

1 **SARS-CoV-2 antibody magnitude and detectability are driven by disease severity,**
2 **timing, and assay**

3

4 Michael J. Peluso^{1*}, Saki Takahashi^{1*}, Jill Hakim^{1*}, J. Daniel Kelly², Leonel Torres^{1,3},
5 Nikita S. Iyer³, Keirstinne Turcios¹, Owen Janson¹, Sadie E. Munter³, Cassandra
6 Thanh³, Christopher C. Nixon³, Rebecca Hoh¹, Viva Tai¹, Emily A. Fehrman¹, Yanel
7 Hernandez¹, Matthew A. Spinelli¹, Monica Gandhi¹, Mary-Ann Palafox⁴, Ana Vallari⁴,
8 Mary A. Rodgers⁴, John Prostko⁴, John Hackett Jr.⁴, Lan Trinh⁵, Terri Wrin⁵, Christos J.
9 Petropoulos⁵, Charles Y. Chiu^{6,7,8}, Philip J. Norris⁹, Clara DiGermanio⁹, Mars Stone⁹,
10 Michael P. Busch^{6,9}, Susanna K. Elledge¹⁰, Xin X. Zhou¹⁰, James A. Wells^{10,11}, Albert
11 Shu⁶, Theodore W. Kurtz⁶, John E. Pak¹², Wesley Wu¹², Peter D. Burbelo¹³, Jeffrey I.
12 Cohen¹⁴, Rachel L. Rutishauser³, Jeffrey N. Martin², Steven G. Deeks¹, Timothy J.
13 Henrich^{3^}, Isabel Rodriguez-Barraquer^{1^}, Bryan Greenhouse^{1^}

14

15 1 Division of HIV, Infectious Diseases, and Global Medicine, University of California,
16 San Francisco, USA

17 2 Department of Epidemiology and Biostatistics, University of California, San Francisco,
18 USA

19 3 Division of Experimental Medicine, University of California, San Francisco, USA

20 4 Abbott Laboratories, Abbott Park, IL, USA

21 5 Monogram Biosciences, Inc., South San Francisco, CA, USA

22 6 Department of Laboratory Medicine, University of California, San Francisco, USA

23 7 Division of Infectious Diseases, University of California, San Francisco, USA

- 1 8 UCSF-Abbott Viral Diagnostics and Discovery Center, San Francisco, USA
2 9 Vitalant Research Institute, San Francisco, CA, USA
3 10 Department of Pharmaceutical Chemistry, University of California, San Francisco,
4 USA
5 11 Department of Cellular & Molecular Pharmacology, University of California, San
6 Francisco, USA
7 12 Chan Zuckerberg Biohub, San Francisco, USA
8 13 National Institute of Dental Research, National Institutes of Health, Bethesda, USA
9 14 Laboratory of Infectious Diseases, National Institute of Allergy and Infectious
10 Diseases, National Institutes of Health, Bethesda, USA

11

12 *,^ These authors contributed equally

13

14 Keywords: SARS-CoV-2, COVID-19, antibody, serosurveillance

15

16

17

18

19

20

21

1 **ABSTRACT**

2 Serosurveillance studies are critical for estimating SARS-CoV-2 transmission and
3 immunity, but interpretation of results is currently limited by poorly defined variability in
4 the performance of antibody assays to detect seroreactivity over time in individuals with
5 different clinical presentations. We measured longitudinal antibody responses to SARS-
6 CoV-2 in plasma samples from a diverse cohort of 128 individuals over 160 days using
7 14 binding and neutralization assays. For all assays, we found a consistent and strong
8 effect of disease severity on antibody magnitude, with fever, cough, hospitalization, and
9 oxygen requirement explaining much of this variation. We found that binding assays
10 measuring responses to spike protein had consistently higher correlation with
11 neutralization than those measuring responses to nucleocapsid, regardless of assay
12 format and sample timing. However, assays varied substantially with respect to
13 sensitivity during early convalescence and in time to seroreversion. Variations in
14 sensitivity and durability were particularly dramatic for individuals with mild infection,
15 who had consistently lower antibody titers and represent the majority of the infected
16 population, with sensitivities often differing substantially from reported test
17 characteristics (e.g., amongst commercial assays, sensitivity at 6 months ranged from
18 33% for ARCHITECT IgG to 98% for VITROS Total Ig). Thus, the ability to detect
19 previous infection by SARS-CoV-2 is highly dependent on the severity of the initial
20 infection, timing relative to infection, and the assay used. These findings have important
21 implications for the design and interpretation of SARS-CoV-2 serosurveillance studies.

22

1 INTRODUCTION

2 Despite advances in severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2)
3 prevention and treatment, the novel coronavirus continues to infect individuals at an
4 unprecedented rate. Because vaccination programs remain limited in scope, millions of
5 individuals worldwide continue to rely upon natural post-infection immunity for protection
6 from reinfection. Serosurveillance studies measuring the prevalence of antibodies to
7 SARS-CoV-2 have been and will continue to be a key means for estimating
8 transmission over time and extrapolating potential levels of immunity in populations,
9 though precise correlates of protection have yet to be established. However, limited
10 available data on the sensitivity of antibody assays to detect prior infection - particularly
11 in appropriately representative populations and over time - make it difficult to accurately
12 interpret results from these studies.¹ For these reasons, longitudinal characterization of
13 antibody responses following SARS-CoV-2 infection with a range of clinical
14 presentations is an important research gap and will be critical to interpreting
15 seroepidemiological data and informing public health responses to the pandemic.

16 Infection with SARS-CoV-2 is associated with substantial variability in disease
17 presentation, with severity ranging from asymptomatic infection to the need for high-
18 level oxygen support and mechanical ventilation.^{2,3} There appear to be important
19 relationships between the severity of illness and the magnitude and durability of the
20 antibody response,⁴⁻¹² but limited data are available evaluating the contributions of
21 demographic factors and clinical features. Numerous platforms are available for the
22 detection of antibody responses to SARS-CoV-2, which rely on different viral antigens
23 and utilize different assay methods, and there is no guarantee that they will provide

1 comparable data. With a few notable exceptions,^{6,7} most studies to date have produced
2 antibody data from a single or limited number of platforms to evaluate antibody
3 responses following infection.^{8–10,13–15} Comparisons across platforms and assay format
4 differences (e.g., direct vs. indirect detection), including the correlation between binding
5 assays and neutralization capacity, have thus far have been limited.^{4,7,16}

6 Here, we characterize the antibody responses to SARS-CoV-2 among a diverse cohort
7 of individuals with documented infection, with a focus on investigating (1) the
8 determinants of the magnitude and durability of humoral immune responses across a
9 spectrum of disease severity, (2) the relationship between antibody responses across a
10 wide variety of binding assay platforms (13 total) and their correlation with neutralization
11 capacity, and (3) the implications of individual, temporal, and assay variability for
12 serosurveillance. Our findings provide insight into the interpretation of antibody test
13 results, and have important implications for our understanding of humoral immunity to
14 natural infection as well as for serosurveillance.

15 **METHODS**

16 *Study cohort*

17 Participants were volunteers in the University of California, San Francisco-based Long-
18 term Impact of Infection with Novel Coronavirus (LIINC) natural history study
19 (NCT04362150). All volunteers signed informed consents for the study. LIINC is an
20 observational cohort that enrolls individuals with SARS-CoV-2 infection documented by
21 clinical nucleic acid amplification testing who have recovered from the acute phase of
22 infection. Volunteers are recruited by clinician- or self-referral. They are eligible to enroll

1 between 14 and 90 days after onset of COVID-19 symptoms and are offered monthly
2 visits until 4 months after illness onset; they are then seen every 4 months thereafter.

3 Clinical data from the initial LIINC study visit was used for this analysis. At this visit,
4 LIINC participants underwent a detailed clinical interview conducted by a study
5 physician or research coordinator using a standardized data collection instrument.

