

## **Title:** Evaluation of serological lateral flow assays for severe acute respiratory syndrome coronavirus-2

### **Author list:**

Bianca A. Trombetta<sup>1\*</sup>, Savannah E. Kandigian<sup>1\*</sup>, Robert R. Kitchen<sup>2,3\*</sup>, Korneel Grauwet<sup>4\*</sup>, Pia Kivisäkk Webb<sup>1,2</sup>, Glenn A. Miller<sup>3</sup>, Charles G. Jennings<sup>4,5</sup>, Sejal Jain<sup>6,7</sup>, Samara Miller<sup>2,8,9,10</sup>, Yikai Kuo<sup>1,4</sup>, Thadryan Sweeney<sup>4</sup>, Tal Gilboa<sup>11,12</sup>, Maia Norman<sup>11,12,13</sup>, Daimon P. Simmons<sup>14</sup>, Christopher E. Ramirez<sup>1</sup>, Melissa Bedard<sup>14</sup>, Catherine Fink<sup>15</sup>, Jina Ko<sup>11,16</sup>, Esmarine J. De León Peralta<sup>17,18,19</sup>, Gerald Watts<sup>14</sup>, Emma Gomez-Rivas<sup>14</sup>, Vanessa Davis<sup>12</sup>, Rocky Barilla<sup>20</sup>, Jianing Wang<sup>8,21</sup>, Pierre Cunin<sup>14</sup>, Samuel Bates<sup>22</sup>, Chevaun Morrison-Smith<sup>12</sup>, Benjamin Nicholson<sup>23</sup>, Edmond Wong<sup>23</sup>, Leena El-Mufti<sup>1</sup>, Michael Kann<sup>23</sup>, Anna Bolling<sup>1</sup>, Brooke Fortin<sup>1</sup>, Hayden Ventresca<sup>21</sup>, Wen Zhou<sup>24</sup>, Santiago Pardo<sup>1</sup>, Megan Kwock<sup>25</sup>, Aditi Hazra<sup>2,26</sup>, Leo Cheng<sup>27</sup>, Q. Rushdy Ahmad<sup>11</sup>, James A. Toombs<sup>11,28</sup>, Rebecca Larson<sup>29,30</sup>, Haley Pleskow<sup>23,31</sup>, Nell Meosky Luo<sup>31</sup>, Christina Samaha<sup>31</sup>, Unnati M. Pandya<sup>2,33</sup>, Pushpamali De Silva<sup>18</sup>, Sally Zhou<sup>34,35,36</sup>, Zakary Ganhadeiro<sup>34,35,36</sup>, Sara Yohannes<sup>28</sup>, Rakiesha Gay<sup>28, 35</sup>, Jacqueline Slavik<sup>28</sup>, Shibani S. Mukerji<sup>1</sup>, Petr Jarolim<sup>7,12</sup>, David R. Walt<sup>11,12</sup>, Becky C. Carlyle<sup>1,2</sup>, Lauren L. Ritterhouse<sup>17\*\*</sup>, Sara Suliman<sup>14\*\*\*#</sup>

\* and \*\* contributed equally

### **#Corresponding Author:**

Sara Suliman, MPH, PhD, Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston MA. Diagnostics Accelerator Working Group and Laboratory of the Mass General and Brigham Center for COVID Innovation (MGBCCI), 149 Navy Yard room: 4325, Charlestown MA, USA. Email: [ssuliman1@bwh.harvard.edu](mailto:ssuliman1@bwh.harvard.edu)

### **Affiliations:**

1. Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Charlestown, Boston, MA
2. Department of Medicine, Harvard Medical School, Boston, MA
3. Mass General Brigham Innovation, Boston, MA
4. Cardiology Division, Massachusetts General Hospital, Charlestown, MA
5. Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
6. Department of Medical Oncology and Center for Cancer-Genome Discovery, Dana-Farber Cancer Institute, Boston, MA
7. Department of Pathology, Harvard Medical School, Boston, MA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

8. Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA
9. Harvard Stem Cell Institute, Cambridge, MA
10. Department of Psychiatry, Massachusetts General Hospital, Boston, MA
11. Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA
12. Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
13. Sackler School of Biomedical Sciences, Tufts University School of Medicine, Boston, MA
14. Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston, MA
15. Medical Diagnostic Technology Evaluation, LLC, Carlisle, MA
16. Center for Systems Biology, Massachusetts General Hospital, Boston, MA
17. Department of Pathology, Massachusetts General Hospital, Boston, MA
18. Wellman Center for Photomedicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA
19. Department of Dermatology, Massachusetts General Hospital, Boston, MA
20. Evergrande Center for Immunologic Diseases, Brigham and Women's Hospital, Boston, MA
21. Mass General Hospital Institute of Health Professions, Charlestown, MA
22. Functional Genomics Laboratory, Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA
23. Center for Cancer Research, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA
24. Division of Nephrology and Endocrine Unit Department of Medicine, Massachusetts General Hospital, Boston, MA
25. Cancer Center Protocol Office, Massachusetts General Hospital, Boston, MA
26. Division of Preventative Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
27. Radiology and pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA
28. Brigham Research Institute, Brigham and Women's Hospital, Boston, MA
29. Immunology Program, Harvard Medical School, Boston, MA
30. Cellular Immunotherapy Program, Cancer Center, Massachusetts General Hospital, Boston, Massachusetts, USA
31. Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, MA
32. Folia Health, Inc., Cambridge, MA
33. Vincent Center for Reproductive Biology, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston, MA
34. Department of Biology, Northeastern University, Boston, MA

35. College of Science, Northeastern University, Boston, MA

36. Mass General Brigham COVID Center for Innovation, Diagnostics Accelerator,  
Boston, MA

**Author's Contributions:**

BAT, SEK, RRK, BCC and SS wrote the first draft of the manuscript

BAT, SEK, and KG led the project management, volunteer coordination and lab operations

BAT, SEK, BCC, PK and LR established the MGB Diagnostics Accelerator lab, acquired samples and established biosafety and IRB protocols

SS, SM, SJ, DS, PJ, CER developed evaluation protocols

SS and LR managed the MGB COVID Diagnostics Accelerator lab

BAT and RRK led sample and data management

RRK and BCC led the analysis and data interpretation

RA, JT and GM corresponded with LFA providers and down-selected assays for evaluation

CER, BN, EP, SEK, BAT, EW, LE, CK, AB, BF, HV, WZ, SP, MK, JK, JT, AH, RA, RL, HP, NL, CS, UP, PD, SZ, ZG, VD, GW, RB, JW, PC, SB, CMS volunteered in the MGB COVID Diagnostics Accelerator lab

YK, TS and RRK developed the interactive web app for interpretation of test results

DW led the Diagnostics Pillar of the Mass General Brigham COVID Center for Innovation

All authors planned and/or performed experiments, edited and approved the final version of the manuscript

**Funding:**

This study was funded by a grant from the Massachusetts Consortium for Pathogen Readiness (MassCPR).

**Running title:** COVID-19 Serological Testing

**Total word count:** 3,500

1 **Abstract (250 words)**

2 **Background:** COVID-19 has resulted in significant morbidity and mortality  
3 worldwide. Lateral flow assays can detect anti-Severe Acute Respiratory Syndrome  
4 Coronavirus-2 (SARS-CoV-2) antibodies to monitor transmission. However,  
5 standardized evaluation of their accuracy and tools to aid in interpreting results are  
6 needed.

7 **Methods:** We evaluated 20 IgG and IgM assays selected from available tests  
8 in April 2020. We evaluated the assays' performance using 56 pre-pandemic negative  
9 and 56 SARS-CoV-2-positive plasma samples, collected 10-40 days after symptom  
10 onset, confirmed by a molecular test and analyzed by an ultra-sensitive immunoassay.  
11 Finally, we developed a user-friendly web app to extrapolate the positive predictive  
12 values based on their accuracy and local prevalence.

13 **Results:** Combined IgG+IgM sensitivities ranged from 33.9% to 94.6%, while  
14 combined specificities ranged from 92.6% to 100%. The highest sensitivities were  
15 detected in Lumiquick for IgG (98.2%), BioHit for both IgM (96.4%), and combined  
16 IgG+IgM sensitivity (94.6%). Furthermore, 11 LFAs and 8 LFAs showed perfect  
17 specificity for IgG and IgM, respectively, with 15 LFAs showing perfect combined  
18 IgG+IgM specificity. Lumiquick had the lowest estimated limit-of-detection (LOD) (0.1  
19  $\mu\text{g/mL}$ ), followed by a similar LOD of 1.5  $\mu\text{g/mL}$  for CareHealth, Cellex, KHB, and  
20 Vivachek.

21 **Conclusion:** We provide a public resource of the accuracy of select lateral flow  
22 assays with potential for home testing. The cost-effectiveness, scalable manufacturing  
23 process, and suitability for self-testing makes LFAs an attractive option for monitoring  
24 disease prevalence and assessing vaccine responsiveness. Our web tool provides an

- 25 easy-to-use interface to demonstrate the impact of prevalence and test accuracy on  
26 the positive predictive values.  
27  
28 **Key words:** SARS-CoV2, Coronavirus, antibodies, COVID-19, lateral flow assays.

## 29 Introduction

30 Coronavirus disease 2019 (COVID-19), caused by infection with the severe  
31 acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was declared a global  
32 pandemic on March 11<sup>th</sup>, 2020<sup>1</sup>. A second wave of the pandemic is well underway<sup>1, 2,</sup>  
33 <sup>3, 4</sup>. However, accurate estimates of transmission rely on accurate and widely available  
34 immunosurveillance tools to measure SARS-CoV-2 infection in diverse community  
35 settings. Among SARS-CoV-2-infected individuals, 40-45% are estimated to remain  
36 asymptomatic<sup>5</sup>, suggesting that prevalence is likely underestimated<sup>6</sup>. Therefore,  
37 detecting prior exposure to SARS-CoV-2 as opposed to other viruses, including other  
38 coronaviruses is crucial<sup>7</sup>.

39 There are different types of clinical SARS-CoV-2 tests<sup>8</sup>. Diagnostic testing  
40 relies on reverse-transcriptase quantitative polymerase chain reaction (RT-qPCR) and  
41 antigen-based immunodiagnosics to detect active infection<sup>9, 10, 11</sup>. Conversely,  
42 serological tests can monitor population prevalence and prior exposure by measuring  
43 antibodies against SARS-CoV-2<sup>12, 13, 14</sup>. These include enzyme-linked immunosorbent  
44 assays (ELISAs), chemiluminescence assays, and lateral flow assays (LFAs)<sup>11, 15, 16</sup>.  
45 LFAs are attractive for home testing and population surveillance, since they are  
46 affordable, scalable, rely on easily accessible specimens such as fingerstick whole  
47 blood and give a result readout within minutes<sup>13</sup>. Since multiple vaccines received  
48 emergency use authorization<sup>17</sup>, LFAs could be used to determine whether vaccines  
49 elicit a detectable and durable immune response<sup>18, 19, 20, 21, 22</sup>. Furthermore, early  
50 seroconversion was shown to predict better clinical outcomes<sup>23, 24</sup>. Hence, easy-to-  
51 use LFAs will have important applications in the upcoming phases of the pandemic.  
52 Since the onset of the COVID-19 epidemic, multiple studies evaluated the accuracy of

53 serological tests<sup>15, 25, 26, 27, 28, 29</sup>. Many of these tests received Emergency Use  
54 Authorization (EUA) through the Food and Drug Administration (FDA)<sup>30</sup>.

55 Despite the utility of SARS-CoV2 antibody tests, misinterpretation of results is  
56 very likely<sup>31</sup>. A negative serological test result does not preclude prior infection since  
57 seroconversion occurs 9-11 days after symptoms onset for IgM and 18-20 days for  
58 IgG antibodies, respectively<sup>16, 32, 33, 34</sup>. Conversely, positive results do not indicate  
59 active infection<sup>31</sup>. Furthermore, the prevalence of SARS-CoV-2 is highly variable<sup>1, 6</sup>,  
60 and known to directly impact the predictive value of a test result. A higher prevalence  
61 increases the likelihood that positive test results indicate a real infection (i.e. higher  
62 positive predictive value)<sup>35</sup>, but will also decrease the negative predictive value,  
63 resulting in more false negative results<sup>35</sup>. Therefore, accessible tools to assist the  
64 public with interpreting results based on test accuracy and different prevalence  
65 scenarios are critical<sup>31, 36</sup>.

66 In April 2020, the Mass General Brigham Center (MGB) for COVID Innovation's  
67 direct-to-consumer working group scanned available serological assays and selected  
68 20 lateral flow assays, based on reported assay characteristics and supply chain  
69 availability<sup>37</sup>. The LFAs were evaluated by blinded operators using the same samples  
70 to standardize the evaluation of their accuracy. Additionally, we developed a user-  
71 friendly web-tool to provide context for the end user to interpret their results. Here, we  
72 report the evaluation data to serve as a public resource to guide implementation of  
73 LFAs, and the tool to aid the interpretation of home testing results.

