

1 Evaluation of SARS-CoV-2 Antibody Point of Care Devices in the Laboratory and Clinical
2 Setting

3
4
5 Kirsty McCance¹, Helen Wise¹, Jennifer Simpson¹, Becky Bachelor², Harriet Hale², Lindsay
6 McDonald¹, Azul Zorzoli¹, Elizabeth Furrie³, Charu Chopra⁴, Frauke Muecksch⁵,
7 Theodora Hatzioannou⁵, Paul D. Bieniasz^{5,6} Kate Templeton¹, Sara Jenks¹

8
9 ¹Department of Biochemistry, Royal Infirmary of Edinburgh, NHS Lothian, Edinburgh,
10 Scotland

11 ²Western General Hospital, NHS Lothian, Edinburgh

12 ³Ninewells Hospital and Medical School, NHS Tayside, Dundee, Scotland

13 ⁴Department of Immunology, Royal Infirmary of Edinburgh, NHS Lothian

14 ⁵Laboratory of Retrovirology, The Rockefeller University, New York, New York, USA

15 ⁶Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA

16
17 *corresponding author: kirsty.mccance@nhslothian.scot.nhs.uk

18
19 **Abbreviations**

20 LFA (lateral flow assay), POC (point of care), PPV (positive predictive value), NPV
21 (negative predictive value), RT-PCR (real time polymerase chain reaction), NT50 (half
22 maximal neutralising titre), hCoV (Human Coronavirus), SARS-CoV-2 (severe acute
23 respiratory syndrome coronavirus 2) , COVID-19 (coronavirus disease 2019).

24

25 **Contributions:** SJ, KT, CC, BB and HW came up with the study idea. HW and AZ collated
26 company details and manufacturer info. KM and SJ assisted with study visits. HW, KM, JS
27 and LM carried out evaluation of the POC devices in the laboratory. KM, SJ, and EF carried
28 out capillary POC testing. FM, TH and PB carried out neutralising antibody assays. HW,
29 KM, JS and SJ carried out data analysis and wrote the first draft of the manuscript. All
30 authors reviewed the manuscript.

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50 **Abstract**

51

52 SARS-CoV-2 Antibody tests have been marketed to diagnose previous SARS-CoV-2
53 infection and as a test of immune status. There is a lack of evidence on the performance and
54 clinical utility of these tests. We aimed to carry out an evaluation of 14 point of care (POC)
55 SARS-CoV-2 antibody tests.

56

57 Serum from participants with previous RT-PCR (Real-Time Polymerase chain reaction)
58 confirmed SARS-CoV-2 infection and pre-pandemic controls were used to determine
59 specificity and sensitivity of each POC device. Changes in sensitivity with increasing time
60 from infection were determined on a cohort of participants. Corresponding neutralising
61 antibody status was measured to establish whether the detection of antibodies by the POC
62 device correlated with immune status. Paired capillary and serum samples were collected to
63 ascertain whether POC devices performed comparably on capillary samples.

64

65 Sensitivity and specificity varied between the POC devices and in general did not meet the
66 manufacturers reported performance characteristics signifying the importance of independent
67 evaluation of these tests. The sensitivity peaked at >20 days following symptoms onset
68 however sensitivity of 3 POC devices evaluated at extended time points showed that
69 sensitivity declined with time and this was particularly marked at >140 days post infection
70 onset. This is relevant if the tests are to be used for sero-prevalence studies. Neutralising
71 antibody data showed positive antibody results on POC devices did not necessarily confer
72 high neutralising antibody titres and these POC devices cannot be used to determine immune
73 status to the SARS-CoV-2 virus. Comparison of paired serum and capillary results showed

74 that there was a decline in sensitivity using capillary blood. This has implications in the
75 utility of the test as they are designed to be used on capillary blood by the general population.

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99 **Introduction**

100 The emergence of SARS-CoV-2 has triggered a race by laboratories and manufacturers to
101 bring commercial diagnostic tests to market. The unusually rapid pace of this development
102 risks potential compromises on quality in the absence of rigorous independent evaluation and
103 validation. It is therefore essential to ensure adequate performance of tests for population
104 wide or individual use to prevent roll out of tests which add no, or at best, minimal clinical
105 value to individual patients or the wider population.

106

107 The current most widely used diagnostic test for SARS-CoV-2 is based on real-time
108 polymerase chain reaction (RT-PCR) amplification of viral RNA from an upper respiratory
109 tract sample (1,2). However, due to a limited time window of active infection, capacity
110 constraints and restriction of access to symptomatic patients, cases determined using this
111 method underestimate the true burden of infection. In contrast, serological assays test for
112 previous infection and are therefore a key additional tool for monitoring prevalence of
113 infection within the population. Antibody tests can also be used as an aid in diagnosis where
114 COVID-19 is suspected clinically but the PCR time window has passed (3) and significant
115 interest exists in the potential for use of these tests at an individual level to provide an
116 indication of immune status and act as an ‘immunity passport’. For countries where vaccine
117 availability is limited pre-screening the population with antibody tests to identify individuals
118 who may either not require vaccination or be suitable for a reduced vaccine dosing regimen
119 may help optimise the use of limited vaccine resources.

120

121 Of particular interest is the use of monoclonal antibody treatment in hospitalised patients with
122 SARS-CoV-2. Ronapreve, a combination of Casirivimab and Imdevimab, is a monoclonal
123 antibody treatment directed at the Spike protein receptor binding domain on SARS-CoV-2.

124 This has been shown to significantly reduce 28 day mortality in seronegative hospitalised
125 patients (4).Therefore a well performing antibody tests has the potential to identify patients
126 who would benefit from this treatment.

127

128 A vast number of commercially available immunoassays have been developed to detect
129 SARS-CoV-2 IgG, IgM, IgA and total antibodies (5). Although the majority of antibody
130 production is directed towards the more abundant N (nucleocapsid) protein, the S (spike)
131 protein contains the receptor-binding domain responsible for host cell attachment and
132 antibodies to the S protein are therefore predicted to be neutralising (6).

133

134 In contrast to laboratory-based immunoassays, which require venous sampling and transport
135 to centralised testing sites, lateral flow immunoassays (LFA) offer the potential to allow
136 rapid, cheap, mass population antibody testing on capillary samples in the home environment.
137 This relief on clinical pressure and its use at a population level relies on the LFA working on
138 capillary whole blood samples and the tests offering sufficient ease of use, interpretation and
139 acceptability to the wider population. If LFAs are to be used for population sero-surveillance
140 they will have to be sensitive enough to detect the presence of antibodies in those who only
141 suffered from mild disease or were asymptomatic. Of even greater importance is the test
142 specificity when testing at a population level where the pre-test probability is low. Without
143 adequate specificity the chance of a positive result being a false positive is considerable and
144 this is of particular concern if these tests were to be used as evidence of immune status or
145 ‘immunity passports’. It is currently unclear if the detection of antibodies to the COVID 19
146 virus by LFA confers immunity to the virus and some manufactures fail to disclose whether
147 the assay is directed towards the N or S protein.

148

149 Worldwide, over 200 LFAs have been produced to date (5), for the majority of which only
150 manufacturer reported performance data is available. There is no requirement for external
151 validation to obtain CE marking and the manufacturer's validation process is often not
152 publically available. The UK Medicines and Healthcare Products Regulatory Agency
153 (MHRA) advise that any SARS-CoV-2 LFA should have sensitivity and specificity > 98%
154 (95 % CI 96-100 %) in samples collected over 20 days post symptom onset irrespective of
155 whether they are performed as a home self test kit or by a trained health care professional (7).
156 Despite claims by many manufactures to have achieved this target (see S1), UK external
157 validations of LFAs have so far failed to reproduce this (8–10).

158

159 As part of a Scottish evaluation of SARS-CoV-2 diagnostic tests, we carried out an
160 independent validation of 14 point of care SARS-CoV-2 antibody tests including 13 LFAs.
161 We assessed their specificity and sensitivity against manufacturer's claims and determined
162 compatibility with MHRA criteria. For tests that performed well against the MHRA criteria,
163 we studied changes in sensitivity with increasing time from infection. To investigate
164 suitability for home use, we evaluated ease of use and interpretation of these tests when
165 performed directly by patients. Neutralising antibody status was determined for samples to
166 establish whether the detection of antibodies by the LFA correlated with immune status.