6 Demographic data collected included age, sex, gender, race, ethnicity, education level,
7 income level, and housing status. Data related to SARS-CoV-2 infection included the
8 date and circumstances of diagnosis, illness, and treatment history. Each participant
9 was asked to estimate the date of symptom onset in relation to the timing of their first
10 SARS-CoV-2 nucleic acid amplification test result. Participants were questioned
11 regarding the presence, duration (in days), and current status of a list of COVID-19
12 symptoms and additional somatic symptoms derived from the Patient Health
13 Questionnaire ¹⁷, as well as measures of quality of life derived from the EQ-5D-5L
14 Instrument ¹⁸. We determined from medical records whether each individual was
15 hospitalized (defined as spending >24 hours in the emergency department or hospital)
16 and whether they required supplemental oxygen, admission to an intensive care unit
17 (ICU), or mechanical ventilation. Past medical history was ascertained and concomitant
18 medications recorded.

19 At each visit, blood was collected by venipuncture. Serum and plasma were isolated via
20 centrifugation of non-anticoagulated and heparinized blood, respectively, and stored at -
21 80C. For the current analyses, we included 128 participants who were enrolled between
22 April and July, 2020 and who had at least one measurement on a binding assay or
23 neutralization platform.

1 *Antibody assays*

2 Supplementary Table 1 describes each of the assays performed as part of this analysis,
3 including their sensitivity and specificity (as reported by the manufacturers for
4 commercial assays, or by validation testing for non-commercial research assays).
5 Commercially available platforms included the Abbott ARCHITECT (IgG), Ortho Clinical
6 Diagnostics VITROS (IgG and total Ig), DiaSorin LIAISON (IgG), Roche Elecsys (total
7 Ig), and Monogram PhenoSense (neutralizing antibodies) assays according to
8 manufacturer specifications. Individuals living with HIV infection were excluded from
9 analyses involving the Monogram PhenoSense assay, which utilizes an HIV backbone
10 for the pseudovirus. Non-commercial research use assays included the Luciferase
11 Immunoprecipitation Systems (LIPS) assay (total Ig) targeting the nucleocapsid (N)
12 protein and the receptor binding domain (RBD) of the spike (S) protein performed in the
13 Burbelo laboratory (additional raw data on responses to full S protein, highly correlated
14 with RBD responses, are included in Supplementary Tables 2 and 3),¹⁹ the split
15 luciferase assay (total Ig) targeting N and S performed in the Wells laboratory,²⁰ and
16 the Luminex assay (IgG) targeting N (one full-length and one fragment), S, and RBD
17 performed in the Greenhouse laboratory.

18 For the research use Luminex assay, we used a published protocol with modifications.²¹
19 Plasma samples were diluted to 1:100 in blocking buffer A (1xPBS, 0.05% Tween, 0.5%
20 bovine serum albumin, 0.02% sodium azide). Antigens were produced using previously
21 described constructs.^{22,23} Antigen concentrations used for COOH-bead coupling were
22 as follows: S, 4 ug/mL; RBD, 2 ug/mL; and N, 3 ug/mL. Concentration values were
23 calculated from the Luminex median fluorescent intensity (MFI) using a plate-specific

1 standard curve consisting of serial dilutions of a pool of positive control samples. Any
2 samples with MFIs above the linear range of the standard curve were serially diluted
3 and rerun until values fell within range to obtain a relative concentration. A cutoff for
4 positivity was established for each antigen above the maximum concentration value
5 observed across 114 pre-pandemic SARS-CoV-2 negative control samples tested on
6 the platform.

7

8 *Statistical Analyses*

9 Comparing individuals across assays and estimating time to seroreversion: For each
10 assay, we fit a linear mixed effects model that included a patient-specific random
11 intercept. Given the longitudinal nature of our data set, we fit mixed effects models to
12 explicitly account for the repeated measurement of individuals over time. We log-
13 transformed the response variable for a subset of the assays based on assessment of
14 their correlations with log-transformed neutralization titers (Supplementary Table 1). For
15 observation h for individual i , we modeled their antibody response Y_{hi} on each assay as
16 follows:

$$17 \quad Y_{hi} = \beta_s + \lambda Time_{hi} + u_{0i} + e_{hi} \quad (\text{Equation 1})$$

18 In Equation 1, β_s represents the overall mean for severity class s , where s was
19 dichotomized into whether an individual was hospitalized or not-hospitalized; λ
20 represents the fixed effect of $Time_{hi}$, where $Time_{hi}$ is data on the time since symptom
21 onset (if symptomatic) or since positive PCR test (if asymptomatic). In addition, u_{0i}
22 represents an individual-level random effect that is normally distributed with a mean of 0

1 and a standard deviation of τ , and e_{hi} represents the residual error that is normally
2 distributed with a mean of 0 and a standard deviation of σ . We also considered models
3 that included additional fixed effects for covariates such as age, ethnicity, sex, and HIV
4 status (Supplementary Table 4). We did not find consistent differences in the slope of
5 the antibody responses (λ) by hospitalization status across the majority of the assays
6 used here; therefore we used a single slope for each assay throughout (Supplementary
7 Table 5).

8 Since the timing of the baseline visit was variable between individuals, in order to
9 directly compare the magnitude of measured responses for individuals on each assay,
10 we used the mixed-effects model to estimate the antibody response that each person
11 would have at 21 days post symptom onset (random intercept).²⁴ We also used the
12 model estimates to calculate the mean time to seroreversion T for severity class s on
13 each assay, given the cutoff value for positivity (Supplementary Table 1), as follows:

$$14 \quad T_s = (cutoff - \beta_s) / \lambda \quad (\text{Equation 2})$$

15 We performed bootstrapping to obtain 95% confidence intervals of T_s for each of the 14
16 assays. We used the time to seroreversion as the outcome here rather than alternative
17 quantities such as the half-life, as the serologic responses obtained here did not all
18 necessarily represent direct measurements of antibody titers. These models were fit
19 using the *lme4* package using the R statistical software (<https://www.R-project.org/>).

20 Random forests modeling of demographic/clinical predictors and antibody responses:

21 For each assay, we used random forests to model antibody responses based on 50
22 demographic and clinical predictors (Supplementary Table 6). We dichotomized the

1 antibody response for each individual on each assay based on whether or not their
2 estimated random intercept was in the top half or the bottom half of all fitted random
3 intercepts on that assay. We first fit models to these dichotomized antibody responses
4 using all available predictors; subsequently, we fit models to these dichotomized
5 antibody responses on a down-selected set of predictors selected based on variable
6 importance (i.e., mean decrease in accuracy). We quantified prediction accuracy using
7 the out-of-bag error rate and the area under the curve (AUC). These models were built
8 using the *randomForest* package using the R statistical software (version 3.5.3).

9 Estimating time-varying assay sensitivity: For each assay, we fit an extension of the
10 linear mixed effects models described in Equation 1 above in a Bayesian hierarchical
11 modeling framework, where we we allowed the standard deviation of the random
12 intercept (τ) to be severity-specific (now referred to as τ_s). For parsimony, we assumed
13 the slopes, λ , to again be shared across all individuals as described above. To estimate
14 changes in assay sensitivity over time, we simulated population distributions of \hat{Y} by
15 iteratively sampling values from each of the posterior distributions of β_s , λ , τ_s , and σ . For
16 each sampled value of τ_s and σ , we then sampled draws of u_{0i} and e_{hi} and determined
17 the assay sensitivity at various time points (0, 60, 120, and 180 days after
18 seroconversion) as the proportion of overall draws where \hat{Y} was above the cutoff value
19 for positivity. We ran 4 Markov chain Monte Carlo (MCMC) chains of length 2,000 each
20 using the Stan programming language (<https://mc-stan.org/>), and assessed
21 convergence using the Gelman-Rubin Rhat statistic. We used uninformative priors for
22 all parameters and hyper-parameters.

1 Data and code to reproduce all analyses are available at:

2 <https://github.com/EPPICenter/liinc-Ab-dynamics>.

3 *Ethical considerations*

4 All participants signed a written informed consent form. The study was approved by the
5 University of California, San Francisco Institutional Review Board.

6 **RESULTS**

7 *Participant demographics and characteristics*

8 As shown in Table 1, the cohort of 128 participants had an average age of 48 years
9 (range 19-85 years), was relatively balanced in terms of sex (45% female at birth), and
10 26% of participants self-identified as being of Latinx ethnicity, a group which has been
11 identified to be at-risk for COVID-19. Common medical comorbidities were hypertension
12 (23%), lung disease (16%), and diabetes (13%). Notably, 18 individuals (14%) were
13 living with HIV infection; 17 of 18 were on antiretroviral therapy. 121 individuals (95%)
14 reported symptoms during their COVID-19 illness, but only 32 (25%) required
15 hospitalization. Among those who had been hospitalized, 84% required supplemental
16 oxygen, 42% required ICU admission, and 13% required mechanical ventilation. A
17 minority of individuals (n=7, 5%) were asymptomatic.