74

## 75 **Methods**

### 76 **Sample procurement**

77 We procured 56 SARS-CoV-2 positive, 46 pre-pandemic negative, and 10 HIV+ EDTA  
78 plasma samples. SARS-CoV-2 positive EDTA plasma samples were obtained from  
79 clinical discards banked within 24-72 hours of collection at the Crimson Core of the  
80 MGB Biobank from hospitalized symptomatic patients. All samples had positive  
81 SARS-CoV-2 PCR results using an EUA approved test at the Brigham and Women's  
82 Hospital (Panther Fusion SARS-CoV-2 assay, Hologic, Inc., San Diego, CA or Xpert  
83 Xpress SARS-CoV-2, Cepheid, Inc., Sunnyvale, CA) or the Clinical Research  
84 Sequencing Platform at the Broad Institute of MIT and Harvard (in house Laboratory  
85 Developed Test) 10-40 days prior to sample collection. The participants' charts were  
86 reviewed by study staff to identify samples collected  $\geq 10$  days after onset of symptoms  
87 and to exclude immunosuppressed participants, after which samples were  
88 anonymized and stripped of protected health information. Pre-pandemic negative  
89 control samples were randomly selected from healthy participants with a Charlson  
90 Age-Comorbidity Index<sup>38</sup> score  $\leq 2$ , with EDTA plasma banked in the MGB Biobank  
91 between Jan 1-Dec 1, 2019 from inpatients. HIV-positive control samples were  
92 obtained from EDTA plasma samples banked prior to January 2020 in a study on  
93 neuropathic pain in HIV. All HIV-positive participants were on antiretroviral therapy.  
94 For 8 out of the 10 HIV-positive samples, viral load quantification was available and  
95 showed 256 copies/ml or less, and 5 showed either undetectable loads or under 20  
96 copies/ml. The study was approved under the Massachusetts General Brigham (MGB)  
97 Institutional Review Board (protocol no. 2020P001204).



## 98 **Lateral Flow Assays (LFAs)**

99 Twenty commercial IgG/IgM lateral flow assays (LFAs) from 18 manufacturers were  
100 evaluated (**Supplementary Table 1**). LFAs were analyzed by blinded operators  
101 according to manufacturer instructions for use (IFU), with the exception of using  
102 micropipettes instead of manufacturer-provided droppers to minimize technical  
103 variability. Samples were thawed on ice, randomized, and brought to room  
104 temperature. Kit components were also brought to room temperature. The IFU-  
105 specified volumes of sample and buffer were added to the cassette. Specified sample  
106 volumes varied for different LFAs but were typically in the 5-20 $\mu$ L range. The cassettes  
107 were run at room temperature on a flat surface and results read immediately after the  
108 time interval defined in the IFU (typically ranging from 10-15 minutes). Each cassette  
109 was independently scored by two blinded raters as either “positive,” “negative,” or  
110 “invalid”. Ratings were designated according to the interpretation guidelines outlined  
111 in each individual IFU. Each cassette was photographed under four standardized  
112 illumination conditions and viewpoints for future analysis.

113

## 114 **Reproducibility testing**

115 For inter-operator reproducibility analysis, separate pools of EDTA plasma were  
116 obtained from >30 pre-pandemic healthy individuals (negative pool) and >30  
117 convalescent participants collected after symptom resolution at the Massachusetts  
118 General Hospital (MGH) respiratory illness clinic (positive pool). Convalescent  
119 samples for the positive pool were confirmed to be positive using the COBAS SARS-  
120 CoV-2 PCR test (Roche Diagnostics, Indianapolis, IN) at MGH. A total of 20 replicates  
121 per pool were run by two independent pairs of blinded operators, alternating between

122 positive and negative pools (10 replicates per pair). Reproducibility was calculated  
123 according to agreement between operator ratings as well as concordance of readout  
124 with sample pool COVID status.

125

### 126 **Sensitivity and specificity testing**

127 Our cohort of 112 EDTA plasma samples was used across all 20 LFAs to evaluate  
128 performance: sensitivity, specificity, positive predictive value (PPV) and negative  
129 predictive value (NPV). Samples were sub-aliquoted throughout the analysis to  
130 minimize freeze-thaw cycles. Binary presence/absence were used, and discordant  
131 calls were resolved by a third operator inspecting photographs taken of the relevant  
132 LFA.

133

### 134 **LFA usability**

135 In addition to initial screening<sup>37</sup>, each LFA kit was assessed for consumer usability  
136 based on complexity of kit materials, sample requirements, and IFU clarity. Supplied  
137 kit components were documented for completeness and examined for ease-of-use.  
138 IFU protocols were rated on a scale from 0-14 according to a predefined rubric  
139 (**Supplementary Table 2**) by three independent raters. Usability evaluations are  
140 shown in **Supplementary Table 3**. Sample input requirements for each LFA are in  
141 **Supplementary Table 1**.

142

### 143 **Ultrasensitive Simoa Serology Assays**

144 Plasma samples were diluted 4000-fold, and the total IgG and IgM levels against the  
145 SARS-CoV2 spike protein were measured using a custom Single Molecule Array

146 (Simoa) assay as described<sup>39</sup> on an automated HD-X Analyzer (Quanterix, Billerica,  
147 MA, USA), to provide a quantitative reference for anti-spike antibody titers in the  
148 plasma samples. Normalized mean Average Enzymes per Bead (AEB) levels were  
149 calculated using a standard set of calibrators produced by serially diluting a large  
150 volume of plasma from seroconverted individuals. Antibody concentrations were  
151 estimated using a calibration curve of recombinant anti-SARS-CoV-2 antibodies<sup>40</sup>.

152

153 **Analysis and Webapp:** Detailed methods are in the supplementary.

154

## 155 **Results**

### 156 **Study Population**

157 We obtained plasma samples from 56 pre-pandemic patients, including 10 HIV+, and  
158 56 symptomatic patients in March and April 2020. HIV-negative samples were  
159 matched for sex and age between the COVID- and COVID+ groups. The overall study  
160 population included 25.9% Blacks, 4.5% Latinx, 9.8% Asian and 43.8% Non-Hispanic  
161 Whites (**Table 1**). COVID+ samples were from individuals between 10-40 days post  
162 symptom onset, with 60.7% samples taken between 2-4 weeks. Of all the COVID+  
163 participants, 7 (12.5%) were symptomatic outpatients and 49 (87.5%) were  
164 hospitalized. Among those hospitalized, 25 (51%) required intensive care unit (ICU)  
165 treatment and 4 (8.2%) were deceased at the time of chart review (**Table 2**). Additional  
166 mortalities were possible after chart review since some participants were in critical  
167 condition in the ICU.

## 168 **LFA Performance: Reproducibility**

169 Four independent operators working in teams of two on separate days applied each  
170 pool 10 times to each LFA, with processing as dictated by the instructions for use  
171 (IFU). API version1 LFAs were not assessed for reproducibility due to limited cassette  
172 availability and high sample volumes required. Of the remaining 19 LFAs, three  
173 (BTNX, Camtech, and Carehealth) had 100% consistent correct outcomes across both  
174 isotypes, with an additional three (BioHit, Zhuhai Livzon, and Phamatech) having no  
175 incorrect or inconsistent outcomes with one or two invalid tests (**Figure 1**). IgG was  
176 the more reproducible isotype. The majority of incorrect consistent calls came from  
177 operators calling a COVID+ sample IgM negative (**Figure 1**).

178

## 179 **LFA Performance: Practicalities of Use**

180 We assessed the LFAs according to this rubric (**Supplementary Table 2**) by three  
181 independent raters and assigned a composite score on a scale of 0-14  
182 (**Supplementary Table 3**). Five LFAs (BioHit, InTec, Lumiquick, Phamatech, and  
183 U2U) received full marks for IFU clarity. LFAs frequently lost points for imprecise  
184 instructions regarding correct usage of disposable droppers as well as optimal time  
185 between adding sample and reading the results. Important kit usability criteria, such  
186 as whether the included pipette droppers show clear volume markings, were also  
187 recorded (**Supplementary Table 3**). While these kits may not yet be intended for the  
188 general public, it will be important to clarify the instructions moving forward and include  
189 clearly marked droppers to minimize potential for sample volume errors.

190

## 191 **LFA Performance: Sensitivity and Specificity**

192 To focus on LFA specificity, given the likely use of these tests in low-prevalence  
193 settings, disagreement between operators was interpreted as a negative call. Across  
194 all but three of the LFAs (Biohit, BTNX, Vivacheck), sensitivity was higher for IgG than  
195 IgM. Sensitivity for IgG ranged from 98.2% (Lumiquick) down to 72.7% (Oranoxis),  
196 and for IgM from 96.4% (BioHit) to 23.2% (Oxo and U2U) (**Table 3**). LFA specificity  
197 was much higher for both isotypes, with 11 LFAs having a specificity of 100% for IgG  
198 (API, API v2, BTNX, Camtech, Genobio, Oranoxis, Phamatech, Ray Biotech, Ray  
199 Biotech\_v2, U2U and Zhuhai Livzon), and 8 LFAs having a specificity of 100% for IgM  
200 (API, CareHealth, Cellex, Lumiquick, Oranoxis, Ray Biotech v2, U2U, Zhuhai Livzon)  
201 (**Table 3**). Under the assumption that the likelihood of two randomly occurring false  
202 positives for any one individual is low, an IgG/IgM composite score (averaged operator  
203 scores, see Methods) was produced to maximize test specificity. Using this composite  
204 score, all LFAs except BioHit, Cellex, Edinburgh, InTec and Vivachek achieved a  
205 specificity of 100%. This result underscores the potential of considering the outcome  
206 in both isotypes to minimize false positives, although it is more likely that a single  
207 isotype will be used in clinical testing.

208

209 We created heatmaps to visualize individual sample outcomes across all LFAs  
210 to assess whether we systematically detected the same miscalls across multiple LFAs  
211 (**Figure 2**). False negatives (blue squares in the COVID+ panel) amongst COVID+  
212 patients were somewhat reproducible, with three COVID+ samples called negative in  
213 both isotypes by all or all but one LFA. These miscalls were not clearly explained by  
214 known demographics (age, sex) or clinical variables (disease severity, weeks post

215 symptom onset) (**Figure 2**, bottom panel). To investigate whether these miscalls were  
216 related to low titers of anti-SARS-CoV-2 antibodies from participants with a  
217 suppressed immune response, all 112 samples were analyzed for anti-spike IgG and  
218 IgM antibodies using a custom quantitative Simoa assay<sup>39</sup>. The three samples that  
219 were called negative across almost all LFAs, which were collected 2-3 weeks after  
220 symptoms onset, had the lowest levels of anti-spike antibodies in COVID+ samples  
221 for both IgG and IgM, suggesting these participants had slower or suppressed immune  
222 responses to SARS-CoV-2 infection.

223 Unlike false negatives, most false positives (**Figure 2**, red squares in COVID-  
224 panel) amongst COVID- individuals appear largely uncorrelated between LFAs.  
225 However, two samples showed IgG false positives across multiple LFAs, which may  
226 suggest long lasting antibodies from exposure to coronaviruses other than SARS-  
227 CoV-2. Nonetheless, this observation is not reflected in the antibody levels from the  
228 Simoa analysis, which showed barely detectable anti-spike antibodies in these two  
229 samples.

230

### 231 **Defining the limit of detection for qualitative LFAs**

232 The custom Simoa anti-spike IgG and IgM antibody assays use a standard curve to  
233 determine standardized antibody concentration in each sample<sup>39</sup>. The results obtained  
234 from this assay can therefore be used to estimate a limit of detection (LOD) for each  
235 of the qualitative LFAs. The cumulative number of false negative LFA calls in COVID+  
236 samples (**Figure 3**, y-axis) were computed as a function of decreasing antibody  
237 concentrations (**Figure 3**, x-axis) separately for IgG and IgM. We define the LOD for  
238 each LFA/antibody as the concentration at which  $\geq 95\%$  of the COVID+ samples are

239 called unambiguously positive (**Supplementary table 4**). Using this definition, for IgG,  
240 all LFAs (except Genobio, Oranoxis, OZO, Ray Biotech, and U2U) have an LOD within  
241 the linear range of the SIMOA assay (1 - 10,000  $\mu\text{g/mL}$ ). Lumiquick has the lowest  
242 LOD at 0.1  $\mu\text{g/mL}$ , which was extrapolated by dilution to be within the linear range of  
243 the Simoa standards, and CareHealth, Cellex, KHB and Vivachek all have an LOD of  
244 1.5  $\mu\text{g/mL}$  (**Supplementary table 4**). Considering the generally lower sensitivity  
245 observed with the IgM assays (**Figure 2**), IgM assays consistently have higher LODs,  
246 with 9 exceeding 1,000  $\mu\text{g/mL}$ . BioHit has the lowest IgM LOD at 0.6  $\mu\text{g/mL}$ , and API  
247 version2, BTNX, CareHealth and Vivachek all have LODs under 10  $\mu\text{g/mL}$ .

