Title: Test performance evaluation of SARS-CoV-2 serological assays

Author Names and Affiliations:
Jeffrey D. Whitman, M.D., M.S.[1]*, Joseph Hiatt, B.A.[2,3,4,5,6,7]*, Cody T. Mowery, B.S.[2,3,5,6,7]*, Brian R. Shy, M.D., Ph.D.[1]*, Ruby Yu, Ph.D.[5,7]*, Tori N. Yamamoto, Ph.D.[5,6,7], Ujjwal Rathore, Ph.D.[4,5,6,7], Gregory M. Goldgof, M.D., Ph.D.[1], Caroline Whitty, B.S.[1,5,7], Jonathan M. Woo, B.S.[5,6,7], Antonia E. Gallman, B.S.[2,5,8], Tyler E. Miller, M.D., Ph.D.[36] , Andrew G. Levine, M.D., Ph.D.[1], David N. Nguyen, M.D., Ph.D.[5,6,9], Sagar P. Bapat, M.D., Ph.D.[1,5,7], Joanna Balcerek, M.D., Ph.D.[1], Sophia A. Bylsma, B.S.[20], Ana M. Lyons, B.S.[21], Stacy Li, B.S.[21], Allison Wai-yi Wong, Ph.D.[2], Eva Mae Gillis-Buck, M.Phil.[10], Zachary B. Steinhart, Ph.D.[5,7], Youjin Lee, Ph.D.[5], Ryan Apathy, B.S.[5,6,7], Mitchell J. Lipke, B.S.[5,7], Jennifer Anne Smith, Ph.D.[7,11], Tina Zheng, B.S.[2,3,12,13], Ian C. Boothby, B.A.[2,14], Erin Isaza, B.S.[2,15], Jackie Chan, B.S.[5], Dante D. Acenas II, B.A.[5], Jinwoo Lee, Ph.D.[2,16], Trisha A. Macrae, Ph.D.[2,16], Than S. Kyaw, B.S.[2,5], David Wu, B.S.[2,3], Dianna L. Ng, M.D.[13,17], Wei Gu, M.D., Ph.D.[1], Vanessa A. York, B.S.[18], Haig Alexander Eskandarian, Ph.D.[18], Perri C. Callaway, B.A.[18,19], Lakshmi Warrier, B.S.[18], Mary E. Moreno, B.S.[18], Justine Levan, Ph.D.[18], Leonel Torres, B.S.[18], Lila A. Farrington, Ph.D.[18], Rita Loudermilk, B.S.[22], Kanishka Koshal, M.P.H.[22], Kelsey C. Zorn, M.H.S.[23], Wilfredo F. Garcia-Beltran, M.D., Ph.D.[36], Diane Yang, Ph.D.[36], Michael G. Astudillo, M.D.[36], Bradley E. Bernstein, M.D., Ph.D.[36], Jeffrey A. Gelfand, M.D.[37], Edward T. Ryan, M.D.[37], Richelle C. Charles, M.D.[37], A. John Iafrate, M.D., Ph.D.[36], Jochen K. Lennerz, M.D., Ph.D.[36], Steve Miller, M.D., Ph.D.[1], Charles Y. Chiu, M.D., Ph.D. [1,9,24], Susan L. Stramer, Ph.D.[25], Michael R. Wilson, M.D. [3,22], Aashish Manglik, M.D., Ph.D.[27], Chun Jimmie Ye, Ph.D.[28,29,30,31,32,33], Nevan J. Krogan, Ph.D.[4,34,35], Mark S. Anderson, M.D., Ph.D.[7], Jason G. Cyster, Ph.D.[5,8], Joel D. Ernst, M.D.[18], Alan H. B. Wu, Ph.D.[1], Kara L. Lynch, Ph.D.[1], Caryn Bern, M.D., M.P.H.[33]**, Patrick D. Hsu, Ph.D.[6,20]**, Alexander Marson, M.D., Ph.D.[4,5,6,7,11,13,28,29,30]**
*Co-first author
**Co-corresponding author

Affiliations:
1. Department of Laboratory Medicine, University of California, San Francisco, San Francisco, CA 94143, USA
2. Medical Scientist Training Program, University of California, San Francisco, CA 94143, USA
3. Biomedical Sciences Graduate Program, University of California, San Francisco, CA 94143, USA
4. J. David Gladstone Institutes, San Francisco, CA 94158, USA
5. Department of Microbiology and Immunology, University of California, San Francisco, CA 94143, USA
6. Innovative Genomics Institute, University of California, Berkeley, CA 94720, USA
7. Diabetes Center, University of California, San Francisco, San Francisco, CA 94143, USA
8. Howard Hughes Medical Institute, University of California, San Francisco
9. Department of Medicine, Division of Infectious Diseases, University of California, San Francisco, San Francisco, CA 94143, USA
10. Department of Surgery, University of California, San Francisco, CA 94143, USA
11. Department of Medicine, University of California, San Francisco, San Francisco, CA 94143, USA

12. Department of Neurology, University of California, San Francisco, CA 94158, USA
13. Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94158, USA
14. Department of Dermatology, University of California, San Francisco, San Francisco, CA 94158, USA
15. Program in Quantitative Biology, University of California, San Francisco, San Francisco, CA 94158, USA
16. School of Medicine, University of California, San Francisco, San Francisco, CA 94158, USA
17. Department of Pathology, University of California, San Francisco, San Francisco, CA, 94158, USA
18. Division of Experimental Medicine, Department of Medicine, University of California San Francisco, San Francisco CA, USA
19. Infectious Diseases and Immunity Graduate Group, University of California Berkeley, Berkeley, CA, USA
20. Department of Bioengineering, University of California, Berkeley, Berkeley CA 94720 USA
21. Department of Integrative Biology, University of California, Berkeley, Berkeley CA 94720 USA
22. Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, San Francisco, CA
23. Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA
24. UCSF-Abbott Viral Diagnostics and Discovery Center, San Francisco, CA, USA
25. Scientific Affairs, American Red Cross, Gaithersburg, MD
26. Department of Anesthesia and Perioperative Care, University of California, San Francisco, San Francisco, CA 94158
27. Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158
28. Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA
29. Chan Zuckerberg Biohub, San Francisco, CA, USA
30. Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA
31. Division of Rheumatology, Department of Medicine, University of California, San Francisco, San Francisco, CA, USA
32. Institute of Computational Health Sciences, University of California, San Francisco, San Francisco, CA, USA
33. Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA
34. Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158, USA
35. Quantitative Biosciences Institute, University of California, San Francisco, CA 94158, USA
36. Department of Pathology, Massachusetts General Hospital/Harvard Medical School, Boston, MA, USA
37. Division of Infectious Diseases, Massachusetts General Hospital/Harvard Medical School, Boston, MA, USA

Correspondence to: alexander.marson@ucsf.edu, pdhsu@berkeley.edu, caryn.bern2@ucsf.edu

ABSTRACT

Background

Serological tests are crucial tools for assessments of SARS-CoV-2 exposure, infection and potential immunity. Their appropriate use and interpretation require accurate assay performance data.

Method

We conducted an evaluation of 10 lateral flow assays (LFAs) and two ELISAs to detect anti-SARS-CoV-2 antibodies. The specimen set comprised 130 plasma or serum samples from 80 symptomatic SARS-CoV-2 RT-PCR-positive individuals; 108 pre-COVID-19 negative controls; and 52 recent samples from individuals who underwent respiratory viral testing but were not diagnosed with Coronavirus Disease 2019 (COVID-19). Samples were blinded and LFA results were interpreted by two independent readers, using a standardized intensity scoring system.

Results

Among specimens from SARS-CoV-2 RT-PCR-positive individuals, the percent seropositive increased with time interval, peaking at 81.8-100.0% in samples taken >20 days after symptom onset. Test specificity ranged from 84.3-100.0% in pre-COVID-19 specimens. Specificity was higher when weak LFA bands were considered negative, but this decreased sensitivity. IgM detection was more variable than IgG, and detection was highest when IgM and IgG results were combined. Agreement between ELISAs and LFAs ranged from 75.8-94.8%. No consistent cross-reactivity was observed.

Conclusion

Our evaluation showed heterogeneous assay performance. Reader training is key to reliable LFA performance, and can be tailored for survey goals. Informed use of serology will require evaluations covering the full spectrum of SARS-CoV-2 infections, from asymptomatic and mild infection to severe disease, and later convalescence. Well-designed studies to elucidate the mechanisms and serological correlates of protective immunity will be crucial to guide rational clinical and public health policies.

INTRODUCTION

As of April 23, 2020, more than 190,000 deaths have been attributed to Coronavirus Disease 2019 (COVID-19).[1] Millions of infections by SARS-CoV-2, the virus responsible for COVID-19, have been reported, though its full extent has yet to be determined due to limited testing.[2] Government interventions to slow viral spread have disrupted daily life and economic activity for billions of people. Strategies to ease restraints on human mobility and interaction, without provoking major resurgence of transmission and mortality, will depend on accurate estimates of population levels of infection and immunity.[3] Current testing for the virus largely depends on labor-intensive molecular techniques.[4] Individuals with positive molecular tests represent only a small fraction of all infections, given limited deployment and the brief time window when PCR testing is sensitive.[5-7] The proportion of undetected cases in the original epidemic focus was estimated to be as high as 86%,[8] and asymptomatic infections are suspected to play a substantial role in transmission.[9-14]

Widely available, reliable antibody detection assays would enable more accurate estimates of SARS-CoV-2 prevalence and incidence. On February 4, 2020, the Secretary of the United States Department of Health and Human Services issued emergency use authorization (EUA) for diagnosis of SARS-CoV-2,[15] allowing nucleic acid detection and immunoassay tests to be offered based on manufacturer-reported data without formal FDA clearance.[16] In response, dozens of companies have begun to market laboratory-based immunoassays and point-of-care tests. Rigorous, comparative performance data are crucial to inform clinical care and public health response.

We conducted a head-to-head comparison of available serology tests – immunochromatographic lateral flow assays (LFAs) and enzyme-linked immunosorbent assays (ELISAs) – including an evaluation of performance by time from symptom onset and disease severity. Our goal is to provide

well-controlled performance data to highlight potentially useful serological assays, and to help guide their development and deployment.

METHODS

Ethical approvals: This study was approved by institutional review boards at the University of California, San Francisco (UCSF)/Zuckerberg San Francisco General Hospital (ZSFG) and Massachusetts General Hospital (MGH).

Study Design: The study population included individuals with symptomatic infection and positive SARS-CoV-2 real-time polymerase chain reaction (RT-PCR) testing of nasopharyngeal or oropharyngeal swabs, who had remnant serum and plasma specimens in clinical laboratories serving the UCSF and ZSFG Medical Center networks. We included multiple specimens per individual, but no more than one sample per time interval (1-5, 6-10, 11-15, 16-20, and >20 days after symptom onset). If an individual had more than one specimen for a given time interval, only the later specimen was included. For specificity, we included 108 pre-COVID-19 specimens from American Red Cross blood donors.[17] We assessed cross-reactivity using 52 specimens from 2020: 50 with test results for other respiratory viruses (Biofire FilmArray; BioFire Diagnostics, Salt Lake City, UT), and 32 from individuals with negative results by SARS-CoV-2 RT-PCR. We based minimum sample size calculations on expected binomial exact 95% confidence limits. A total of 290 samples were included in the final analysis, including 130 from 80 SARS-CoV-2 RT-PCR-positive individuals. Some specimens were exhausted during the analysis and were not included in all tests. Data obtained from samples that did not conform to our study design were excluded.

Clinical data were extracted from electronic health records and entered in a HIPAA-secure REDCap database hosted by UCSF. Data included demographic information, major co-morbidities, patient-reported symptom onset date, symptoms and indicators of severity.

Independent data from testing efforts at MGH, with slight deviations in methods, are included as Supplementary Data. Briefly, 57 heat-inactivated serum/plasma samples from 44 SARS-CoV-2 RT-PCR-positive individuals were included. For specificity, the MGH study included 60 heat-inactivated, pre-COVID-19 samples from 30 asymptomatic adults and 30 individuals admitted with febrile and/or respiratory illness with a confirmed pathogen.

