

1 **The sensitivity improved two-test algorithm "SIT²": a** 2 **universal optimization strategy for SARS-CoV-2 serology**

3 Thomas Perkmann MD*¹, Thomas Koller MD*¹, Nicole Perkmann-Nagele MD¹, Maria Oszvar-Kozma
4 MSc¹, David W Eyre DPhil², Philippa Matthews DPhil³, Abbie Bown MD⁴, Nicole Stoesser DPhil³, Marie-
5 Kathrin Breyer MD PhD⁵, Robab Breyer-Kohansal MD⁵, Otto C Burghuber MD⁶, Sylvia Hartl MD^{5,6}, Daniel
6 Aletaha MD⁷, Daniela Sieghart PhD⁷, Peter Quehenberger MD¹, Rodrig Marculescu MD¹, Patrick Mucher
7 MSc¹, Astrid Radakovics MSc², Miriam Klausberger PhD⁸, Mark Duerkop PhD⁹, Barbara Holzer PhD¹⁰,
8 Boris Hartmann PhD¹⁰, Robert Strassl MD¹, Gerda Leitner MD¹¹, Florian Grebien PhD¹², Wilhelm Gerner
9 PhD^{13,14,15}, Reingard Grabherr PhD⁸, Oswald F Wagner MD¹, Christoph J Binder MD PhD¹, and Helmuth
10 Haslacher MD PhD^{1*}

11 ¹Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria

12 ²Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK

13 ³Nuffield Department of Medicine, University of Oxford, Oxford, UK

14 ⁴Public Health England (PHE) Porton Down, Salisbury, UK

15 ⁵Department of Respiratory and Critical Care Medicine and Ludwig Boltzmann Institute for Lung Health, Otto Wagner Hospital,
16 Vienna, Austria

17 ⁶Sigmund Freud University, Medical School and Ludwig Boltzmann Institute for Lung Health, Vienna, Austria

18 ⁷Division of Rheumatology, Department of Medicine III, Medical University of Vienna, Vienna, Austria

19 ⁸Institute of Molecular Biotechnology, Department of Biotechnology, University of Natural Resources and Life Sciences (BOKU)
20 Vienna, Vienna, Austria

21 ⁹Institute of Bioprocess Science and Engineering, Department of Biotechnology, University of Natural Resources and Life Sciences
22 (BOKU) Vienna, Vienna, Austria

23 ¹⁰Institute for Veterinary Disease Control, Austrian Agency for Health and Food Safety (AGES), Moedling, Austria

24 ¹¹Department of Blood Group Serology and Transfusion Medicine, Medical University of Vienna, Vienna, Austria

25 ¹²Institute for Medical Biochemistry, University of Veterinary Medicine, Vienna, Austria

26 ¹³Institute of Immunology, University of Veterinary Medicine Vienna, Austria

27 ¹⁴Christian Doppler Laboratory for an Optimized Prediction of Vaccination Success in Pigs, University of Veterinary Medicine,
28 Vienna, Austria

29 ¹⁵The Pirbright Institute, Pirbright, United Kingdom (current)

30 **both authors contributed equally*

31 Address all correspondence to:

32 Helmuth HASLACHER, LECT MD PhD MSc BSc BA

33 Medical University of Vienna

34 Department of Laboratory Medicine

35 Waehringer Guertel 18-20

36 1090 Vienna

37 Austria

38 Tel: +43 1 40400 53190

39 Fax: +43 1 40495 15547

40 Mail: helmuth.haslacher@meduniwien.ac.at

41 **Background**

42 Reliable antibody tests are an essential tool to identify individuals who have developed
43 an adaptive immune response to SARS-CoV-2. However, attempts to maximize the
44 specificity of SARS-CoV-2 antibody tests have come at the cost of sensitivity,
45 exacerbating the total test error with increasing seroprevalence. Here, we present a
46 novel method to maximize specificity while maintaining or even increasing sensitivity: the
47 "Sensitivity Improved Two-Test" or "SIT²" algorithm.

48 **Methods**

49 SIT² involves confirmatory re-testing of samples with results falling in a predefined
50 retesting-zone of an initial screening test, with adjusted cut-offs to increase sensitivity.
51 We verified and compared the performance of SIT² to single tests and orthogonal testing
52 (OTA) in an Austrian cohort (1,117 negative, 64 post-COVID positive samples) and
53 validated the algorithm in an independent British cohort (976 negatives, 536 positives).

54 **Results**

55 The specificity of SIT² was superior to single tests and non-inferior to OTA. The
56 sensitivity was maintained or even improved using SIT² when compared to single tests
57 or OTA. SIT² allowed correct identification of infected individuals even when a live virus
58 neutralization assay could not detect antibodies. Compared to single testing or OTA,
59 SIT² significantly reduced total test errors to 0·46% (0·24-0·65) or 1·60% (0·94-2·38) at
60 both 5% or 20% seroprevalence.

61 **Conclusion**

62 SIT² proved to be the best diagnostic choice at both 5% and 20% seroprevalence in all
63 tested scenarios. It is an easy algorithm to apply to different available SARS-CoV-2

- 64 antibody testing systems and can potentially be helpful for the serology of other
65 infectious diseases.

66 **Introduction**

67 Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has led to a
68 worldwide pandemic confronting us with unprecedented epidemiological, therapeutic,
69 but not at last diagnostic challenges.