248

#### 249 **Interpreting test positivity with low prevalence in the general population**

250 Low prevalence places a high burden on specificity<sup>35</sup>. Given the high proportion of true  
251 negative individuals in the population being studied, prevalence increases the ratio of  
252 false positive to true positive test outcomes<sup>41</sup>. Positive predictive values (PPV)  
253 correspond to the likelihood that a positive test result reflect true positivity as measured  
254 by a gold standard PCR result. We computed PPVs as a function of the cumulative  
255 fraction of the population infected with (and assumed to have produced antibodies to)  
256 SARS-CoV-2 (**Figure 4**). Here we see that even with the conservative interpretation  
257 of LFA outcome (requiring the majority of operators to see a band to call a sample  
258 positive for either IgG or IgM), when the population seroprevalence is ~5% the PPV  
259 for these assays spans a large range (from ~30% to 100%). As the population  
260 prevalence increases, the burden on specificity is decreased, and at 50% prevalence  
261 the PPV of all LFAs is above 87.5%. Posterior PPV can be improved for most LFAs

262 by requiring both IgG and IgM to be read as positive in order to count an individual as  
263 positive (**Figure 4, right panels**).

264 To visualize the effect of changing population prevalence on the PPV, we  
265 created an interactive webapp (<https://covid.omics.kitchen>; **Figure 5**), which allows  
266 the user to extrapolate the likelihood that they do in fact have SARS-CoV-2 antibodies  
267 if they have a positive LFA result, given the infection prevalence and test accuracy.  
268 The app includes benchmark performance of all 20 LFAs evaluated here, as well as  
269 those reported in Whitman et al<sup>16</sup> (filtered to remove samples taken under 10 days  
270 post symptom-onset). In an effort to further generalize the utility of this tool, we allow  
271 the user to explore the effects of assay performance under difference prevalence  
272 scenarios, and to input the reported prevalence data from specific geographic  
273 locations within the US<sup>1</sup>, based on national, state, and county records.

274

## 275 **Discussion**

276 In this study, we report a standardized cross-evaluation of LFAs on the same  
277 pre-pandemic SARS-CoV-2-negative and PCR-confirmed SARS-CoV-2-positive  
278 samples, and rate their reproducibility, usability, and performance characteristics.  
279 Overall, the LFAs showed a higher propensity for false negative than false positive  
280 readings. Results are public: <https://covidinnovation.partners.org/evaluation/>. We use  
281 the Simoa technology<sup>39</sup> to measure the concentrations of anti-spike protein IgG and  
282 IgM antibodies and extrapolate the assays' limits of detection. We also established a  
283 web tool to aid users in understanding the likelihood they have anti-SARS-CoV-2  
284 antibodies given a positive test result. This resource of performance characteristics of



285 several LFAs and a tool for result interpretation, can both be used for  
286 immunosurveillance and future home testing applications.

287 LFAs are tractable tools to estimate community seroprevalence, especially with  
288 anticipated seasonal fluctuations in the transmission dynamics of SARS-CoV-2 and  
289 other viruses that cause the common cold, which confound the symptomatologic  
290 diagnosis of COVID-19<sup>42</sup>. As new waves of the COVID-19 pandemic resurge around  
291 the globe<sup>1</sup>, and with commencing vaccinations against SARS-CoV-2 infections<sup>18, 19, 20,</sup>  
292 <sup>22</sup>, there is a renewed interest in serological tests to detect anti-SARS-CoV-2  
293 antibodies<sup>28, 43</sup>. Affordable LFAs offer an attractive option for monitoring the presence  
294 and longevity of anti-SARS-CoV-2 antibodies, and determining population-level herd  
295 immunity<sup>44</sup>. LFAs also obviate the need for complex laboratories to process the  
296 samples<sup>45</sup>. As the pandemic expands to previously unexposed communities, it is  
297 critical to use simple tools to monitor exposure dynamics and seroconversion in SARS-  
298 CoV-2-exposed individuals, as well as vaccine-induced immunity<sup>22</sup>. We tested a  
299 mixture of LFAs targeting SARS-CoV2 nucleocapsid, spike proteins, or both.  
300 Moderna's mRNA-1273 and Pfizer-BioNTech COVID-19 vaccines encode SARS-  
301 CoV2 spike proteins to induce anti-spike antibodies<sup>20, 46</sup>. Therefore, LFAs targeting the  
302 spike and nucleocapsid proteins of SARS-CoV2 could be used to differentiate vaccine-  
303 and infection-induced antibody responses, respectively.

304 Rigorous evaluation of these LFAs by manufacturer-independent parties is  
305 important. The US FDA independently reviews medical products before  
306 commercialization. The FDA used Emergency Use Authorization (EUA) authority to  
307 accelerate the implementation of diagnostic products during the pandemic.  
308 Commercial manufacturers were required to submit a completed EUA request<sup>30</sup>.

309 Unfortunately, the rush to market introduced many tests that did not meet typical US  
310 or international standards<sup>47</sup>. Therefore, the FDA and international regulatory agencies  
311 continue to update guidelines for authorization of new serological tests. Our evaluation  
312 plan mirrored the FDA guidelines for evaluating serological tests<sup>30</sup>. We included 10  
313 HIV-positive samples to test whether they have higher false positive results in SARS-  
314 CoV2-negative samples<sup>48</sup>, and did not detect higher false positive results.

315 The mere detection of IgG or IgM responses does not guarantee that  
316 neutralizing antibodies are present at protective titers<sup>13, 49</sup>. The study demographics  
317 suggest a slight over-representation of African Americans among cases, as reported<sup>50</sup>.  
318 However, the sample size was underpowered to formally determine the effect of race  
319 on test performance. In our analysis, IgM detectability was less sensitive and  
320 reproducible than IgGs across multiple LFAs, possibly due to both lower IgM titres,  
321 and lower limits of detection for the assays. Waning antibody responses have been  
322 reported in some SARS-CoV-2-infected individuals<sup>51, 52, 53, 54</sup>. Furthermore, reported  
323 cases of re-infection with SARS-CoV-2 suggest that prior exposure, and even  
324 seroconversion, do not universally protect against SARS-CoV-2 infection<sup>54, 55, 56</sup>. This  
325 could result from low antibody titers as shown in an immunocompromised patient<sup>57</sup>,  
326 low durability of infection-induced antibodies<sup>52, 53, 54</sup>, or low neutralizing potential of  
327 SARS-CoV-2 antibodies in some individuals<sup>44</sup>. IgG and IgM antibodies may also target  
328 irrelevant epitopes outside the spike and receptor binding domains, and consequently  
329 be less efficient at intercepting infection by the virus<sup>58</sup>. The WHO cautions against  
330 interpreting presence of antibodies, even neutralizing ones, as lower risk of re-  
331 infection and transmission<sup>59</sup>. The presence of antibodies could, however, be used for

332 rapid immunosurveillance to monitor extent of population transmission, particularly in  
333 asymptomatic but SARS-CoV2-exposed individuals<sup>6, 43</sup>.

334 One major concern about the deployment of these tests is the misinterpretation  
335 of positive results<sup>13, 49</sup>. As more tests move towards FDA clearance for home use,  
336 clear scientific communication about the result interpretation becomes more crucial<sup>31,</sup>  
337 <sup>49</sup>. A positive serological result does not necessarily mean active infection<sup>31, 41, 60</sup>.  
338 Although combined use of molecular and seroconversion results can be used to  
339 confirm the diagnosis of symptomatic and hospitalized individuals<sup>45</sup>, a positive  
340 serological test in the absence of symptoms dissociates the presence of the antibodies  
341 from the time of infection<sup>54</sup>. Additionally, it is important to understand the implication  
342 of false positive and false negative results, particularly in the context of a low-  
343 prevalence disease such as COVID-19<sup>35</sup>. Low prevalence decreases the negative  
344 predictive value of a test, but increases the rate of false positives<sup>35</sup>. A false positive  
345 serological test result may prematurely instill confidence that one has immunity against  
346 SARS-CoV-2 infection, thus resulting in behavioral changes that increase risk of  
347 transmission<sup>61</sup>. Hence, the probability that a person without antibodies will test  
348 negative on a serological test is more important than test sensitivity<sup>59, 60, 61</sup>.

349 Our study presents a few limitations. Although we successfully benchmarked the  
350 performance of the LFAs to a quantitative assay<sup>39</sup>, we did not determine the  
351 neutralizing potential of these antibodies. Secondly, samples were acquired when  
352 PCR testing was restricted to severely ill patients. For epidemiological studies and  
353 population surveillance, it will be important to evaluate assay performance on  
354 asymptomatic individuals.

355           In conclusion, our study provides a public resource to aid researchers,  
356 healthcare providers, public health professionals, and industries impacted by the  
357 pandemic such as airlines, in choosing the appropriate LFAs for their intended use  
358 cases.

359

360 **Acknowledgments**

361 This study was possible through the Diagnostics Pillar of the Mass General Brigham  
362 COVID Center for Innovation (MGBCCI) (<https://covidinnovation.partners.org/>),  
363 predominantly staffed by volunteers. The LFA horizon scan and selection was  
364 performed by the Direct-to-Consumer working group, and the Diagnostics Accelerator  
365 and Validation working group developed the protocols and performed the evaluations.  
366 The study is supported by funding from the Massachusetts Consortium on Pathogen  
367 Readiness (MassCPR) (<https://masscpr.hms.harvard.edu/>).

368

369

370 **Figure legends**

371 **Figure 1: LFA reproducibility.** Colors represent LFA outcome performed across two  
372 days (A & B) by two independent pairs of operators on a COVID positive (COVID+)  
373 pooled sample and a COVID negative (COVID-) pooled sample. On each day, ten  
374 technical replicates of each pool were performed. Blue represents replicates where  
375 both operators that day agreed on the outcome, and the outcome was correctly called  
376 positive or negative. Red represents occasions where both operators agreed, but the  
377 outcome was incorrectly called. Orange represents replicates where the operators did  
378 not agree on the LFA outcome. Light yellow represents an invalid test where control  
379 bands did not meet criteria for a valid test. IgG was the most reproducible isotype for  
380 all LFAs except Oranoxis, and consistent false negatives were common occurrences  
381 for IgM for the majority of LFAs.

382

383 **Figure 2: Per-individual LFA performance.** Colors represent scores computed for  
384 IgG (top panel) and IgM (lower panel), where dark red=+1 (operators all agree a band  
385 is present) and dark blue=-1 (operators all agree there is no band) with intermediate  
386 colors (pink and light blue) representing varying degrees of operator disagreement.  
387 Grey represents invalid runs. Samples are ordered within the COVID-HIV-, COVID-  
388 HIV+, and COVID+ (all HIV-) groups in order of decreasing average score across all  
389 LFAs for both antibodies. Clinical variables include age (<50yrs: light blue, 50-69yrs:  
390 intermediate blue, ≥70yrs: dark blue), sex (blue: male, pink: female), disease severity  
391 (hospitalized: light green, ICU: orange, ICU+respirator: red, deceased: black), and  
392 weeks post symptom-onset (dark green: 1-2 weeks, lightest green: ≥5 weeks).

393

394 **Figure 3: Determining the limit of detection of qualitative LFAs.** Samples were  
395 ranked from highest concentration of anti-spike antibody (determined by Simoa) on  
396 the right, to lowest concentration (x-axis). As the sample concentration decreases to  
397 the left, a cumulative count of false negatives is shown on the y-axis. IgG is shown on  
398 the left, IgM on the right. Note the difference in magnitude of the y-axes between IgG  
399 and IgM; these LFAs are generally more sensitive to IgG than IgM.

400

401 **Figure 4: Effect of a changing disease prevalence (x axis) on the population-**  
402 **level PPV of each test** for scores derived from IgG (left), IgM (middle), and the  
403 average score of both antibodies (right). As disease prevalence increases, the high  
404 burden on specificity of LFAs is reduced. Color coding: Grey: no improvement; Light  
405 blue: rescues one of the two antibodies; Dark blue: rescues both of the antibodies;  
406 Orange: performs worse than one of the antibodies.

407

408 **Figure 5: Example image of the interactive webapp** hosted at  
409 <https://covid.omics.kitchen>. Using the sliders, one is able to visualize the effect of  
410 changing the disease prevalence (and the test performance) on the resulting  
411 probability that, given a positive LFA test, the individual does in fact have SARS-CoV-  
412 2 antibodies. The figure shows an example of positive predictive value using the US  
413 national prevalence of 5% (on November 27<sup>th</sup>, 2020) and the performance  
414 characteristics of the InTec LFA.