Sample Preparation: Samples from UCSF and ZSFG were assigned a random well position in one of four 96-well plates. Samples were thawed at 37°C, and up to 200uL was transferred to the assigned well without heat inactivation. Samples were then sub-aliquoted (12.5uL) to replica plates for testing. Replica plates were stored at -20°C until needed, then thawed for ten minutes at room temperature and briefly centrifuged before testing. All sample handling followed UCSF Biosafety Committee-approved Biosafety Level 2 (BSL2) practices.

For the MGH study, samples were heat-inactivated at 56°C for 60 minutes, aliquoted, and stored at 4°C and -20°C. Samples stored at 4°C were used within 7 days. Frozen aliquots were stored until needed with only a single freeze-thaw cycle for any sample. All samples were brought to room temperature and briefly centrifuged prior to adding the recommended volume to the LFA cartridge.

Immunochromatographic Lateral Flow Assays (LFAs): Ten lateral flow assays were evaluated on samples from UCSF and ZSFG (Supplementary Table 1). At the time of testing, cartridges were labeled by randomized sample location (plate, well). The appropriate sample volume was transferred from the plate to the indicated sample port, followed immediately by provided diluent, following

manufacturer instructions. The lateral flow cartridges were incubated for the recommended time at room temperature before readings. Each cartridge was assigned a semi-quantitative score (0 for negative, 1 to 6 for positive) for test line intensity by two independent readers blinded to specimen status and to each other's scores (Supplementary Figure 1).[17] For some cartridges (DeepBlue, UCP, Bioperfectus), the positive control indicator failed to appear after addition of diluent in a significant fraction of tests. For these tests, two further drops of diluent were added to successfully recover control indicators in all affected tests. These results were included in analyses. During testing, two plates were transposed 180˚ and assays were run in the opposite order from the wells documented on cartridges. These data were corrected and accuracy was confirmed by empty well position and verification of a subset of results.

Enzyme-Linked Immunosorbent Assays (ELISAs): Epitope Diagnostics ELISAs were performed according to manufacturer specifications. Cutoffs for IgG and IgM detection were calculated as the package insert described (see Supplementary Methods). Values greater than the cutoff were considered positive.

An in-house ELISA was performed with minor deviations from a published protocol.[18] SARS-CoV-2 Receptor Binding Domain (RBD) protein was produced from the published construct (NR-52306, BEI Resources). The positive cutoff was equal to the mean of the OD values of the negative control wells on the respective plate plus three times the standard deviation of the OD value distribution from the 108 pre-COVID-19 plasma. For both ELISAs, background-corrected OD values were divided by the cutoff to generate signal-to-cutoff (S/CO) ratios. Samples with S/CO values greater than 1.0 were considered positive.

Data Analysis: For LFA testing, the second reader's scores were used for performance calculations, and the first reader's score was used to calculate inter-reader agreement statistics. Percent seropositivity among RT-PCR-confirmed cases was calculated by time interval from symptom onset. Specificity was based on results in pre-COVID-2019 samples. Binomial exact 95% confidence intervals were calculated for all estimates. Analyses were conducted in R (3.6.3) and SAS (9.4).

RESULTS

Study population: SARS-CoV-2-positive individuals in the UCSF/ZSFG study ranged from 22 to >90 years of age (Table 1). The majority of SARS-CoV-2-positive individuals were Hispanic/Latinx (68.7%), reflecting the ZSFG patient population and demographics of the epidemic in San Francisco.[19,20] Most presented with cough (91.2%) and fever (86.2%). Chronic medical conditions, such as hypertension, type 2 diabetes mellitus, obesity, and chronic kidney disease, were frequent. Of the 80 cases, 18.7% were outpatients, 45.0% inpatients without ICU care, and 36.2% required ICU care; there had been no reported deaths at the time of chart review.

Test Performance: The percentage of specimens testing positive rose with increasing time from symptom onset (Table 2, Figure 1A), reaching the highest levels in the 16-20 and >20 day time intervals. The highest detection rate was achieved by combining IgM and IgG results (Figure 1B). However, 95% confidence intervals for later time intervals showed substantial overlap with those for earlier intervals (Figure 1B). Four assays (Bioperfectus, Premier, Wondfo, in-house ELISA) achieved >80% positivity in the latest two time intervals (16-20 and >20 days) while maintaining >95% specificity. Some tests were not performed on a subset of specimens due exhausted sample material, which may have affected reported percent positivity. IgM detection was less consistent than IgG for nearly all assays. Kappa agreement statistic ranged from 0.95 to 0.99 for IgG and 0.81-1.00 for IgM

for standardized intensity score and training (Supplementary Table 2 and Supplementary Figure 2). Although variability in mean band intensities exists among different assays, the rate of sample positivity was generally consistent (Figure 2).

We observed a trend towards higher percent positivity by LFA for patients admitted to ICU compared to those with milder disease, but the specimen numbers per time interval were low, limiting statistical power (Supplementary Figure 3).

Test specificity ranged from 84.3%-100.0%, with 39/108 samples demonstrating false positive results by at least one LFA (Table 2 and Figure 2B). Of the false positive results, 61.5% (24/39) had a weak positive score (1). Intensity scores of 2-3 were seen in 30.8% (12/39) and scores of 4-6 were seen in 7.7% (3/39) of the positives from the pre-COVID-19 samples.

We evaluated the tradeoff between percent positivity and specificity as a function of LFA reader score. Changing the positive LFA threshold from 1 to 2 decreased the mean overall percent positivity across tests from 66.0% (range: 56.9%-74.2%) to 56.7% (range: 44.0%-64.6%) and increased the average specificity from 94.6% (range: 84.3%-100.0%) to 98.2% (range: 94.4%-100.0%) (Supplementary Figure 4). An independent study at MGH compared three LFAs, of which BioMedomics was also assessed in the current study (Supplementary Table 3). Overall, both studies showed a trend for increased detection of SARS-CoV-2 specific antibodies with increased time from symptom onset. However, the MGH study displayed increased specificity with lower percent positivity at early timepoints after symptom onset. MGH positivity thresholds were set higher to prioritize test specificity (Supplementary Figure 4B-C).

A set of specimens obtained during the COVID-19 outbreak that had negative SARS-CoV-2 RT-PCR testing and/or alternative respiratory pathogen diagnoses demonstrated higher numbers of positive results compared to the pre-COVID-19 sample set (Figure 2C). Five specimens had positives results by >3 tests, all with respiratory symptoms and concurrent negative SARS-CoV-2 RT-PCR testing (Figure 2C, arrows). One patient was positive on 8 different tests including the in-house ELISA. In this limited panel, no consistent pattern of cross-reactivity was identified in samples from individuals with non-SARS-CoV-2 respiratory viruses, including 2 strains of seasonal coronavirus (1 coronavirus OC43, 3 coronavirus HKU1).

Agreement between results of LFAs with those of IgG and IgM Epitope ELISAs ranged from 75.8%-85.7%, while agreement with the in-house ELISA ranged from 83.6%-94.8% (Figure 3A). LFA band intensity scores showed a direct correlation with ELISA S/CO values (Figure 3B).

DISCUSSION

This study describes test performance for 12 COVID-19 serology assays on a panel of 130 samples from 80 individuals with PCR-confirmed SARS-CoV-2 infection and 108 pre-COVID-19 specimens. For each test, we quantified detection of IgM and/or IgG antibodies by time period from onset of symptoms and assessed specificity and cross-reactivity. We hope these data will inform the medical community, public health efforts, and governmental institutions in planning for SARS-CoV-2 serological testing. This study also seeks to provide feedback to manufacturers about areas of success and necessary improvement. There is no “gold standard” to identify true seropositive blood samples. The extent and time-course of antibody development are not fully understood as yet, and may vary between different populations, even among RT-PCR-confirmed cases.

We focused on comparisons of percent positivity by time interval, rather than reporting the "sensitivity" of each assay. As expected, percent positivity rose with time after symptom onset.[5,6,21-23] High rates of positive results were not reached until at least 2 weeks into clinical illness; diagnosis at time of symptom onset thus remains dependent on viral detection methods. The assays showed a trend to higher positive rates within time intervals for more severe disease, but this finding should be interpreted with caution, due to the limited data from ambulatory cases. The majority of samples >20 days post-symptom onset had detectable anti-SARS-CoV-2 antibodies, suggesting good to excellent sensitivity for all evaluated tests in hospitalized patients three or more weeks into their disease course. However, well-powered studies testing ambulatory or asymptomatic individuals, including performance with capillary blood, will be essential to guide appropriate use of serology.

Our data demonstrate specificity greater than 95% for the majority of tests evaluated and >99% for 2 LFAs (Wondfo, Sure Biotech) and the in-house ELISA (adapted from Amanat et al, 2020)[18]. We observed moderate-to-strong positive bands in several pre-COVID-19 blood donor specimens, some of them positive by multiple assays, suggesting the possibility of non-specific binding of plasma proteins, non-specific antibodies, or cross-reactivity with other viruses. Three of the pre-COVID-19 specimens (2.3%) were scored positive by more than three assays. Intriguingly, the fraction of positive tests was higher in a set of recent specimens obtained during the COVID-19 outbreak from individuals undergoing respiratory viral workup, many with negative SARS-CoV-2 RT-PCR. Five of these (9.6%) had positive results by more than three assays, without relation to a specific viral pathogen, suggesting non-specific reactivity and/or missed COVID-19 diagnoses. One specimen was positive by 8 of 12 assays, including the in-house ELISA. The patient was >90 years old and presented with altered mental status, fever, and ground glass opacities on chest radiological imaging. SARS-CoV-2 RT-PCR was negative and ancillary laboratory testing suggested a urinary tract infection. This case

could represent COVID-19 not detected by RT-PCR, reinforcing the importance of caution in interpreting negative molecular results as ruling out the infection. Moreover, this suggests potential utility of serological testing as a supplemental diagnostic tool, given the decline in sensitivity of current molecular tests in the second week of illness.[5] Appropriate clinical algorithms for serology testing, including confirmatory or reflexive testing, have yet to be determined. These algorithms will be affected by test performance characteristics and prevalence of disease across all presentations of infected individuals.

Importantly, we still do not know the extent to which positive results by serology reflect a protective immune response.[24] Future functional studies are critical to determine whether specific antibody responses predict virus neutralization and protection against re-infection. Until this is established, conventional antibody assays should not be used as predictors of future infection risk.

High specificity testing is crucial in low-prevalence settings. One approach to increase specificity would employ confirmatory testing with an independent assay (perhaps recognizing a distinct epitope or antigen). Our comparison of UCSF and MGH data suggests that reclassifying faint bands as negative or inconclusive would increase specificity, albeit at the expense of sensitivity. Detection cutoffs should be selected recognizing the likely tradeoff between specificity and sensitivity, the population prevalence, and the goals of testing. Accurate LFA use also depends on adequate reader training and standardized intensity cutoffs. Our findings highlight the need for rigorous evaluation of alternative implementation algorithms in multiple settings.

Our study also reinforces the need for assay validation using standardized sample sets with: 1) known positives from individuals with a range of clinical presentations at multiple time points after onset of symptoms, 2) pre-COVID-19 outbreak samples for specificity, and 3) samples from individuals with

other viral and inflammatory illnesses as cross-reactivity controls. Coordinated efforts to validate and ensure widespread availability of such standardized sample sets would facilitate effective utilization. We will continue to evaluate serology assays, and provide updated data on a dedicated website (https://covidtestingproject.org). Current and future studies by our group and others will provide an essential evidence base to guide serological testing during the COVID-19 pandemic.