70 While case identification focuses primarily on polymerase chain reaction (PCR) or
71 antigen testing, detecting antibodies to SARS-CoV-2 is essential to identify individuals
72 who develop a SARS-CoV-2-specific adaptive immune response following infection or
73 vaccination. So, reliable determination of SARS-CoV-2 antibodies can provide critical
74 information for healthcare decision-making (1). Accurate measurement of antibody
75 levels is paramount in estimating seroprevalences because subsequent attempts to
76 compensate for antibody test errors can result in highly variable estimates of the actual
77 seroprevalence rates(2). Moreover, antibody positivity has also been shown to correlate
78 with protective immunity(3, 4) and several quantitative assays have been used to
79 determine vaccine-induced antibody levels(5, 6).

80 Current test systems measure antibodies against either the SARS-CoV-2 nucleocapsid
81 (NC) or spike (S) protein or fragments thereof(7). For correct serological results, both
82 high specificity and sensitivity are crucial. The suboptimal specificities of some of the
83 early SARS-CoV-2 antibody tests were problematic in the context of the low
84 seroprevalence at the onset of the pandemic. This has changed rapidly, and currently
85 available SARS-CoV-2 antibody tests are amongst the most specific serological tests
86 available. In addition, despite marked regional differences, global seroprevalence has
87 increased significantly(8). Unfortunately, the claim by many test manufacturers that
88 assay sensitivities are close to 100% has proven to be inaccurate and too optimistic in

89 practice(7, 9). Two major possible reasons underlying this are that: i) lower antibody
90 levels in non-hospitalized mild or asymptomatic cases sometimes fall below
91 manufacturers' cut-offs, which are set to optimize specificity to limit false-positive
92 results(10, 11), and ii) the decline in antibody titers over time towards the positivity
93 threshold increases the demands on assay sensitivity(12, 13). The latter phenomenon is
94 dependent on the test system used(14, 15) and can be a considerable problem, as
95 shown by an epidemiological study from Brazil(2).

96 Consequently, maximizing both the sensitivity and specificity of SARS-CoV-2 serology is
97 essential. However, there has been no attempt to optimize both parameters
98 simultaneously through a systematic approach to our knowledge. Generally, there is a
99 trade-off between sensitivity and specificity for a given test depending on the selected
100 threshold for positivity; increasing specificity to near 100% comes at the expense of
101 sensitivity, and vice versa. Attempts to increase sensitivity by lowering thresholds for
102 positivity below the recommended values set by manufacturers lead to unacceptably low
103 specificities. Orthogonal testing algorithms (OTA), as recommended by the Center for
104 Disease Control (CDC) for low seroprevalence settings, increase specificity at the
105 expense of sensitivity(16). Although this recommendation was justified at the beginning
106 of the pandemic, it is no longer appropriate given the significant increase in
107 seroprevalence in most areas of the world and the known sensitivity problem of SARS-
108 CoV-2 antibody tests and contributes to the rise in the overall error rate.

109 In this paper, we present a new, simple, and widely applicable method, the Sensitivity-
110 Improved Two-Test (SIT²) algorithm, which, for the first time, allows the maximization of
111 specificity while maintaining or even increasing sensitivity.

112 **Methods**

113 **Study design and cohorts**

114 Sera used in this non-blinded prospective cross-sectional study were either residual
115 clinical specimens or samples stored in the MedUni Wien Biobank (n=1,181), a facility
116 specialized in the preservation and storage of human biomaterial, which operates within
117 a certified quality management system (ISO 9001:2015)(17).

118 For derivation of the SIT² algorithm, sample sets from individuals known to be negative
119 and positive were established for testing. As previously described(11), samples collected
120 before 01.01.2020 (i.e., assumed SARS-CoV-2 negative) were used as a specificity
121 cohort (n=1117): a cross-section of the Viennese population (the LEAD study)(18),
122 preselected for samples collected between November and April to enrich for seasonal
123 infections (n=494); a collection of healthy voluntary donors (n= 265); a disease-specific
124 collection of samples from patients with rheumatic diseases (n=358); (see also Tables
125 S1 and S2).

126 Of the SARS-CoV-2 positive cohort (n=64 samples from n=64 individuals), five
127 individuals were asymptomatic, 42 had mild-moderate symptoms, four reported severe
128 symptoms, and 13 were admitted to the Intensive Care Unit (ICU). The timing of
129 symptom onset was determined by a questionnaire for convalescent donors and by
130 reviewing individual health records for patients. For asymptomatic donors (n=5), SARS-
131 CoV-2 RT-PCR confirmation time was used instead (for more details, see Tables S1 and
132 S3). All included participants gave written informed consent to donate their samples for
133 research purposes. The overall evaluation plan conformed with the Declaration of

134 Helsinki as well as relevant regulatory requirements. It was reviewed and approved by
135 the ethics committee of the Medical University of Vienna (1424/2020).

136 For validation of the SIT² algorithm, we used data from an independent United Kingdom
137 cohort(19), including 1,512 serum/plasma samples (536 PCR confirmed SARS-CoV-2
138 positive cases; 976 negative cases, collected earlier than 2017).