415

416 **Tables:**

417

	COVID- HIV-	COVID- HIV+	COVID+ HIV-	Overall
	n = 46	n = 10	n = 56	n=112
<b>Sex</b>				
Female	28 (60.9%)	1 (10%)	31 (55.4%)	60
Male	18 (39.1%)	9 (90%)	25 (44.6%)	52
<b>Age</b>				
Mean (SD)	46.7 (13.4)	58.6 (5.3)	58.7 (20.4)	53.8 (17.8)
Median [Min, Max]	48.0 [19.0, 68.0]	58.0 [48.0,65.0]	57.5 [24.0, 98.0]	54.5 [19.0,98.0]
<b>Race</b>				
Asian	8 (17.4%)	0 (0%)	3 (5.4%)	11 (9.8%)
Black	10 (21.7%)	4 (40%)	16 (28.6%)	30 (26.8%)
LatinX	0 (0%)	0 (0%)	5 (8.9%)	5 (4.5%)
Unknown/Other	8 (17.4%)	1 (10%)	8 (14.3%)	17 (15.2%)
White	20 (43.5%)	5 (50%)	24 (42.9%)	49 (43.8%)

418 **Table 1:** Demographic information of all individuals whose plasma samples were used

419 for this study. Individuals are broken out by COVID+/- and HIV+/- status. Black

420 includes one mixed African American in the COVID-HIV+ group.

421



Interval between symptom onset and blood collection	7-14 days	15-21 days	22-28 days	29+ days	Overall
	n = 9	n = 18	n = 16	n = 13	n=56
<b>Sex</b>					
Female	6 (66.7%)	10 (55.6%)	11 (68.8%)	4 (30.8%)	31 (55.4%)
Male	3 (33.3%)	8 (44.4%)	5 (31.2%)	9 (69.2%)	25 (44.6%)
<b>Age</b>					
Mean (Standard Deviation)	72.0 (24.0)	53.0 (15.8)	52.9 (19.9)	64.7 (20.2)	58.7 (20.4)
Median [Min, Max]	85 [30, 98]	49 [30, 84]	54.5 [24, 88]	67 [29, 94]	57.5 [24, 98]
<b>Race</b>					
Asian	1 (11.1%)	1 (5.6%)	1 (6.2%)	0 (0%)	3 (5.4%)
Black	5 (55.6%)	5 (27.8%)	2 (12.5%)	4 (30.8%)	16 (28.6%)
LatinX	1 (11.1%)	3 (16.7%)	1 (6.2%)	0 (0%)	5 (8.9%)
Unknown/Other	1 (11.1%)	1 (5.6%)	2 (12.5%)	4 (30.8%)	8 (14.3%)
White	1 (11.1%)	8 (44.4%)	10 (62.5%)	5 (38.5%)	24 (42.9%)
<b>Primary COVID PCR test</b>					
Cepheid Xpert Xpress SARS-CoV-2	1 (11.1%)	0 (0%)	1 (6.2%)	0 (0%)	2 (3.6%)
Panther Fusion SARS-CoV-2	8 (88.9%)	17 (94.4%)	12 (75.0%)	12 (92.3%)	49 (87.5%)
In house LDT at the Broad Institute	0 (0%)	1 (5.6%)	3 (18.8%)	1 (7.7%)	5 (8.9%)
<b>Severity</b>					
Deceased	0 (0%)	0 (0%)	4 (25%)	0 (0%)	4 (7.1%)
Hospitalized, not ICU	7 (77.8%)	13 (72.2%)	2 (12.5%)	2 (15.4%)	24 (42.9%)
ICU	0 (0%)	0 (0%)	2 (12.5%)	2 (15.4%)	4 (7.1%)
ICU (intubated)	0 (0%)	1 (5.6%)	1 (6.2%)	1 (7.7%)	3 (5.4%)
ICU (intubated) - Recovered	0 (0%)	3 (16.7%)	1 (6.2%)	7 (53.8%)	11 (19.6%)
ICU (no intubation) - Recovered	1 (11.1%)	0 (0%)	2 (12.5%)	0 (0%)	3 (5.4%)
Outpatient	1 (11.1%)	1 (5.6%)	4 (25%)	1 (7.7%)	7 (12.5%)

422 **Table 2:** Clinical information for the COVID-positive individuals whose plasma was  
423 used in this study. Individuals are broken out by the number of days between symptom  
424 onset and blood collection.

	IgG lateral flow assay performance													
	Manufacturer	N	TP	TN	FP	FN	sensitivity	specificity	PPV	NPV	sensitivity_CI95	specificity_CI95	PPV_CI95	NPV_CI95
	API	112	52	56	0	4	0.929	1	1	0.933	(0.853-1)	(0.991-1)	(0.99-1)	(0.861-1)
	API (v2)	112	52	56	0	4	0.929	1	1	0.933	(0.853-1)	(0.991-1)	(0.99-1)	(0.861-1)
	BioHit	110	52	50	4	4	0.929	0.926	0.929	0.926	(0.853-1)	(0.847-1)	(0.853-1)	(0.847-1)
	BTNX	110	48	56	0	6	0.889	1	1	0.903	(0.796-0.982)	(0.991-1)	(0.99-1)	(0.821-0.985)
	Camtech	111	48	55	0	8	0.857	1	1	0.873	(0.756-0.958)	(0.991-1)	(0.99-1)	(0.783-0.963)
	CareHealth	112	53	53	3	3	0.946	0.946	0.946	0.946	(0.878-1)	(0.878-1)	(0.878-1)	(0.878-1)
	Cellex	112	53	52	4	3	0.946	0.929	0.93	0.945	(0.878-1)	(0.853-1)	(0.855-1)	(0.876-1)
	Edinburgh	108	47	51	2	8	0.855	0.962	0.959	0.864	(0.753-0.957)	(0.901-1)	(0.893-1)	(0.768-0.96)
	Genobio	111	42	56	0	13	0.764	1	1	0.812	(0.643-0.885)	(0.991-1)	(0.988-1)	(0.713-0.911)
	InTec	111	49	54	1	7	0.875	0.982	0.98	0.885	(0.779-0.971)	(0.938-1)	(0.931-1)	(0.797-0.973)
	KHB	112	53	54	2	3	0.946	0.964	0.964	0.947	(0.878-1)	(0.906-1)	(0.906-1)	(0.88-1)
	Lumiquick	111	54	53	3	1	0.982	0.946	0.947	0.981	(0.938-1)	(0.878-1)	(0.88-1)	(0.935-1)
	Oranoxis	110	40	55	0	15	0.727	1	1	0.786	(0.6-0.854)	(0.991-1)	(0.988-1)	(0.683-0.889)
	OZO	112	43	54	2	13	0.768	0.964	0.956	0.806	(0.649-0.887)	(0.906-1)	(0.885-1)	(0.704-0.908)
	Phamatech	95	38	47	0	10	0.792	1	1	0.825	(0.667-0.917)	(0.989-1)	(0.987-1)	(0.718-0.932)
	Ray Biotech	112	43	56	0	13	0.768	1	1	0.812	(0.649-0.887)	(0.991-1)	(0.988-1)	(0.713-0.911)
	Ray Biotech (v2)	111	51	55	0	5	0.911	1	1	0.917	(0.827-0.995)	(0.991-1)	(0.99-1)	(0.839-0.995)
	U2U	112	43	56	0	13	0.768	1	1	0.812	(0.649-0.887)	(0.991-1)	(0.988-1)	(0.713-0.911)
	Vivachek	111	52	55	1	3	0.945	0.982	0.981	0.948	(0.876-1)	(0.938-1)	(0.935-1)	(0.882-1)
	Zhuhai Livzon	109	47	54	0	8	0.855	1	1	0.871	(0.753-0.957)	(0.991-1)	(0.989-1)	(0.779-0.963)
	IgM lateral flow assay performance													
	Manufacturer	N	TP	TN	FP	FN	sensitivity	specificity	PPV	NPV	sensitivity_CI95	specificity_CI95	PPV_CI95	NPV_CI95
	API	112	27	56	0	29	0.482	1	1	0.659	(0.342-0.622)	(0.991-1)	(0.981-1)	(0.552-0.766)
	API (v2)	112	51	55	1	5	0.911	0.982	0.981	0.917	(0.827-0.995)	(0.938-1)	(0.934-1)	(0.839-0.995)
	BioHit	110	54	49	5	2	0.964	0.907	0.915	0.961	(0.906-1)	(0.82-0.994)	(0.835-0.995)	(0.898-1)
	BTNX	110	48	53	3	6	0.889	0.946	0.941	0.898	(0.796-0.982)	(0.878-1)	(0.867-1)	(0.812-0.984)
	Camtech	111	42	52	3	14	0.75	0.945	0.933	0.788	(0.628-0.872)	(0.876-1)	(0.849-1)	(0.682-0.894)
	CareHealth	112	50	56	0	6	0.893	1	1	0.903	(0.803-0.983)	(0.991-1)	(0.99-1)	(0.821-0.985)
	Cellex	112	31	56	0	25	0.554	1	1	0.691	(0.415-0.693)	(0.991-1)	(0.984-1)	(0.584-0.798)
	Edinburgh	108	29	49	4	26	0.527	0.925	0.879	0.653	(0.386-0.668)	(0.845-1)	(0.753-1)	(0.539-0.767)
	Genobio	111	31	52	4	24	0.564	0.929	0.886	0.684	(0.424-0.704)	(0.853-1)	(0.766-1)	(0.573-0.795)
	InTec	111	43	54	1	13	0.768	0.982	0.977	0.806	(0.649-0.887)	(0.938-1)	(0.921-1)	(0.704-0.908)
	KHB	112	31	55	1	25	0.554	0.982	0.969	0.688	(0.415-0.693)	(0.938-1)	(0.893-1)	(0.58-0.796)
	Lumiquick	111	37	56	0	18	0.673	1	1	0.757	(0.54-0.806)	(0.991-1)	(0.986-1)	(0.653-0.861)
	Oranoxis	110	48	55	0	7	0.873	1	1	0.887	(0.776-0.97)	(0.991-1)	(0.99-1)	(0.8-0.974)
	OZO	112	13	54	2	43	0.232	0.964	0.867	0.557	(0.113-0.351)	(0.906-1)	(0.662-1)	(0.453-0.661)
	Phamatech	95	38	46	1	10	0.792	0.979	0.974	0.821	(0.667-0.917)	(0.927-1)	(0.911-1)	(0.712-0.93)
	Ray Biotech	112	40	55	1	16	0.714	0.982	0.976	0.775	(0.587-0.841)	(0.938-1)	(0.917-1)	(0.671-0.879)
	Ray Biotech (v2)	111	47	55	0	9	0.839	1	1	0.859	(0.734-0.944)	(0.991-1)	(0.989-1)	(0.766-0.952)
	U2U	112	13	56	0	43	0.232	1	1	0.566	(0.113-0.351)	(0.991-1)	(0.962-1)	(0.463-0.669)
	Vivachek	111	51	55	1	4	0.927	0.982	0.981	0.932	(0.849-1)	(0.938-1)	(0.934-1)	(0.859-1)
	Zhuhai Livzon	109	46	54	0	9	0.836	1	1	0.857	(0.729-0.943)	(0.991-1)	(0.989-1)	(0.763-0.951)
	Combined IgG + IgM performance													
	Manufacturer	N	TP	TN	FP	FN	sensitivity	specificity	PPV	NPV	sensitivity_CI95	specificity_CI95	PPV_CI95	NPV_CI95
	API	112	36	56	0	20	0.643	1	1	0.737	(0.509-0.777)	(0.991-1)	(0.986-1)	(0.631-0.843)
	API (v2)	112	51	56	0	5	0.911	1	1	0.918	(0.827-0.995)	(0.991-1)	(0.99-1)	(0.841-0.995)
	BioHit	110	53	50	4	3	0.946	0.926	0.93	0.943	(0.878-1)	(0.847-1)	(0.855-1)	(0.871-1)
	BTNX	110	47	56	0	7	0.87	1	1	0.889	(0.771-0.969)	(0.991-1)	(0.989-1)	(0.803-0.975)
	Camtech	111	44	55	0	12	0.786	1	1	0.821	(0.67-0.902)	(0.991-1)	(0.989-1)	(0.722-0.92)
	CareHealth	112	51	56	0	5	0.911	1	1	0.918	(0.827-0.995)	(0.991-1)	(0.99-1)	(0.841-0.995)
	Cellex	112	46	55	1	10	0.821	0.982	0.979	0.846	(0.712-0.93)	(0.938-1)	(0.927-1)	(0.751-0.941)
	Edinburgh	108	35	51	2	20	0.636	0.962	0.946	0.718	(0.5-0.772)	(0.901-1)	(0.86-1)	(0.606-0.83)
	Genobio	111	39	56	0	16	0.709	1	1	0.778	(0.58-0.838)	(0.991-1)	(0.987-1)	(0.675-0.881)
	InTec	111	46	54	1	10	0.821	0.982	0.979	0.844	(0.712-0.93)	(0.938-1)	(0.927-1)	(0.747-0.941)
	KHB	112	39	56	0	17	0.696	1	1	0.767	(0.567-0.825)	(0.991-1)	(0.987-1)	(0.663-0.871)
	Lumiquick	111	40	56	0	15	0.727	1	1	0.789	(0.6-0.854)	(0.991-1)	(0.988-1)	(0.687-0.891)
	Oranoxis	110	43	55	0	12	0.782	1	1	0.821	(0.664-0.9)	(0.991-1)	(0.988-1)	(0.722-0.92)
	OZO	112	19	56	0	37	0.339	1	1	0.602	(0.206-0.472)	(0.991-1)	(0.974-1)	(0.497-0.707)
	Phamatech	95	36	47	0	12	0.75	1	1	0.797	(0.617-0.883)	(0.989-1)	(0.986-1)	(0.686-0.908)