Acknowledgements

We thank all members of the Marson lab and the Hsu lab, Peter Kim, Scott Boyd, Joe DeRisi, Steve Quake, Bryan Greenhouse, Christina Tato, Jennifer Doudna, Fyodor Urnov, David Friedberg, David Neeleman, John Hering, Cindy Cheng, Neal Khosla, Matt Krisiloff, Lachy Groom, Chenling Xu, Dave Fontenot, Jim Karkanias, Gajus Worthington, Bill Burkholder, Charlie Craik, XPrize Pandemic Alliance, Warris Bokhari, Zem Joaquin, Siavash Sarlati, Scott Nesbit, William Poe, Sam Broder, Verily, Charlie Kim, Aleksandra Kijac, Marc Solit and the Coronavirus Standards Working Group, Diane Havlir, Joanne Engel, Peter Farley, Jeff MacGregor, Kimberly Hou, Bob Sanders, Sarah Yang, and Sean Parker. We thank Yagahira Elizabeth Castro-Sesquen for sharing her semi-quantitative LFA scale, which was adapted for use in our current study. The work was supported by gifts from Anthem Blue Cross Blue Shield, the Chan Zuckerberg Biohub, and anonymous philanthropy. We thank the following sources for donation of test kits: the manufacturers of Bioperfectus, Decombio, Sure-Bio, UCP Biosciences; David Friedberg; John Hering; Henry Schein (Melville, NY). The Wilson Lab has received support from the Rachleff Family Foundation. The Hsu lab has received support from S. Altman, V. and N. Khosla, D. and S. Deb, the Curci Foundation, and Emergent Ventures. P.D.H. holds a Deb Faculty Fellowship from the UC Berkeley College of Engineering and is the recipient of the Rainwater Foundation Prize for Innovative Early-Career Scientist. The Marson lab has received gifts from J. Aronov, G. Hoskin, K. Jordan, B. Bakar, the Caufield family and funds from the Innovative Genomics Institute (IGI), the Northern California JDRF Center of Excellence and the Parker Institute for Cancer Immunotherapy (PICI). We thank the National Institutes of Health for their support (J.D.W. R38HL143581; A.E.G. F30AI150061; D.N.N. L40 AI140341; S.P.B. NHLBI R38HL143581; T.A.M. 1F30HD093116; D.W NIH 1F31NS106868-01; J.G.C. R01 AI40098; MSTP students supported by T32GM007618). R.Y. was supported by an AP Giannini

Postdoctoral Fellowship. J.A.S. was supported by the Larry L. Hillblom Foundation (2019-D-006-FEL). A.M. holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund, is an investigator at the Chan–Zuckerberg Biohub and is a recipient of The Cancer Research Institute (CRI) Lloyd J. Old STAR grant.

Competing Interests

This work was supported by gifts from Anthem Blue Cross Blue Shield, the Chan Zuckerberg Biohub, and anonymous philanthropy. C.Y.C. is the director of the UCSF-Abbott Viral Diagnostics and Discovery Center, receives research support funding from Abbott Laboratories and is on the Scientific Advisory Board of Mammoth Biosciences, Inc. C. J. Y. is cofounder of DropPrint Genomics and serves as an advisor to them. M.S.A. holds stock in Medtronic and Merck. P.D.H. is a cofounder of Spotlight Therapeutics and serves on the board of directors and scientific advisory board, and is an advisor to Serotiny. P.D.H. holds stock in Spotlight Therapeutics and Editas Medicine. A.M. is a cofounder of Spotlight Therapeutics and Arsenal Biosciences and serves on their boards of directors and scientific advisory boards. A.M. has served as an advisor to Juno Therapeutics, is a member of the scientific advisory board at PACT Pharma, and is an advisor to Trizell. A.M. owns stock in Arsenal Biosciences, Spotlight Therapeutics and PACT Pharma. RY owns stock in Abbvie, Bluebird Bio, Bristol Myers Squibb, Cara Therapeutics, Editas Medicine, Esperion, and Gilead Sciences. Unrelated to this current work, the Marson lab has received sponsored research support from Juno Therapeutics, Epinomics and Sanofi, and a gift from Gilead.

REFERENCES

1. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020. (Accessed April 12, 2020, at https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.)

2. IDSA Statement on COVID-19 Testing. 2020. (Accessed April 20th, 2020, at https://www.idsociety.org/globalassets/idsa/public-health/covid-19-idsa-testing-intro.pdf.)

3. Weitz JS, Beckett SJ, Coenen AR, et al. Intervention Serology and Interaction Substitution: Modeling the Role of 'Shield Immunity' in Reducing COVID-19 Epidemic Spread. medRxiv 2020:2020.04.01.20049767.

4. World Health Organization. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020. Geneva: World Health Organization; 2020 2020.

5. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis 2020.

6. Wolfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. Nature 2020.

7. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020.

8. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science 2020.

9. Bai Y, Yao L, Wei T, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA 2020.

10. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerg Infect Dis 2020;26.

11. Kimball A, Hatfield KM, Arons M, et al. Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. MMWR Morb Mortal Wkly Rep 2020;69:377-81.

12. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill 2020;25.

13. Qiu H, Wu J, Hong L, Luo Y, Song Q, Chen D. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. Lancet Infect Dis 2020.

14. Tong ZD, Tang A, Li KF, et al. Potential Presymptomatic Transmission of SARS-CoV-2, Zhejiang Province, China, 2020. Emerg Infect Dis 2020;26:1052-4.

15. DETERMINATION OF A PUBLIC HEAL TH EMERGENCY AND DECLARATION THAT CIRCUMSTANCES EXIST JUSTIFYING AUTHORIZATIONS PURSUANT TO SECTION 564(b) OF THE FEDERAL FOOD, DRUG, AND COSMETIC ACT, 21 U.S.C. § 360bbb-3 2020. (Accessed April 20th, 2020, at https://www.fda.gov/media/135010/download.)

16. Policy for Diagnostic Testsfor Coronavirus Disease-2019 during the Public Health Emergency. 2020. (Accessed April 14, 2020, at https://www.fda.gov/media/135659/download.)

17. Whitman JD, Bulman CA, Gunderson EL, et al. Chagas disease serological test performance in United States blood donor specimens. Journal of clinical microbiology 2019.

18. Amanat F, Stadlbauer D, Strohmeier S, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. medRxiv 2020:2020.03.17.20037713.

19. San Francisco Department of Public Health Annual Report 2017-2018. 2018. (Accessed April 19th, 2020, at https://www.sfdph.org/dph/files/reports/PolicyProcOfc/DPHAnnualReportFY17-18.pdf.)

20. San Francisco COVID-19 Data Tracker. 2020. (Accessed April 20th, 2020, at https://data.sfgov.org/stories/s/San-Francisco-COVID-19-Data-Tracker/fjki-2fab/.)

21. Lassaunière R, Frische A, Harboe ZB, et al. Evaluation of nine commercial SARS-CoV-2 immunoassays. medRxiv 2020:2020.04.09.20056325.

22. Tan W, Lu Y, Zhang J, et al. Viral Kinetics and Antibody Responses in Patients with COVID-19. medRxiv 2020:2020.03.24.20042382.

23. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis 2020.

24. Jiang S, Hillyer C, Du L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. Trends Immunol 2020.

Variable	All Patients (N=80)	1-5d (N=27)	6-10d (N=36)	11-15d (N=35)	16-20d (N=21)	>20d (N=11)
Age (yr)	52.7±15.1	49.1±14.7	53.3±15.1	57.6±15.2	54.3±14.5	55.5±14.8
Male sex	55 (68.7%)	14 (51.8%)	24 (66.6%)	22 (62.8%)	14 (66.6%)	8 (72.7%)
Racial or ethnic group						
Hispanic	55 (68.7%)	17 (62.9%)	29 (80.5%)	24 (68.5%)	14 (66.6%)	7 (63.6%)
Asian	7 (8.75%)	3 (11.1%)	2 (5.55%)	4 (11.4%)	3 (14.2%)	0 (0.0%)
White	7 (8.75%)	3 (11.1%)	1 (2.77%)	2 (5.71%)	2 (9.52%)	0 (0.0%)
Black	6 (7.5%)	2 (7.40%)	3 (8.33%)	4 (11.4%)	1 (4.76%)	2 (18.1%)
Other/not reported	5 (6.25%)	2 (7.40%)	1 (2.77%)	1 (2.85%)	1 (4.76%)	2 (18.1%)
Presenting symptoms						
Cough	73 (91.2%)	23 (85.1%)	33 (91.6%)	32 (91.4%)	19 (90.4%)	9 (81.8%)
Fever	69 (86.2%)	22 (81.4%)	30 (83.3%)	30 (85.7%)	19 (90.4%)	9 (81.8%)
Myalgia	30 (37.5%)	8 (29.6%)	12 (33.3%)	14 (40.0%)	9 (42.8%)	3 (27.2%)
Chest pain	21 (26.2%)	5 (18.5%)	8 (22.2%)	8 (22.8%)	6 (28.5%)	4 (36.3%)
Headache	21 (26.2%)	4 (14.8%)	11 (30.5%)	10 (28.5%)	7 (33.3%)	4 (36.3%)
Chills	20 (25.0%)	5 (18.5%)	9 (25.0%)	8 (22.8%)	8 (38.0%)	2 (18.1%)
Sore throat	20 (25.0%)	4 (14.8%)	11 (30.5%)	9 (25.7%)	6 (28.5%)	3 (27.2%)
Malaise	18 (22.5%)	4 (14.8%)	7 (19.4%)	10 (28.5%)	5 (23.8%)	1 (9.09%)
Diarrhea	13 (16.2%)	4 (14.8%)	7 (19.4%)	6 (17.1%)	4 (19.0%)	1 (9.09%)
Anorexia	9 (11.2%)	2 (7.40%)	1 (2.77%)	3 (8.57%)	5 (23.8%)	1 (9.09%)
Nausea and/or vomiting	9 (11.2%)	2 (7.40%)	2 (5.55%)	3 (8.57%)	3 (14.2%)	1 (9.09%)
Anosmia and/or dysgeusia	4 (5.0%)	1 (3.70%)	1 (2.77%)	2 (5.71%)	0 (0.0%)	1 (9.09%)
Chronic medical conditions						
Hypertension	36 (45.0%)	11 (40.7%)	17 (47.2%)	21 (60.0%)	11 (52.3%)	6 (54.5%)
T2DM	33 (41.2%)	11 (40.7%)	17 (47.2%)	19 (54.2%)	8 (38.0%)	6 (54.5%)
Obesity	19 (23.7%)	7 (25.9%)	9 (25.0%)	11 (31.4%)	6 (28.5%)	6 (54.5%)
CKD	10 (12.5%)	4 (14.8%)	3 (8.33%)	6 (17.1%)	4 (19.0%)	3 (27.2%)
Hypothyroid	6 (7.5%)	3 (11.1%)	3 (8.33%)	3 (8.57%)	0 (0.0%)	0 (0.0%)
Solid organ transplant	6 (7.5%)	2 (7.40%)	0 (0.0%)	2 (5.71%)	2 (9.52%)	2 (18.1%)
CAD	5 (6.25%)	1 (3.70%)	1 (2.77%)	2 (5.71%)	2 (9.52%)	3 (27.2%)
Asthma	4 (5.0%)	1 (3.70%)	1 (2.77%)	3 (8.57%)	2 (9.52%)	0 (0.0%)
CHF	3 (3.75%)	2 (7.40%)	2 (5.55%)	2 (5.71%)	1 (4.76%)	0 (0.0%)
Liver disease	3 (3.75%)	0 (0.0%)	1 (2.77%)	2 (5.71%)	1 (4.76%)	1 (9.09%)
Malignancy	3 (3.75%)	1 (3.70%)	2 (5.55%)	1 (2.85%)	2 (9.52%)	0 (0.0%)
Emphysema	2 (2.5%)	0 (0.0%)	1 (2.77%)	1 (2.85%)	1 (4.76%)	1 (9.09%)
Prior stroke	2 (2.5%)	1 (3.70%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (9.09%)
HIV	1 (1.25%)	1 (3.70%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Other immune compromised condition*	5 (6.25%)	1 (3.70%)	1 (2.77%)	3 (8.57%)	2 (9.52%)	1 (9.09%)
Highest-level of care						
Ambulatory**	15 (18.7%)	8 (29.6%)	2 (5.55%)	4 (11.4%)	2 (9.52%)	0 (0.0%)
Admitted	36 (45.0%)	11 (40.7%)	19 (52.7%)	12 (34.2%)	5 (23.8%)	4 (36.3%)
ICU	29 (36.2%)	8 (29.6%)	15 (41.6%)	19 (54.2%)	14 (66.6%)	7 (63.6%)

*Other immune compromised condition includes rheumatology patients (rheumatoid arthritis, psoriasis, Crohn's disease, ankylosing spondylitis, and reactive arthritis), all of whom were taking immune modulating/suppressing therapies.
**Ambulatory care includes outpatient as well as patients seen in ED and not admitted.