18 The baseline visit for participants occurred at a median of 63 days (range: 22 - 157)
19 after symptom onset (Supplementary Figure 1). Participants contributed a median of 2
20 samples each (range: 1 - 4) and were followed-up for a median of 110 days after
21 symptom onset (range: 22 - 157). A total of 267 samples from the 128 enrolled

1 participants were tested using at least one of the 14 assays evaluated; 171 samples
2 from 88 individuals were tested using all 14 assays (Supplementary Table 1,
3 Supplementary Figure 2). Individuals with HIV were excluded from the Neut-Monogram
4 as noted above.

5 *Substantial heterogeneity in antibody responses across individuals and assays*

6 We observed substantial heterogeneity in measured antibody responses in individuals
7 at baseline and throughout follow-up across all assays (Figure 1, raw data available in
8 Supplementary Tables 2 and 3). We observed variable trajectories of antibody
9 responses between assays, with some (N-Abbott, N-Split Luc, S-Ortho IgG, and Neut-
10 Monogram) showing a clear decrease over time, other assays (S-Ortho Ig and N-
11 Roche) showing a clear increase, and the remainder with more stable values (Figure 1,
12 Supplementary Table 7). When comparing antibody levels between individuals,
13 responses were very heterogeneous, with some individuals mounting strong responses
14 for all assays and others with weak responses even at the initial visit (below the
15 positivity cutoff for some assays).

16 *Strong correlation between binding and neutralization assays*

17 We observed high levels of correlation between estimated antibody levels at 21 days
18 post symptom onset (random intercept) for all assays, with Spearman correlations
19 ranging between 0.55 and 0.96 (Figure 2A, Supplementary Figure 3). Rank correlations
20 were consistently higher between binding assays using the same antigenic target
21 (S/RBD vs. N) than between those using different targets, despite the variety in
22 platforms used and the measurement of responses to both targets on some platforms

1 (LIPS, Luminex, split luciferase). Titers of neutralizing antibodies correlated well with all
2 binding assays (range: 0.60 to 0.88) and correlated most highly with responses to the S
3 protein (range: 0.76 to 0.88), as might be expected given the expression of S protein on
4 the pseudovirus used in the neutralization assay (Figure 2B, Supplementary Figure 3).
5 We found no substantive differences in correlations between binding and neutralization
6 assays at timepoints before vs. after 90 days, suggesting these relationships did not
7 appreciably change over the duration of observed follow-up (Supplementary Table 8).

8 *Disease severity is strongly associated with the magnitude of antibody responses*

9 Baseline antibody responses for each study participant showed remarkably consistent
10 patterns across all assays when stratified by severity class, with asymptomatic
11 individuals having the lowest responses, hospitalized individuals having the highest, and
12 symptomatic but not hospitalized individuals having intermediate responses (Figure 3).
13 While the number of asymptomatic individuals was small, responses were significantly
14 lower in these individuals than those who were symptomatic but not hospitalized for
15 multiple assays; hospitalized individuals had significantly higher responses than both
16 other groups for all assays with the notable exception of the neutralization assay
17 (Supplementary Table 9). Despite these consistent patterns, there was still substantial
18 variation in the magnitude of responses between participants within each severity
19 category. Notably, age, sex, HIV status, and Latinx ethnicity showed little association
20 with antibody responses after adjusting the analysis for hospitalization (Supplementary
21 Table 4).

22 *Need for hospitalization, cough, and fever are key predictors of antibody responses*

1 We next examined which of 50 individual demographic and clinical variables were the
2 strongest predictors of the magnitude of the antibody response (top vs. bottom half of
3 responders for each assay, Supplementary Figure 4) using a random forests algorithm.
4 Among the entire cohort (n=128), the presence and duration of cough and fever, and
5 need for hospitalization and supplemental oxygen during the initial illness, were the
6 most important predictors of the antibody response (Figure 4A). The ranks of their
7 importance varied subtly but were largely consistent across the 14 assays evaluated,
8 and random forests models including only these 6 variables were able to predict high
9 versus low magnitude of response on each assay with reasonably high accuracy (AUCs
10 ranging from 0.74 to 0.86, Supplementary Table 10). Among those individuals who were
11 not hospitalized (n=96), presence and duration of fever and cough remained the most
12 important predictors of a high antibody response (Figure 4B). These four variables
13 alone were predictive of high vs. low responses within this subset of individuals with
14 modest accuracy (all except RBD-LIPS with AUCs above 0.6).

15 *Time to seroreversion varies considerably across platforms and by infection severity*

16 Using the mixed effect model described above, we estimated the expected time to
17 seroreversion (when antibodies would become undetectable for an average individual)
18 for each platform, assuming that antibody responses changed linearly over time. We
19 estimated separate times to seroreversion for hospitalized and non-hospitalized
20 individuals, since hospitalization status was a strong predictor of baseline antibody
21 status. Estimated time to seroreversion was substantially shorter for non-hospitalized
22 vs. hospitalized individuals for all assays, consistent with the lower initial antibody titers
23 in those individuals. For those assays where antibody levels decreased over time, we

1 also observed marked variation in times to seroreversion between assays, ranging from
2 96 days for N(frag)-Lum to 925 days for S-DiaSorin; the estimated time to seroreversion
3 is infinity for RBD-LIPS, S-Ortho Ig, and N-Roche, which exhibited increasing mean
4 antibody responses over time in this dataset (Figure 5A, Supplementary Table 7).

5 *Sensitivity of assays to detect prior infection varies as a function of time and infection*
6 *severity*

7 We next assessed how sensitivity of each platform varied as a function of time and
8 disease severity. Because the sample size of asymptomatic individuals was small, they
9 were grouped with other non-hospitalized individuals for this analysis (Supplementary
10 Figure 5). We found considerable heterogeneity in sensitivity between assays and as a
11 function of illness severity and time since infection. Across all fourteen assays,
12 sensitivity at each time point was higher in the hospitalized subset of the cohort than in
13 the non-hospitalized subset, the latter group representing the majority of infections in
14 the general population (Figure 5B, Supplementary Table 11). The magnitude of this
15 difference varied between assays and also over time, and was often considerable
16 (Supplementary Figure 6). Estimated sensitivity declined over time for eleven of the
17 fourteen assays, but increased for RBD-LIPS, S-Ortho Ig, and N-Roche, consistent with
18 the observed increase in magnitude of response over time for these assays. Overall,
19 RBD-LIPS showed the most consistently high sensitivity over time and the smallest
20 difference between hospitalized and non-hospitalized individuals, ranging from 88%
21 (95% credible interval: 81% - 94%) at month 0 to 99% (95% credible interval: 96% -
22 100%) at month 6 in non-hospitalized individuals. Of the remaining research-use
23 assays, N-LIPS, RBD- and N-Split Luc, and RBD- and S-Lum showed a similar pattern,

1 with high sensitivity initially followed by a decline in sensitivity amongst non-hospitalized
2 individuals over time; N(full)-Lum and N(frag)-Lum had consistently poor sensitivity for
3 non-hospitalized individuals. All commercial assays performed similarly well during early
4 convalescence, albeit with lower sensitivity for non-hospitalized individuals for S-
5 DiaSorin and S-Ortho Ig. N-Abbott showed the greatest decline in sensitivity with time
6 and the greatest difference between hospitalized and non-hospitalized participants, with
7 sensitivity varying from 100% (95% credible interval: 100% - 100%) in hospitalized
8 individuals soon after infection to 33% (95% credible interval: 24% - 42%) in non-
9 hospitalized individuals at 6 months. Of note, neutralization titers remained detectable
10 for nearly all hospitalized individuals up to six months, but were estimated to become
11 undetectable on this assay for nearly half of non-hospitalized individuals by this time
12 point.

13 *Varying assay sensitivity affects interpretation of individual antibody test results*

14 Lastly, negative predictive values were calculated for the commercial assays to illustrate
15 potential effects of these changes in sensitivity if the assays were used to ascertain
16 prior infection in individuals (assuming values of specificity as reported by the
17 manufacturers). As expected, negative predictive values decreased with increasing
18 prevalence and, except for S-Ortho Ig and N-Roche, time (Supplementary Figure 7).