Ray Biotech	112	41	56	0	15	0.732	1	1	0.789	(0.607-0.857)	(0.991-1)	(0.988-1)	(0.687-0.891)
Ray Biotech (v2)	111	48	55	0	8	0.857	1	1	0.873	(0.756-0.958)	(0.991-1)	(0.99-1)	(0.783-0.963)
U2U	112	22	56	0	34	0.393	1	1	0.622	(0.256-0.53)	(0.991-1)	(0.977-1)	(0.516-0.728)
Vivachek	111	52	55	1	3	0.945	0.982	0.981	0.948	(0.876-1)	(0.938-1)	(0.935-1)	(0.882-1)
Zhuhai Livzon	109	45	54	0	10	0.818	1	1	0.844	(0.707-0.929)	(0.991-1)	(0.989-1)	(0.747-0.941)

425

426 **Table 3:** LFA performance in terms of sensitivity and specificity. N: Total number of  
 427 valid assays; TP: True Positive; TN: True Negative; FP: False Positive; FN: False  
 428 Negative; PPV: Positive Predictive Value; NPV: Negative Predictive Value; and CI95:  
 429 95% confidence interval.

430

431

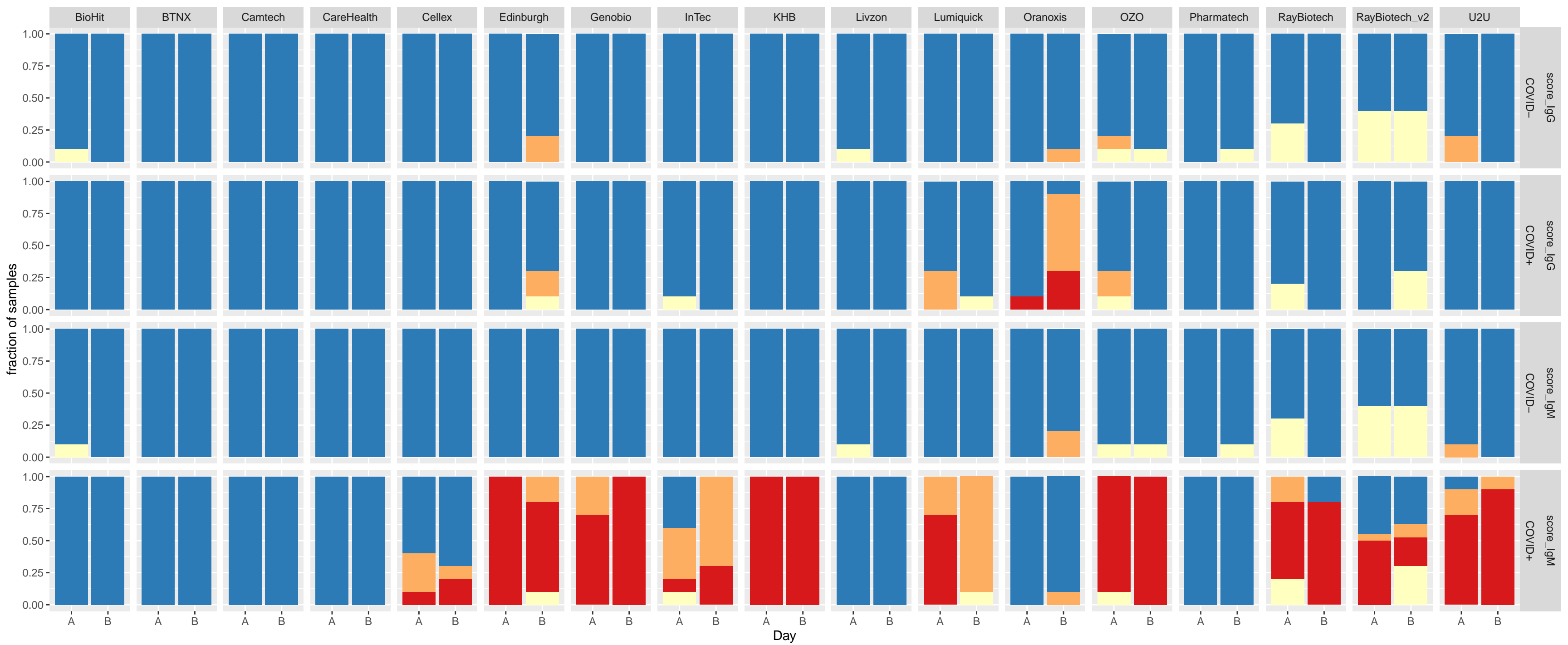
432 **References:**

- 433  
434 1. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real  
435 time. *Lancet Infect Dis* **20**, 533-534 (2020).  
436  
437 2. Editors. Dying in a Leadership Vacuum. *N Engl J Med* **383**, 1479-1480 (2020).  
438  
439 3. Looi, M.K. Covid-19: Is a second wave hitting Europe? *BMJ* **371**, m4113 (2020).  
440  
441 4. Wiersinga, W.J., Rhodes, A., Cheng, A.C., Peacock, S.J. & Prescott, H.C. Pathophysiology,  
442 Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review.  
443 *JAMA* **324**, 782-793 (2020).  
444  
445 5. Oran, D.P. & Topol, E.J. Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative  
446 Review. *Ann Intern Med* **173**, 362-367 (2020).  
447  
448 6. Havers, F.P. *et al.* Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United  
449 States, March 23-May 12, 2020. *JAMA Intern Med* (2020).  
450  
451 7. Ehrlich, H., Boneva, D. & Elkbuli, A. The intersection of viral illnesses: A seasonal influenza  
452 epidemic amidst the COVID-19 pandemic. *Ann Med Surg (Lond)* **60**, 41-43 (2020).  
453  
454 8. Tahmasebi, S., Khosh, E. & Esmaeilzadeh, A. The outlook for diagnostic purposes of the 2019-  
455 novel coronavirus disease. *J Cell Physiol* **235**, 9211-9229 (2020).  
456  
457 9. Kumar, R., Nagpal, S., Kaushik, S. & Mendiratta, S. COVID-19 diagnostic approaches: different  
458 roads to the same destination. *Virusdisease* **31**, 97-105 (2020).  
459  
460 10. Hellou, M.M. *et al.* Nucleic-acid-amplification tests from respiratory samples for the diagnosis  
461 of coronavirus infections: systematic review and meta-analysis. *Clin Microbiol Infect* (2020).  
462  
463 11. D'Cruz, R.J., Currier, A.W. & Sampson, V.B. Laboratory Testing Methods for Novel Severe  
464 Acute Respiratory Syndrome-Coronavirus-2 (SARS-CoV-2). *Front Cell Dev Biol* **8**, 468 (2020).  
465  
466 12. Pecora, N.D. & Zand, M.S. Measuring the Serologic Response to Severe Acute Respiratory  
467 Syndrome Coronavirus 2: Methods and Meaning. *Clin Lab Med* **40**, 603-614 (2020).  
468  
469 13. Peeling, R.W. *et al.* Serology testing in the COVID-19 pandemic response. *Lancet Infect Dis*  
470 **20**, e245-e249 (2020).  
471  
472 14. La Marca, A. *et al.* Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide  
473 to molecular and serological in-vitro diagnostic assays. *Reprod Biomed Online* **41**, 483-499  
474 (2020).  
475  
476 15. Pickering, S. *et al.* Comparative assessment of multiple COVID-19 serological technologies  
477 supports continued evaluation of point-of-care lateral flow assays in hospital and community  
478 healthcare settings. *PLoS Pathog* **16**, e1008817 (2020).  
479  
480 16. Whitman, J.D. *et al.* Evaluation of SARS-CoV-2 serology assays reveals a range of test  
481 performance. *Nat Biotechnol* **38**, 1174-1183 (2020).  
482  
483 17. Oliver, S.E. *et al.* The Advisory Committee on Immunization Practices' Interim  
484 Recommendation for Use of Pfizer-BioNTech COVID-19 Vaccine - United States, December  
485 2020. *MMWR Morb Mortal Wkly Rep* **69**, 1922-1924 (2020).  
486  
487 18. Jackson, L.A. *et al.* An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. *N Engl J*  
488 *Med* **383**, 1920-1931 (2020).

- 489  
490 19. Anderson, E.J. *et al.* Safety and Immunogenicity of SARS-CoV-2 mRNA-1273 Vaccine in Older  
491 Adults. *N Engl J Med* (2020).  
492  
493 20. Walsh, E.E. *et al.* Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine  
494 Candidates. *N Engl J Med* (2020).  
495  
496 21. Dagotto, G., Yu, J. & Barouch, D.H. Approaches and Challenges in SARS-CoV-2 Vaccine  
497 Development. *Cell Host Microbe* **28**, 364-370 (2020).  
498  
499 22. Poland, G.A., Ovsyannikova, I.G. & Kennedy, R.B. SARS-CoV-2 immunity: review and  
500 applications to phase 3 vaccine candidates. *Lancet* **396**, 1595-1606 (2020).  
501  
502 23. Zohar, T. *et al.* Compromised Humoral Functional Evolution Tracks with SARS-CoV-2 Mortality.  
503 *Cell* **183**, 1508-1519 e1512 (2020).  
504  
505 24. Lucas, C. *et al.* Kinetics of antibody responses dictate COVID-19 outcome. *MedRxiv* (2020).  
506  
507 25. Wang, H., Ai, J., Loeffelholz, M.J., Tang, Y.W. & Zhang, W. Meta-analysis of diagnostic  
508 performance of serology tests for COVID-19: impact of assay design and post-symptom-onset  
509 intervals. *Emerg Microbes Infect* **9**, 2200-2211 (2020).  
510  
511 26. Caini, S. *et al.* Meta-analysis of diagnostic performance of serological tests for SARS-CoV-2  
512 antibodies up to 25 April 2020 and public health implications. *Euro Surveill* **25** (2020).  
513  
514 27. Lisboa Bastos, M. *et al.* Diagnostic accuracy of serological tests for covid-19: systematic review  
515 and meta-analysis. *BMJ* **370**, m2516 (2020).  
516  
517 28. Moura, D.T.H. *et al.* Diagnostic Characteristics of Serological-Based COVID-19 Testing: A  
518 Systematic Review and Meta-Analysis. *Clinics (Sao Paulo)* **75**, e2212 (2020).  
519  
520 29. Gutierrez-Cobos, A. *et al.* Evaluation of diagnostic accuracy of 10 serological assays for  
521 detection of SARS-CoV-2 antibodies. *Eur J Clin Microbiol Infect Dis* (2020).  
522  
523 30. FDA. EUA Authorized Serology Test Performance. *United States Food and Drug Administration*  
524 [https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-](https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance)  
525 [use-authorizations-medical-devices/eua-authorized-serology-test-performance](https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance) (2020).  
526  
527 31. Bermingham, W.H., Wilding, T., Beck, S. & Huissoon, A. SARS-CoV-2 serology: Test, test, test,  
528 but interpret with caution! *Clin Med (Lond)* **20**, 365-368 (2020).  
529  
530 32. Long, Q.X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* **26**,  
531 845-848 (2020).  
532  
533 33. Lou, B. *et al.* Serology characteristics of SARS-CoV-2 infection after exposure and post-  
534 symptom onset. *Eur Respir J* **56** (2020).  
535  
536 34. Zhao, J. *et al.* Antibody Responses to SARS-CoV-2 in Patients With Novel Coronavirus  
537 Disease 2019. *Clinical infectious diseases : an official publication of the Infectious Diseases*  
538 *Society of America* **71**, 2027-2034 (2020).  
539  
540 35. Tenny, S. & Hoffman, M.R. Prevalence. *StatPearls: Treasure Island (FL)*, 2020.  
541  
542 36. Theel, E.S. *et al.* The Role of Antibody Testing for SARS-CoV-2: Is There One? *J Clin Microbiol*  
543 **58** (2020).  
544