139 [Antibody testing](#)

140 For the derivation analyses, SARS-CoV-2 antibodies were either measured according to
141 the manufacturers' instructions on three different commercially available automated
142 platforms (Roche Elecsys® SARS-CoV-2 [nucleocapsid total antibody assay, further
143 referred to as Roche NC], Abbott SARS-CoV-2-IgG assay [nucleocapsid IgG assay,
144 Abbott NC], DiaSorin LIASION® SARS-CoV-2 S1/S2 assay [S1/S2 combination antigen
145 IgG assay, DiaSorin S1/S2]) or using 96-well enzyme-linked immunosorbent assays
146 (ELISAs) (Technoclone Technozym® RBD and Technozym® NP) yielding quantitative
147 results(20) (for details see Supplement, Supplemental Methods). The antibody assays
148 used in the validation cohort were Abbott NC, DiaSorin S1/S2, Roche NC, Siemens RBD
149 total antibody, and a novel 384-well trimeric spike protein ELISA (Oxford
150 Immunoassay)(19), resulting in 20 evaluable combinations. All samples from the
151 Austrian SARS-CoV-2-positive cohort also underwent live virus neutralization testing
152 (VNT), and neutralization titers (NT) were calculated, as is described in detail in the
153 Supplemental Methods.

154

155 Sensitivity improved two-test method (SIT²)

156 Our newly developed sensitivity improved two-test (SIT²) method consists of the
157 following key components: i) sensitivity improvement by cut-off modification and ii)
158 specificity rescue by a second, confirmatory test (Fig. 1A).

159 For the first component of the SIT² algorithm, positivity thresholds were optimized for
160 sensitivity according to the first published alternative thresholds for the respective
161 assays(21-23). Additionally, a high cut-off, above which a result can be reliably regarded
162 as true positive without the need for further confirmation, was defined. These levels were
163 based on in-house observations(11). The lowering of positivity thresholds improves
164 sensitivity; the high cut-off prevents unnecessary re-testing of clearly positive samples.
165 Moreover, the high cut-off avoids possible erroneous exclusion by the confirmatory test.
166 The newly defined interval between the reduced threshold for positivity and the high cut-
167 off is the re-testing zone (Fig. 1A). The initial antibody test (screening test) is then
168 followed by a confirmatory test, whereby positive samples from the re-testing zone of the
169 screening test are re-tested. Also, for the confirmatory test, sensitivity-adapted assay
170 thresholds are needed (Figs.1A, 1B). As false-positive samples are usually only positive
171 in one test system (Fig. S1), false positives can be identified, and specificity markedly
172 restored with minimal additional testing as most samples do not fall within the re-testing
173 zone(11, 24). A flowchart of the testing strategy and the applied cut-off levels and their
174 associated quality criteria are presented in Figs. 1B, 1C.

175 Test strategy evaluation

176 On the derivation cohort, we compared the overall performance of the following SARS-
177 CoV-2 antibody testing strategies: i) testing using single assays; ii) simple lowering of

178 thresholds; iii) classical orthogonal testing (OTA), and iv) our newly developed SIT²
179 algorithm at assumed seroprevalences of 5% and 20%. As part of the derivation, we
180 then compared the performance of OTAs and SIT² against the results of a virus
181 neutralization assay. On the validation cohort, we then compared the performance of
182 OTAs and SIT². Finally, we used data from this cohort to evaluate the performance of
183 SIT² versus single tests at seroprevalences of 5%, 10%, 20%, and 50% if the Abbott and
184 DiaSorin assays (i.e., assays with varying degrees of discrepancies in sensitivity and
185 specificity) were used.

186 [Statistical analysis](#)

187 Unless otherwise indicated, categorical data are given as counts (percentages), and
188 continuous data are presented as median (interquartile range). Total test errors were
189 compared by Mann-Whitney tests or, in case they were paired, by Wilcoxon tests. 95%
190 confidence intervals (CI) for sensitivities and specificities were calculated according to
191 Wilson, 95% CI for predictive values were computed according to Mercaldo-Wald unless
192 otherwise indicated. Sensitivities and specificities were compared using z-scores. P
193 values <0.05 were considered statistically significant. All calculations were performed
194 using Analyse-it 5.66 (Analyse-it Software, Leeds, UK) and MedCalc 19.6 (MedCalc
195 bvba, Ostend, Belgium). Graphs were drawn using Microsoft Visio (Armonk, USA) and
196 GraphPad Prism 7.0 (La Jolla, USA).

197 **Results**

198 Based on the derivation cohort of 1,117 pre-pandemic sera and 64 sera from
199 convalescent COVID-19 patients (80% non-hospitalized, 20% hospitalized), the Roche
200 NC, Abbott NC, and DiaSorin S1/S2 antibody assays gave rise to 545, 780, and 860
201 false-negative results per 100,000 tests, and 285, 760 and 1,710 false-positive results
202 per 100,000 tests respectively, assuming a seroprevalence of 5% (Fig. 2A, left panel).

203 **Effects of threshold lowering on Sensitivity and Specificity**

204 Lowering the positivity thresholds for the Roche NC, Abbott NC, and Diasorin S1/S2 to
205 17%, 38% and 75% of the manufacturers cut-offs increased the sensitivity significantly
206 and reduced false-negative results to 80, 155, and 545 per 100,000 tests, respectively,
207 but substantially increased false-positive results to 1,520, 2,280 and 2,660 per 100,000
208 tests respectively (all at an assumed seroprevalence of 5%; Table S4, Fig. 2A, left
209 panel).

210 **Classical OTA compared to SIT²**

211 Subsequently, we evaluated 12 OTA combinations using the fully automated SARS-
212 CoV-2 antibody tests from Roche NC, Abbott NC, and DiaSorin S1/S2 as screening
213 tests, each combined with one of the other fully automated assays or a commercially
214 available NC or RBD-specific ELISA as a confirmation test. Combining these tests as
215 classical OTAs significantly increased specificity and reduced false positives to 0 (0-95)
216 per 100,000 tests. However, the rate of false negatives was 1,095 (955-1,230) per
217 100,000 tests, and therefore considerably higher than for single testing strategies. In
218 contrast, the SIT² algorithm minimized false positives to 0 (0-166·25) per 100,000 tests

219 while also reducing false negatives to 390 (235-607.5) per 100,000 tests (Fig. 2A left
220 panel; Table S5).