167

168

169 **Methods**

170

171 **Point of care (POC) device selection and the evaluation process**

172 Companies approached National Services Scotland (NSS) with their POC device. The
173 manufacturer claims and costing were reviewed and manufacturers who passed these initial

174 checks were invited to send up to 100 POC test kits for initial evaluation. 14 point-of-care
175 devices designed to detect antibodies to the SARS-CoV-2 virus were evaluated. The tests
176 included in the evaluation were AbC-19, Alpha Pharma, Biomerica, Biozek, Fortress,
177 Jiangsu, Lepu, Menarini Healgen, Mologic, Pharmact, Roche, Wuhan Easy Diagnosis,
178 Wuhan Life Origin Biotech (Syzbio) and LumiraDX. 13 of the POC devices were LFAs,
179 consisting of immunochromatography based cassettes, with the LumiraDX assay using a
180 point of care microfluidic immunofluorescence assay. Each device measured IgG and IgM
181 except Abc19 which only measured IgG and LumiraDX which gave an antibody result of
182 unspecified subclass. Target antigens (either S or N) were not always disclosed by the
183 manufacturers. The tests were performed as per the manufacturer's instructions. Typically,
184 2.5µL to 10µL of serum or capillary sample was pipetted into the sample well followed by a
185 pre-specified volume of buffer supplied by the manufacturer (see S1). Serum samples were
186 run at room temperature in the laboratory at the Royal Infirmary of Edinburgh and the read
187 time varied from 10 to 20 minutes depending on manufacturer's instructions. Photographic
188 records of serum results were taken. Capillary samples were run as part of a research clinic.
189 Antibody results were visually read and recorded as positive, weak positive, or negative for
190 IgG and IgM. Positive and weak positive results were deemed an overall positive result.

191
192 In the first stage of the process 50 serum samples from RT-PCR confirmed SARS-CoV-2
193 positive participants and 50 negative controls were run. Initial sensitivity and specificity were
194 calculated and if the test kit performed well (typically defined as an IgG specificity of >98%
195 and sensitivity of >95% at >day 21 post infection) the company was taken forward to the
196 second stage of evaluation and asked to provide further test kits. Further stages of the
197 evaluation process which were carried out for a subset of kits included more stringent
198 specificity testing, determination of batch to batch variation, sensitivity as time from infection

199 increased, comparison of capillary to serum results and evaluation of ease of use by study
200 participants in a research clinic.

201

202 Overview of participants and samples

203 Serum samples from patients with RT-PCR confirmed SARS-CoV-2 were collected between
204 March and November 2020. Excess serum was stored at the point of discard from
205 hospitalised COVID-19 patients using ethical permissions obtained through NRS
206 BioResource.

207

208 Out-patients with previous RT-PCR confirmed SARS-CoV-2 infection were also recruited to
209 two COVID convalescent research studies where participants were invited to donate serial
210 blood samples (SR1407 BioResource study, n=112 participants) and to provide capillary and
211 serum blood samples (COVID-19 Antibody Test Evaluation (CATE) Study, n= 82
212 participants). For both studies the date of symptom onset and positive RT-PCR test result
213 were recorded. Inclusion criteria were those aged >16 years old with previous RT-PCR
214 confirmed diagnosis of COVID-19. Only 5 study participants were diagnosed with COVID-
215 19 in hospital with the remainder being diagnosed in the community. Due to the limited
216 access to COVID-19 testing at the start of the pandemic when most study recruitment
217 occurred many study participants were health care workers. Exclusion criteria were any frail
218 or shielding individuals. Each participant provided between 1 and 5 serum samples. For
219 participants providing serial samples these were collected between 24 and 216 days post PCR
220 confirmed infection. Capillary samples were collected as part of the CATE study research
221 clinic between September and November 2020. Capillary samples were tested in the clinic
222 setting with paired serum samples subsequently run on the same POC devices in the
223 laboratory to assess concordance on capillary and serum samples. A subset of CATE study

224 participants were asked to carry out one of the POC tests themselves including finger prick
225 testing with the aid of the manufacturers printed instructions. Health care workers intervened
226 if the participant needed support in completing the testing and any intervention was recorded.
227 Afterwards participants were asked to complete a questionnaire on their experience of
228 performing the test. For a subset of other test kits capillary results were interpreted by both
229 the participant (with the aid of a labelled diagram) and the researcher.

230

231 Negative controls consisted of venous samples collected prior to December 2019. For more
232 stringent specificity testing samples positive for rheumatoid factor, seasonal coronaviruses,
233 other respiratory pathogens, CMV (cytomegalovirus) or EBV (Epstein-Barr Virus) were used
234 as negative controls along with samples from the national antenatal screening programme and
235 BTS (The British transfusion service).

236

237 Wherever possible the same panel of samples were run on each POC device but due to
238 limitations in sample volume not all devices were assessed with the same panel of samples.
239 However, a similar variety of samples types were used to test each kit.

240

241 Neutralisation assay

242 Neutralising antibody assays were performed on serum samples using a pseudotyped
243 chimeric virus, expressing the SARS-CoV-2 spike protein, as described in (11).

244

245 Statistical analysis

246 The primary outcome was the sensitivity and specificity of each POC test. For
247 sensitivity, tests were compared against PCR-confirmed SARS-CoV-2 infection.

248 Specificity of each POC test was evaluated against pre-pandemic negative samples,

249 with all positives counting as false positives. The analysis included all available data
250 for the relevant outcome and are presented with the corresponding 95% confidence
251 intervals (Clopper Pearson).

252 For samples where neutralising antibody levels are available positive predictive value
253 (PPV) and negative predictive value (NPV) are calculated at two NT₅₀ thresholds; ≥ 50
254 and ≥ 160 .

255 For the comparison of the performance of each POC test between clinic capillary and
256 serum samples we calculated the sensitivity and 95% confidence interval. The
257 McNemar test was used to assess for significant difference between dependent groups.
258 Agreement between serum and capillary samples was measured using the Cohen's
259 Kappa. Results of the Kappa were interpreted as previously described (<0 , poor
260 agreement; 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60,
261 moderate agreement; 0.61–0.80, substantial agreement; and >0.8 , almost perfect
262 agreement) (12).

263 All data were analysed using PRISM version 9 and SPSS, and a p value <0.05 was
264 considered significant.

265

266 Ethics

267 Ethical approval for access to pre-pandemic stored samples and hospitalised COVID-19
268 patient samples was obtained through the NRS BioResource (Ref: SR1410). Outpatient
269 COVID-19 convalescent patient samples were obtained through the NRS BioResource (Ref:
270 SR1407) and the COVID-19 Antibody Test Evaluation (CATE) study, with ethical approval
271 from London-Brent Research Ethics Committee (REC ref: 20/HRA/3764 IRAS: 286538

272 **Results**

273

274 **Comparison of POC device performance using serum samples**

275 14 POC devices were evaluated for sensitivity and specificity including 13 LFAs and one
 276 microfluidic immunofluorescence assay (Lumira DX). In the first stage the sensitivity of the
 277 LFA was assessed using a panel of 50 serum samples from RT-PCR positive individuals. The
 278 specificity of the LFA was assessed using a panel of 50 serum samples collected prior to
 279 December 2019 for routine virological investigations. Adequately performing LFA then
 280 progressed to stage 2, where they were evaluated against an expanded panel of serum
 281 samples. In particular, a more stringent specificity panel was used, including serum samples
 282 taken from individuals with a positive PCR result for the endemic coronaviruses, other
 283 respiratory pathogens, and samples with positive rheumatoid factor, CMV or EBV serology.

284

285 The results are summarised in **Table 1** and **Fig 1**. For each sample, and each device, the IgG
 286 and IgM result was scored as negative, weak positive or positive (the Abc19 LFA only had an
 287 IgG window, and the LumiraDX platform only returned an antibody result). For the purposes
 288 of the sensitivity and specificity calculations, weak positive and positive results were counted
 289 as positive.