Table 1: Baseline demographic characteristics, presenting symptoms, chronic medical conditions, initial disposition and highest-level outcome for all participants whose samples were included in each time interval for serological testing. Only one sample per patient was included in each time interval, and some individuals are represented by multiple samples in different time intervals. In total, we tested 130 samples taken from 80 SARS-CoV-2 RT-PCR-positive cases.

Percentage of positive specimens from patients with positive SARS-CoV-2 RT-PCR by days since symptom onset												
	IgM				IgG				IgM or IgG			
	Total N	Positive	%	95% CI	Total N	Positive	%	95% CI	Total N	Positive	%	95% CI
Immunochromatographic Lateral Flow Assays												
BioMedomics												
1-5 days	26	7	**26.92**	11.57-47.79	26	6	**23.08**	8.97-43.65	26	8	**30.77**	14.33-51.79
6-10 days	36	22	**61.11**	43.46-76.86	36	19	**52.78**	35.49-69.59	36	23	**63.89**	46.22-79.18
11-15 days	34	25	**73.53**	55.64-87.12	34	23	**67.65**	49.47-82.61	34	26	**76.47**	58.83-89.25
16-20 days	21	16	**76.19**	52.83-91.78	21	14	**66.67**	43.03-85.41	21	17	**80.95**	58.09-94.55
>20 days	11	9	**81.82**	48.22-97.72	11	9	**81.82**	48.22-97.72	11	9	**81.82**	48.22-97.72
Bioperfectus												
1-5 days	27	11	**40.74**	22.39-61.20	27	7	**25.93**	11.11-46.28	27	11	**40.74**	22.39-61.20
6-10 days	35	26	**74.29**	56.74-87.51	35	23	**65.71**	47.79-80.87	35	27	**77.14**	59.86-89.58
11-15 days	35	28	**80.00**	63.06-91.56	35	27	**77.14**	59.86-89.58	35	30	**85.71**	69.74-95.19
16-20 days	21	16	**76.19**	52.83-91.78	21	14	**66.67**	43.03-85.41	21	17	**80.95**	58.09-94.55
>20 days	10	10	**100.00**	69.15-100.00	10	9	**90.00**	55.50-99.75	10	10	**100.00**	69.15-100.00
DecomBio												
1-5 days	25	8	**32.00**	14.95-53.50	25	7	**28.00**	12.07-49.39	25	8	**32.00**	14.95-53.50
6-10 days	36	24	**66.67**	49.03-81.44	36	24	**66.67**	49.03-81.44	36	24	**66.67**	49.03-81.44
11-15 days	34	29	**85.29**	68.94-95.05	34	29	**85.29**	68.94-95.05	34	29	**85.29**	68.94-95.05
16-20 days	20	14	**70.00**	45.72-88.11	20	14	**70.00**	45.72-88.11	20	14	**70.00**	45.72-88.11
>20 days	11	10	**90.91**	58.72-99.77	11	10	**90.91**	58.72-99.77	11	10	**90.91**	58.72-99.77
DeepBlue												
1-5 days	27	12	**44.44**	25.48-64.67	27	6	**22.22**	8.62-42.26	27	12	**44.44**	25.48-64.67
6-10 days	36	28	**77.78**	60.85-89.88	36	18	**50.00**	32.92-67.08	36	28	**77.78**	60.85-89.88
11-15 days	35	28	**80.00**	63.06-91.56	35	21	**60.00**	42.11-76.13	35	28	**80.00**	63.06-91.56
16-20 days	21	16	**76.19**	52.83-91.78	21	15	**71.43**	47.82-88.72	21	17	**80.95**	58.09-94.55
>20 days	11	10	**90.91**	58.72-99.77	11	9	**81.82**	48.22-97.72	11	10	**90.91**	58.72-99.77
Innovita												
1-5 days	27	4	**14.81**	4.19-33.73	27	7	**25.93**	11.11-46.28	27	7	**25.93**	11.11-46.28
6-10 days	36	12	**33.33**	18.56-50.97	36	17	**47.22**	30.41-64.51	36	20	**55.56**	38.10-72.06
11-15 days	32	12	**37.50**	21.10-56.31	33	25	**75.76**	57.74-88.91	33	25	**75.76**	57.74-88.91
16-20 days	14	4	**28.57**	8.39-58.10	14	9	**64.29**	35.14-87.24	14	9	**64.29**	35.14-87.24
>20 days	6	1	**16.67**	0.42-64.12	6	4	**66.67**	22.28-95.67	6	5	**83.33**	35.88-99.58
Premier												
1-5 days	27	10	**37.04**	19.40-57.63	27	6	**22.22**	8.62-42.26	27	10	**37.04**	19.40-57.63
6-10 days	35	25	**71.43**	53.70-85.36	35	18	**51.43**	33.99-68.62	35	25	**71.43**	53.70-85.36
11-15 days	35	28	**80.00**	63.06-91.56	35	22	**62.86**	44.92-78.53	35	29	**82.86**	66.35-93.44
16-20 days	21	16	**76.19**	52.83-91.78	21	14	**66.67**	43.03-85.41	21	17	**80.95**	58.09-94.55
>20 days	11	10	**90.91**	58.72-99.77	11	9	**81.82**	48.22-97.72	11	10	**90.91**	58.72-99.77
Sure												
1-5 days	27	3	**11.11**	2.35-29.16	27	5	**18.52**	6.3-38.08	27	5	**18.52**	6.30-38.08
6-10 days	35	15	**42.86**	26.32-60.65	35	19	**54.29**	36.65-71.17	35	19	**54.29**	36.65-71.17
11-15 days	35	22	**62.86**	44.92-78.53	35	25	**71.43**	53.70-85.36	35	25	**71.43**	53.7-85.36
16-20 days	21	14	**66.67**	43.03-85.41	21	14	**66.67**	43.03-85.41	21	15	**71.43**	47.82-88.72
>20 days	11	8	**72.73**	39.03-93.98	11	10	**90.91**	58.72-99.77	11	10	**90.91**	58.72-99.77
UCP												
1-5 days	27	7	**25.93**	11.11-46.28	27	7	**25.93**	11.11-46.28	27	7	**25.93**	11.11-46.28
6-10 days	36	21	**58.33**	40.76-74.49	36	18	**50.00**	32.92-67.08	36	21	**58.33**	40.76-74.49
11-15 days	35	26	**74.29**	56.74-87.51	35	25	**71.43**	53.70-85.36	35	27	**77.14**	59.86-89.58
16-20 days	21	15	**71.43**	47.82-88.72	21	14	**66.67**	43.03-85.41	21	15	**71.43**	47.82-88.72
>20 days	11	10	**90.91**	58.72-99.77	11	9	**81.82**	48.22-97.72	11	10	**90.91**	58.72-99.77
VivaChek												
1-5 days	24	7	**29.17**	12.62-51.09	24	7	**29.17**	12.62-51.09	24	7	**29.17**	12.62-51.09
6-10 days	35	22	**62.86**	44.92-78.53	35	22	**62.86**	44.92-78.53	35	22	**62.86**	44.92-78.53
11-15 days	31	26	**83.87**	66.27-94.55	31	25	**80.65**	62.53-92.55	31	26	**83.87**	66.27-94.55
16-20 days	21	15	**71.43**	47.82-88.72	21	14	**66.67**	43.03-85.41	21	15	**71.43**	47.82-88.72
>20 days	10	9	**90.00**	55.50-99.75	10	9	**90.00**	55.50-99.75	10	9	**90.00**	55.50-99.75

Wondfo												
1-5 days									25	10	**40.00**	21.13-61.33
6-10 days									36	24	**66.67**	49.03-81.44
11-15 days									33	27	**81.82**	64.54-93.02
16-20 days									21	17	**80.95**	58.09-94.55
>20 days									11	9	**81.82**	48.22-97.72
ELISAs												
Epitope												
1-5 days	27	5	**18.52**	6.30-38.08	27	11	**40.74**	22.39-61.20	27	11	**40.74**	22.39-61.20
6-10 days	36	19	**52.78**	35.49-69.59	36	28	**77.78**	60.85-89.88	36	29	**80.56**	63.98-91.81
11-15 days	35	27	**77.14**	59.86-89.58	35	31	**88.57**	73.26-96.80	35	31	**88.57**	73.26-96.80
16-20 days	21	14	**66.67**	43.03-85.41	21	16	**76.19**	52.83-91.78	21	17	**80.95**	58.09-94.55
>20 days	11	9	**81.82**	48.22-97.72	11	10	**90.91**	58.72-99.77	11	10	**90.91**	58.72-99.77
In-House												
1-5 days									27	10	**37.04**	19.40-57.63
6-10 days									36	26	**72.22**	54.81-85.80
11-15 days									35	32	**91.43**	76.94-98.20
16-20 days									21	17	**80.95**	58.09-94.55
>20 days									11	9	**81.82**	48.22-97.72
Specificity in 108 blood donor plasma specimens collected before July 2018												
	IgM				**IgG**				**IgM or IgG**			
	Total N	**Positive**	**%**	**95% CI**	**Total N**	**Positive**	**%**	**95% CI**	**Total N**	**Positive**	**%**	**95% CI**
Immunochromatographic Lateral Flow Assays												
BioMedomics	107	13	**87.85**	80.12-93.37	107	4	**96.26**	90.70-98.97	107	14	**86.92**	79.02-92.66
Bioperfectus	104	3	**97.12**	91.80-99.40	104	2	**98.08**	93.23-99.77	104	5	**95.19**	89.14-98.42
DecomBio	107	10	**90.65**	83.48-95.43	107	9	**91.59**	84.63-96.08	107	11	**89.72**	82.35-94.76
DeepBlue	108	17	**84.26**	76.00-90.55	108	1	**99.07**	94.95-99.98	108	17	**84.26**	76.00-90.55
Innovita	108	4	**96.30**	90.79-98.98	108	0	**100.00**	96.64-100.00	108	4	**96.30**	90.79-98.98
Premier	108	2	**98.15**	93.47-99.77	108	1	**99.07**	94.95-99.98	108	3	**97.22**	92.10-99.42
Sure	108	0	**100.00**	96.64-100.00	108	0	**100.00**	96.64-100.00	108	0	**100.00**	96.64-100.00
UCP	107	2	**98.13**	93.41-99.77	107	2	**98.13**	93.41-99.77	107	2	**98.13**	93.41-99.77
VivaChek	99	5	**94.95**	88.61-98.34	99	4	**95.96**	89.98-98.89	99	5	**94.95**	88.61-98.34
Wondfo									106	1	**99.06**	94.86-99.98
ELISAs												
Epitope	108	3	**97.22**	92.10-99.42	108	10	**90.74**	83.63-95.47	108	11	**89.81**	82.51-94.80
In-House									108	1	**99.07**	94.95-99.98
Percentage of positive specimens from individuals who were positive for non-SARS-CoV-2 viral infections and/or tested negative for SARS-CoV-2 by RT-PCR												
	IgM				**IgG**				**IgM or IgG**			
	Total N	**Positive**	**%**	**95% CI**	**Total N**	**Positive**	**%**	**95% CI**	**Total N**	**Positive**	**%**	**95% CI**
Immunochromatographic Lateral Flow Assays												
BioMedomics	52	8	**15.38**	6.88-28.08	52	4	**7.69**	2.14-18.54	52	11	**21.15**	11.06-34.70
Bioperfectus	45	5	**11.11**	3.71-24.05	45	6	**13.33**	5.05-26.79	45	8	**17.78**	8.00-32.05
DecomBio	52	5	**9.62**	3.20-21.03	52	2	**3.85**	0.47-13.21	52	6	**11.54**	4.35-23.44
DeepBlue	52	14	**26.92**	15.57-41.02	52	7	**13.46**	5.59-25.79	52	14	**26.92**	15.57-41.02
Innovita	28	2	**7.14**	0.88-23.50	28	2	**7.14**	0.88-23.50	28	3	**10.71**	2.27-28.23
Premier	52	1	**1.92**	0.05-10.26	52	1	**1.92**	0.05-10.26	52	2	**3.85**	0.47-13.21
Sure	52	0	**0.00**	0.00-6.85	52	0	**0.00**	0.00-6.85	52	0	**0.00**	0.00-6.85
UCP	52	3	**5.77**	1.21-15.95	52	2	**3.85**	0.47-13.21	52	3	**5.77**	1.21-15.95
VivaChek	49	4	**8.16**	2.27-19.60	49	1	**2.04**	0.05-10.85	49	4	**8.16**	2.27-19.60
Wondfo									41	0	**0.00**	0.00-8.60
ELISAs												
Epitope	52	2	**3.85**	0.47-13.21	52	8	**15.38**	6.88-28.08	52	9	**17.31**	8.23-30.33
In-House									52	7	**13.46**	5.59-25.79