- 545 37. Jorfi, M. *et al.* Diagnostic technology for COVID-19: comparative evaluation of antigen and  
546 serology-based SARS-CoV-2 immunoassays, and contact tracing solutions for potential use as  
547 at-home products. *MedRxiv* (2020).  
548
- 549 38. Charlson, M.E., Pompei, P., Ales, K.L. & MacKenzie, C.R. A new method of classifying  
550 prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* **40**,  
551 373-383 (1987).  
552
- 553 39. Norman, M. *et al.* Ultrasensitive high-resolution profiling of early seroconversion in patients with  
554 COVID-19. *Nat Biomed Eng* (2020).  
555
- 556 40. Tian, X. *et al.* Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-  
557 specific human monoclonal antibody. *Emerg Microbes Infect* **9**, 382-385 (2020).  
558
- 559 41. Hernaez, R. & Thrift, A.P. High Negative Predictive Value, Low Prevalence, and Spectrum  
560 Effect: Caution in the Interpretation. *Clin Gastroenterol Hepatol* **15**, 1355-1358 (2017).  
561
- 562 42. Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H. & Lipsitch, M. Projecting the transmission  
563 dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860-868 (2020).  
564
- 565 43. Alter, G. & Seder, R. The Power of Antibody-Based Surveillance. *N Engl J Med* **383**, 1782-  
566 1784 (2020).  
567
- 568 44. Neagu, M. The bumpy road to achieve herd immunity in COVID-19. *J Immunoassay*  
569 *Immunochem*, 1-18 (2020).  
570
- 571 45. Nilles, E.J. *et al.* Evaluation of two commercial and two non-commercial immunoassays for the  
572 detection of prior infection to SARS-CoV-2. *medRxiv* (2020).  
573
- 574 46. Corbett, K.S. *et al.* SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen  
575 preparedness. *Nature* **586**, 567-571 (2020).  
576
- 577 47. Badnjevic, A., Pokvic, L.G., Dzemic, Z. & Becic, F. Risks of emergency use authorizations for  
578 medical products during outbreak situations: a COVID-19 case study. *Biomed Eng Online* **19**,  
579 75 (2020).  
580
- 581 48. Tan, S.S., Chew, K.L., Saw, S., Jureen, R. & Sethi, S. Cross-reactivity of SARS-CoV-2 with  
582 HIV chemiluminescent assay leading to false-positive results. *J Clin Pathol* (2020).  
583
- 584 49. Gill, D. & Ponsford, M.J. Testing for antibodies to SARS-CoV-2. *BMJ* **371**, m4288 (2020).  
585
- 586 50. Kullar, R. *et al.* Racial Disparity of Coronavirus Disease 2019 in African American Communities.  
587 *J Infect Dis* **222**, 890-893 (2020).  
588
- 589 51. Long, Q.X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2  
590 infections. *Nat Med* **26**, 1200-1204 (2020).  
591
- 592 52. Ibarondo, F.J. *et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-  
593 19. *N Engl J Med* **383**, 1085-1087 (2020).  
594
- 595 53. Kutsuna, S., Asai, Y. & Matsunaga, A. Loss of Anti-SARS-CoV-2 Antibodies in Mild Covid-19.  
596 *N Engl J Med* **383**, 1695-1696 (2020).  
597
- 598 54. Milani, G.P. *et al.* Serological follow-up of SARS-CoV-2 asymptomatic subjects. *Sci Rep* **10**,  
599 20048 (2020).  
600

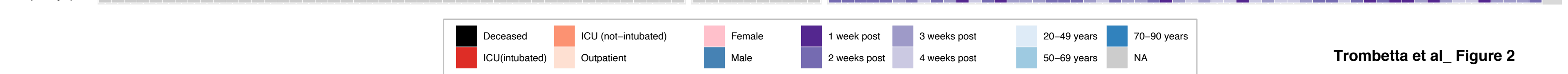
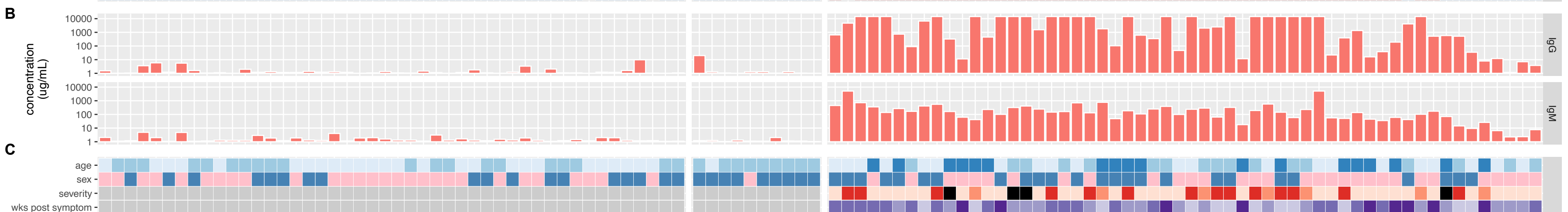
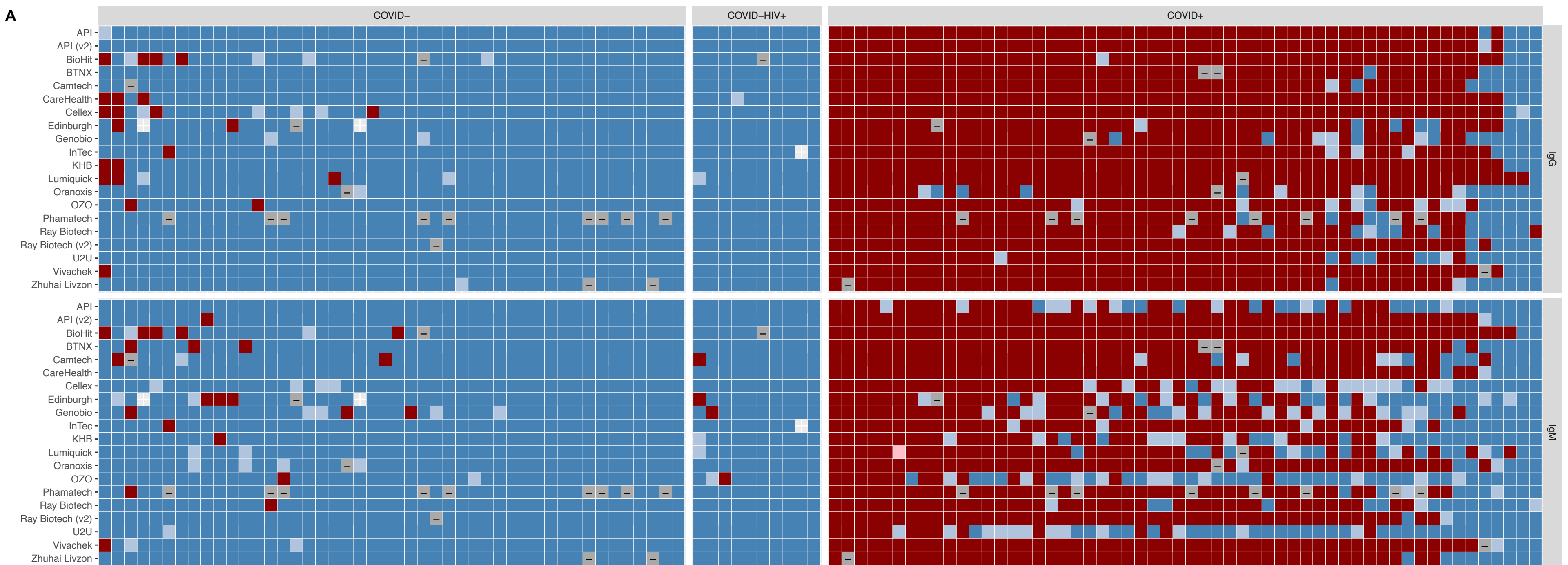
- 601 55. To, K.K. *et al.* COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain  
602 confirmed by whole genome sequencing. *Clinical infectious diseases : an official publication of*  
603 *the Infectious Diseases Society of America* (2020).  
604
- 605 56. Cohen, J.I. & Burbelo, P.D. Reinfection with SARS-CoV-2: Implications for Vaccines. *Clinical*  
606 *infectious diseases : an official publication of the Infectious Diseases Society of America* (2020).  
607
- 608 57. Mulder, M. *et al.* Reinfection of SARS-CoV-2 in an immunocompromised patient: a case report.  
609 *Clinical infectious diseases : an official publication of the Infectious Diseases Society of*  
610 *America* (2020).  
611
- 612 58. Hussain, A. *et al.* Targeting SARS-CoV2 Spike Protein Receptor Binding Domain by  
613 Therapeutic Antibodies. *Biomed Pharmacother* **130**, 110559 (2020).  
614
- 615 59. WHO. "Immunity passports" in the context of COVID-19. *World Health Organization*  
616 *commentaries* [https://www.who.int/news-room/commentaries/detail/immunity-](https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19)  
617 [passports-in-the-context-of-covid-19](https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19). (2020).  
618
- 619 60. Waller, J., Rubin, G.J., Potts, H.W.W., Mottershaw, A.L. & Marteau, T.M. 'Immunity Passports'  
620 for SARS-CoV-2: an online experimental study of the impact of antibody test terminology on  
621 perceived risk and behaviour. *BMJ Open* **10**, e040448 (2020).  
622
- 623 61. Brown, R.C.H., Kelly, D., Wilkinson, D. & Savulescu, J. The scientific and ethical feasibility of  
624 immunity passports. *Lancet Infect Dis* (2020).  
625  
626