290

291 **Table 1: Summary of POC test sensitivity and specificity**

	Stage 1 only														Stage 2 onwards																	
	Alpha Pharma		Jiangsu		Lepu		Menarini batch#2005156		Mologic		Pharmact		Wuhan Jifei Origin/Biotech		Wuhan Easydiagnosis		Abc19	Blomera	Biozek	Fortress	Menarini batch#2003288	Roche	LumiraDx									
Overall Sensitivity	n	50	50	40	100	50	25	50	50	51	76.1-94.8	50	288	140	50	99	151	262														
IgG	%	90.0	78.2-96.0	94.0	83.0-98.0	95.0	83.1-99.1	74.0	64.3-81.1	72.0	57.5-82.2	68.0	46.5-83	92.0	80.8-97.2	88.2	76.1-94.8	76.0	61.8-85.5	89.6	85.5-92.4	92.1	86.4-95.5	72.0	57.7-82.2	88.9	81.0-94.0	65.6	57.0-92.0	N/A	N/A	
IgM	%	90.0	78.2-96.0	78.0	64.0-87.1	87.5	73.2-94.9	63.0	52.8-71.1	14.0	5.8-24.7	68.0	46.5-83.0	72.0	57.5-82.2	90.0	78.2-96.0	N/A	N/A	19.1	14.7-23.3	38.6	30.5-45.8	78.0	64.0-87.1	73.7	63.9-80.9	61.6	53.3-68.2	N/A	N/A	
Ab	%	94.0	83.5-98.3	96.0	86.3-99.3	95.0	83.1-99.1	75.0	65.3-82.0	72.0	57.5-82.2	76.0	54.9-89.0	96.0	86.3-99.3	94.0	83.5-98.3	N/A	N/A	90.1	86.0-96.0	93.6	88.1-96.6	80.0	66.3-88.7	88.9	81.0-93.6	71.5	63.6-77.5	94.3	90.7-96.4	
z20d Sensitivity	n	45	45	35	60	25	11	45	45	45	25	238	120	25	59	66																
IgG	%	91.1	78.8-96.9	95.6	85.0-99.0	97.1	85.1-99.9	76.7	64.0-85.3	84.0	63.9-94.3	81.8	48.2-96.7	93.3	81.7-98.2	91.1	78.8-96.9	96.0	79.6-99.8	93.3	89.3-95.7	95.0	89.4-97.8	92.0	74.0-98.6	94.9	86.0-99.0	95.5	87.0-99.0	N/A	N/A	
IgM	%	88.9	75.9-95.5	77.8	62.9-87.4	91.4	76.9-97.6	65.0	51.6-75.2	4.0	0.1-17.6	81.8	48.2-96.7	73.3	58.1-83.8	83.2	81.3-98.1	N/A	N/A	16.4	11.9-20.9	35.8	27.3-43.7	92.0	74.0-98.6	76.3	63.4-85.0	77.3	65.3-85.4	N/A	N/A	
Ab	%	93.3	81.7-98.2	97.8	88.2-99.9	97.1	85.1-99.9	76.7	64.0-85.3	84.0	63.9-94.3	90.9	58.7-99.5	97.8	88.2-99.9	97.7	88.0-99.9	N/A	N/A	94.1	90.2-96.5	95.8	90.9-98.3	96.0	79.6-99.8	94.9	85.9-98.6	97.0	89.5-99.5	98.1	95.3-99.3	
Specificity	n	50	50	40	100	92	25	50	50	94	327	313	100	100	160	201																
IgG	%	100.0	92.9-100	96.0	86.3-99.3	95.0	83.1-99.1	100.0	96.4-100	98.9	94.1-99.9	76.0	54.9-89.0	98.0	89.4-99.9	98.0	89.4-99.9	96.8	91.0-99.1	99.1	97.3-99.7	98.1	95.9-99.2	100.0	96.4-100	99.0	94.6-99.9	99.4	96.6-100	N/A	N/A	
IgM	%	100.0	92.9-100	96.0	86.3-99.3	90.0	76.3-96.5	100.0	96.4-100	98.9	94.1-99.9	84.0	63.9-94.3	98.0	89.4-99.9	###	92.9-100	N/A	N/A	96.9	94.4-98.3	90.6	82.3-95.2	98.0	92.9-99.6	99.0	94.6-99.9	98.1	94.6-99.5	N/A	N/A	
Ab	%	100.0	92.9-100	92.0	80.8-97.2	85.0	70.2-93.3	100.0	96.4-100	98.9	94.1-99.9	72.0	50.6-86.1	96.0	86.3-99.3	98.0	89.4-99.9	N/A	N/A	96.3	93.7-97.9	94.8	91.0-97.1	98.0	92.9-99.6	98.0	93.0-99.6	97.5	93.7-99.1	99.5	95.0-99.0	
Early Sensitivity																																
Late Sensitivity																																
NT50 correlation																																
Capillary testing																																

292

293 *For each manufacturer the n, overall sensitivity, sensitivity at ≥ 20 days post symptom onset,*
294 *and specificity is shown, with 95% confidence intervals (Clopper-Pearson). The LFA are*
295 *grouped according to which stage of the evaluation they reached. For kits that progressed*
296 *further than stage 1, the additional analyses performed are indicated (Y, performed). For*
297 *Abc19, only IgG data was available, and for LumiraDX, only the overall antibody result was*
298 *available. The results from two separate batches of Menarini kits are also shown.*

299

300 **Fig 1: Sensitivity vs specificity at >20 days**

301 *Sensitivity against specificity for the kits tested (IgG only, and positive or negative result for*
302 *the LumiraDx assay). For each kit the ≥ 20 days post symptom onset sensitivity is shown. The*
303 *MHRA targets of 98% sensitivity and specificity are shown as dotted lines. Pharmact was not*
304 *included on the graph due to low specificity. The LFA that did not progress beyond this stage*
305 *are shown in red.*

306

307 Only the LumiraDX POCT assay reached the MHRA criteria for sensitivity and specificity (>
308 98% for both), although a number of the LFA (Biomerica, Biozek, Fortress, Menarini and
309 Roche) surpassed the specificity criteria, even when tested against the high stringency
310 specificity serum samples. One empirical finding was that there was extensive variation in the
311 strength of the staining between the kits, which could affect ease of interpretation. Another
312 more worrying observation was that batch-to-batch variation in performance was observed
313 for the Menarini assay with batch 2005156 giving a sensitivity of 76.3% on samples ≥ 20 days
314 post infection versus batch 2003288 giving a sensitivity of 94.9% on the same samples
315 (n=99).

316

317 The Abc-19, Biomerica, Biozek, Fortress, Menarini (batch 2003288) and Roche LFA and the
 318 LumiraDX assay underwent further analysis using an expanded panel of RT-PCR positive
 319 samples. Initially, sensitivity was examined at proximal time points to infection for the
 320 Biomerica, Biozek, Menarini, Roche and LumiraDX LFAs and the results are shown in
 321 **Table 2** and **Fig 2**. All the LFA showed an initial increase in sensitivity, peaking at ≥ 21 days
 322 post symptom onset, with sensitivities of $\geq 90\%$ obtained for all LFA from this time point
 323 onwards. Between 8-20 days post symptom onset, LFA performance was more variable, with
 324 Roche showing the lowest sensitivity, and Menarini the highest.

325

326 **Table 2 Sensitivity against time**

Post symptom onset (days)		Biomerica	Biozek	Menarini	Roche	Lumira
0-7	Sensitivity	N/A	N/A	N/A	11.5	26.1
	95% CI	N/A	N/A	N/A	2.4-27.2	10.2-45.1
	n	N/A	N/A	N/A	26	23
8-20	Sensitivity	77.2	81.8	83.3	59.4	71.6
	95% CI	64.2-85.9	59.7-93.5	69.8-91.4	46.4-69.7	60.0-80.1
	n	57	22	48	64	74
21-40	Sensitivity	93.0	95.6	100	93.8	92.9
	95% CI	83.0-97.6	87.6-98.8	84.6-100	79.2-98.9	82.7-97.5
	n	57	68	22	32	56
41-60	Sensitivity	95.2	95.1	89.7	96.6	98.7
	95% CI	88.3-98.4	83.1-98.3	72.6-97.1	82.2-99.8	93.1-99.9
	n	84	49	29	29	78

327 *Sensitivity against time post symptom onset for selected kits (IgG for all LFA, with the*
 328 *exception of LumiraDX, where overall antibody result was used). n and 95% confidence*
 329 *intervals (Clopper Pearson) are shown. N/A, not done.*

330

331 **Fig 2 Sensitivity against time**

332 *Sensitivity against time post symptom onset for selected kits (IgG for all LFA, with the*
 333 *exception of LumiraDX, where overall antibody result was used).*

334 An equally important question was what the sensitivity of the POCT LFA was at later time
 335 points post symptom onset, since it has been shown that antibody levels wane with time
 336 following the primary antibody response to SARS-CoV-2 (13–15). Sensitivity was analysed
 337 across three time windows from 20 days post symptom onset for the Abc19, Biomerica,
 338 Roche and LumiraDX predominantly using samples from the convalescent cohort, and the
 339 results are shown in **Table 3** and **Fig 3**.

