Christoph Urach, Melanie Zechmeister, Dominik Brunmeir, Wolfgang Huf, and Niki Popper. Model based estimation of the SARS-CoV-2 immunization level in Austria and consequences for herd immunity effects. *medRxiv*, 2021. [1](#)
- [5] Nazmul Hassan, Shahriar Mahmud, Kaniz Fatema Nipa, and MD Kamrujjaman. Mathematical modeling and Covid-19 forecast in Texas, usa: a prediction model analysis and the probability of disease outbreak. *Disaster Medicine and Public Health Preparedness*, pages 1–27, 2021. [1](#)
- [6] Saad B Omer, Inci Yildirim, and Howard P Forman. Herd immunity and implications for SARS-CoV-2 control. *Jama*, 324(20):2095–2096, 2020. [2](#)
- [7] Roy M Anderson, Carolin Vegvari, James Truscott, and Benjamin S Collyer. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. *The Lancet*, 396(10263):1614–1616, 2020. [2](#)
- [8] Alicia T Widge, Nadine G Roupheal, Lisa A Jackson, Evan J Anderson, Paul C Roberts, Mamodikoe Makhene, James D Chappell, Mark R Denison, Laura J Stevens, Andrea J Pruijssers, et al. Durability of responses after SARS-CoV-2 mRNA-1273 vaccination. *New England Journal of Medicine*, 384(1):80–82, 2021. [2](#)
- [9] Ania Wajnberg, Fatima Amanat, Adolfo Firpo, Deena R Altman, Mark J Bailey, Mayce Mansour, Meagan McMahon, Philip Meade, Damodara Rao Mendu, Kimberly Muellers, et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science*, 370(6521):1227–1230, 2020. [2](#)
- [10] Lauren B Rodda, Jason Netland, Laila Shehata, Kurt B Pruner, Peter A Morawski, Christopher D Thouvenel, Kennedy K Takehara, Julie Eggenberger, Emily A Hemann, Hayley R Waterman, et al. Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell*, 184(1):169–183, 2021. [2](#)
- [11] Valeria De Giorgi, Kamille A West, Amanda N Henning, Leonard N Chen, Michael R Holbrook, Robin Gross, Janie Liang, Elena Postnikova, Joni Trenbeath, Sarah Pogue, Tania Scinto, Harvey J Alter, and Cathy Conry Cantilena. Naturally acquired SARS-CoV-2 immunity persists for up to 11 months following infection. *The Journal of Infectious Diseases*, 06 2021. [2](#)
- [12] CDC. Interim clinical considerations for use of COVID-19 vaccines. Available at <https://www.cdc.gov/vaccines/covid-19/clinical-considerations/covid-19-vaccines-us.html> (2021/06/15). [2](#)
- [13] Bridget M Kuehn. Israel’s real-life evidence that vaccine can prevent severe COVID-19. *JAMA*, 325(16):1603–1603, 2021. [2](#)
- [14] Josh Holder. Tracking coronavirus vaccinations around the world. Available at <http://www.nytimes.com/interactive/2021/world/covid-vaccinations-tracker.html> (2021/06/22). [2](#)