221 At an assumed seroprevalence of 20%, the consequences of reduced test sensitivities
222 led to an even more substantial rise in false-negative results for all testing strategies
223 (Fig. 2A, right panel). Both SIT² and OTA reduced false-positive results substantially
224 compared to single testing strategies (SIT²: 0 [0-140] and OTA: 0 [0-80] false-positives
225 per 100,000 tests). However, SIT² achieved a substantially lower rate of false-negative
226 results when compared to OTA (1,560 [940-2,420] vs 4,380 [3,820-4,920] per 100,000
227 tests). The reduction in false-positive results using SIT² compared to single tests with
228 the manufacturers' recommended thresholds (Table S6) was accompanied by a lower
229 average false-negative rate for SIT² than for single tests (Fig. 2A, right panel). Even at
230 20% seroprevalence, single tests resulted in a non-negligible number of false positives
231 (240, 640, and 1,440 per 100,000 tests).

232 Reduction of total error rates by the Sensitivity-Improved Two-Test

233 Of all the methods assessed, SIT² reached the lowest total error rates per 100,000 tests
234 under both 5% and 20% assumed seroprevalence (455 [235-685] and 1,600 [940-2,490]
235 per 100,000 tests) (Fig. 2B). At a seroprevalence of 5 %, OTA on average performed
236 better than individual tests, and the total error rates of the single tests were higher for
237 the Abbott NC and DiaSorin S1/S2 assay (OTA 1,095 [955-1,325] vs. 830 [Roche NC],
238 1,540 [Abbott NC] and 2,570 [DiaSorin S1/S2] per 100,000 tests). But with a
239 seroprevalence of 20 %, performance of OTAs, worsened compared to single tests
240 (OTA 4,380 [3,820-5,000] vs 1,600 [Roche], 2,540 [Abbott] and 4,420 [DiaSorin] per
241 100,000 tests) (Fig. 2B). Therefore, at both 5% and 20% seroprevalence, SIT² resulted

242 in the lowest overall errors. Compared to OTAs, SIT² yielded a similar improvement in
243 specificity while not suffering from the significant sensitivity reduction (Fig. S2). Since the
244 better overall performance of SIT² compared to OTAs was not due to increased
245 specificity but improved sensitivity, we subsequently set out to examine these
246 differences in more detail.

247 Sensitivities of single tests, OTA and SIT² in relation to Neutralization Testing

248 Next, we compared the sensitivities of the three screening tests as single tests and in
249 both two-test methods (OTA and SIT²), benchmarking them using the Austrian
250 sensitivity cohort (n=64) simultaneously evaluated with an authentic SARS-CoV-2 virus
251 neutralization test (VNT). Regardless of the screening test used (Roche NC, Abbott NC,
252 or DiaSorin S1/S2), OTAs had lower sensitivities than single tests (80·5% [78·5-83·6],
253 78·1% [75·8-82·8], or 75·8% [71·5-78·9] vs. 89·1%, 84·4%, or 82·8% respectively), and
254 SIT² showed the best sensitivities of all methods (95·3% [93·0-96·5], 93·8% [92·2-96·5],
255 or 87·5% [85·1-88·7]) (Fig. 3). SIT² algorithms, including the Roche NC and Abbott NC
256 assays, achieved similar or even higher sensitivities than VNT (Fig. 3, VNT reference
257 line), made possible by the unique re-testing zone of SIT² (Fig. S3).

258 Validation of the Sensitivity-Improved Two-Test using an independent cohort

259 To confirm the improved sensitivity of SIT² compared to OTA, we analyzed the
260 sensitivities of OTAs and SIT² in an independent validation cohort of 976 pre-pandemic
261 samples and 536 post-COVID samples. Out of 20 combinations using the assays Roche
262 NC (total antibody), Abbott NC (IgG), DiaSorin S1/S2 (IgG), Siemens RBD (total
263 antibody), and Oxford trimeric-S (IgG), a statistically significant improvement in
264 sensitivities over OTAs was shown for SIT² in 18 combinations (Fig. 4). The

265 performance was comparable for the remaining two combinations (Siemens RBD with
266 Oxford trimeric-S and vice versa). Still, no statistically significant improvement could be
267 achieved due to the high pre-existing sensitivities of these assays on this particular
268 sample cohort.

269 To further illustrate the effect of SIT² on the outcome of SARS-CoV-2 antibody testing,
270 we compared single testing versus SIT² with the Abbott and DiaSorin assays at varying
271 assumed seroprevalences (5, 10, 20, and 50%), given that the Abbott NC assay is a
272 highly specific (99.9%), but moderately sensitive test (92.7%), and the DiaSorin S1/S2
273 assay has the most limited specificity (98.7%) of all evaluated assays but an acceptable
274 sensitivity (96.3%). Regardless of whether a lack of specificity (DiaSorin S1/S2) or
275 sensitivity (Abbott NC) had to be compensated for, SIT² improved the overall error rate
276 compared to the individual tests in all four combinations and at all four assumed
277 seroprevalence levels (Fig. 5).