Table 2: Summary statistics for immunochromatographic lateral flow assays (LFAs) and Enzyme-Linked Immunosorbent Assays (ELISAs). Samples are binned by time after patient-reported symptom onset for SARS-CoV-2 RT-PCR-positive cases. Percent of seropositivity assessed by each assay in SARS-CoV-2 RT-PCR-positive samples is reported with 95% Confidence Intervals (95% CI). The column "IgM or IgG" refers to positivity of either isotype. Specificity is determined relative to pre-COVID-19 negative control serum samples. Percent of seropositivity assessed by each assay is reported with 95% Confidence Intervals for samples from individuals who were positive for non-SARS-CoV-2 viral infections and/or tested negative for SARS-CoV-2 by RT-PCR.

Figure 1

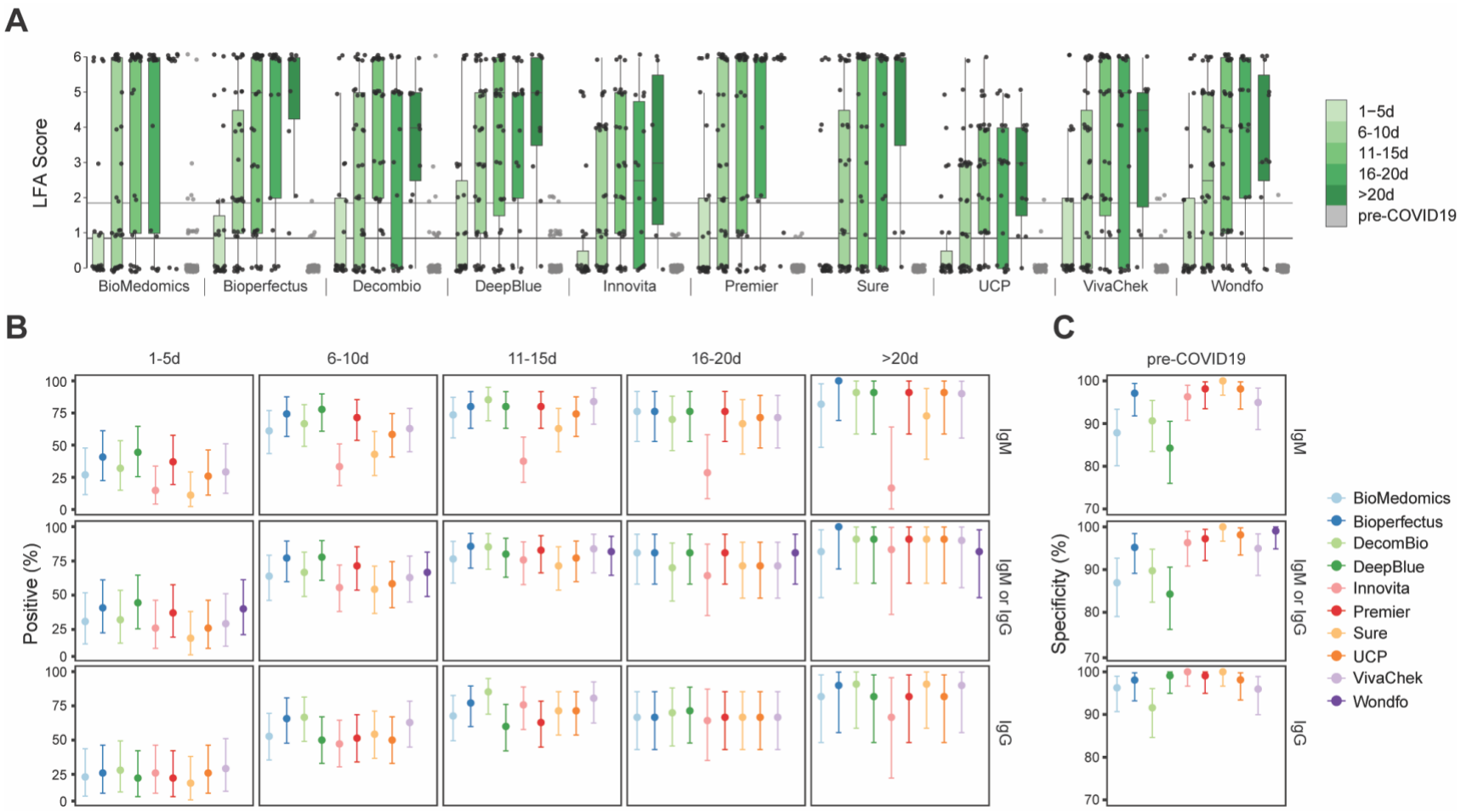


Figure 1: Performance data for immunochromatographic lateral flow assays (LFAs). **A.** The reader 2 score (0-6 based on band intensity) is reported for each assay, binned by time after patient-reported symptom onset. For tests with separate IgG and IgM bands, the higher score is reported. Joint IgM/IgG signal is represented by a single band in Wondfo. The lower, dark grey line refers to the positivity threshold (Score greater than or equal to 1) used in this study. The upper, light grey line refers to an alternative positivity threshold (Score greater than or equal to 2) discussed in the text and Supplementary Figure 4. **B.** Percent of SARS-CoV-2 RT-PCR-positive samples testing positive by each LFA are plotted relative to time after patient-reported symptom onset. The "IgM or IgG" category refers to positivity of either isotype. **C.** Specificity is plotted for each test using pre-COVID-19 negative control samples. All error bars signify 95% confidence intervals.

Figure 2.

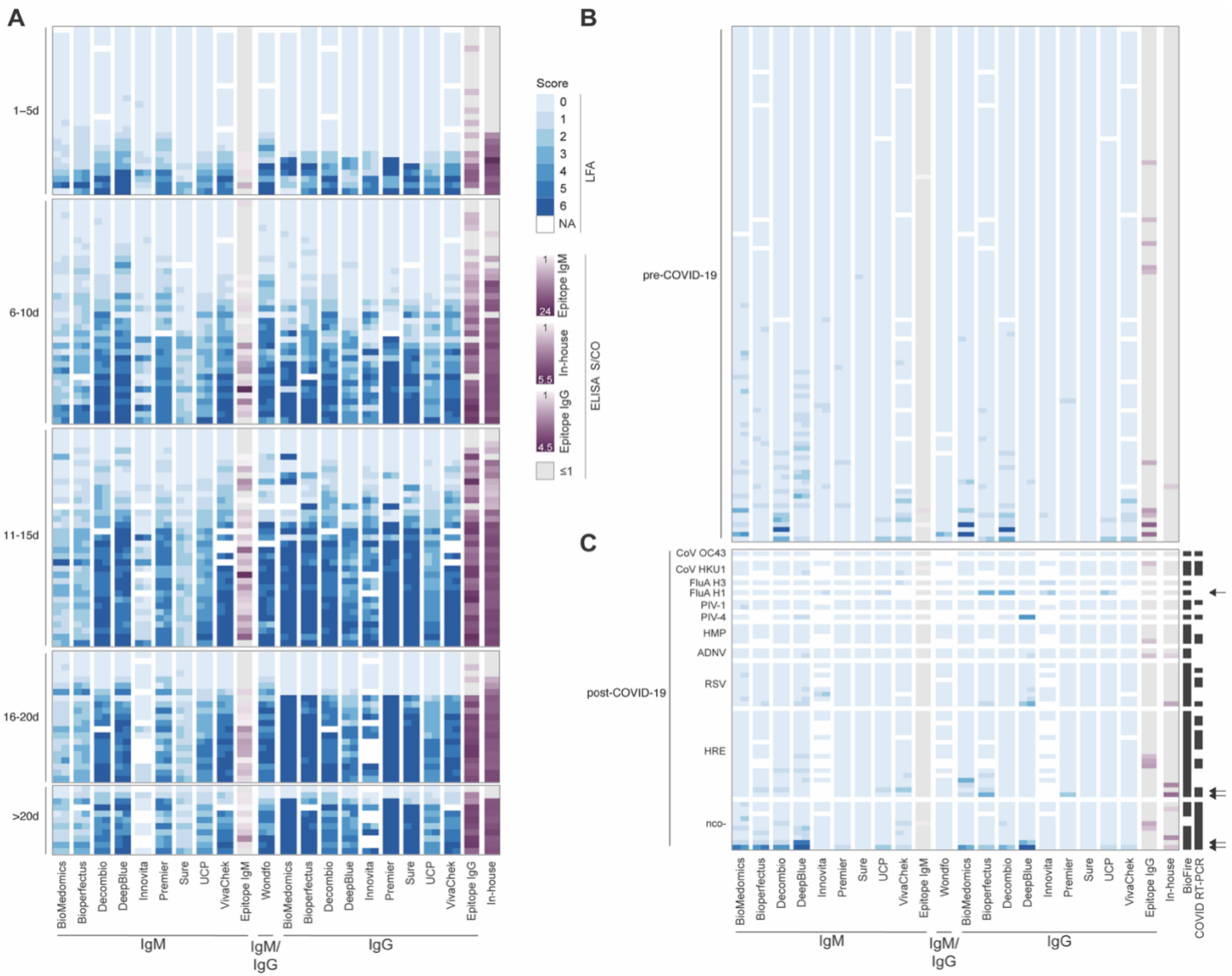


Figure 2: LFA and ELISA values by serological assay. **A.** LFA scores for each of two readers (blue) and mean ELISA Signal/Cutoff Ratio (S/CO, purple) for each specimen are grouped by binned time after patient-reported symptom onset and plotted by assay. White cells indicate samples not run with the corresponding assay. For ELISAs, grey indicates S/CO less than or equal to 1. The same legend applies to Panels B and C. The F(ab')2 specific secondary antibody used in our in-house ELISA preferentially binds the IgG light chain but has some reactivity for other isotypes (IgM, IgA). **B.** LFA score and ELISA S/CO values are plotted for pre-COVID-19 historical control serum samples to determine assay specificity. **C.** LFA score and ELISA S/CO values are plotted for serum samples obtained from 52 individuals after the emergence of COVID-19 (post-COVID-19), some of which received Biofire FilmArray (BioFire Diagnostics, Salt Lake City, UT) and/or SARS-CoV-2 RT-PCR testing (all negative) as indicated (black cells) in the appropriate columns. Arrows highlight specimens from five individuals with moderate to strong band intensity further discussed in the text. Specimens are grouped by positive testing for Coronavirus HKU1 (CoV HKU1), Coronavirus OC43 (CoV OC43), Influenza A Virus A/H3 (FluA H3), Influenza A Virus A/H1 2009 (FluA H1), Parainfluenza Type 1 Virus (PIV-1), Parainfluenza Type 4 Virus (PIV-4), Human Metapneumovirus (HMP), Adenovirus (ADNV), Respiratory Syncytial Virus (RSV), Human Rhinovirus/Enterovirus (HRE), or negative testing for SARS-CoV-2 and other viruses (nco-).