340

341 **Table 3 Sensitivity against time post symptom onset**

Post symptom onset (days)		Abc19	Biomerica	Roche	Lumira
20-79	Sensitivity	80.3	91.1	98.1	98.0
	95% CI	73.0-85.4	85.4-94.5	94.4-99.5	94.4-99.5
	n	152	157	154	153
80-139	Sensitivity	81.4	88.6	93.2	97.7
	95% CI	66.6-90.4	75.4-95.4	81.3-98.12	87.7-99.9
	n	43	44	44	43
≥ 140	Sensitivity	68.5	54.4	82.3	98.2
	95% CI	56.6-77.4	42.8-64.0	72.1-88.9	90.6-99.9
	n	73	79	79	57

342 *Sensitivity against time post symptom onset for selected kits (IgG for all LFA, with the*
 343 *exception of LumiraDX, where overall antibody result was used). n and 95% confidence*
 344 *intervals (Clopper Pearson) are shown. N/A, not done.*

345

346 **Fig 3 Sensitivity against time post symptom onset**

347 *Sensitivity against time post symptom onset for selected kits (IgG for all LFA, with the*
 348 *exception of LumiraDX, where overall antibody result was used).*

349

350 The LumiraDX assay was the only assay evaluated that maintained its sensitivity at 98%
 351 across the time interval studied (up to 224 days post symptom onset). The three LFA showed
 352 variable drops in sensitivity with increasing time from symptom onset, with Biomerica
 353 dropping from 91% to 54% over the time points studied.

354

355 For the convalescent cohort samples, neutralising titres were also available, and the POCT
356 result (for Abc19, Biomerica, Roche and LumiraDX) could be compared with the half
357 maximal neutralising titre (NT_{50}). Initially, the NT_{50} result for each serum sample was
358 plotted, dividing each LFA by positive and negative IgG (or for LumiraDX, antibody) result
359 (**Fig 4**).

360

361 **Fig 4 Neutralising antibody titres compared to POC test result**

362 *NT_{50} against binary POC test result. The NT_{50} results for the positive and negative LFA*
363 *result are plotted by kit (the mean and SEM are denoted by black lines). Dotted lines*
364 *represent the $NT_{50} \geq 50$ and ≥ 160 thresholds, and the p values from a Kruskal-Wallis test*
365 *are shown.*

366

367 There was a statistically significant difference in NT_{50} level between the positive and
368 negative LFA result for all assays. However, for the LFA positive result, a large range of
369 NT_{50} values was observed. To examine this further, the sensitivity and specificity of the LFA
370 was evaluated at two NT_{50} thresholds; ≥ 50 and ≥ 160 (**Table 4**). These two thresholds were
371 selected based on a detectable NT_{50} (≥ 50) and a threshold which may confer protection from
372 reinfection (≥ 160) (16).

373

374

375

376

377

378

379 **Table 4. Performance of POC devices compared to $NT_{50} \geq 50$ and ≥ 160**

	NT50 \geq 50				NT50 \geq 160			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
Abc19	86.6	61.5	90.3	52.5	95.7	42.3	63.8	90.2
	81.3-90.2	47.0-72.9	85.5-93.5	39.3-63.5	90.8-98.1	33.7-49.9	56.8-69.3	79.8-95.6
Biomerica	89.2	53.4	88.0	56.4	95.8	35.5	60.4	89.1
	84.3-92.4	39.9-64.7	83.0-91.4	42.3-67.8	91.0-98.1	27.6-42.8	53.7-65.9	77.8-95.1
Roche	98.2	27.6	83.7	80.0	100.0	14.7	54.9	100.0
	95.4-99.4	16.7-38.8	78.6-87.3	56.3-92.9	97.4-100	9.2-20.6	48.6-60.1	83.2-100
LumiraDX	98.6	4.7	83.5	40.0	100.0	4.3	55.2	100.0
	95.9-99.6	0.6-13.9	78.2-87.2	5.3-81.1	97.3-100	1.4-8.8	48.8-60.6	47.8-100

380 *Sensitivity, specificity, PPV and NPV for each LFA to predict NT_{50} of ≥ 50 or ≥ 160 . 95%*

381 *confidence intervals (Clopper-Pearson) are also shown.*

382

383 The sensitivity of all the POCT for detecting an $NT_{50} \geq 50$ (of the 280 serum samples, 222
 384 had an $NT_{50} \geq 50$ and 58 had an $NT_{50} < 50$) was at least 90%, and for Roche and LumiraDX it
 385 was $> 98\%$. However, the NPV varied between 40 and 80%, as a result of the small number
 386 of samples that were negative by the POCT, relative to the number of positive samples (**Fig**
 387 **4**). Specificity performance was also poor with a large number of false positive results
 388 (POCT positive, but $NT_{50} < 50$). In particular, the LumiraDX assay had a specificity of $< 5\%$,
 389 whereas the Abc19 assay had the highest specificity of 62%. These results impacted the PPV
 390 in this high prevalence setting, which ranges from 90% for the Abc19 assay and 84% for the
 391 LumiraDX assay.

392

393 When the NT_{50} threshold was raised to ≥ 160 (where 142 of the serum samples had an $NT_{50} \geq$
 394 160 and 138 had an $NT_{50} < 160$), the sensitivity for all the POCT rose to at least 96%, with a
 395 consequential increase in NPV to at least 90%. However, the specificity of all of the POCT
 396 fell, with the Abc19 kit having the highest specificity at 42%. This affected the PPV, which
 397 also fell from the lower NT_{50} threshold, to between 55 and 64%.

398

399 **Comparison of POCT device performance in serum and capillary samples**

400 A major advantage of the POCT devices over immunoassays performed within accredited
401 laboratories is the lack of requirement for phlebotomy. However, many evaluations of POCT
402 performance have not compared serum and capillary samples. Furthermore, it has been
403 suggested that LFA could be suitable for home use, outside of a healthcare setting, and this
404 would require the participant to interpret their test result.

405

406 In this work, paired serum and capillary samples were available for a number of participants
407 in the convalescent cohort, and the performance of seven POCT assays (Abc19, Biomerica,
408 Biozek, Fortress, Menarini, Roche and LumiraDX) was compared. The serum sample was
409 processed in the laboratory, and read by a health care worker, and the capillary sample was
410 read by the study participant (with the aid of a labelled diagram or instruction leaflet) and a
411 second health care worker in the research clinic. Sensitivity findings for the 3 matched
412 interpretations are shown in **Table 5 and Fig 5**.

413

414

415

416

417

418

419

420

421

422

423

424 **Table 5. Comparison of paired serum and capillary results on selected POC devices**

	ABC 19	Roche	Menarini	Fortress	Biomerica	Biozek
Matched capillary (participant) capillary (HCW) and serum (laboratory) results	N= 45	N= 42	N=38	N=41	N=11	N=41
Sensitivity (%) (95% CI)						
Capillary (participant)	33.3 (20.0-49.0)	64.3 (48.0-78.4)	63.2 (46.0-78.2)	61.0 (44.5-75.8)	18.2 (2.3-51.8)	39.0 (24.2-55.5)
Capillary (HCW)	53.3 (37.9-68.3)	69.0 (52.9-82.4)	71.1 (54.1-84.6)	58.5 (42.1-73.7)	9.1 (0.2-41.3)	46.3 (30.7-62.6)
Serum (laboratory)	73.3 (58.1-85.4)	88.1 (74.4-96.0)	76.3 (59.8-88.6)	46.3 (30.7-62.6)	36.4 (10.9-69.2)	73.2 (57.1-85.8)
Kappa						
Serum Vs Capillary (HCW)	0.31 (0.06-0.57) p= 0.35	0.46 (0.18-0.74) p=0.008	0.59 (0.30-0.88) p=0.697	0.47 (0.21-0.73) p=0.227	-0.17 (-0.46-0.12) p=0.38	0.29 (0.05-0.54) p=0.007
Serum Vs Capillary (participant)	0.21 (0.02- 0.40) P<0.001	0.39 (0.13-0.65) p=0.002	0.57 (0.30-0.84) p=0.125	0.42 (0.16-0.68) p=0.146	0.12 (-0.43-0.67) p=0.63	0.38 (0.17-0.59) P<0.001
Capillary (HCW) Vs Capillary (participant)	0.52 (0.29- 0.75) p= 0.12	0.89 (0.75-1.03) p=0.50	0.82 (0.63-1.01) p=0.250	0.95 (0.85-1.05) p=1.000	0.62 (-0.04-1.28) p=1.000	0.85 (0.69-1.01) p=0.25

425

426 *Sensitivity for LFA devices comparing serum, and HCW and participant interpreted capillary*
 427 *data. Sensitivity and 95% CI (Clopper-Pearson). P values calculated using McNemar's test*
 428 *comparing serum and capillary sensitivity. Kappa describing extent of agreement between*
 429 *results.*

430

431 **Fig 5 Sensitivity comparison between paired serum and capillary samples**

432 *Sensitivity for LFA devices comparing serum, and HCW and participant interpreted capillary*
 433 *data. Sensitivity and 95% CI (Clopper-Pearson) were plotted.*

434

435 A common feature of the LFA results in capillary samples was a reduction in sensitivity
 436 compared to the serum samples. This was seen in five out of the 7 LFA analysed, however,
 437 the magnitude of the effect varied. Both Biozek and Roche showed significant decreases
 438 between serum and capillary sensitivity when interpreted by either HCW or participant. This

439 decrease in sensitivity ranged from 19.1 to 34.2%. ABC 19 showed a significant decrease in
440 sensitivity between serum and participant read capillary results decreasing from 73.3% to
441 33.3%. The Fortress LFA showed a small increase in sensitivity in capillary samples
442 compared to serum but this was not statistically significant. The LumiraDX assay showed
443 100% sensitivity for both samples (n=11 but a large number of test fails (n=18) occurred on
444 capillary samples at the research clinic.