278 Discussion

279 The simplest and most practical way to characterize the immune response to prior
280 SARS-CoV-2 infection and vaccination is to measure SARS-CoV-2-specific antibodies.
281 As long as the COVID-19 pandemic continues, serology-based assays' ongoing use and
282 development will play a role in controlling and surveillance of this disease. The
283 importance and value of SARS-CoV-2 antibody assays are well established(1, 3, 25-27),
284 although some performance limitations may affect their applicability. To date, variable
285 seroprevalence and very low antibody levels, especially in asymptomatic or mild
286 disease(10), have challenged the performance of available tests(9, 11).

287 In the present work, we first sought to overcome these limitations by lowering positivity
288 thresholds to improve test sensitivity(21-23, 28) or by using conventional orthogonal
289 testing to maximize specificity(16, 29-32). As recommended by the CDC and used in
290 several seroprevalence studies(31, 33), OTAs indeed maximized specificity and thus
291 minimized false-positive results in the two sample cohorts evaluated, regardless of
292 whether the assumed seroprevalence was 5% or 20%. However, this was accompanied
293 by a further reduction of the a priori sub-optimal sensitivities of individual tests and
294 increased false negatives, as shown previously(16). In seroprevalence studies,
295 insufficient sensitivities can lead to apparent discrepancies. This is exemplified in a large
296 study in Iceland, where only slightly more than 90% of individuals testing positive for
297 SARS-CoV-2 by RT-PCR were seropositive(33), resulting from a priori low sensitivities
298 of the single tests used or the OTA applied, or a combination of both.

299 Reducing the cut-offs for positivity leads to a significant increase in sensitivity, also
300 shown previously(2, 21, 22), but is associated with a dramatic increase in false-positive

301 results at low seroprevalence. This effect only gradually decreases with increasing
302 seroprevalence (Fig. 5).

303 The rates of false positives, false negatives, and overall errors of test systems depend
304 on their specificity and sensitivity and the respective seroprevalence in the population. At
305 a population level, at low seroprevalences, a test system with suboptimal specificity
306 combined with insufficient sensitivity can add up to a correctly determined
307 seroprevalence rate since false-positive, and false-negative results can compensate for
308 each other (Fig. S2, left panel). However, at an individual level, the determination of
309 serostatus in such a context is highly flawed, making individual-level results unreliable
310 and unsuitable for management and subgroup analyses. For single tests, specificity and
311 sensitivity always behave in opposite ways since an increase in the positivity threshold
312 increases the specificity but lowers the sensitivity and vice versa. Low test specificity
313 leads to a significant increase in the overall error rate at low seroprevalence. Low test
314 sensitivity leads to a substantial increase in the overall error rate at high seroprevalence.
315 However, since seroprevalence is often neither known and varies widely from region to
316 region, it is difficult to judge whether a less specific or less sensitive test is the lesser of
317 two evils.

318 To overcome these limitations, we propose a new, universally adaptable two-test system
319 that can perform better than any other known approach regardless of the actual
320 seroprevalence: the sensitivity-improved Two-Test or SIT². Its generalizability can be
321 inferred from the following features: i) the adapted cut-offs used to optimize sensitivity
322 were determined in various independent studies and were not explicitly calculated for
323 our cohort(21-23); ii) SIT² was effective, albeit with different efficiencies, in a total of 32

324 different test combinations; and iii) SIT² was successfully validated in an independent
325 cohort which was profoundly different from the derivation cohort. Therefore, SIT² does
326 not require a particular infrastructure or the availability of high-performance individual
327 test systems but achieves the best performance out of an available test.

328 Our SIT² strategy can rescue the specificity with minimal repeat testing required (see
329 Table S6). For example, when applying the Roche NC as a screening test to our cohort,
330 only 27 out of 1,181 samples needed confirmation testing with the Abbott NC test to
331 correctly identify 62/64 true positives. Simultaneously, all false-positive results were
332 eliminated, including those added by lowering the cut-offs (Table S4 and Fig. S1).
333 Additionally, it was more sensitive than virus neutralization testing, which identified only
334 60/64 clinical positives (Fig. 3). This result is not completely surprising as it is known that
335 not all patients who recovered from COVID-19 show detectable levels of neutralizing
336 antibodies(34). Nevertheless, it should be noted that although antibody binding assays
337 may have a higher sensitivity than neutralization assays, they only partially reflect the
338 functional activity of SARS-CoV-2-reactive antibodies(35, 36).

339 In addition to the low or absent antibody formation shortly after COVID-19 infection(10,
340 34), the antibody levels also decline over time(14), and this is dependent on the assay
341 used. For example, an antibody-positivity half-life of ~85 days has been specifically
342 described for the Abbott NC test(15). The decrease in antibody levels or antibody
343 reactivity measured in specific assays can lead to a status of "seroreversion" in post-
344 COVID-19 individuals. While some of this disappearance may reflect a genuine loss of
345 antibody (i.e., true seroreversion), some of it reflects an artifact caused by the decline of
346 antibody levels below assay positivity thresholds, as remarkably illustrated in a recent

347 Brazilian study. The seroprevalence (measured by Abbott NC) apparently decreased
348 from 46% to just under 21% between June 2020 and October 2020, when the
349 manufacturer's recommended cut-off of 1.4 S/C was applied. In contrast, the same test
350 results with a threshold reduced to 0.4 S/C yielded ~54% seroprevalence in June 2020
351 and around 45% in October 2020(2).