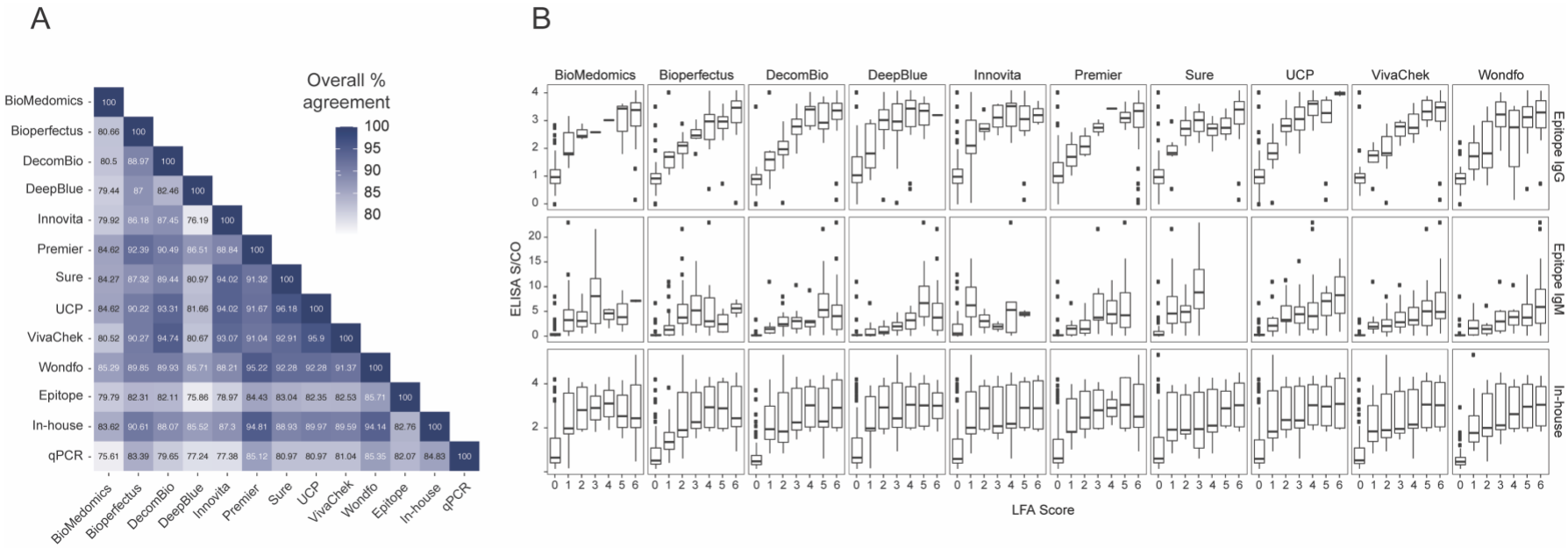


Figure 3: Agreement of serological assays for SARS-CoV-2. **A.** Percent agreement is plotted across all assay combinations, and values signify the binomial regression of the two assays across all tests. Samples were labeled "positive" if any one isotype was detected (LFA score ≥ 1, S/CO > 1) for each assay. **B.** IgM or IgG LFA scores for each assay are compared to Signal/Cutoff Ratios (S/CO) from three different ELISAs for all SARS-CoV-2 RT-PCR-positive samples. Joint IgM/IgG signal is represented by a single band in Wondfo, so data were plotted as IgM or IgG depending on ELISA comparison. The F(ab')2 specific secondary antibody used in our in-house ELISA preferentially binds the IgG light chain but contains some reactivity for other isotypes (IgM, IgA). Error bars signify 95% confidence intervals.

Supplementary Materials

Title: Test performance evaluation of SARS-CoV-2 serological assays

Table of Contents

Author Names and Affiliations:
Jeffrey D. Whitman, M.D., M.S.[1]*, Joseph Hiatt, B.A.[2,3,4,5,6,7]*, Cody T. Mowery, B.S.[2,3,5,6,7]*, Brian R. Shy, M.D., Ph.D.[1]*, Ruby Yu, Ph.D.[5,7]*, Tori N. Yamamoto, Ph.D.[5,6,7], Ujjwal Rathore, Ph.D.[4,5,6,7], Gregory M. Goldgof, M.D., Ph.D.[1], Caroline Whitty, B.S.[1,5,7], Jonathan M. Woo, B.S.[5,6,7], Antonia E. Gallman, B.S.[2,5,8], Tyler E. Miller, M.D., Ph.D.[36] , Andrew G. Levine, M.D., Ph.D.[1], David N. Nguyen, M.D., Ph.D.[5,6,9], Sagar P. Bapat, M.D., Ph.D.[1,5,7], Joanna Balcerek, M.D., Ph.D.[1], Sophia A. Bylsma, B.S.[20], Ana M. Lyons, B.S.[21], Stacy Li, B.S.[21], Allison Wai-yi Wong, Ph.D.[2], Eva Mae Gillis-Buck, M.Phil.[10], Zachary B. Steinhart, Ph.D.[5,7], Youjin Lee, Ph.D.[5], Ryan Apathy, B.S.[5,6,7], Mitchell J. Lipke, B.S.[5,7], Jennifer Anne Smith, Ph.D.[7,11], Tina Zheng, B.S.[2,3,12,13], Ian C. Boothby, B.A.[2,14], Erin Isaza, B.S.[2,15], Jackie Chan, B.S.[5], Dante D. Acenas II, B.A.[5], Jinwoo Lee, Ph.D.[2,16], Trisha A. Macrae, Ph.D.[2,16], Than S. Kyaw, B.S.[2,5], David Wu, B.S.[2,3], Dianna L. Ng, M.D.[13,17], Wei Gu, M.D., Ph.D.[1], Vanessa A. York, B.S.[18], Haig Alexander Eskandarian, Ph.D.[18], Perri C. Callaway, B.A.[18,19], Lakshmi Warrier, B.S.[18], Mary E. Moreno, B.S.[18], Justine Levan, Ph.D.[18], Leonel Torres, B.S.[18], Lila A. Farrington, Ph.D.[18], Rita Loudermilk, B.S.[22], Kanishka Koshal, M.P.H.[22], Kelsey C. Zorn, M.H.S.[23], Wilfredo F. Garcia-Beltran, M.D., Ph.D.[36], Diane Yang, Ph.D.[36], Michael G. Astudillo, M.D.[36], Bradley E. Bernstein, M.D., Ph.D.[36], Jeffrey A. Gelfand, M.D.[37], Edward T. Ryan, M.D.[37], Richelle C. Charles, M.D.[37], A. John Iafrate, M.D., Ph.D.[36], Jochen K. Lennerz, M.D., Ph.D.[36], Steve Miller, M.D., Ph.D.[1], Charles Y. Chiu, M.D., Ph.D. [1,9,24], Susan L. Stramer, Ph.D.[25], Michael R. Wilson, M.D. [3,22], Aashish Manglik, M.D., Ph.D.[27], Chun Jimmie Ye, Ph.D.[28,29,30,31,32,33], Nevan J. Krogan, Ph.D.[4,34,35], Mark S. Anderson, M.D., Ph.D.[7], Jason G. Cyster, Ph.D.[5,8], Joel D. Ernst, M.D.[18], Alan H. B. Wu, Ph.D.[1], Kara L. Lynch, Ph.D.[1], Caryn Bern, M.D., M.P.H.[33]**, Patrick D. Hsu, Ph.D.[6,20]**, Alexander Marson, M.D., Ph.D.[4,5,6,7,11,13,28,29,30]**

*Co-first author

**Co-corresponding author

SUPPLEMENTARY METHODS:

Enzyme-Linked Immunosorbent Assays (ELISAs): Epitope Diagnostics assays were performed in singlicate (due to the number of plates available for this study) and carried out according to manufacturer's instructions. Briefly, for IgM detection, 100uL of control samples or 10ul of patient serum and 100ul of sample diluent were added to indicated wells. Plates were incubated for thirty minutes at 37°C and manually washed 5x in provided Wash Buffer. Each well received 100uL of HRP-labeled COVID-19 antigen, was incubated for thirty minutes at 37°C, and was manually washed 5x in provided Wash Buffer. Each well then received 100uL of colorimetric substrate, was incubated for twenty minutes, and then received 100uL of Stop Solution. The absorbance at 450nm (OD450) was measured using a BioTek Synergy H1 Microplate Reader within ten minutes of adding Stop Solution. Positive cutoff for IgM detection were calculated as described in the Epitope Diagnostics protocol: IgM Positive cutoff = 1.1 * ((average of negative control readings) + 0.10). Values less than or equal to the Positive cutoff were interpreted as Negative. For IgG detection, 1uL of serum was diluted 1:100 in Sample Diluent and loaded into designated wells. Plates were incubated for thirty minutes at room temperature and manually washed 5x in provided Wash Buffer. Each well received 100uL of provided HRP-labeled COVID-19 Tracer Antibody, plates were incubated for thirty minutes at room temperature, and manually washed 5x in provided Wash Buffer. Then, each well received 100uL of Substrate, was incubated for twenty minutes, and then received 100uL of Stop Solution. The absorbance at 450nm (OD450) was measured using a BioTek Synergy H1 Microplate Reader within ten minutes of adding Stop Solution. Positive cutoffs for IgG detection were calculated as described in the Epitope Diagnostics protocol: IgG Positive cutoff = 1.1 * ((average of negative control readings) + 0.18). Values less than or equal to the Positive cutoff were interpreted as Negative.

An in-house RBD-based ELISA was performed with minor deviations from a published protocol (Amanat *et al.* 2020, Krammer Lab, MSSM, New York, NY, USA). SARS-CoV-2

Receptor Binding Domain (RBD) protein was produced using the published construct (NR-52306, BEI Resources) by Aashish Manglik (UCSF). 96-well plates (3855, Thermo Scientific) were coated with 2ug/ml RBD protein and stored at 4°C for up to five days before use. Test serum aliquots (12uL), as well as pre-July 2018 historical Negative Control serum from two donors and Positive Control serum from a patient with confirmed anti-SARS-CoV-2 IgG, were diluted 1:5 in 1X PBS (10010-023, Gibco), mixed, and heat inactivated at 56°C for one hour. RBD-treated plates were washed 3x with PBS-Tween (PBST, BP337-500, Fisher Bioreagents) using a BioTek 405 TS Microplate Washer and blocked with PBST-Milk (3% w/v, AB10109-01000, AmericanBio) for one hour at 20°C. Samples were further diluted 1:10 (1:50 final) in PBST-Milk (1% w/v) and 100uL was transferred to the blocked ELISA plates in duplicate plates. Samples were incubated for two hours at 20°C and washed 3x with PBST. The peroxidase AffiniPure Goat Anti-human IgG ($F(ab')_2$ specific) secondary antibody (109-035-097, Lot 146576, Jackson ImmunoResearch) used in this study binds the IgG light chain and has some reactivity for other isotypes (IgM, IgA). This secondary antibody was diluted 1:750 in PBST-Milk (1% w/v), 50ul was added to each sample well, and samples were incubated for one hour at 20°C. Plates were subsequently washed 3x with PBST. We dispensed 100uL of 1x SigmaFast OPD Solution (P9187, Sigma-Aldrich) to each sample well and incubated plates for ten minutes at room temperature. We added 50ul of 3M HCl (A144-212, Fisher Chemical) to stop the reaction and immediately read the optical density at 490nm (OD490) using a BioTek Synergy H1 Microplate Reader. OD490 values were corrected for each plate by subtracting the mean value of each plate's blank wells. To determine a cutoff for positive values, we calculated the mean value of negative wells for each plate, plus three standard deviations.