- [15] Mohamed Arafkas, Tanja Khosrawipour, Piotr Kocbach, Kacper Zielinski, Justyna Schubert, Agata Mikolajczyk, Maria Celinska, and Veria Khosrawipour. Current meta-analysis does not support the possibility of COVID-19 reinfections. *Journal of medical virology*, 93(3):1599–1604, 2021. 2
- [16] David S Khoury, Deborah Cromer, Arnold Reynaldi, Timothy E Schlub, Adam K Wheatley, Jennifer A Juno, Kanta Subbarao, Stephen J Kent, James A Triccas, and Miles P Davenport. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nature medicine*, pages 1–7, 2021. 2
- [17] Ricardo da Silva Antunes, Suresh Pallikkuth, Erin Williams, Esther Dawen Yu, Jose Mateus, Lorenzo Quiambao, Eric Wang, Stephen A Rawlings, Daniel Stadlbauer, Kaijun Jiang, et al. Differential t cell reactivity to seasonal coronaviruses and SARS-CoV-2 in community and health care workers. *medRxiv*, 2021. 2
- [18] Deborah Cromer, Jennifer A Juno, David Khoury, Arnold Reynaldi, Adam K Wheatley, Stephen J Kent, and Miles P Davenport. Prospects for durable immune control of SARS-CoV-2 and prevention of reinfection. *Nature Reviews Immunology*, 21(6):395–404, 2021. 2
- [19] Melissa A Valerio-Shewmaker, SM DeSantis, MD Swartz, and et al. Estimation of total immunity to Sars-Cov-2 in Texas. 2021. 2
- [20] Sarah Messiah, Melissa A. Valerio-Shewmaker, Stacia M. DeSantis, Michael D. Swartz, Ashraf Yaseen, Frances A. Brito, Harold W. Kohl, III, Steven H. Kelder, Kimberly Aguillard, Onyinye S. Omega-Njemnobi, Camille Breaux, Jessica A. Ross, Michael O. Gonzalez, Shiming Zhang, Leqing Wu, David Lakey, Jennifer A. Shuford, Stephen J. Pont, and Boerwinkle. Estimated prevalence of SARS-CoV-2 antibodies in the Texas pediatric population. Available at SSRN: <https://ssrn.com/abstract=3868061> or <http://dx.doi.org/10.2139/ssrn.3868061>, 2021. 2
- [21] Texas CARES UTHealth, School of Public Health. Texas CARES Dashboard. <https://sph.uth.edu/projects/texascares/dashboard>. Accessed: 2021-07-15. 2
- [22] Timothy Cardozo and Ronald Veazey. Informed consent disclosure to vaccine trial subjects of risk of COVID-19 vaccines worsening clinical disease. *International journal of clinical practice*, 75(3):e13795, 2021. 3
- [23] Nawal Al Kaabi, Yuntao Zhang, Shengli Xia, Yunkai Yang, Manaf M Al Qahtani, Najiba Abdulrazzaq, Majed Al Nusair, Mohamed Hassany, Jaleela S Jawad, Jehad Abdalla, et al. Effect of 2 inactivated SARS-CoV-2 vaccines on symptomatic COVID-19 infection in adults: A randomized clinical trial. *JAMA*, 2021. 3
- [24] Benjamin J Tepping. *Elementary Sampling Theory*. Taylor & Francis, Englewood Cliffs, New Jersey, 1968. 4, 1
- [25] Paul S Levy and Stanley Lemeshow. *Sampling of Populations*. Wiley, Hoboken, 2008. 4, 1
- [26] School of Public Health UTHealth. Texas COVID-19 dashboard. <https://sph.uth.edu/dept/bads/covid19-dashboard>. Accessed: 2021-07-15. 5
- [27] Pinja Jalkanen, Pekka Kolehmainen, Hanni K Häkkinen, and et al. Covid-19 mRNA vaccine induced antibody responses against three SARS-CoV-2 variants. *Nature Communications*, 12(3991), 2021. 6
- [28] Emanuele Amodio, Giuseppina Capra, Alessandra Casuccio, Simona De Grazia, Dario Genovese, Stefano Pizzo, Giuseppe Calamusa, Donatella Ferraro, Giovanni Maurizio Giammanco, Francesco Vitale, et al. Antibodies responses to SARS-CoV-2 in a large cohort of vaccinated subjects and seropositive patients. *Vaccines*, 9(7):714, 2021. 6
- [29] Kristen W Cohen, Susanne L Linderman, Zoe Moodie, Julie Czartoski, Lilin Lai, Grace Mantus, Carson Norwood, Lindsay E Nyhoff, Venkata Viswanadh Edara, Katharine Floyd, et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Medrxiv*, 2021. 6