352 Our study has both strengths and limitations. One strength is the size of the cohorts
353 examined, both in deriving the SIT² algorithm (N=1,181) and validating it (N=1,512). The
354 composition of our specificity cohort is also unique: it consists of three sub-cohorts with
355 selection criteria to further challenge analytical specificity. The lower cut-offs used to
356 increase sensitivity were not modeled within our datasets but were derived from ROC-
357 analyses data of independent studies(21-23). Furthermore, we were able to test the
358 performance of the two-test systems in a total of 32 combinations, 12 in the derivation
359 cohort and another 20 combinations in the validation cohort. As a limitation, in the
360 Austrian cohort, only samples ≥ 14 days after symptom onset were included. Therefore,
361 no conclusions on the sensitivity of the early seroconversion phase can be made from
362 these data. Furthermore, mild and asymptomatic cases were underrepresented in the
363 British cohort, perhaps leading to an observed higher sensitivity of the test systems.

364 In conclusion, we describe the novel two-test algorithm SIT² that, in contrast to classical
365 orthogonal testing, can simultaneously maximize the specificity and the sensitivity of
366 SARS-CoV-2 antibody tests. Poor sensitivity remains a problem for SARS-CoV-2
367 antibody tests. Low a priori antibody levels in many post-COVID-19 individuals
368 combined with the steady antibody decline over time reduce effective sensitivity and
369 lead to many errors in the assignment of SARS-CoV-2 serostatus. In the early post-

370 vaccine era of the pandemic, reliable assignation of SARS-CoV-2 serostatus will
371 become even more relevant to public health and infection control policy. With SIT², we
372 present a platform-independent and straightforward solution to minimize the overall error
373 rate in determining SARS-CoV-2 antibody presence at the individual and population
374 level, thereby increasing the accuracy of subsequent decisions.

375 **Acknowledgments**

376 We sincerely thank Marika Gerdov, Susanne Keim, Karin Mildner, Elisabeth Ponweiser,
377 Manuela Repl, Ilse Steiner, Christine Thun, and Martina Trella for excellent technical
378 assistance. Finally, we want to thank all the donors of the various study cohorts. The
379 MedUni Wien Biobank is funded to participate in the biobank consortium BBMRI.at
380 (www.bbmri.at) by the Austrian Federal Ministry of Science, Research and Technology.
381 There was no external funding received for the work presented. However, test kits for
382 the Technoclone ELISAs were kindly provided by the manufacturer.

383 **Declaration of interests**

384 NP received a travel grant from DiaSorin. DWE reports lecture fees from Gilead outside
385 the submitted work. OCB reports grants from GSK, grants from Menarini, grants from
386 Boehringer Ingelheim, grants from Astra, grants from MSD, grants from Pfizer, and
387 grants from Chiesi, outside the submitted work. SH does receive unrestricted research
388 grants (GSK, Boehringer, Menarini, Chiesi, Astra Zeneca, MSD, Novartis, Air Liquide,
389 Vivisol, Pfizer, TEVA) for the Ludwig Boltzmann Institute of COPD and Respiratory
390 Epidemiology, and is on advisory boards for G. SK, Boehringer Ingelheim, Novartis,
391 Menarini, Chiesi, Astra Zeneca, MSD, Roche, Abbvie, Takeda, and TEVA for respiratory
392 oncology and COPD. PQ is an advisory board member for Roche Austria and reports
393 personal fees from Takeda outside the submitted work. The Dept. of Laboratory
394 Medicine (Head: OWF) received compensations for advertisement on scientific
395 symposia from Roche, DiaSorin, and Abbott and holds a grant for evaluating an in-vitro
396 diagnostic device from Roche. CJB is a Board Member of Technoclone. HH receives

397 compensations for biobank services from Glock Health Science and Research and
398 BlueSky immunotherapies.

399 **References**

- 400 1. Alter G, Seder R. The Power of Antibody-Based Surveillance. *N Engl J Med* 2020
401 Oct 29;383 18:1782-4. Epub 2020/09/02 as doi: 10.1056/NEJMe2028079.
- 402 2. Buss LF, Prete CA, Jr., Abraham CMM, Mendrone A, Jr., Salomon T, de Almeida-
403 Neto C, Franca RFO, et al. Three-quarters attack rate of SARS-CoV-2 in the
404 Brazilian Amazon during a largely unmitigated epidemic. *Science* 2021 Jan
405 15;371 6526:288-92. Epub 2020/12/10 as doi: 10.1126/science.abe9728.
- 406 3. Lumley SF, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, Hatch SB,
407 Marsden BD, et al. Antibody Status and Incidence of SARS-CoV-2 Infection in
408 Health Care Workers. *N Engl J Med* 2020 Dec 23. Epub 2020/12/29 as doi:
409 10.1056/NEJMoa2034545.
- 410 4. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, Angus B,
411 et al. Single-dose administration and the influence of the timing of the booster
412 dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine:
413 a pooled analysis of four randomised trials. *Lancet* 2021;397 10277:881—91 as
414 doi: 10.1016/s0140-6736(21)00432-3 PMID - 33617777.
- 415 5. Eyre DW, Lumley SF, Wei J, Cox S, James T, Justice A, Jesuthasan G, et al.
416 Quantitative SARS-CoV-2 anti-spike responses to Pfizer-BioNTech and Oxford-
417 AstraZeneca vaccines by previous infection status. *Clin Microbiol Infect* 2021 Jun
418 7. Epub 2021/06/11 as doi: 10.1016/j.cmi.2021.05.041.