19 **DISCUSSION**

20 This study documents the large heterogeneity in longitudinal antibody responses to
21 SARS-CoV-2 across a large number of commercial and research assays in a diverse
22 cohort of individuals. Measured responses in all binding assays correlated well with

1 each other and, particularly for those measuring responses to spike protein, with
2 pseudovirus neutralization. For all assays, we found a consistent, strong, and dose-
3 dependent effect of disease severity on antibody magnitude. Despite these similarities,
4 assays performed quite differently in terms of sensitivity to detect prior infection and in
5 the durability of measured responses, leading to large discrepancies in sensitivity
6 between assays in the months following infection. Thus, the ability to detect previous
7 infection by SARS-CoV-2 using an antibody test is highly dependent on the severity of
8 the initial infection, when the sample is obtained relative to infection, and the assay
9 used.

10 Prior work has shown that antibody responses in individuals with symptomatic COVID-
11 19 have in some cases been associated with disease severity.⁴⁻¹² We observed
12 significant variability in antibody responses between study participants which was
13 largely explained by the self-reported symptom constellation and the severity of the
14 acute illness. A few simple variables consistently predicted the magnitude of the
15 antibody responses across multiple assay platforms and antigen targets; these
16 symptoms (e.g., fever, cough) are similar to those recently described in a population-
17 based Icelandic cohort.⁶ Importantly, in contrast to that cohort, characteristics like age
18 and sex were not predictive of these responses, after accounting for disease severity.

19 We also observed substantial heterogeneity between assays in terms of overall
20 sensitivity, particularly over time. These findings build upon prior work showing
21 differences in sensitivity during early convalescence.^{7,10} This finding may also provide
22 insight into apparent discrepancies in previous studies that have reported different
23 durabilities of antibody responses.^{25,26} Differences in performance were particularly

1 pronounced amongst non-hospitalized individuals, who have lower antibody responses
2 and comprise the majority of those infected with SARS-CoV-2. Notably, antibody
3 responses in such individuals are expected to be reliably detectable over 6 months in
4 only 2 of the commercial binding assays tested – S-Ortho Ig and N-Roche – which were
5 the only to employ direct detection format of antibodies (to different viral targets).
6 Interestingly, these 2 assays, along with a research-use assay employing direct
7 detection (RBD-LIPS), were the only ones to demonstrate increasing rather than
8 decreasing antibody signal over time, possibly due to continued affinity maturation of
9 the antibody response²⁷ playing a larger role in detection with this format. This is in
10 contrast to the indirect format assays and neutralization assay, all of which
11 demonstrated waning over time. Another possible explanation for discrepancies
12 between assays is that only assays with the highest signal to noise ratio are able to
13 consistently detect antibodies above background in those with the lowest titers, i.e.
14 those multiple months out from mild infections. This variation in sensitivity is relevant
15 for several reasons. First, it provides further evidence that use-cases need to be
16 considered when evaluating performance of antibody tests. While all evaluated assays
17 had near perfect sensitivity for detecting antibody responses among hospitalized
18 individuals and therefore could be useful as an indirect diagnostic tool in that setting,
19 their sensitivity to detect responses in the general population, where most infections are
20 mild, is much lower and quite variable.¹ Second, it implies that using assays with
21 decreasing sensitivity over time for population seroprevalence studies will
22 underestimate the true proportion of previously infected individuals, and that this
23 underestimation will be more substantial as the amount of time that has passed since

1 infection increases. Third, it shows dramatic differences between the sensitivity reported
2 by test manufacturers, often limited to validation sample sets from hospitalized and/or
3 recently infected individuals that were readily available early in the pandemic, and the
4 expected sensitivity in the general population. Because of this, our study provides
5 information that could be useful for assay selection when planning future serosurveys,
6 and could help correct the interpretation of large-scale population-based seroprevalence
7 studies that have used some of these assays. Finally, individual patients or providers
8 utilizing these assays to assess the presence or absence of prior infection and/or
9 immune status should take these considerations into account, given the poor negative
10 predictive value of some tests.

11 This analysis has several notable strengths, including the utilization of a broad array of
12 antibody platforms at multiple time points in a diverse cohort of individuals with varying
13 degrees of illness severity and rich clinical phenotyping. However, there are several
14 limitations. The cohort was not population-based and therefore not truly representative
15 of all individuals with SARS-CoV-2 infection. Despite this, we endeavored to recruit a
16 cohort that spanned the spectrum of SARS-CoV-2 infection. Second, this analysis
17 included only a small number of asymptomatic individuals, which are likely to differ from
18 symptomatic patients in terms of initial and possibly longitudinal responses based on
19 our limited data and prior results;¹¹ additional asymptomatic individuals are being
20 recruited for future analyses. Third, the current analysis is limited to samples obtained
21 up to four months after infection. Data from longer follow-up times will be very useful in
22 order to estimate the longer term kinetics of antibody responses in all platforms with
23 more certainty. The assumption of linearity, while empirically appropriate on the time-

1 scales of data that we have here, may not be accurate for extrapolation into the distant
2 future as antibody responses often follow more complex dynamics of boosting and
3 waning over time.²⁸ Finally, it is important to recognize that assays optimally suited for
4 serosurveillance may not be equally suitable for other use-cases, such as identifying
5 recent infection, detecting reinfection, determining protective capacity, or determining
6 potency of COVID-19 convalescent plasma.²⁹ Evaluating the performance of assays for
7 each of these use-cases will require different study designs and sample sets.

8 As SARS-CoV-2 vaccination becomes a reality, many serosurveillance efforts will need
9 to increasingly rely upon assays that can distinguish vaccination from natural infection,
10 especially when it is not possible to obtain additional information on the vaccination
11 status of participants. Currently distributed vaccines in the US are based on S protein,
12 thus responses to S-based assays will likely reflect a combination of natural infection
13 and vaccination, whereas assays based on N will reflect only natural infection. While
14 several of the N-based assays we evaluated performed well, one of the two
15 commercially available N-based assays demonstrated substantial waning of sensitivity
16 over time which will affect estimates of the seroprevalence of natural infection, but could
17 be a more useful indicator of recent infection along with other potential markers such as
18 IgM. Further studies will be needed in order to characterize the performance of these
19 and other assays for serosurveillance in the presence of vaccine induced immunity.

20 *Conclusions*

21 In this study, we demonstrated substantial differences in the detectability of antibody
22 responses to SARS-CoV-2 related to illness severity, time since infection, and assay

1 platform. These results will be important in choosing and interpreting serologic assays
2 for evaluating infection and immunity in population surveillance studies.

3 **ACKNOWLEDGEMENTS**

4 We are grateful to the LIINC study participants and to the clinical staff who provided
5 care to these individuals during their acute illness period. We acknowledge LIINC study
6 team members Tamara Abualhsan, Mireya Arreguin, Jennifer Bautista, Monika Deswal,
7 Heather Hartig, Marian Kerbleski, Lynn Ngo, Fatima Ticas, and Meghann Williams for
8 their contributions to the program. We are grateful to Khamal Anglin, Grace Bronstone,
9 Jessica Chen, Michelle Davidson, Kevin Donohue, Peyton Ellis, Sarah Goldberg, Scott
10 Lu, Jonathan Massachi, Sujata Mathur, Irum Mehdi, Victoria Wong Murray, Enrique
11 Martinez Ortiz, Jesus Pineda-Ramirez, Mariela Romero, Paulina Rugart, Hannah Sans,
12 Joshua Shak, Jaqueline Tavs, and Jacob Weiss for assistance with data entry and
13 validation. We thank Ms. Heather Tanner (Vitalant Research Institute) for technical
14 assistance.