445

446 The concordance between serum, HCW and participant read capillary results are displayed in
447 table 5. Concordance between LFA varied. For example of the 45 Abc19 results, 24 were
448 interpreted as positive by a HCW and 15 by the participant. In contrast, for the Fortress LFA,
449 there was a single discrepant positive result between the HCW and participant interpretation,
450 where the participant reported a positive, and the HCW a negative, result.

451 Participants who performed self-testing using the AbC-19 LFA in the clinic also completed a
452 questionnaire that addressed aspects of self testing using the kits (n = 35 participants). Their
453 responses are summarised in **Fig 6**. Overall, 89% of participants reported it was “very easy”
454 or “easy’ to understand the leaflet explaining the results, and 75% reported it was “very easy”
455 or “easy” to understand the instructions. Any issues appeared to be predominantly associated
456 with taking the capillary samples and using the test kit, with 20% of participants reporting it
457 was “difficult” or very difficult’ to take the sample.

458

459 **Fig 6 Ease of use**

460 *Summary of ease of use questionnaire. Participant ease of use responses to the indicated*
461 *categories were expressed as a percentage of responses.*

462

463

464 **Discussion**

465

466 The SARS-CoV-2 pandemic has generated unprecedented demand for testing both to confirm
467 acute infection and past virus exposure. In particular, serological assays measure prior
468 exposure to the virus and are being used in high volume settings to measure seroprevalence at
469 a population level. The role for these tests at an individual level is still poorly understood, and
470 there are huge risks of using poorly performing tests. For example, using a test with a poor
471 specificity in a low prevalence population would mean a high ratio of people with a false
472 positive result compared to a true positive result, leading people to believe they were
473 antibody positive, when they are not.

474

475 This study was designed as part of a national evaluation of SARS-CoV-2 antibody POCTs
476 within NHS Scotland, and aimed to address several questions regarding the performance of
477 the POCT. To be effective, POCT should have similar performance in samples from patients
478 who experience relatively mild illness compared to patients requiring hospitalisation. This
479 was addressed using a cohort of convalescent patients, the majority of whom did not require
480 hospital treatment during their illness. This cohort was also used to assess the performance of
481 the POCT with increasing time from infection, and to examine how well they were able to
482 predict neutralising antibody titres.

483

484 Another major requirement is for POCT to perform well in capillary samples compared to
485 serum. This point is critical since a major strength of the POCT devices is that they could be
486 used on capillary samples, thus reducing the requirement for phlebotomy.

487

488 In this work, 14 POCTs were evaluated on serum samples, and a number of these tests
489 underwent further evaluation on capillary samples. From the serum evaluation, it became
490 clear that specificity performance of many of the POCT was good – with eleven out of
491 fourteen reaching the MHRA specificity target of > 98% for IgG, or in the case of LumiraDX
492 total Ab. However, only a single device (LumiraDX) reached the MHRA sensitivity criteria
493 of >98% \geq 20 days post symptom onset. The sensitivity panel that was used to assess these
494 POCT in the first instance consisted predominantly of hospitalised patients, who tend to have
495 higher antibody titres than patients with milder disease (17–22), making disease severity a
496 less likely explanation for these observations. The reduced sensitivity of the LFA compared
497 to manufacturer’s claims is not unique to the LFA; the sensitivity of a laboratory analyser
498 immunoassay on this cohort was 93% at \geq 20 days post symptom onset. Therefore at
499 relatively early time points post symptom onset, a negative result cannot rule out SARS-CoV-
500 2 exposure. Indeed, for the subgroup of LFA where sensitivity with time was examined it was
501 clear that, at early time points, sensitivity increased with time reaching maximal sensitivity
502 after 21 days post symptom onset. This was in keeping with other reports that have studied
503 the time to reach maximal antibody titres (17,23–25).

504

505 An equally important question regarding timing of testing was how long post infection the
506 POCT could detect an antibody response. The issue of sero-reversion is important for
507 seroprevalence studies, and the ideal test for this purpose would have a high sensitivity for a
508 long period of time. Otherwise, sero-reversion may result in underestimation of
509 seroprevalence, and models to account for this trend of waning antibodies themselves risk
510 | over- or under-correcting for this effect (26,27). To examine this, a cohort of SARS-CoV-2
511 | convalescent individuals for whom longitudinal serum samples were available were used.
512 Crucially, these patients were most representative of the majority of SARS-CoV-2 infections

513 in so far as only 5 participants providing serial samples required hospitalisation, and none
514 required treatment in intensive care. In this work, the LumiraDX assay was the only assay to
515 maintain a high sensitivity at time points from 21 to 244 days post symptom onset; the other
516 assays showed variable levels of declining sensitivity. This was most marked for the
517 Biomerica LFA, which had a sensitivity of 91% at 21-79 days post symptom onset, dropping
518 to 54% at ≥ 140 days post symptom onset.

519
520 The relative importance of specificity compared to sensitivity at least in part depends on the
521 intended use. For example, for serosurveillance, particularly in low prevalence populations,
522 specificity should be maximised to reduce the false positive rate. However, in clinical
523 settings, drops in specificity with increased sensitivity might be acceptable if there were
524 adequate follow up tests to identify true positives.

525
526 A major question remains over what a positive antibody test result means in terms of
527 protection from reinfection, or protection from infection following vaccination. This is
528 pertinent given the continued discussions regarding the issue of “immunity passports” and a
529 possible role for antibody tests in assisting populations where vaccine supply is scarce
530 through identifying antibody negative individuals who should be prioritise for vaccination
531 and/or antibody positive individuals who may be suitable for reduced vaccine dosing
532 regimens. In this work, we compared the LFA result with the NT₅₀ level (**Fig 4** and **Table 4**).
533 It was apparent that a positive LFA was associated with a wide range of neutralising antibody
534 levels, and that a number of LFA positive samples had low NT₅₀ levels (< 50). This adds to
535 the concern that a positive LFA result may not inform about the likelihood of protection from
536 infection. Alongside specificity concerns and the increased circulation of variants of concern
537 that may be able to evade existing humoral responses (28,29), this highlights the risk of using

538 these tests for the purpose of immunity passports. Whilst some of the better performing LFA
539 may be suitable for sero-prevalence studies, they should not be used to draw conclusions on
540 an individual's protection from reinfection.

541

542 A major advantage of POCT stems from their potential to relieve requirement for phlebotomy
543 and to be performed outside of a healthcare setting. However, for this potential to be realised,
544 additional factors including POCT performance in capillary samples and ease of use should
545 be considered. In this work, a head to head comparison of POCT performance on serum and
546 capillary samples was performed, and for the LFAs HCW and participant interpretation of the
547 test result was also compared. There were differences in performance on serum and capillary
548 samples, with the majority of LFA showing reduced sensitivity on capillary samples
549 compared to serum. Concordance between serum and capillary results was sometimes poor
550 indicating that that the performance of a test on serum under laboratory conditions cannot be
551 assumed to equate to performance on a capillary sample. Only the LumiraDX device was
552 equally sensitive on both capillary and serum samples.

553

554 There was good agreement between capillary results interpreted by HCW and participants for
555 the Roche, Menarini, Biozek and Fortress kits studied indicating that these LFA tests may be
556 best suited for use as home test kits as interpretation at home is likely to match interpretation
557 in a hospital setting. For other LFAs, such as AbC-19, many participants had difficulty
558 noticing the faint lines produced by the test kit and interpreted the result as being negative
559 when it was interpreted as positive by a HCW.

560

561 A disadvantage of the POCT devices compared to laboratory immunoassays is that they are
562 strictly qualitative (with the exception of LumiraDX where a numeric value is converted to a

563 qualitative result), meaning the possibility of introducing equivocal zones is not available.
564 Further issues that came to light during the course of this evaluation work included the
565 possibility of batch-to-batch variation in kits from the same manufacturer. For example two
566 different batches of Menarini kits showed marked differences in sensitivity (76.7% on one
567 batch versus 94.9% on another batch). Whilst we did not observe any failed tests during our
568 evaluation of the LFA using serum or capillary samples, a high failure rate was seen using the
569 LumiraDX device with capillary samples with 18 out of 29 participants having a failed test
570 result with lumiraDX. Unpublished data available from another POCT evaluation site in
571 Scotland (NHS Tayside) where a newer version of the LumiraDX device was being trialled
572 indicated that there were no issues with failed SARS-CoV-2 antibody capillary tests with the
573 newer version of the LumiraDX device suggesting that this issue has now been resolved by
574 the company. However, strategies to identify such issues should continue to be employed by
575 the manufacturer and/or the end user to ensure accurate performance on capillary samples.