- [30] Jennifer M Dan, Jose Mateus, Yu Kato, Kathryn M Hastie, Esther Dawen Yu, Caterina E Faliti, Alba Grifoni, Sydney I Ramirez, Sonya Haupt, April Frazier, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science*, 371(6529), 2021. [6](#)
- [31] Jackson S Turner, Wooseob Kim, Elizaveta Kalaidina, Charles W Goss, Adriana M Rauseo, Aaron J Schmitz, Lena Hansen, Alem Haile, Michael K Klebert, Iskra Pusic, et al. SARS-CoV-2 infection induces long-lived bone marrow plasma cells in humans. *Nature*, pages 1–5, 2021. [6](#)
- [32] Zijun Wang, Frauke Muecksch, Dennis Schaefer-Babajew, Shlomo Finkin, Charlotte Viant, Christian Gaebler, Hans-Heinrich Hoffmann, Christopher O Barnes, Melissa Cipolla, Victor Ramos, et al. Naturally enhanced neutralizing breadth to SARS-CoV-2 after one year. *bioRxiv*, 2021. [6](#)
- [33] World Health Organization. COVID-19 natural immunity: scientific brief, 10 May 2021. Technical report, World Health Organization, 2021. [6](#)

## Supplementary Material to, “Estimation of Total Immunity to SARS-CoV-2 in Texas”

### S.1. Order Restricted (Isotonic) Estimation of Probabilities

The algorithm below is retrieved from Algorithm 3 in <https://core.ac.uk/download/pdf/33107977.pdf>. The maximum likelihood estimate of  $\pi = (\pi_1, \pi_2, \dots, \pi_H)$  under the restriction of  $\pi_1 \leq \pi_2 \leq \dots \leq \pi_H$ , is calculated in the following way: Let  $n_h$  the number of observations in group  $h$ ,

Step 1: Do  $\tilde{\pi}_h$  equal to the sample proportion in group  $h$ .

Step 2: While not  $\pi_h \leq \pi_{h+1}$ , for  $h = 1, \dots, H - 1$ , do

For  $h = 1, \dots, H$

If  $\tilde{\pi}_h \not\leq \tilde{\pi}_{h+1}$  do

$$\tilde{\pi}_h = \frac{n_h}{n_h + n_{h+1}} \tilde{\pi}_h + \frac{n_{h+1}}{n_h + n_{h+1}} \tilde{\pi}_{h+1}, \quad \text{and} \quad \tilde{\pi}_{h+1} = \tilde{\pi}_h$$

### S.2. Confidence Interval for the Seroprevalence

The construction of the confidence interval for  $\widetilde{SPR}_t$  is based on the confidence interval for a proportion under a stratified sampling design [e.g., 24, 25]. Recall, the weight  $w_h = N_h/N$  and  $N_h$  denote the proportion of and the number of individuals in the population in the age group  $h$ , respectively, and  $N$  the population total. An estimate of the sampling variance of the sample proportion  $\eta_{ht}$  is

$$\widetilde{\text{Var}}(\hat{\eta}_{ht}) = \left[ \frac{N_h - n_{ht}}{N_h - 1} \right] \frac{\hat{\eta}_{ht}(1 - \hat{\eta}_{ht})}{n_{ht}} \approx \frac{\hat{\eta}_{ht}(1 - \hat{\eta}_{ht})}{n_{ht}}$$

where  $n_{ht}$  is the number of participants in age group  $h$  (at week  $t$ ). The approximation in the above equation is valid since  $N_h \gg n_{ht}$ . We base the confidence interval on this equation plugging in  $\tilde{\eta}_{ht}$  instead of  $\hat{\eta}_{ht}$

$$\widetilde{\text{Var}}(\tilde{\eta}_{ht}) \approx \frac{\tilde{\eta}_{ht}(1 - \tilde{\eta}_{ht})}{n_{ht}}$$

Then the sampling variance of  $\widetilde{SPR}_t$  is estimated with

$$\widetilde{\text{Var}}(\widetilde{SPR}_t) = \sum_h w_h^2 \widetilde{\text{Var}}(\tilde{\eta}_{ht}) \approx \sum_h w_h^2 \frac{\tilde{\eta}_{ht}(1 - \tilde{\eta}_{ht})}{n_{ht}}$$

The asymptotic  $1 - \alpha$  confidence interval for  $SPR_t$  is then

$$\widetilde{SPR}_t \pm z_{1-\alpha/2} \sqrt{\widetilde{\text{Var}}(\widetilde{SPR}_t)}$$

where  $z_\alpha$  is the percentile  $\alpha$  of the standard normal distribution, so when  $1 - \alpha = 0.95$ ,  $z_{1-\alpha/2} = 1.96$ .