15 **FUNDING**

16 This work was supported by the National Institute of Allergy and Infectious Diseases
17 (NIH/NIAID 3R01AI141003-03S1 [to TJ Henrich], NIH/NIAID R01AI158013 [to M
18 Gandhi and M Spinelli] and by the Zuckerberg San Francisco Hospital Department of
19 Medicine and Division of HIV, Infectious Diseases, and Global Medicine. MJP is
20 supported on NIH T32 AI60530-12 and by the UCSF Resource Allocation Program. ST
21 is supported by the Schmidt Science Fellows, in partnership with the Rhodes Trust. IRB
22 and ST acknowledge research funding from the MIDAS Coordination Center COVID-19

1 Urgent Grant Program (MIDASNI2020-5). PDB is supported by the Intramural Research
2 Program of the National Institute of Dental Research and JIC is supported by the
3 National Institute of Allergy and Infectious Diseases. CYC is supported in part by NIH
4 grants R01 HL105704 from the National Heart, Lung, and Blood Institute, R33- 29077
5 (C.Y.C.), and Abbott Laboratories. BG and JAW are supported in part by the Chan
6 Zuckerberg Biohub Investigator Fund, XZ was supported by Damon-Runyon Cancer
7 Research Fellowship.

8 **DATA AVAILABILITY**

9 Raw antibody data at the participant and visit level are provided in Supplementary
10 Tables 2 and 3.

11 **COMPETING INTERESTS**

12 M-AP, AV, MAR, JP, JH are employees of Abbott Laboratories. LT, TW, and CJP are
13 employees of Monogram Biosciences, Inc., a division of LabCorp. C.Y.C. is the director
14 of the UCSF-Abbott Viral Diagnostics and Discovery Center (VDDC) and receives
15 research support funding from the Abbott Laboratories Inc. TJH reports grants from
16 Merck and Co., Gilead Biosciences, and Bristol-Myers Squibb outside the submitted
17 work. MJP, ST, JH, JDK, LT, NSI, KT, OJ, SEM, CT, CCN, RH, VT, EF, YH, MAS, MG,
18 SKE, XXZ, JAW, AS, TWK, JEP, WW, RLR, JNM, SGD, IRB, and BG have no interests
19 to declare.

20 **AUTHOR CONTRIBUTIONS**

1 MJP, JDK, RLR, NJM, SGD, TJH, IRB, and BG oversee the LIINC cohort, which is
2 funded by MAS, MG, SGD, and TJH. MJP, ST, JH, TJH, IRB, and BG designed the
3 study and analysis plan. MJP, LT, RH, VT, EF, and YH collected clinical data using
4 study instruments designed by MJP, JDK, RH, SGD, and JNM. LT, NSI, KT, OJ, SEM,
5 CT, CCN processed and stored biological specimens. LT, NSI, KT, OJ, SEM, CT, CCN,
6 M-AP, AV, MAR, JP, JH, LT, TW, CJP, CYC, PJN, CDG, MS, MPB, SKE, XXZ, JAW,
7 AS, TWK, JEP, WW, PDB, JIC, TJH, IRB, and BG performed or oversaw performance
8 of antibody measurements in their respective laboratories. MJP, ST, JH, TJH, IRB, and
9 BG analyzed and interpreted the raw data and drafted the manuscript. JDK, M-AP, AV,
10 MAR, JP, JH, TW, CJP, CYC, PJN, CD, MS, MPB, SKE, JAW, TWK, PDB, RLR, JNM,
11 SGD provided substantial scientific critiques and recommendations regarding the
12 interpretation of the data. All authors edited and approved the final version of the
13 manuscript.

1 REFERENCES

- 2 1. Takahashi, S., Greenhouse, B. & Rodríguez-Barraquer, I. Are Seroprevalence
3 Estimates for Severe Acute Respiratory Syndrome Coronavirus 2 Biased? *J. Infect.*
4 *Dis.* **222**, 1772–1775 (2020).
- 5 2. Guan, W.-J. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *N.*
6 *Engl. J. Med.* **382**, 1708–1720 (2020).
- 7 3. Richardson, S. *et al.* Presenting Characteristics, Comorbidities, and Outcomes
8 Among 5700 Patients Hospitalized With COVID-19 in the New York City Area.
9 *JAMA* **323**, 2052–2059 (2020).
- 10 4. Seow, J. *et al.* Longitudinal observation and decline of neutralizing antibody
11 responses in the three months following SARS-CoV-2 infection in humans. *Nat*
12 *Microbiol* **5**, 1598–1607 (2020).
- 13 5. Long, Q.-X. *et al.* Clinical and immunological assessment of asymptomatic SARS-
14 CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020).
- 15 6. Gudbjartsson, D. F. *et al.* Humoral Immune Response to SARS-CoV-2 in Iceland.
16 *N. Engl. J. Med.* **383**, 1724–1734 (2020).
- 17 7. Naaber, P. *et al.* Evaluation of SARS-CoV-2 IgG antibody response in PCR positive
18 patients: Comparison of nine tests in relation to clinical data. *PLoS One* **15**,
19 e0237548 (2020).
- 20 8. Chen, Y. *et al.* Quick COVID-19 Healers Sustain Anti-SARS-CoV-2 Antibody
21 Production. *Cell* **183**, 1496–1507.e16 (2020).
- 22 9. Chen, X. *et al.* Disease severity dictates SARS-CoV-2-specific neutralizing antibody
23 responses in COVID-19. *Signal Transduct Target Ther* **5**, 180 (2020).

- 1 10. Kowitdamrong, E. *et al.* Antibody responses to SARS-CoV-2 in patients with
2 differing severities of coronavirus disease 2019. *PLoS One* **15**, e0240502 (2020).
- 3 11. Lei, Q. *et al.* Antibody dynamics to SARS-CoV-2 in asymptomatic COVID-19
4 infections. *Allergy* (2020) doi:10.1111/all.14622.
- 5 12. Zhao, J. *et al.* Antibody Responses to SARS-CoV-2 in Patients With Novel
6 Coronavirus Disease 2019. *Clin. Infect. Dis.* **71**, 2027–2034 (2020).
- 7 13. Ibarondo, F. J. *et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with
8 Mild Covid-19. *N. Engl. J. Med.* **383**, 1085–1087 (2020).
- 9 14. Muecksch, F. *et al.* Longitudinal analysis of clinical serology assay performance
10 and neutralising antibody levels in COVID19 convalescents. *medRxiv* (2020)
11 doi:10.1101/2020.08.05.20169128.
- 12 15. Ainsworth, M. *et al.* Performance characteristics of five immunoassays for SARS-
13 CoV-2: a head-to-head benchmark comparison. *Lancet Infect. Dis.* (2020)
14 doi:10.1016/S1473-3099(20)30634-4.
- 15 16. Lau, E. H. Y. *et al.* Neutralizing antibody titres in SARS-CoV-2 infections. *Nat.*
16 *Commun.* **12**, 63 (2021).
- 17 17. Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The PHQ-15: validity of a new
18 measure for evaluating the severity of somatic symptoms. *Psychosom. Med.* **64**,
19 258–266 (2002).
- 20 18. Herdman, M. *et al.* Development and preliminary testing of the new five-level
21 version of EQ-5D (EQ-5D-5L). *Qual. Life Res.* **20**, 1727–1736 (2011).
- 22 19. Burbelo, P. D. *et al.* Sensitivity in Detection of Antibodies to Nucleocapsid and
23 Spike Proteins of Severe Acute Respiratory Syndrome Coronavirus 2 in Patients

- 1 With Coronavirus Disease 2019. *J. Infect. Dis.* **222**, 206–213 (2020).
- 2 20. Elledge, S. K. *et al.* Engineering luminescent biosensors for point-of-care SARS-
3 CoV-2 antibody detection. *medRxiv* (2020) doi:10.1101/2020.08.17.20176925.
- 4 21. Wu, L. *et al.* Optimisation and standardisation of a multiplex immunoassay of
5 diverse *Plasmodium falciparum* antigens to assess changes in malaria
6 transmission using sero-epidemiology. *Wellcome Open Res* **4**, 26 (2019).
- 7 22. Amanat, F. *et al.* A serological assay to detect SARS-CoV-2 seroconversion in
8 humans. *Nat. Med.* **26**, 1033–1036 (2020).
- 9 23. Pilarowski, G. *et al.* Performance Characteristics of a Rapid Severe Acute
10 Respiratory Syndrome Coronavirus 2 Antigen Detection Assay at a Public Plaza
11 Testing Site in San Francisco. *J. Infect. Dis.* (2021) doi:10.1093/infdis/jiaa802.
- 12 24. Long, Q.-X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19.
13 *Nat. Med.* **26**, 845–848 (2020).
- 14 25. Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N.*
15 *Engl. J. Med.* **383**, e74 (2020).
- 16 26. Iyer, A. S. *et al.* Persistence and decay of human antibody responses to the
17 receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Sci*
18 *Immunol* **5**, (2020).
- 19 27. Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* (2021)
20 doi:10.1038/s41586-021-03207-w.
- 21 28. Huang, A. T. *et al.* A systematic review of antibody mediated immunity to
22 coronaviruses: kinetics, correlates of protection, and association with severity. *Nat.*
23 *Commun.* **11**, 4704 (2020).