- 419 6. Perkmann T, Perkmann-Nagele N, Koller T, Mucher P, Radakovics A,
420 Marculescu R, Wolzt M, et al. Anti-Spike Protein Assays to Determine SARS-
421 CoV-2 Antibody Levels: a Head-to-Head Comparison of Five Quantitative Assays.
422 *Microbiol Spectr* 2021 Jun 30:e0024721. Epub 2021/07/01 as doi:
423 10.1128/Spectrum.00247-21.
- 424 7. Wang H, Ai J, Loeffelholz MJ, Tang YW, Zhang W. Meta-analysis of diagnostic
425 performance of serology tests for COVID-19: impact of assay design and post-
426 symptom-onset intervals. *Emerg Microbes Infect* 2020 Dec;9 1:2200-11. Epub
427 2020/09/24 as doi: 10.1080/22221751.2020.1826362.
- 428 8. Bobrovitz N, Arora RK, Cao C, Boucher E, Liu M, Donnici C, Yanes-Lane M, et al.
429 Global seroprevalence of SARS-CoV-2 antibodies: A systematic review and
430 meta-analysis. *Plos One* 2021;16 6:e0252617. Epub 2021/06/24 as doi:
431 10.1371/journal.pone.0252617.
- 432 9. Oved K, Olmer L, Shemer-Avni Y, Wolf T, Supino-Rosin L, Prajgrod G, Shenhar
433 Y, et al. Multi-center nationwide comparison of seven serology assays reveals a
434 SARS-CoV-2 non-responding seronegative subpopulation. *EClinicalMedicine*
435 2020 Dec;29:100651. Epub 2020/11/26 as doi: 10.1016/j.eclinm.2020.100651.
- 436 10. Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, Hu JL, et al. Clinical and
437 immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med*
438 2020 Aug;26 8:1200-4. Epub 2020/06/20 as doi: 10.1038/s41591-020-0965-6.
- 439 11. Perkmann T, Perkmann-Nagele N, Breyer MK, Breyer-Kohansal R, Burghuber
440 OC, Hartl S, Aletaha D, et al. Side-by-Side Comparison of Three Fully Automated
441 SARS-CoV-2 Antibody Assays with a Focus on Specificity. *Clin Chem* 2020 Nov
442 1;66 11:1405-13. Epub 2020/08/11 as doi: 10.1093/clinchem/hvaa198.

- 443 12. Choe PG, Kim KH, Kang CK, Suh HJ, Kang E, Lee SY, Kim NJ, et al. Antibody
444 Responses 8 Months after Asymptomatic or Mild SARS-CoV-2 Infection. *Emerg*
445 *Infect Dis*;27 3:928-31 as doi: 10.3201/eid2703.204543 PMID - 33350923.
- 446 13. Seow J, Graham C, Merrick B, Acors S, Pickering S, Steel KJA, Hemmings O, et
447 al. Longitudinal observation and decline of neutralizing antibody responses in the
448 three months following SARS-CoV-2 infection in humans. *Nat Microbiol*;5
449 12:1598-607 as doi: 10.1038/s41564-020-00813-8 PMID - 33106674.
- 450 14. Muecksch F, Wise H, Batchelor B, Squires M, Semple E, Richardson C, McGuire
451 J, et al. Longitudinal analysis of serology and neutralizing antibody levels in
452 COVID19 convalescents. *J Infect Dis* 2020 Nov 3. Epub 2020/11/04 as doi:
453 10.1093/infdis/jiaa659.
- 454 15. Lumley SF, Wei J, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, Hatch
455 SB, et al. The duration, dynamics and determinants of SARS-CoV-2 antibody
456 responses in individual healthcare workers. *Clin Infect Dis* 2021 Jan 6. Epub
457 2021/01/06 as doi: 10.1093/cid/ciab004.
- 458 16. Turbett SE, Anahtar M, Dighe AS, Garcia Beltran W, Miller T, Scott H, Durbin SM,
459 et al. Evaluation of Three Commercial SARS-CoV-2 Serologic Assays and Their
460 Performance in Two-Test Algorithms. *J Clin Microbiol* 2020 Dec 17;59 1. Epub
461 2020/10/07 as doi: 10.1128/JCM.01892-20.
- 462 17. Eskandary F, Regele H, Baumann L, Bond G, Kozakowski N, Wahrmann M,
463 Hidalgo LG, et al. A Randomized Trial of Bortezomib in Late Antibody-Mediated
464 Kidney Transplant Rejection [eng]. *J Am Soc Nephrol* 2018 Feb;29 2:591-605.
465 Epub 2017/12/16 as doi: 10.1681/asn.2017070818.