576

577 The strength of this work lies in its contribution to our understanding of POCT performance –
578 in particular performance at different times post infection, confirmed evidence of batch to
579 batch variation with some POCT kits and through demonstrating the general lack of
580 correlation between a positive POCT result and the presence of neutralising antibodies. Our
581 data also demonstrates that the results obtained in the lab using serum cannot automatically
582 be assumed to be representative of finger-prick capillary test results. Furthermore, a good
583 proportion of this work used a convalescent cohort of patients with relatively mild disease,
584 making the findings applicable to community based studies. However, the study does have
585 limitations. Due to sample availability constraints and the large number of POCTs evaluated
586 not all could be evaluated on the same panel of positive and negative samples. In particular,
587 the numbers of samples where paired serum and capillary sample results were available to

588 assess concordance was limited for the majority of the POCT examined. For one of the test
589 kits ease of use was assessed but as many of the study participants providing capillary
590 samples in this study were healthcare workers the data obtained through this may not be
591 representative of the general population.

592

593 In summary, our results highlight a wide variation in performance of SARS-CoV-2
594 antibody test kits and illustrates the importance of evaluating multiple different
595 aspects of test performance including checking for batch to batch variation, changes in
596 sensitivity as time from infection increases, correlation with neutralising antibodies
597 and performance on capillary samples. Thorough evaluation of all these aspects is
598 essential prior to considering the utilisation of these tests for antibody or ‘immunity’
599 passports and identification of hospitalised patients who would benefit from
600 monoclonal antibody treatment.

601

602

603

604

605 **Acknowledgements**

606 Menarini test kits (batch batch 2003288) were supplied by Annette Thomas from Cardiff &
607 Vale University Health board. Funding and support for this study was provided by Lothian
608 NRS BioResource. We would also like to thank members of the National Service Scotland
609 team for their support with co-ordinating this work.

610

611

612 **References**

- 613
- 614 1. Guidance and Standard Operating Procedure. COVID-19 Virus Testing in NHS
615 Laboratories [Internet]. NHS England and NHS Improvement; Available from:
616 [https://www.rcpath.org/uploads/assets/90111431-8aca-4614-](https://www.rcpath.org/uploads/assets/90111431-8aca-4614-b06633d07e2a3dd9/Guidance-and-SOP-COVID-19-Testing-NHS-Laboratories.pdf)
617 [b06633d07e2a3dd9/Guidance-and-SOP-COVID-19-Testing-NHS-Laboratories.pdf](https://www.rcpath.org/uploads/assets/90111431-8aca-4614-b06633d07e2a3dd9/Guidance-and-SOP-COVID-19-Testing-NHS-Laboratories.pdf)
- 618 2. Interim Guidelines for Collecting, Handling, and Testing Clinical Specimens for
619 COVID-19 [Internet]. Centers for Disease Control and Prevention; Available from:
620 <https://www.cdc.gov/coronavirus/2019-nCoV/lab/guidelines-clinicalspecimens.html>
- 621 3. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. Profiling Early Humoral
622 Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clin Infect Dis*. 2020
623 Mar 21;ciaa310.
- 624 4. RECOVERY Collaborative Group, Horby PW, Mafham M, Peto L, Campbell M,
625 Pessoa-Amorim G, et al. Casirivimab and imdevimab in patients admitted to hospital
626 with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial
627 [Internet]. *Infectious Diseases (except HIV/AIDS)*; 2021 Jun [cited 2021 Oct 20].
628 Available from: <http://medrxiv.org/lookup/doi/10.1101/2021.06.15.21258542>
- 629 5. SARS-CoV-2 diagnostic pipeline [Internet]. Available from:
630 <https://www.finddx.org/covid-19/pipeline/>
- 631 6. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2.
632 *JAMA*. 2020 Jun 9;323(22):2249.
- 633 7. Target product profile antibody tests to help determine if people have immunity to
634 SARS-CoV-2, 2020. [Internet]. MHRA; Available from:
635 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/883897/Target_Product_Profile_antibody_tests_to_help_determine_if_people_have_immunity_to_SARS-CoV-2_Version_2.pdf)
636 [_data/file/883897/Target_Product_Profile_antibody_tests_to_help_determine_if_people](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/883897/Target_Product_Profile_antibody_tests_to_help_determine_if_people_have_immunity_to_SARS-CoV-2_Version_2.pdf)
637 [_have_immunity_to_SARS-CoV-2_Version_2.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/883897/Target_Product_Profile_antibody_tests_to_help_determine_if_people_have_immunity_to_SARS-CoV-2_Version_2.pdf)
- 638 8. Flower B, Brown JC, Simmons B, Moshe M, Frise R, Penn R, et al. Clinical and
639 laboratory evaluation of SARS-CoV-2 lateral flow assays for use in a national COVID-
640 19 seroprevalence survey. *Thorax*. 2020 Aug 12;thoraxjnl-2020-215732.
- 641 9. Adams ER, Anand R, Andersson MI, Auckland K, Baillie JK, Barnes E, et al.
642 Evaluation of antibody testing for SARS-Cov-2 using ELISA and lateral flow
643 immunoassays [Internet]. *Infectious Diseases (except HIV/AIDS)*; 2020 Apr [cited 2020
644 Apr 20]. Available from: <http://medrxiv.org/lookup/doi/10.1101/2020.04.15.20066407>
- 645 10. Pickering S, Betancor G, Galão RP, Merrick B, Signell AW, Wilson HD, et al.
646 Comparative assessment of multiple COVID-19 serological technologies supports
647 continued evaluation of point-of-care lateral flow assays in hospital and community
648 healthcare settings. Fouchier RAM, editor. *PLOS Pathog*. 2020 Sep 24;16(9):e1008817.
- 649 11. Muecksch F, Wise H, Batchelor B, Squires M, Semple E, Richardson C, et al.
650 Longitudinal analysis of serology and neutralizing antibody levels in COVID19
651 convalescents. *J Infect Dis*. 2020 Nov 3;jiaa659.

- 652 12. Landis JR, Koch GG. The measurement of observer agreement for categorical data.
653 Biometrics. 1977 Mar;33(1):159–74.
- 654 13. Antonin Bal, Trabaud M-A, Fassier J-B, Rabilloud M, Saker K, Langlois-Jacques C, et
655 al. Six-month antibody response to SARS-CoV-2 in healthcare workers assessed by
656 virus neutralisation and commercial assays. Clin Microbiol Infect. 2021
657 Jan;S1198743X21000057.
- 658 14. L’Huillier AG, Meyer B, Andrey DO, Arm-Vernez I, Baggio S, Didierlaurent A, et al.
659 Antibody persistence in the first six months following SARS-CoV-2 infection among
660 hospital workers: a prospective longitudinal study. Clin Microbiol Infect. 2021
661 Jan;S1198743X21000318.
- 662 15. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, et al. Immunological memory
663 to SARS-CoV-2 assessed for up to 8 months after infection. Science. 2021 Feb
664 5;371(6529):eabf4063.
- 665 16. Addetia A, Crawford KHD, Dingens A, Zhu H, Roychoudhury P, Huang M-L, et al.
666 Neutralizing Antibodies Correlate with Protection from SARS-CoV-2 in Humans during
667 a Fishery Vessel Outbreak with a High Attack Rate. McAdam AJ, editor. J Clin
668 Microbiol. 2020 Aug 21;58(11):e02107-20, /jcm/58/11/JCM.02107-20.atom.
- 669 17. Long Q-X, Liu B-Z, Deng H-J, Wu G-C, Deng K, Chen Y-K, et al. Antibody responses
670 to SARS-CoV-2 in patients with COVID-19. Nat Med [Internet]. 2020 Apr 29 [cited
671 2020 Apr 29]; Available from: <http://www.nature.com/articles/s41591-020-0897-1>
- 672 18. The UCSF COMET Consortium, Combes AJ, Courau T, Kuhn NF, Hu KH, Ray A, et
673 al. Global absence and targeting of protective immune states in severe COVID-19.
674 Nature. 2021 Mar 4;591(7848):124–30.
- 675 19. Tan AT, Linster M, Tan CW, Le Bert N, Chia WN, Kunasegaran K, et al. Early
676 induction of functional SARS-CoV-2 specific T cells associates with rapid viral
677 clearance and mild disease in COVID-19 patients. Cell Rep. 2021 Jan;108728.
- 678 20. Guthmiller JJ, Stovicek O, Wang J, Changrob S, Li L, Halfmann P, et al. SARS-CoV-2
679 Infection Severity Is Linked to Superior Humoral Immunity against the Spike. Schultz-
680 Cherry S, editor. mBio. 2021 Jan 19;12(1):e02940-20, /mbio/12/1/mBio.02940-20.atom.
- 681 21. Wang P, Liu L, Nair MS, Yin MT, Luo Y, Wang Q, et al. SARS-CoV-2 neutralizing
682 antibody responses are more robust in patients with severe disease. Emerg Microbes
683 Infect. 2020 Jan 1;9(1):2091–3.
- 684 22. Wajnberg A, Amanat F, Firpo A, Altman DR, Bailey MJ, Mansour M, et al. Robust
685 neutralizing antibodies to SARS-CoV-2 infection persist for months. Science. 2020 Oct
686 28;eabd7728.
- 687 23. Borremans B, Gamble A, Prager K, Helman SK, McClain AM, Cox C, et al.
688 Quantifying antibody kinetics and RNA shedding during early-phase SARS-CoV-2
689 infection [Internet]. Open Science Framework; 2020 May [cited 2020 May 19].
690 Available from: <https://osf.io/evy4q>