Assay	Supplier	Product	Antigen *	Format**	Lot(s)	Product Number	Distributor	Kit Acquisition for Study	Performance Notes
LFAs	**BioMedomics** Inc, Morrisville, NC, USA	COVID-19 IgM and IgG Rapid Test	RBD	1	2020032103	51-002-20	Henry Schein, Melville, NY, USA	Provided by Distributor Free of Charge	Some control band inconsistency
	Bioperfectus Technologies Co Ltd, Jiangsu, China	PerfectPOC Novel Corona Virus (SARS-CoV-2) IgM/IgG Rapid Test Kit	NP, SP	1	20200313, 20200313, 20210312	SC30201W	--------	Provided by Supplier Free of Charge	Extra diluent necessary
	Decombio Biotechnology Co Ltd, Beijing, China	Novel Coronavirus (SARS-CoV-2) IgM/IgG Combo Rapid Test-Cassette	--------	1	--------	--------	--------	Provided by Supplier Free of Charge	Some control band inconsistency
	DeepBlue Medical Technology Co Ltd, Anhui, China	COVID-19 (SARS-CoV-2) IgG/IgM Antibody Test Kit (Colloidal Gold)	--------	1	20200305	--------	--------	Donated by John Hering, who purchased from supplier	Extra diluent necessary, Some control band inconsistency
	Innovita Biological Technology Co Ltd, Qian'an, China	Novel Coronavirus (2019-nCoV) Ab Test (Colloidal Gold)	NP, SP	2	20200304	--------	20/20 GeneSystems, Rockville, MD, USA	Purchased from Distributor	Some band smearing
	Premier Biotech, Minneapolis, MN, USA	COVID-19 IgG/IgM Rapid Test Cassette	--------	1	COV20030071	INGM-MC42S	--------	Purchased from Supplier	Some band smearing
	Sure Biotech, New York, NY, USA; Wan Chai, Hong Kong	SARS-CoV-2 IgM/IgG Antibody Rapid Test	NP, SP	1	COV1252003B	VC012103	--------	Provided by Supplier Free of Charge	--------
	UCP Biosciences, San Jose, CA, USA	Coronavirus IgG/IgM Antibody (COVID-19) Test Cassette	--------	1	SMP20200312, SMP20200313	U-CoV-102	--------	Provided by Supplier Free of Charge	Extra diluent necessary
	VivaChek Biotech Co, Hangzhou, China	VivaDiag™ SARS-CoV-2 IgM/IgG Rapid Test (COVID-19 IgM/IgG Rapid Test)	--------	1	E2003002	VID35-08-011	Everest Links Pte Ltd, Singapore	Purchased from Distributor	Some band smearing
	Wondfo Biotech Co Ltd, Guangzhou, China	SARS-CoV-2 Antibody Test (Lateral Flow Method)	--------	3	W19500318	W195	--------	Donated by David Friedberg, who purchased from supplier	Some band smearing
MGH LFAs	**SD Biosensor**, Suwon-si, Gyeonggi-doz, Republic of Korea	STANDARD Q COVID-19 IgM/IgG Duo	NP	2	QCO1020006	Q-NCOV-01D	Henry Schein, Melville, NY, USA	Provided by Distributor Free of Charge	--------
	Biolidics Limited, Mapex, Singapore	2019-nCoV IgG/IgM antibody detection kit	NP, RBD	1	V20200330	CBB-F015016-V	--------	Purchased from Supplier	--------
	Biomedomics Inc, Morrisville, NC, USA	COVID-19 IgM and IgG Rapid Test	RBD	1	2020022702 2020032103	51-002-20	Henry Schein, Melville, NY, USA	Lot 1 provided by Distributor Free of Charge; Lot 2 purchased from Supplier	--------

ELISAs	**Epitope** Diagnostics, San Diego, CA, USA	KT-1033 EDI™ Novel Coronavirus COVID-19 **IgM** ELISA Kit	NP	--	P630C	KT-1032	--------	Purchased from Supplier	--------
	Epitope Diagnostics, San Diego, CA, USA	KT-1032 EDI™ Novel Coronavirus COVID-19 **IgG** ELISA Kit	NP	--	P637U	KT-1033	--------	Purchased from Supplier	--------
	In-House ELISA	Peroxidase AffiniPure Goat Anti-human IgG (F(ab')2 specific) secondary antibody (Jackson ImmunoResearch)	RBD	--	146576	109-035-097	Adapted from Krammer Lab, Icahn School of Medicine at Mt. Sinai, New York, NY, USA	Lab-developed test	--------

***Antigen:**
NP = Nucleocapsid protein
SP = Spike protein
RBD = Receptor binding domain, Spike protein

****LFA Test Cartridge Format:**
1: Single lane, separate IgM and IgG bands
2: Separate IgM and IgG lanes
3: Single lane, single band for both IgM and IgG

Supplementary Table 1. Immunoassay Kit and Manufacturer Information. **Bold** signifies labels used in text and figures.

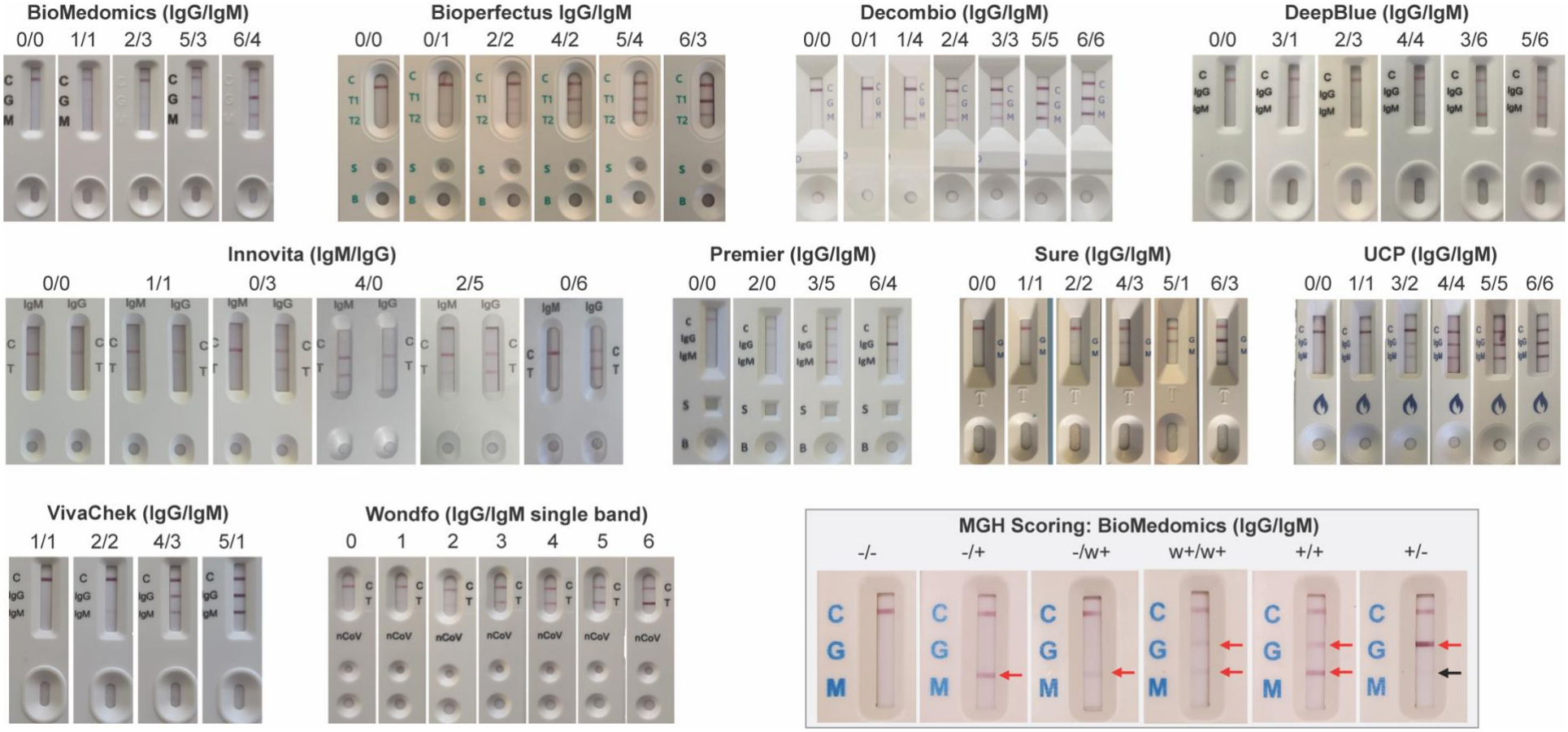
Supplier	IgG			IgM		
	n	Positive Kappa Correlation	Weighted Kappa Correlation	n	Positive Kappa Correlation	Weighted Kappa correlation
BioMedomics	287	0.9651	0.9581	287	0.8247	0.8258
Bioperfectus	277	0.9587	0.9489	277	0.9134	0.8634
DecomBio	285	0.9763	0.9531	285	0.9846	0.9661
DeepBlue	290	0.9549	0.8974	290	0.9218	0.9380
Innovita	252	0.9590	0.8493	251	0.8087	0.8031
Premier	289	0.9719	0.9881	289	0.9681	0.9342
Sure	289	0.9908	0.9666	289	0.9302	0.7971
UCP	289	0.9566	0.9575	289	1.0000	0.9485
VivaChek	269	0.9912	0.9670	269	0.9336	0.9441
Wondfo	273	0.9916	0.9543	-	-	-

Supplementary Table 2. Reader Agreement on Immunochromatographic Lateral Flow Assays (LFAs). Cohen's Kappa correlations were calculated for scores of the IgG band (left) and IgM band (right) of each LFA. The LFA manufactured by Wondfo has a single band for IgG and IgM detection and is displayed here as IgG for convenience. Positive Kappa Correlation: unweighted inter-reader agreement on positive (LFA score > 0) vs. negative (LFA score = 0) reads. Weighted Kappa Correlation: inter-reader agreement on LFA score (0-6), weighted by the square of the difference in reads. All correlations were calculated with the *irr* package version 0.84.1 in R version 3.6.1 using RStudio.