- 1 29. Denise M. Hinton, Chief Scientist, Food and Drug Administration, Office of the
- 2 Assistant Secretary for Preparedness and Response, Office of the Secretary, U.S.
- 3 Department of Health and Human Services. Letter of Authorization, Reissuance of
- 4 Convalescent Plasma EUA February 4, 2021. (2021).

5

6

1 TABLE AND FIGURE LEGENDS

2 **Table 1: Demographic and clinical characteristics of the study participants.**

3 **Figure 1: Longitudinal antibody kinetics.** Time since symptom onset is shown on the
4 x-axis versus the measured antibody response for each assay. For asymptomatic
5 individuals, the time since the first positive PCR test was used. Black points indicate
6 individual time points, and longitudinal samples are connected with grey lines. Y-axes
7 are transformed as indicated in Supplementary Table 1. Red dotted lines indicate cutoff
8 values for positivity, as indicated in Supplementary Table 1.

9 **Figure 2: Correlation of responses between assays. (A)** Spearman correlation of
10 random intercepts derived from a mixed effects model, representing responses at day
11 21 for each individual from the longitudinal data. Assays are sorted by hierarchical
12 clustering using average distance clustering. Darker blue indicates higher correlation;
13 colored label box indicates antigen for each binding assay and the neutralizing assay.
14 **(B)** Pairwise scatterplots showing the random intercepts for the neutralizing assay (x-
15 axis) versus the random intercepts for each of the other assays (y-axis).

16 **Figure 3: Severity-stratified antibody response at baseline.** Swarm plot of antibody
17 response at the baseline visit for each study participant by assay, stratified into
18 individuals who experienced no symptoms, individuals who experienced symptoms but
19 were not hospitalized, and those who experienced symptoms and were hospitalized. Y-
20 axes are transformed as indicated in Supplementary Table 1.

1 **Figure 4: Clinical predictors of antibody responses.** Rank of variable importance
2 (1=highest rank, 50=lowest rank) in a random forest classifier of top half vs. bottom half
3 of responders for each assay, based on random intercepts, including **(A)** all individuals
4 (n=128) and **(B)** only individuals who were not hospitalized (n=96), as hospitalization is
5 a strong predictor of antibody response. Variable importance was determined as the
6 reduction in classification error averaged across 10 runs of the algorithm. Variables only
7 relevant to hospitalized individuals (i.e., whether or not the individual was hospitalized,
8 whether or not oxygen was required, whether or not the individual was in the ICU, and
9 whether or not the individual required a ventilator) were omitted from the classifier in
10 panel B and shown in grey. In addition, HIV status is excluded as a predictor for the
11 Neut-Monogram assay for reasons described in the main text. The dependent variable
12 (individual-level random intercepts derived from a mixed effects model) is dichotomized
13 into “high” and “low”, determined by the random intercept being in the upper or lower
14 half of all random intercepts for that assay, respectively. Full labels of the predictor
15 variables are provided in Supplementary Table 6.

16 **Figure 5: Estimated time to seroreversion and assay sensitivity by time and**
17 **hospitalization status. (A)** Mean time to seroreversion for individuals tested on each
18 assay, stratified by hospitalization status, with 95% confidence intervals derived from
19 bootstrapping. Symbol (‡) indicates increasing antibody responses over time (95%
20 confidence interval for time to seroreversion was negative and did not cross 0), and
21 symbol (*) indicates antibody responses for which 95% confidence interval of time to
22 seroreversion crossed 0. **(B)** Estimated sensitivity of each assay (showing posterior
23 median estimates and 95% credible intervals), stratified by hospitalization status at 2

1 month intervals, from 0 months to 6 months after seroconversion. Seroconversion was
2 assumed to occur (if at all) 21 days after symptom onset (if symptomatic) or 21 days
3 after positive PCR test (if asymptomatic).

4 **SUPPLEMENTARY INFORMATION**

5 **Supplementary Table 1: Description of each assay.** Unit abbreviations: S/C =
6 sample result to calibrator result index; COI = cutoff index; AU/mL = arbitrary unit per
7 mL; ID50 = 50% inhibitory dilution; RLU = relative light unit; LU = light unit; conc =
8 relative concentration. Symbol (*) indicates that the cutpoint for Neut-Monogram is the
9 lower limit of detection for the assay. Antigen abbreviations: N = nucleocapsid; S =
10 spike; RBD = receptor binding domain.

11 **Supplementary Table 2: Raw data at the patient level.** Patient ID, severity class,
12 binned age in years, and sex.

13 **Supplementary Table 3: Raw data at the sample level.** Patient ID, time since
14 symptom onset (or for asymptomatic individuals, time since the first positive PCR test),
15 and antibody response for each of the 15 assays (including S-LIPS, which was highly
16 correlated with RBD-LIPS).

17 **Supplementary Table 4: Outputs of regression models evaluating the association**
18 **between demographic variables and antibody levels, controlling and not**
19 **controlling for hospitalization.** Severity is characterized by hospitalization status
20 (reference group: hospitalized). Demographic covariates considered for inclusion as
21 population-level intercepts are HIV status (reference group: HIV+), sex (reference
22 group: male), ethnicity (reference group: Hispanic), and age (reference group: older

1 than or equal to 44). Log transformations of the data performed if indicated in

2 Supplementary Table 1.

3 **Supplementary Table 5: Outputs of models testing for interaction between**

4 **hospitalization status and slope in the linear mixed effects models.** Columns

5 indicate the assay and estimate of the additional contribution of being hospitalized to the

6 slope ($\Delta\lambda_{\text{hosp}}$), with 95% confidence intervals and p-values.

7 **Supplementary Table 6: Clinical variable shorthands and corresponding REDCap**

8 **questions.** Variable name, question asked, and categories (if applicable).

9 **Supplementary Table 7: Outputs of the linear mixed effects models by assay.**

10 Point estimates of the parameters from the linear mixed effects models: population-level

11 slope (λ), population-level intercept for hospitalization status s (β_s), days to

12 seroreversion (T_s) stratified by hospitalization status, τ , and σ . Log transformations of

13 the data performed if indicated in Supplementary Table 1.

14 **Supplementary Table 8: Correlations of binding assays with pseudotype virus**

15 **neutralization.** Spearman rank correlation coefficients and 95% confidence intervals

16 derived from bootstrapping, comparing (A) random intercepts, (B) all raw data, (C) all

17 raw data from less than 90 days after symptom onset ('early'), and (D) all raw data from

18 greater than 90 days after symptom onset ('late').

19 **Supplementary Table 9: Outputs of pairwise t-tests of antibody responses at**

20 **baseline between severity groups.** P-values < 0.05 depicted in bold.

1 **Supplementary Table 10: Classification accuracy of the random forest model**
2 **when including a down-selected set of predictors.** The outcome for each individual
3 is dichotomized as being in the top half vs. bottom half of random intercepts by assay.
4 Variables included in the down-selected set for the model including all patients include:
5 hospitalization status (binary), needing oxygen (binary), fever (binary), duration of fever
6 (continuous), cough (binary), and duration of cough (continuous). Variables included in
7 the down-selected set for the model including only non-hospitalized patients include:
8 fever (binary), duration of fever (continuous), cough (binary), and duration of cough
9 (continuous). The area under the curve (AUC) estimates are averaged across 10 runs
10 of the algorithm.

11 **Supplementary Table 11: Assay sensitivity by time, severity, and assay.** Columns
12 include the assay, hospitalization status, time, and assay sensitivity (posterior mean,
13 standard error of the posterior mean, 2.5th quantile of the posterior distribution,
14 posterior median, and 97.5th quantile of the posterior distribution).

15 **Supplementary Figure 1: Distribution of time since infection at baseline visit, by**
16 **severity.** Days since symptom onset (if symptomatic) or positive PCR test (if
17 asymptomatic) of the baseline visit, stratified by severity.