- 691 24. Lou B, Li T-D, Zheng S-F, Su Y-Y, Li Z-Y, Liu W, et al. Serology characteristics of
692 SARS-CoV-2 infection after exposure and post-symptom onset. *Eur Respir J*. 2020
693 Aug;56(2):2000763.
- 694 25. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody Responses to SARS-
695 CoV-2 in Patients With Novel Coronavirus Disease 2019. *Clin Infect Dis*. 2020 Nov
696 19;71(16):2027–34.
- 697 26. Sabino EC, Buss LF, Carvalho MPS, Prete CA, Crispim MAE, Fraiji NA, et al.
698 Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *The Lancet*.
699 2021 Jan;S0140673621001835.
- 700 27. Kadelka S, Bouman JA, Ashcroft P, Regoes RR. Comment on Buss et al., *Science* 2021:
701 An alternative, empirically-supported adjustment for sero-reversion yields a 10
702 percentage point lower estimate of the cumulative incidence of SARS-CoV-2 in Manaus
703 by October 2020. *ArXiv210313951 Q-Bio* [Internet]. 2021 Mar 25 [cited 2021 May 25];
704 Available from: <http://arxiv.org/abs/2103.13951>
- 705 28. Hoffmann M, Arora P, Groß R, Seidel A, Hörnich BF, Hahn AS, et al. SARS-CoV-2
706 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell*. 2021
707 Mar;S0092867421003676.
- 708 29. Vidal SJ, Collier AY, Yu J, McMahan K, Tostanoski LH, Ventura JD, et al. Correlates
709 of Neutralization Against SARS-CoV-2 Variants of Concern by Early Pandemic Sera. *J*
710 *Viol*. 2021 Apr 23;JVI.00404-21, jvi;JVI.00404-21v1.

711
712
713
714
715
716

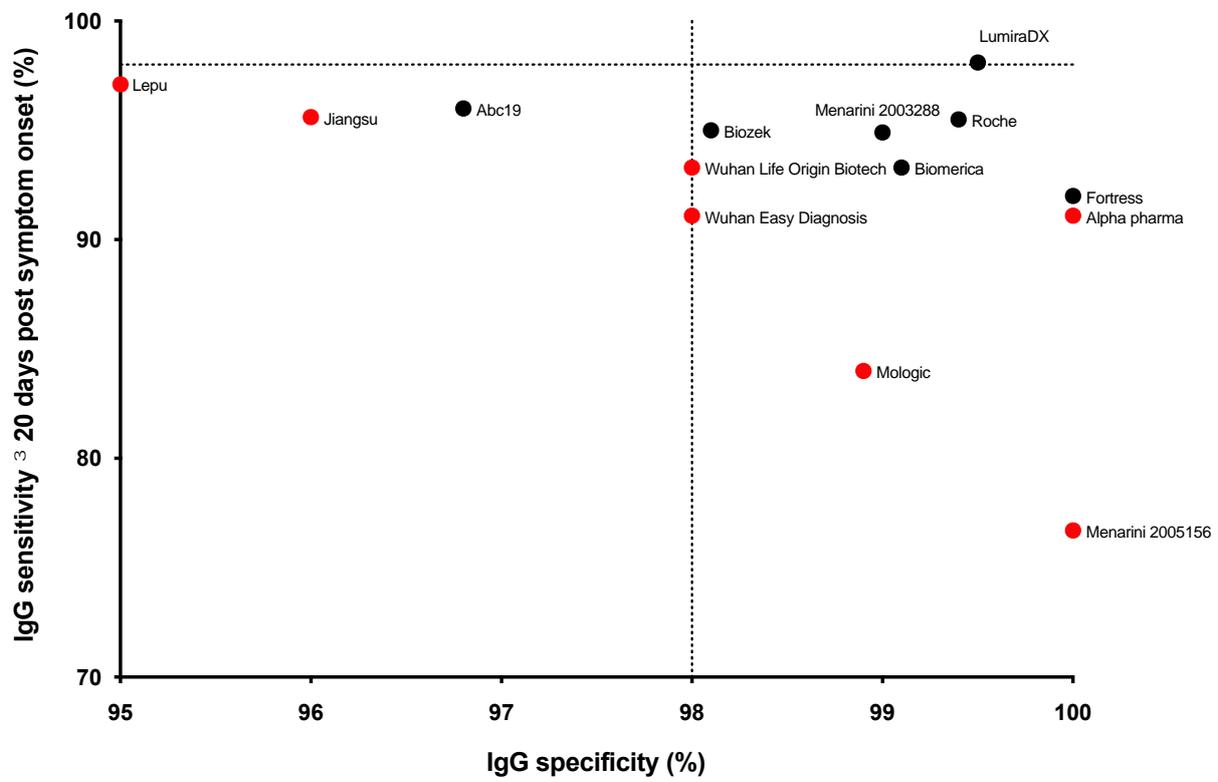
717 Supplementary information

718
719
720

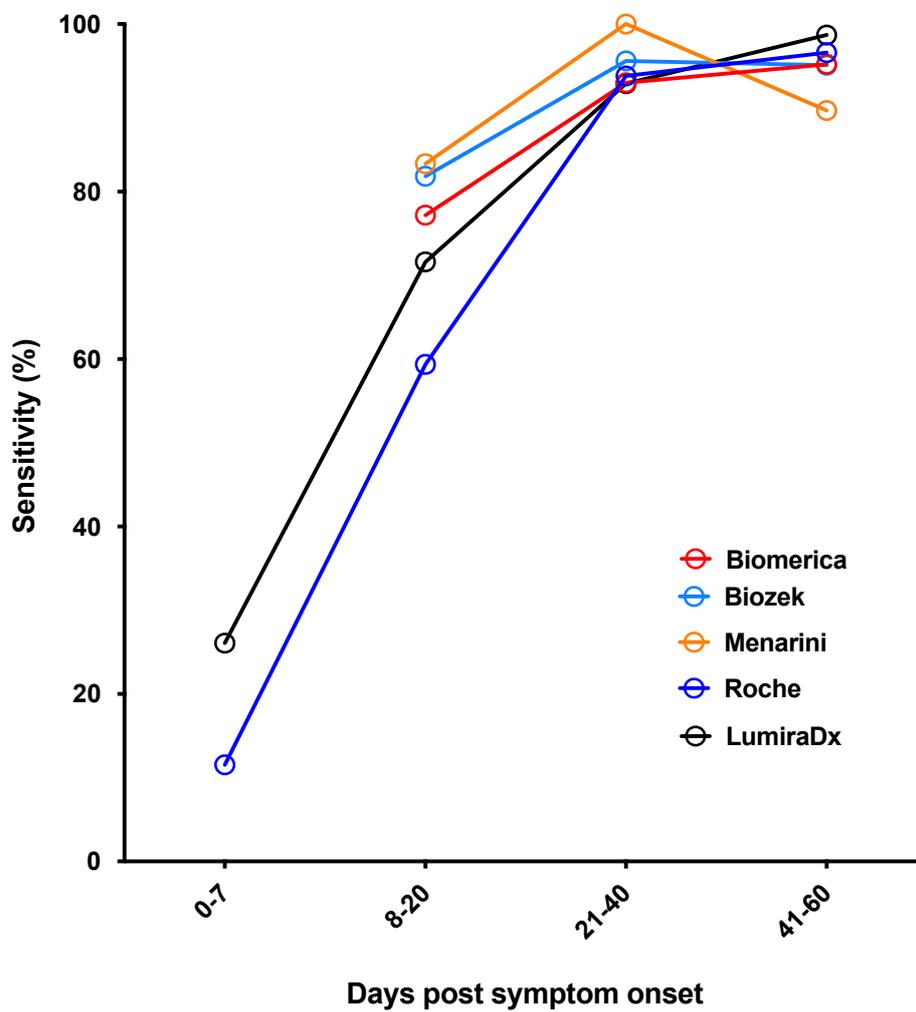
S1 Table. **Manufacturers published product specifications and performance data**