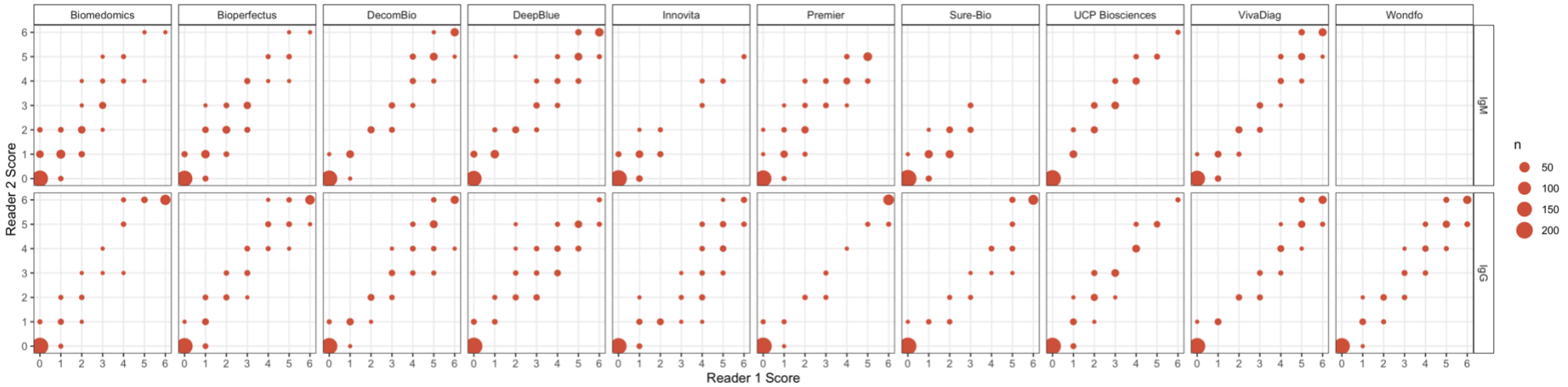
MGH Serology Test Performance Evaluation

	IgM			95% CI		IgG			95% CI		IgM or IgG			95% CI	
Assay	**Total N**	**positive**	**%**	**Lower**	**Upper**	**Total N**	**positive**	**%**	**Lower**	**Upper**	**Total N**	**positive**	**%**	**Lower**	**Upper**
LFAs															
SD Biosensor															
1-5 days	7	0	0.00	0.00	40.96	7	1	14.29	0.36	57.87	7	1	14.29	0.36	57.87
6-10 days	15	6	40.00	16.34	67.71	15	5	33.33	11.82	61.62	15	7	46.67	21.27	73.41
11-15 days	19	15	78.95	54.43	93.95	19	16	84.21	60.42	96.62	19	17	89.47	66.86	98.70
>16 days	7	6	85.71	42.13	99.64	7	6	85.71	42.13	99.64	7	6	85.71	42.13	99.64
Pre-COVID-19	60	0				60	1				60	1			
Biolidics															
1-5 days	7	0	0.00	0.00	40.96	7	0	0.00	0.00	40.96	7	0	0.00	0.00	40.96
6-10 days	15	2	13.33	1.66	40.46	15	8	53.33	26.59	78.73	15	8	53.33	26.59	78.73
11-15 days	19	9	47.37	24.45	71.14	19	16	84.21	60.42	96.62	19	16	84.21	60.42	96.62
>16 days	7	4	57.14	18.41	90.10	7	6	85.71	42.13	99.64	7	6	85.71	42.13	99.64
Pre-COVID-19	60	0				60	0				60	0			
BioMedomics															
1-5 days	7	1	14.29	0.36	57.87	7	0	0.00	0.00	40.96	7	1	14.29	0.36	57.87
6-10 days	15	6	40.00	16.34	67.71	15	6	40.00	16.34	67.71	15	7	46.67	21.27	73.41
11-15 days	19	14	73.68	48.80	90.85	19	14	73.68	48.80	90.85	19	15	78.95	54.43	93.95
>16 days	7	6	85.71	42.13	99.64	7	6	85.71	42.13	99.64	7	6	85.71	42.13	99.64
Pre-COVID-19	60	0				60	0				60	0			

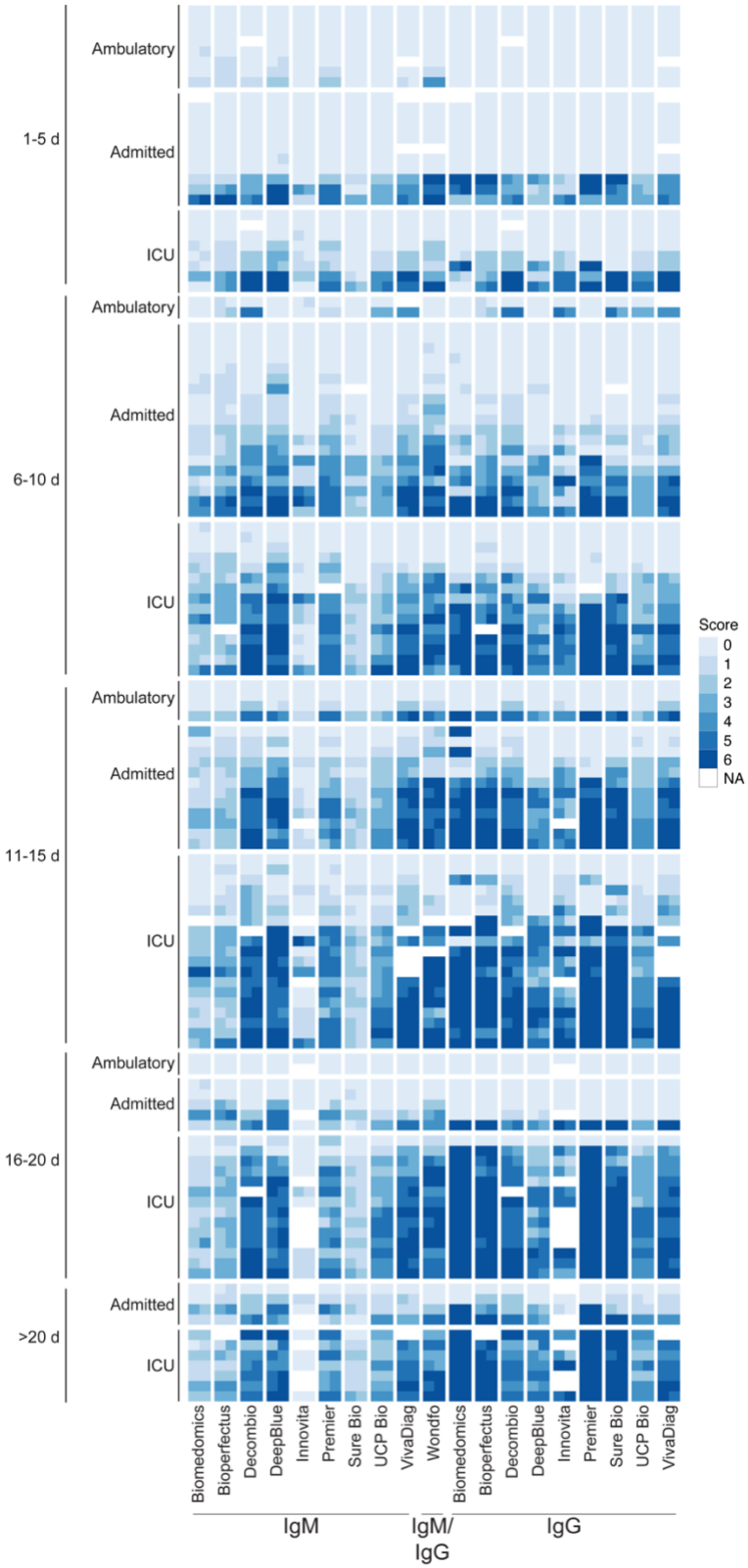
Supplementary Table 3. Assay performance on validation cohort performed at MGH using positivity thresholds based on concordance studies to an MGH-group in-house ELISA. Comparison of MGH and UCSF percent positivity at different positivity thresholds is performed in Supplementary Figure 4. Note, the one negative patient included in the >16-day timepoint was immunocompromised.



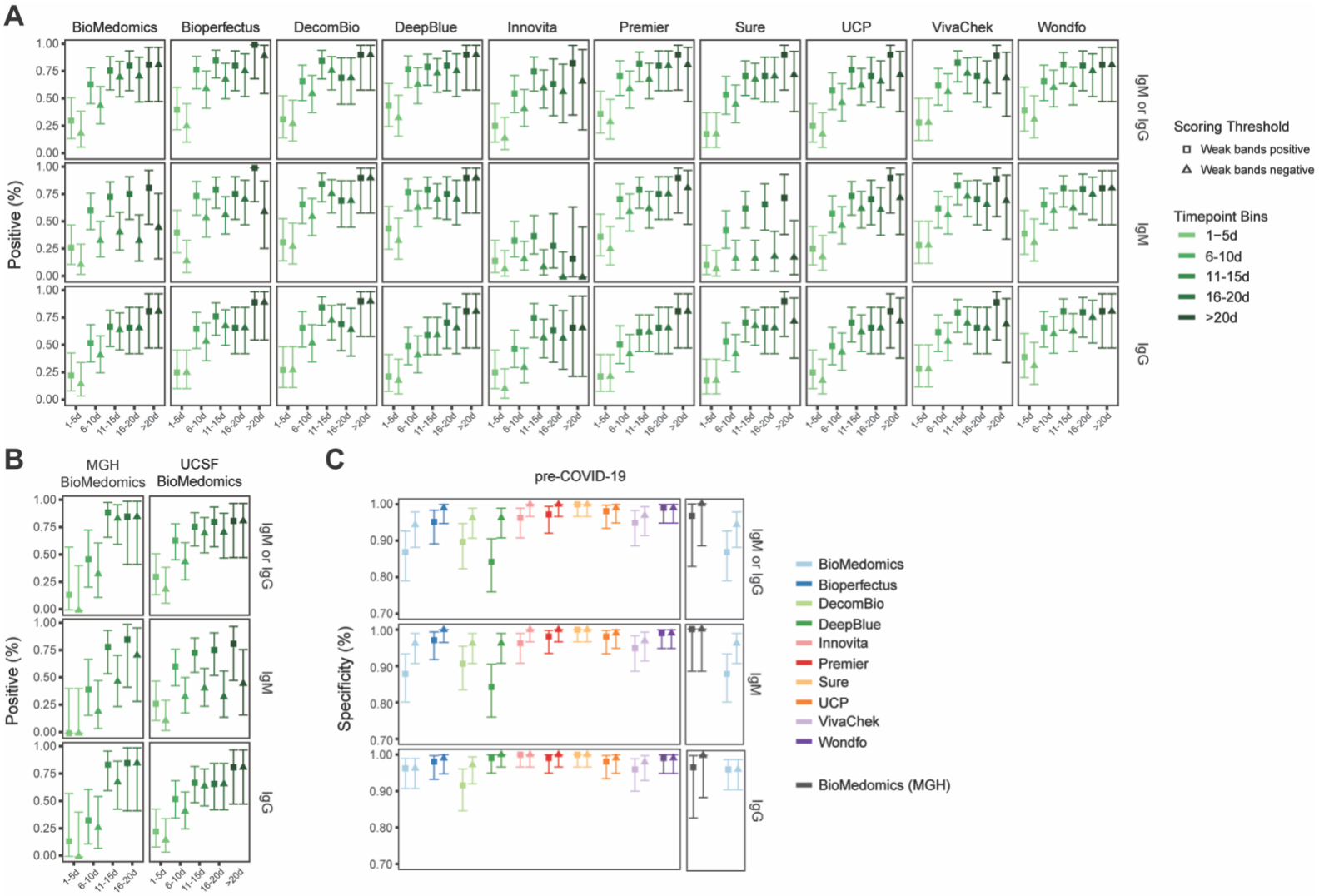
Supplementary Figure 1: Representative images of LFA scoring.



Supplementary Figure 2: Comparison of Reader 1 and Reader 2 LFA scores. The size of each point signifies the number of tests with the indicated reader 1-to-reader 2 score combination. The LFA manufactured by Wondfo has a single band for IgG and IgM detection and is displayed here as IgG for convenience.



Supplementary Figure 3: LFA scores by serological assay according to highest-level clinical care received by the patient by the patient.



Supplementary Figure 4. Comparison of the effect of different positivity thresholds on percent positivity and specificity. **A.** The percent positivity of each assay tested on serum from SARS-CoV-2 RT-PCR-positive patients is plotted by time after patient-reported symptom onset. Squares indicate

percent positivity using Reader Score > 0 (“Weak bands positive”) as the positivity threshold. Triangles indicate percent positivity using Reader Score > 1 (“Weak bands negative”) as the positivity threshold. “IgM or IgG” signifies detection of either isotype. Wondfo reports a single band for IgM and IgG together, and the results are plotted here as both “IgM” and “IgG” for horizontal comparison across assays. **B.** Comparison of percent positivity at each timepoint for BioMedomics assay at either the MGH (left) or UCSF (right) study site using low (square) or high (triangle) positivity thresholds. Note that a weak score at MGH is not directly equivalent to a 1 at UCSF due to difference in reader training. **C.** The specificity of all assays on historical pre-COVID-19 serum using low (square) or high (triangle) positivity thresholds. UCSF BioMedomics data is plotted again in the right subpanel column for direct comparison to MGH BioMedomics data. All error bars indicate 95% confidence intervals.