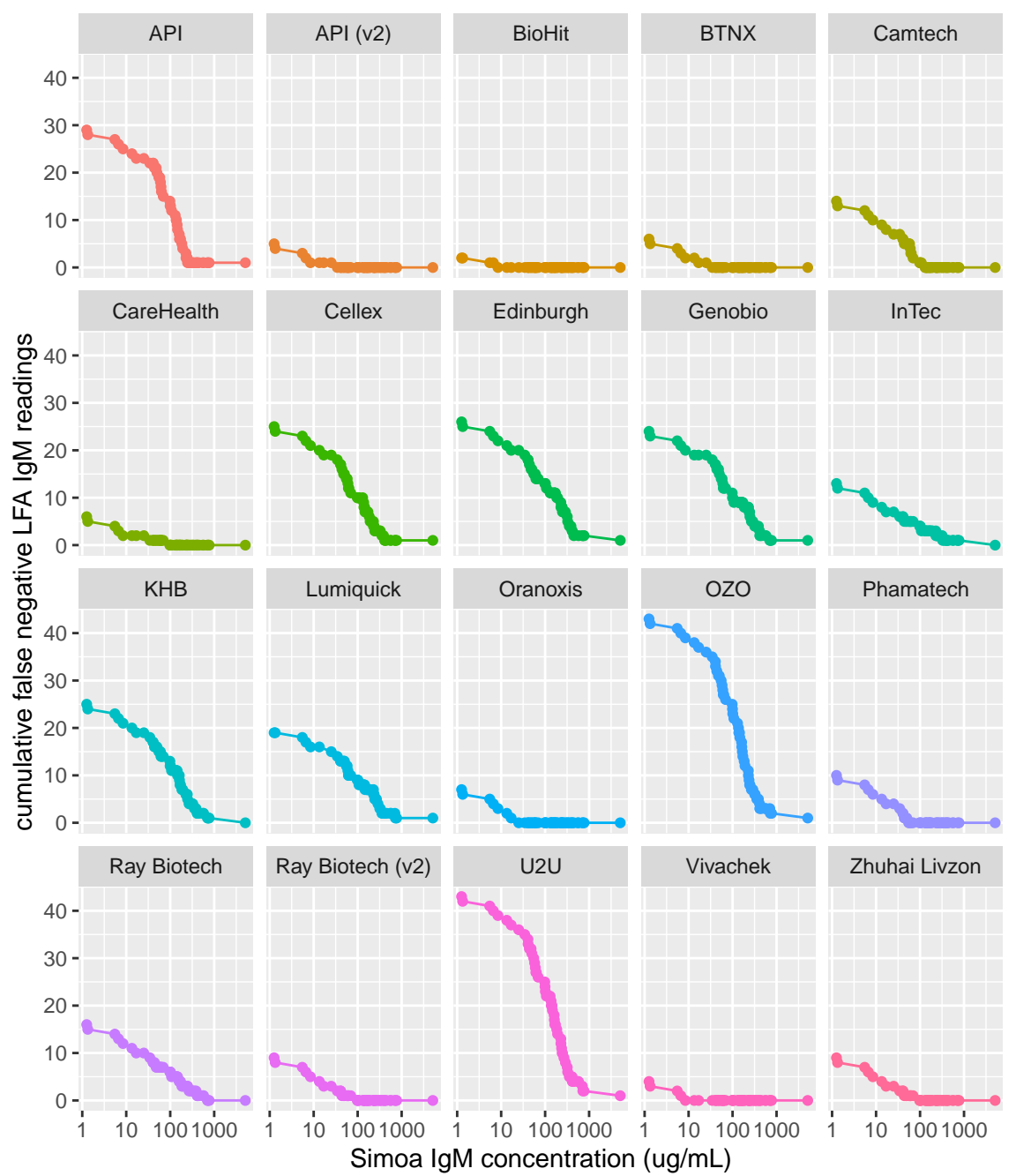
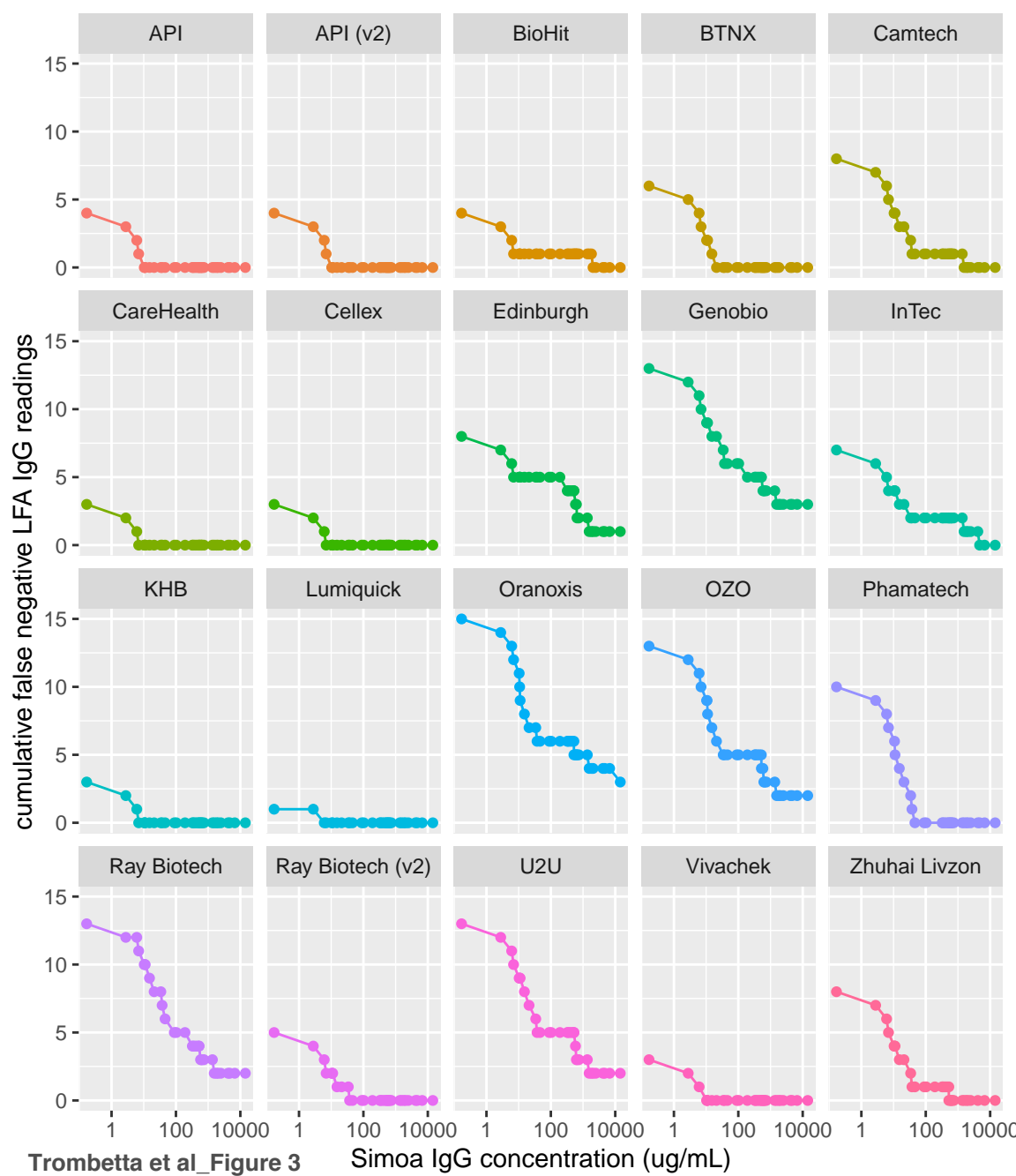
outcome ■ Correct Consistent ■ Inconsistent ■ Incorrect Consistent ■ Invalid

Trombetta et al\_Figure 1

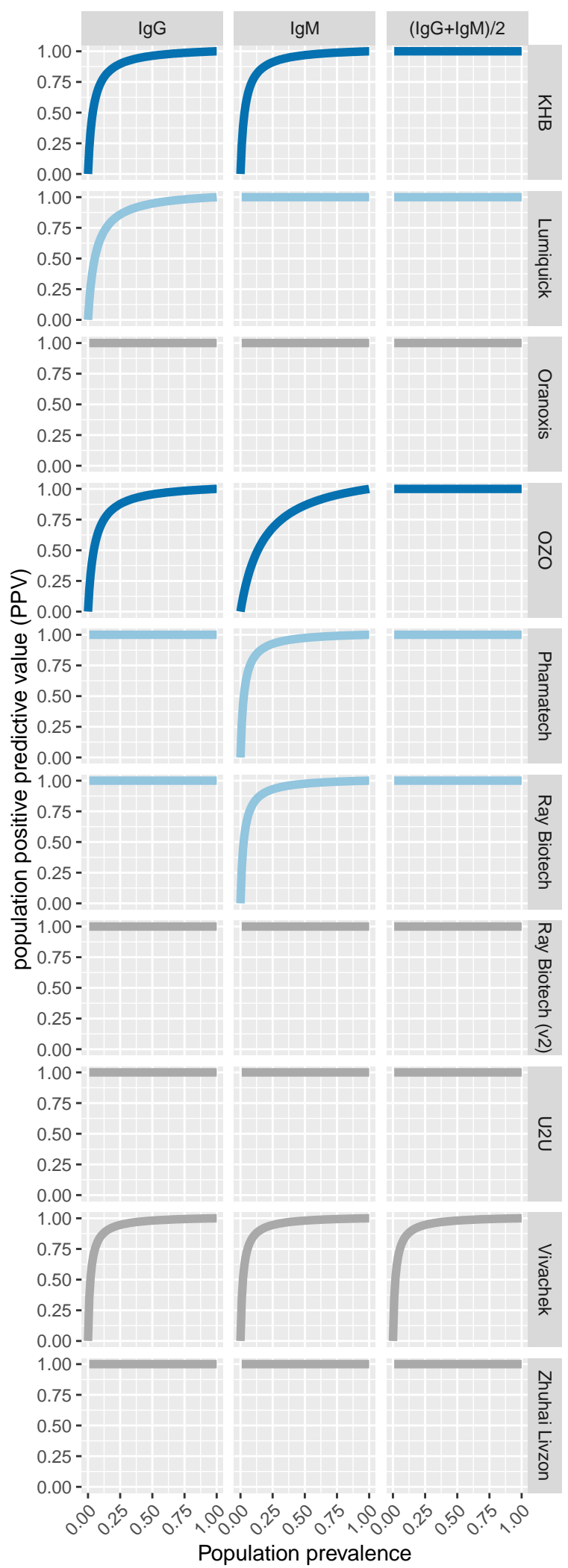
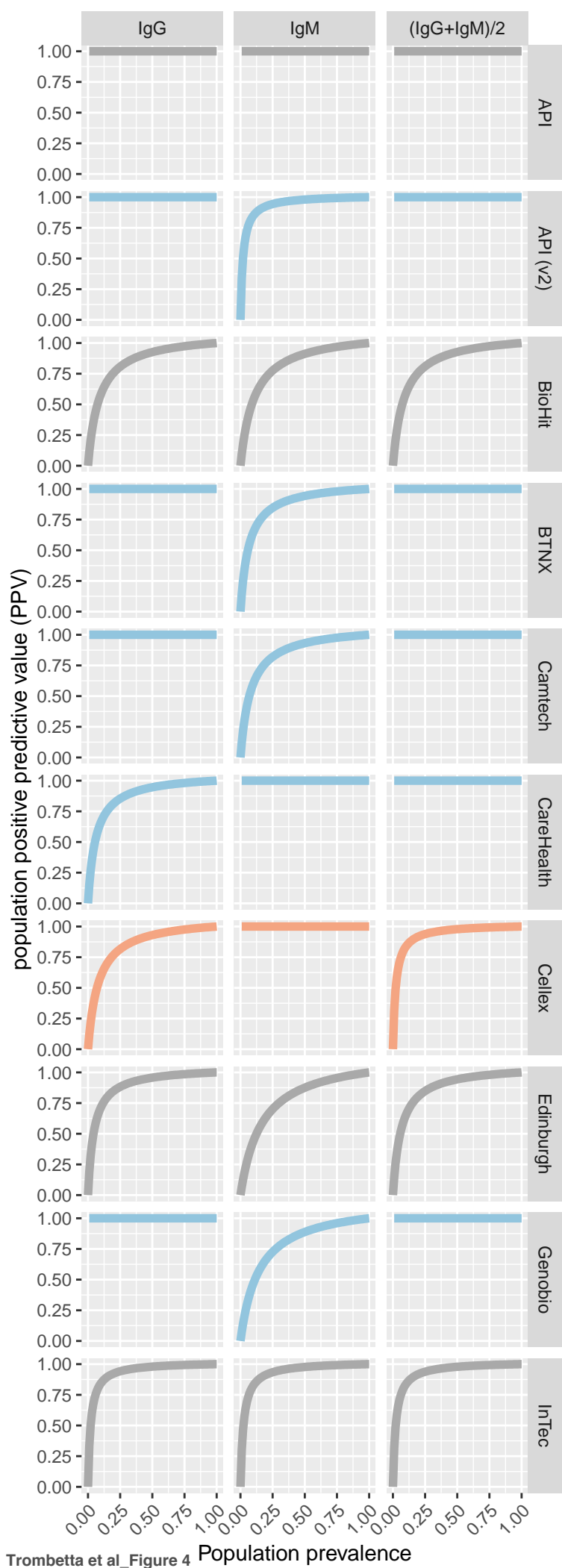




Trombetta et al\_ Figure 2



Trombetta et al\_Figure 3

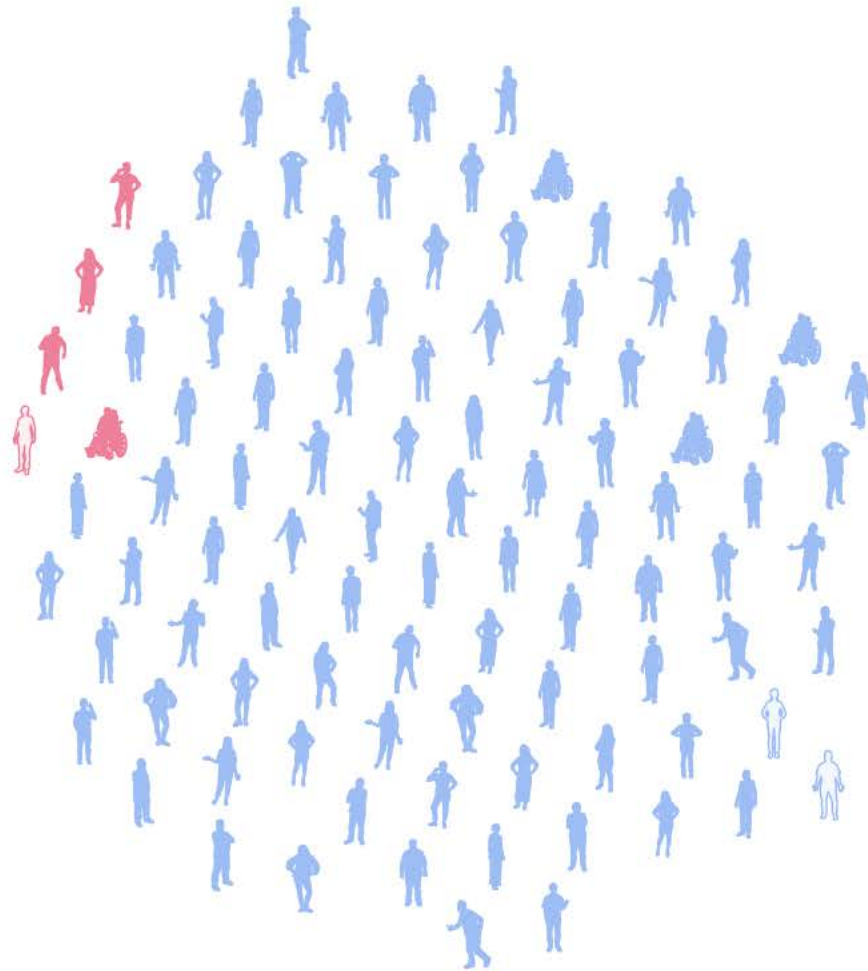


Trombetta et al\_Figure 4 Population prevalence

Population prevalence

# If my antibody test is **positive**, what is the probability that I have SARS-CoV-2 antibodies?

# 70%



The lateral flow assay made by InTec for IgG or IgM antibodies has a sensitivity of **82%** and specificity of **98%**.