18 **Supplementary Figure 2: Sample inclusion for each platform. (A)** Of the 128 total
19 individuals included, how many individuals had at least one sample tested on each
20 assay. **(B)** Of the 267 total samples tested, how many samples were tested on each
21 assay.

1 **Supplementary Figure 3: Correlation between all assays.** The values of the **(A)**
2 random intercepts, **(B)** all raw data, **(C)** all raw data from less than 90 days after
3 symptom onset ('early'), and **(D)** all raw data from greater than 90 days after symptom
4 onset ('late') are shown on the lower triangle. The Spearman rank correlation
5 coefficients and p-values of the observed coefficients are shown on the upper triangle.

6 **Supplementary Figure 4: Distribution of the random intercept of antibody values**
7 **for each individual by assay.** X-axis is the raw value of the random intercept from the
8 linear mixed effects model. Red signifies random intercept values in the bottom half of
9 that assay, and blue signifies random intercept values in the top half of that assay.
10 These binary values were used as outcome variables in the random forests modeling.

11 **Supplementary Figure 5: Raw data by time, hospitalization status, and assay.** Raw
12 antibody response data are either log-transformed or not transformed, according to
13 Supplementary Table 1. The x-axis represents time since seroconversion in days,
14 where seroconversion was assumed to occur (if at all) 21 days after symptom onset (if
15 symptomatic) or 21 days after positive PCR test (if asymptomatic). The cutoff for
16 positivity on that assay is shown by the dashed black line.

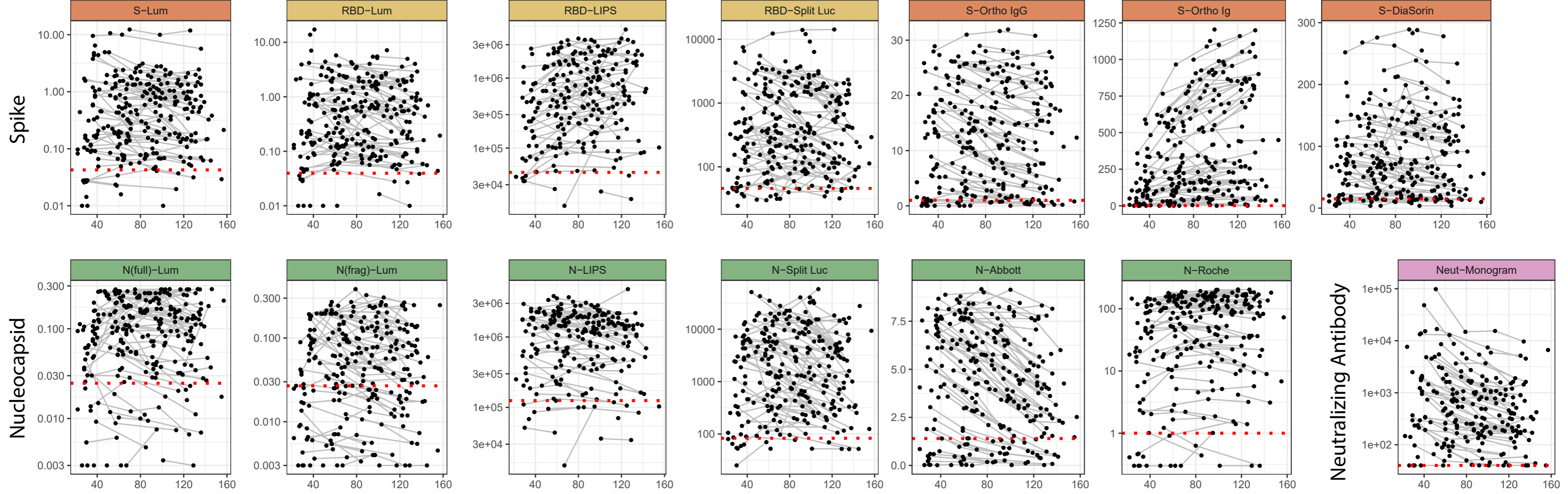
17 **Supplementary Figure 6: Ratio of sensitivity in non-hospitalized individuals to**
18 **hospitalized individuals over time.** Posterior median estimates and 95% credible
19 intervals shown.

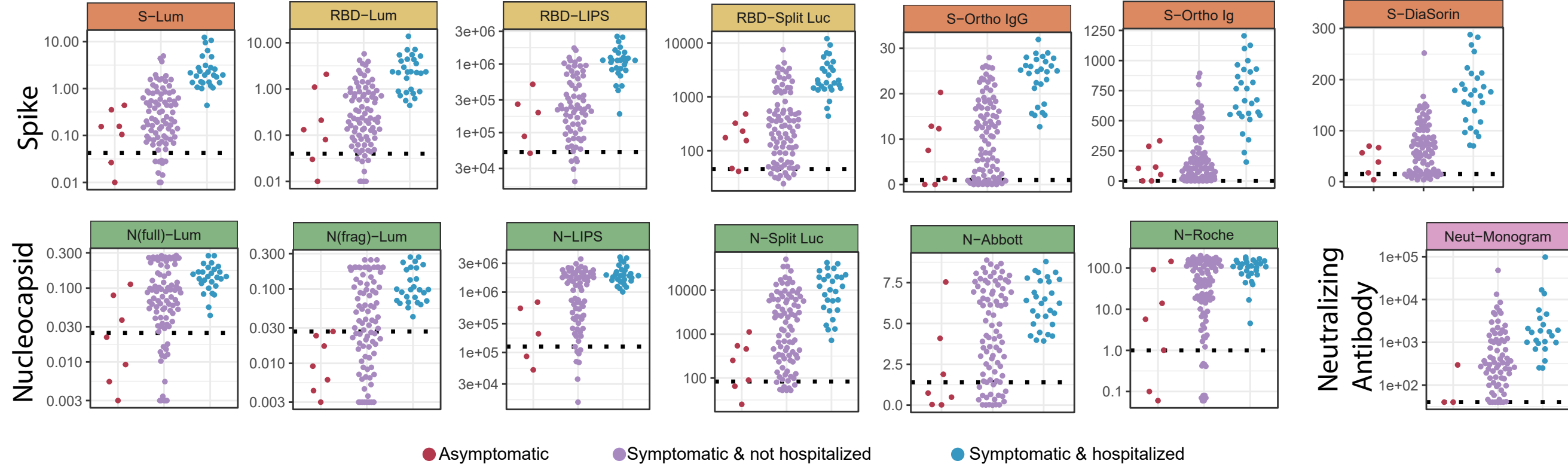
20 **Supplementary Figure 7: Negative predictive values of the commercial assays.**
21 Negative predictive values shown are based on the estimated assay sensitivities for
22 non-hospitalized individuals in Figure 5B, for a range of prevalence between 5% and