721
722
723

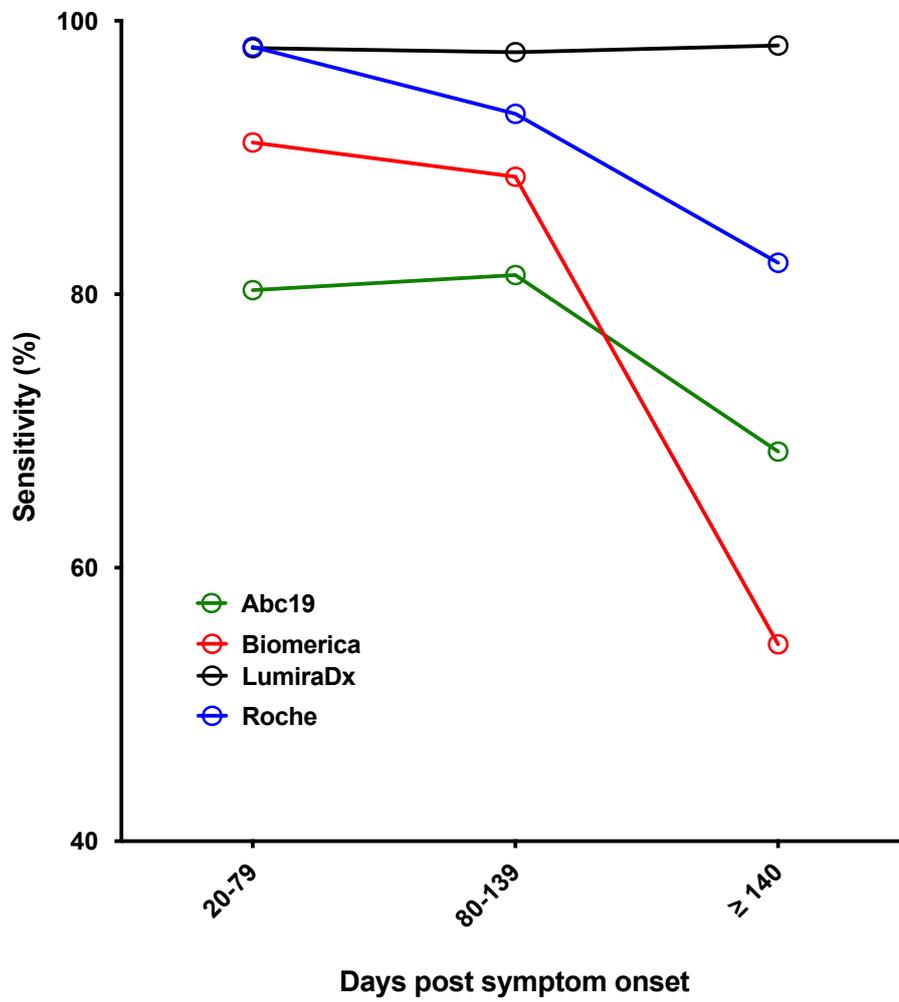
Figure 1



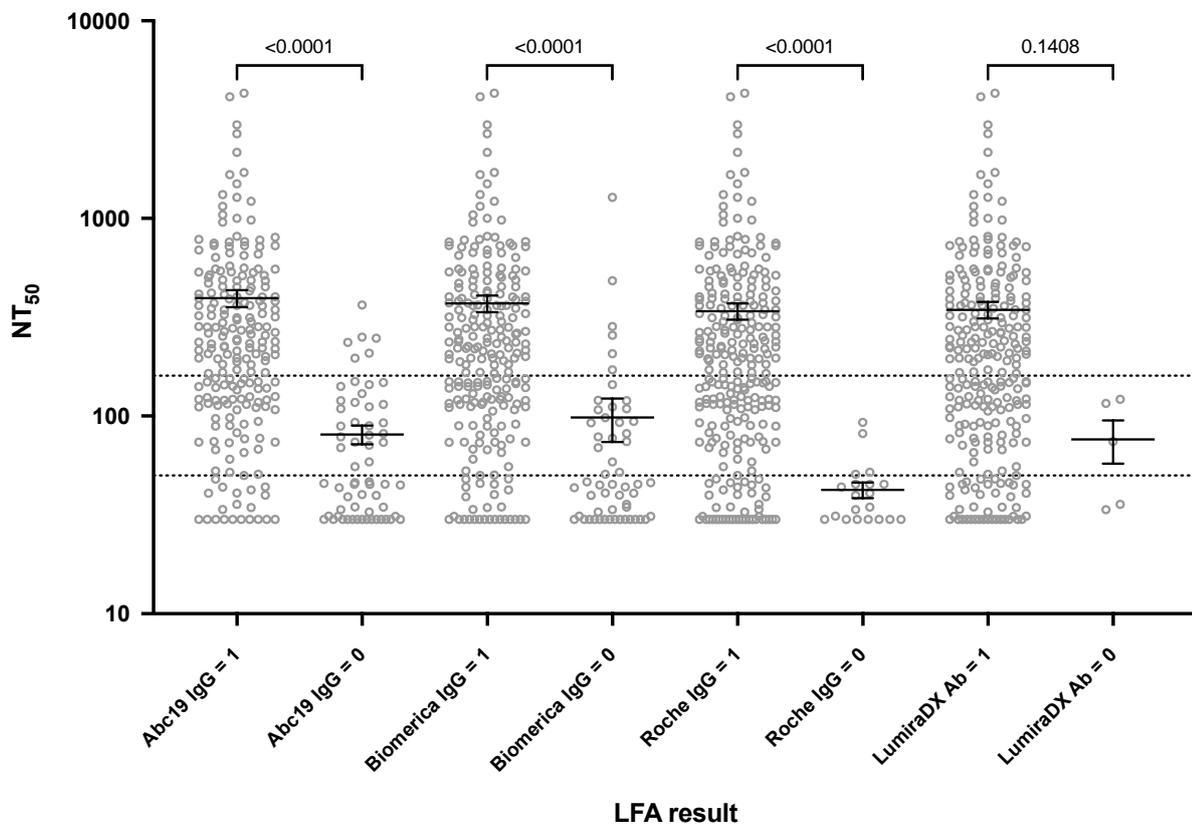
724 **Figure 2**
725



726 **Figure 3**
727



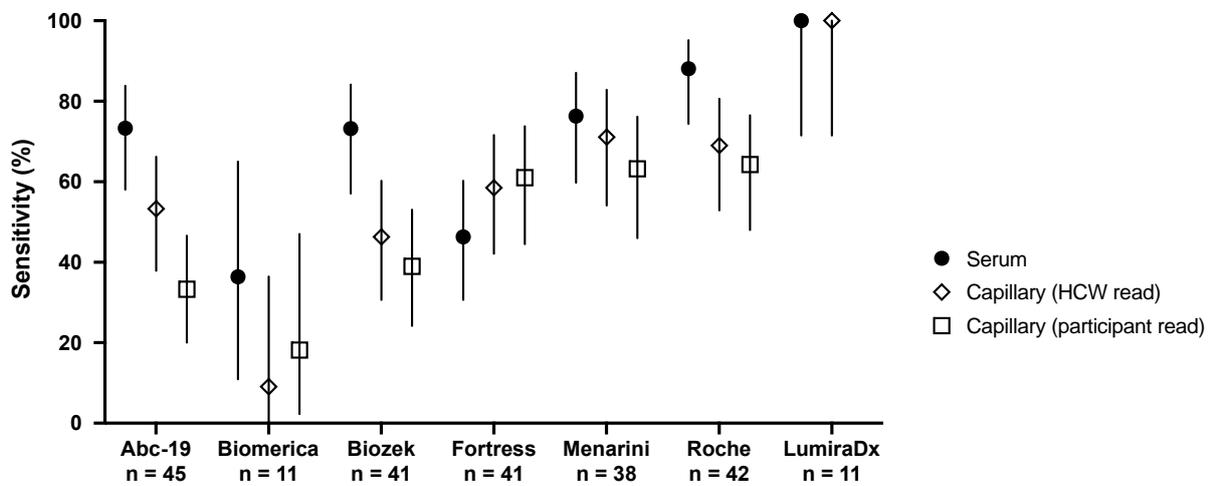
728 **Figure 4**
729



730 **Figure 5**

731

732



733

734

735

736

737

Figure 6