Based on COVID-19 prevalence nationwide and this test's performance metrics, there is a **70%** probability that you actually have SARS-CoV-2 antibodies if you test positive.

Nationwide, it's estimated that **5%** of the population has had COVID-19 caused by the SARS-CoV-2 virus. If everybody were to take **InTec's** antibody test, around **6%** of results would come back positive, including healthy individuals.

Use the sliders below to see how a test's specificity and sensitivity impacts the probability of having SARS-CoV-2 antibodies given a positive result.

5%



Covid-19 Prevalence in:

Sensitivity: 82% <sup>?</sup>



Specificity: 98% <sup>?</sup>



## Online Supplement:

### Supplementary Methods:

#### Analysis

Computational analyses were performed in R/markdown. LFA evaluation reports are available (<http://publicdata.omics.kitchen/Projects/MGBCCI/LFA/VendorReports/>).

Formulae for computing performance metrics and confidence intervals are provided below. IgG/IgM bands were called via visual inspection by two experienced human operators. Invalid LFAs (e.g. missing a control band) were excluded entirely from further analysis and the sample was not re-run.

A score was computed for each sample/assay/antibody combination, using the following algorithm:

- Each of N operators reading the LFA were assigned a weight of  $1/N$  to final score
- If an operator observed a band, the score would increase by  $1/N$ . If the operator saw no band, the score would decrease by  $1/N$

This process was performed separately for the IgG and IgM channels, with an overall score produced for each antibody. This per-sample/per-antibody score thus has the following values: -1 (all operators agree: no band), 0 (operators disagree), and +1 (all operators agree: band observed). Given that COVID-19 remains, even in late 2020, a low-prevalence disease, we apply a conservative case definition (score  $\geq 0$ ), where discordant operator readings (score=0) are classed as negative for presence of the

antibody, in order to favor specificity. A combined IgG and IgM score was computed as the average of the two individual scores:  $(\text{IgG} + \text{IgM})/2$ .

Reproducibility analysis used the same scoring system above. Tests of the COVID+ pool that received a score  $>0$  were called “correct”. COVID- tests were ones that received a score  $<0$ . A score of 1 or -1 was called “consistent”, with all other scores indicating disagreement between operators called as “inconsistent.” These two classifications were concatenated to produce the outcomes “correct consistent”, where both operators agreed on the correct outcome, “incorrect consistent,” where both operators agreed on the incorrect outcome, and “inconsistent,” where operators disagreed on the outcome. The outcome of each individual test was plotted as a proportion of total tests of each pool on each of the two days of testing.

## **Webapp**

We developed an interactive web-application (<https://covid.omics.kitchen>) to help visualize false positive and false negative LFA readings based on test accuracy and disease prevalence. We incorporate data from the 20 LFAs reported here as well as the 6 LFAs evaluated by Whitman, et al<sup>16</sup>. County, State, and US-wide disease rates (cumulative numbers of individuals confirmed to have been infected since the start of the outbreak) are pulled dynamically from the New York Times github repository (<https://github.com/nytimes/covid-19-data>). The webapp is implemented in d3.js JavaScript and hosted on Amazon Web Services AWS/Amplify.

## Formula:

### Sensitivity

- $n = TP + FN$
- $sens = \frac{TP}{n}$
- $SE_{sens} = \sqrt{\frac{sens(1-sens)}{n}}$
- $CI_{95\%} = (sens - z_{\frac{\alpha}{2}} * SE_{sens} - \frac{1}{2n}, sens + z_{\frac{\alpha}{2}} * SE_{sens} + \frac{1}{2n})$

### Specificity

- $n = TN + FP$
- $spec = \frac{TN}{n}$
- $SE_{spec} = \sqrt{\frac{spec(1-spec)}{n}}$
- $CI_{95\%} = (spec - z_{\frac{\alpha}{2}} * SE_{spec} - \frac{1}{2n}, spec + z_{\frac{\alpha}{2}} * SE_{spec} + \frac{1}{2n})$

### PPV

- $n = TP + FP$
- $PPV = \frac{TP}{n}$
- $SE_{PPV} = \sqrt{\frac{PPV(1-PPV)}{n}}$
- $CI_{95\%} = (PPV - z_{\frac{\alpha}{2}} * SE_{PPV} - \frac{1}{2n}, PPV + z_{\frac{\alpha}{2}} * SE_{PPV} + \frac{1}{2n})$

### NPV

- $n = TN + FN$
- $NPV = \frac{TN}{n}$
- $SE_{NPV} = \sqrt{\frac{NPV(1-NPV)}{n}}$
- $CI_{95\%} = (NPV - z_{\frac{\alpha}{2}} * SE_{NPV} - \frac{1}{2n}, NPV + z_{\frac{\alpha}{2}} * SE_{NPV} + \frac{1}{2n})$

**Supplementary Table 1:** LFA commercial kit information, sample requirements, and protocol details

Manufacturer	Lot Number	Biological Target	Sample Requirement	IgG/IgM Cassette Structure	Total input volume (uL)	Incubation time (minutes )
API_V1	CoV1252004C	N and S	whole blood, plasma, serum	Same strip	10	15-20
API_V2	COV1252008B	N and S	whole blood, plasma, serum	Same strip	10	15-20
BioHit	SA200401	N and S	venous whole blood, plasma, serum	Same strip	10	15-20
BTNX	I2004027	N and S	whole blood, plasma, serum	Same strip	5	15
Camtech	CAM240420	S	plasma, serum	same strip	10	15
CareHealth	G070720	N	venous whole blood, plasma, serum	Same strip	5	10-15
Cellex	20200503WI5515C025	N and S	venous whole blood, plasma, serum	Same strip	10	15-20
Edinburgh	2000798A	N and S	whole blood (capillary and venous)	Same strip	20	10
Genobio	VMG200331	N and S	venous whole blood, plasma, serum	Same cassette, separate strips	20	10
InTec	ITP6002-TC25/GJ20030288	N and S	whole blood (venous and fingerstick), plasma, serum	Same cassette, separate strips	20	15-20
KHB	423200332	N	fresh whole blood, plasma, serum	Same strip	10	15
Livzon	CK2004240410	N and S	venous whole blood, plasma, serum	Separate cassettes	20	10-15
Lumiquick	2004219	N and S	venous whole blood, plasma, serum	Same strip	2	15
Oranoxis	RC-0220	S	plasma, serum	Same strip	10	15
Ozo	P2002	S	whole blood (capillary and venous), plasma, serum	Same cassette, separate strips	20	10
Pharmatech	D00647	N	whole blood (venous and fingerstick), plasma, serum	Same strip	10	10
RayBiotech_V1	501202955	N	whole blood (venous and fingerstick), serum	Same cassette, separate strips	25	10
RayBiotech_V2	715202954	N	whole blood (venous and fingerstick), serum	Same cassette, separate strips	25	10
U2U	172004-01	N and S	whole blood (venous and fingerstick), plasma, serum	Same strip	10	15
VivaCheck	SU2005001	N and S	whole blood (venous and fingerstick), serum, heparin plasma	Same strip	10	15



**Supplementary Table 2:** IFU clarity rubric. Kits were assigned one point for each criteria listed if met

	<b>Criteria for IFU clarity</b>
1	Includes explanation of intended use
2	Storage conditions of provided kit components are listed
3	Sample requirements and collection methods (e.g. whole blood, serum, plasma, equilibration temperature, etc.)
4	Storage information for specimens
5	Precautions/provides warnings and conditions to avoid (e.g. do not use expired cassettes, etc.)
6	Test procedure: temperature specified for running sample and cassette
7	Test procedure: clearly states when cassette should be read
8	Test procedure: clearly states valid time window (e.g. test results invalid after a certain interval of time)
9	Test procedure: provides visual diagrams for protocol steps
10	Provides instruction on how to proceed if results are invalid (e.g. repeat test, or increase incubation time)
11	Includes a rubric for interpretation of results
12	Provides visual diagrams or cassette images to assist with interpretation of results (has to include at least an example of a positive and negative readout)
13	Clearly describes correct volume to pipette (e.g. marked pipette, markings explained in the IFU, or dropwise measurement)
14	Complete instructions provided for lancets or other accessories included for use (e.g. instructions present for all kit components)

Supplementary Table 3. Usability ratings and description of kit components.

Manufacturer	Pipette type	If pipette provided, material	If pipette provided: usability (e.g. volume markings?)	Buffer type (dropper, aliquot)	Rater 1	Rater 2	Rater 3	IFU Clarity Final Ratings	Additional materials provided	Notes
APL_V1	Plastic dropper	Plastic	no markings	dropper	11	11	12	12		
APL_V2	None included	NA	NA	dropper	12	13	12	12	lancets, alcohol pads, bandaids	Cassette is double packaged in 2 sealed pouches, no pipette is provided
BioHit	Plastic dropper	Plastic	no markings	dropper	13	14	14	14		
BTNX	Plastic dropper	Plastic	volume markings	dropper	12	13	13	12		Outcome call requires interpretation of color - difficult for color blind?
Camtech	Plastic dropper	Plastic	no markings	dropper	9	12	10	12		Provides useful information on how long it takes materials to reach room temp.
CareHealth	Plastic dropper	Plastic	no markings	dropper	13	14	13	13		instructions say add 10ul of whole blood, but doesn't show what that looks like in the provided pipette
Cellex	Plastic dropper	Plastic	no markings	dropper	11	13	13	13		Provides visual ref to show that even faint bands can be positive.
										Requires 10ul whole blood but doesn't show what that looks like in the plastic pipette
Edinburgh	Plastic dropper	Plastic	volume markings	aliquot	10	13	10	10	cotton swab, lancet, band-aid, alcohol pad	Requires 10ul whole blood doesn't show how much that is in plastic pipette
Genobio	Plastic dropper	Plastic	no markings	dropper	8	11	9	10		Not clear if requires 10ul sample in total, or if the cassette should be "standing"
InTec	Plastic dropper	Plastic	no markings	dropper	14	14	14	14	lancet, alcohol pads	
KHB	Plastic dropper	Plastic	no markings	dropper	13	12	12	11	lancet, alcohol pads	The procedure contains typos, including: "add 10ul of plasma and perum specimen"
Livzon	Capillary pipette	Glass	volume markings	dropper	13	12	13	13		Large range of time in which to interpret samples (1-15 mins)
Lumiquick	Capillary pipette	Plastic	volume markings	dropper	14	14	14	14		
Oranoxis	Capillary pipette	Plastic	no markings	dropper	8	9	8	7		Conflicting guidance on time to read, no specific mentions of sample requirements/ storage
Ozo	Plastic dropper	Plastic	no markings	dropper	11	12	12	12	cotton swab, lancet, alcohol pads	No visual reference
										for invalid cassette
Phamatech	Plastic dropper	Plastic	no markings	dropper	14	14	14	14	lancets, alcohol pads	
RayBiotech_V1					13	13	13			
RayBiotech_V2	Plastic dropper	Plastic	volume markings	aliquot	14	13	13	13		
U2U	Capillary pipette	Plastic	volume markings	dropper	14	14	14	14		
VivaCheck	Plastic dropper	Plastic	no markings	dropper	11	12	11	10		No separate section for sample requirements or collection

**Supplementary Table 4.** LFA Limits of detection estimated from anti-spike antibodies concentrations measured by Simoa; amounts reflect lowest antibody concentrations in plasma samples that show a positive band

LFA manufacturer	LoD IgG ( $\mu\text{g/mL}$ )	LoD IgM ( $\mu\text{g/mL}$ )
API	4.4	5111.7
API (v2)	4.4	6.1
BioHit	4.4	0.6
BTNX	8.7	7.5
Camtech	27.5	83.8
CareHealth	1.5	7.5
Cellex	1.5	5111.7
Edinburgh	1453.8	5111.7
Genobio	14318.7	5111.7
InTec	27.5	2940.5
KHB	1.5	2940.5
Lumiquick	0.1	5111.7
Oranoxis	14318.7	10.9
OZO	14318.7	5111.7
Phamatech	27.5	42.4
Ray Biotech	14318.7	638.4
Ray Biotech (v2)	6.5	29.8
U2U	14318.7	5111.7
Vivachek	1.5	3.4
Zhuhai Livzon	27.5	29.8