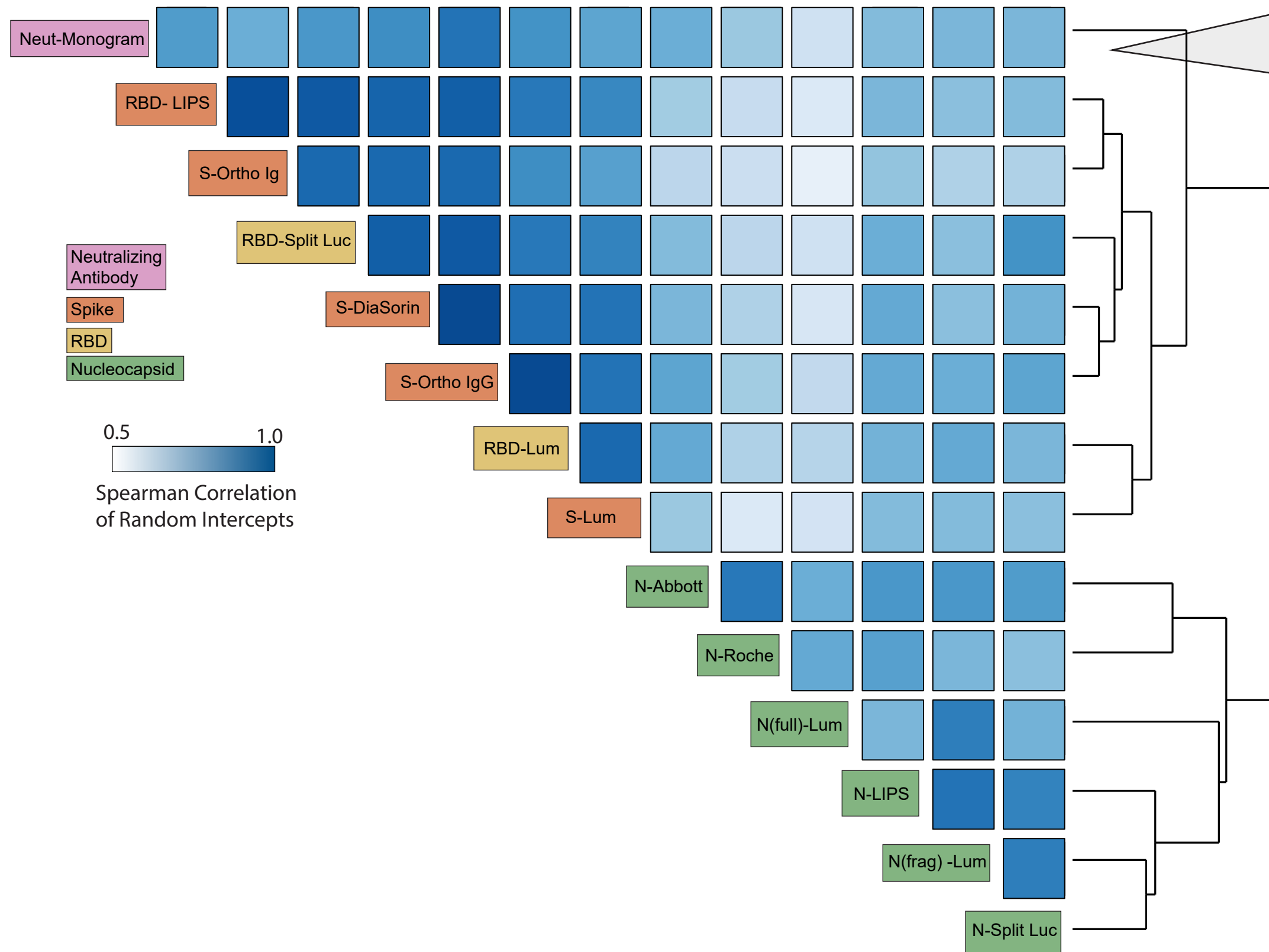
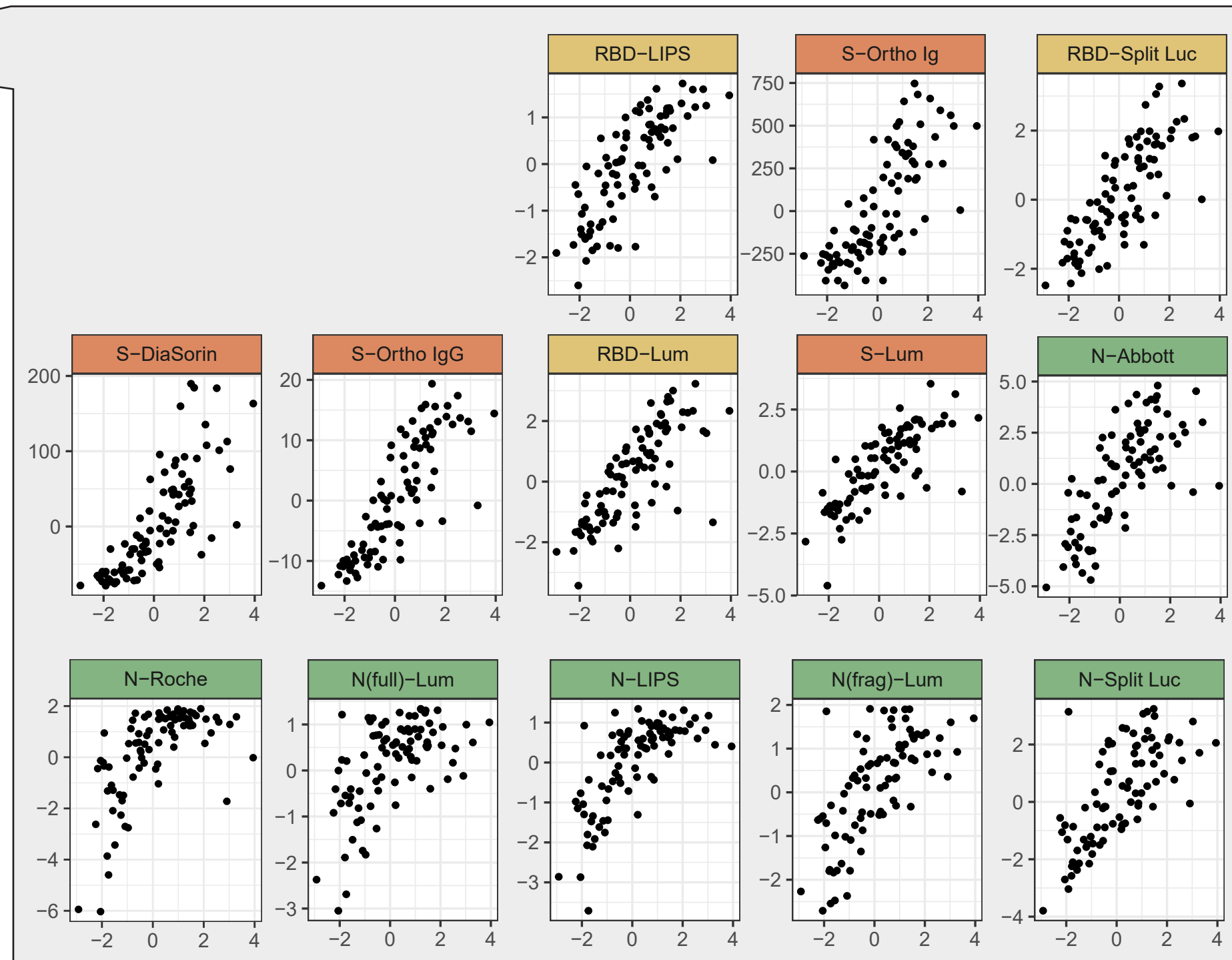
- 1 50% (x-axis). Lower panels show the same data with a smaller range in the y axis to
- 2 visualize small differences.

Characteristic	N=128 (%)
Age (years)	47.8 (range: 19-85)
Female sex at birth	57 (44.5)
Race	
American Indian or Alaska Native	4 (3.1)
Asian	15 (11.7)
Black or African American	7 (5.5)
Native Hawaiian or Other Pacific Islander	3 (2.3)
White	81 (63.4)
Declined	20 (15.6)
Latinx Ethnicity	33 (25.8)
Medical Comorbidities	
Autoimmune disease	9 (7.0)
Active cancer	3 (2.3)
Diabetes	17 (13.3)
HIV	18 (14.1)
Heart disease	3 (2.3)
Hypertension	29 (22.7)
Lung disease	21 (16.4)
Kidney disease	2 (1.6)
Obesity (BMI \geq 30)	38 (29.7)
Clinical Manifestations of COVID-19	
Asymptomatic	7 (5.5)
Symptomatic	121 (94.5)
Fever	86 (70.5)
Chills	75 (61.5)
Fatigue	110 (90.2)
Cough	89 (73.0)
Shortness of breath	77 (63.1)
Rhinorrhea	58 (47.5)
Sore throat	56 (45.9)
Myalgias	84 (68.9)
Nausea	36 (29.5)
Vomiting	12 (9.8)
Diarrhea	50 (41.0)

Anosmia or dysgeusia	82 (67.2)
Headache	77 (63.1)
Hospitalized	31 (24.2)
Required supplemental oxygen	26 (83.9)
Required ICU admission	13 (41.9)
Required mechanical ventilation	4 (12.9)
Numbers of symptoms reported	
1 to 3	13 (10.2%)
4 to 6	34 (26.6%)
7 to 9	47 (36.7%)
10 to 13	27 (21.1%)
Enrollment and follow-up	
Baseline visit, days since onset (median)	63 (range: 22-157)
Follow-up time, days since onset (median)	110 (range: 22-157)
Time points contributed (median)	2 (range: 1-4)





A**B**

A

All participants

B

Non-hospitalized participants