- 466 18. Breyer-Kohansal R, Hartl S, Burghuber OC, Urban M, Schrott A, Agusti A,
467 Sigsgaard T, et al. The LEAD (Lung, Heart, Social, Body) Study: Objectives,
468 Methodology, and External Validity of the Population-Based Cohort Study. *J*
469 *Epidemiol* 2019 Aug 5;29 8:315-24. Epub 2018/10/23 as doi:
470 10.2188/jea.JE20180039.
- 471 19. National S-C-SAEG. Performance characteristics of five immunoassays for
472 SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis* 2020
473 Dec;20 12:1390-400. Epub 2020/09/27 as doi: 10.1016/S1473-3099(20)30634-4.
- 474 20. Klausberger M, Duerkop M, Haslacher H, Wozniak-Knopp G, Cserjan-
475 Puschmann M, Perkmann T, Lingg N, et al. A comprehensive antigen production
476 and characterisation study for easy-to-implement, specific and quantitative SARS-
477 CoV-2 serotests. *Ebiomedicine* 2021 Apr 24;67:103348. Epub 2021/04/28 as doi:
478 10.1016/j.ebiom.2021.103348.
- 479 21. Favresse J, Eucher C, Elsen M, Tre-Hardy M, Dogne JM, Douxfils J. Clinical
480 Performance of the Elecsys Electrochemiluminescent Immunoassay for the
481 Detection of SARS-CoV-2 Total Antibodies. *Clin Chem* 2020 Aug 1;66 8:1104-6.
482 Epub 2020/06/03 as doi: 10.1093/clinchem/hvaa131.
- 483 22. Lau CS, Oh HML, Hoo SP, Liang YL, Phua SK, Aw TC. Performance of an
484 automated chemiluminescence SARS-CoV-2 IG-G assay. *Clin Chim Acta* 2020
485 Nov;510:760-6. Epub 2020/09/11 as doi: 10.1016/j.cca.2020.09.005.
- 486 23. Bonelli F, Sarasini A, Zierold C, Calleri M, Bonetti A, Vismara C, Blocki FA, et al.
487 Clinical and Analytical Performance of an Automated Serological Test That
488 Identifies S1/S2-Neutralizing IgG in COVID-19 Patients Semiquantitatively. *J Clin*
489 *Microbiol* 2020 Aug 24;58 9. Epub 2020/06/26 as doi: 10.1128/JCM.01224-20.

- 490 24. Pfluger LS, Bannasch JH, Brehm TT, Pfefferle S, Hoffmann A, Norz D, van der
491 Meirschen M, et al. Clinical evaluation of five different automated SARS-CoV-2
492 serology assays in a cohort of hospitalized COVID-19 patients. *J Clin Virol* 2020
493 Sep;130:104549. Epub 2020/08/09 as doi: 10.1016/j.jcv.2020.104549.
- 494 25. Anand S, Montez-Rath M, Han J, Bozeman J, Kerschmann R, Beyer P,
495 Parsonnet J, et al. Prevalence of SARS-CoV-2 antibodies in a large nationwide
496 sample of patients on dialysis in the USA: a cross-sectional study. *Lancet* 2020
497 Sep 25. Epub 2020/09/29 as doi: 10.1016/S0140-6736(20)32009-2.
- 498 26. Pollan M, Perez-Gomez B, Pastor-Barriuso R, Oteo J, Hernan MA, Perez-Olmeda
499 M, Sanmartin JL, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a
500 nationwide, population-based seroepidemiological study. *Lancet* 2020 Aug
501 22;396 10250:535-44. Epub 2020/07/10 as doi: 10.1016/S0140-6736(20)31483-
502 5.
- 503 27. Stringhini S, Wisniak A, Piumatti G, Azman AS, Lauer SA, Baysson H, De Ridder
504 D, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva,
505 Switzerland (SEROCoV-POP): a population-based study. *Lancet* 2020 Aug 1;396
506 10247:313-9. Epub 2020/06/15 as doi: 10.1016/S0140-6736(20)31304-0.
- 507 28. Horber S, Soldo J, Relker L, Jurgens S, Guther J, Peter S, Lehmann R, et al.
508 Evaluation of three fully-automated SARS-CoV-2 antibody assays. *Clin Chem*
509 *Lab Med* 2020 Aug 3;58 12:2113-20. Epub 2020/08/04 as doi: 10.1515/cclm-
510 2020-0975.
- 511 29. Xu G, Emanuel AJ, Nadig S, Mehrotra S, Caddell BA, Curry SR, Nolte FS, et al.
512 Evaluation of Orthogonal Testing Algorithm for Detection of SARS-CoV-2 IgG

- 513 Antibodies. Clin Chem 2020 Sep 7. Epub 2020/09/08 as doi:
514 10.1093/clinchem/hvaa210.
- 515 30. Ripperger TJ, Uhrlaub JL, Watanabe M, Wong R, Castaneda Y, Pizzato HA,
516 Thompson MR, et al. Orthogonal SARS-CoV-2 Serological Assays Enable
517 Surveillance of Low-Prevalence Communities and Reveal Durable Humoral
518 Immunity. Immunity 2020 Nov 17;53 5:925-33 e4. Epub 2020/11/02 as doi:
519 10.1016/j.immuni.2020.10.004.
- 520 31. Bolotin S, Tran V, Osman S, Brown KA, Buchan SA, Joh E, Deeks SL, et al.
521 SARS-CoV-2 seroprevalence survey estimates are affected by anti-nucleocapsid
522 antibody decline. J Infect Dis 2021 Jan 5. Epub 2021/01/06 as doi:
523 10.1093/infdis/jiaa796.
- 524 32. Irsara C, Egger AE, Prokop W, Nairz M, Loacker L, Sahanic S, Pizzini A, et al.
525 Evaluation of four commercial, fully automated SARS-CoV-2 antibody tests
526 suggests a revision of the Siemens SARS-CoV-2 IgG assay. Clin Chem Lab Med
527 2021 Jan 14;0 0. Epub 2021/02/09 as doi: 10.1515/cclm-2020-1758.
- 528 33. Gudbjartsson DF, Norddahl GL, Melsted P, Gunnarsdottir K, Holm H, Eythorsson
529 E, Arnthorsson AO, et al. Humoral Immune Response to SARS-CoV-2 in Iceland.
530 N Engl J Med 2020 Oct 29;383 18:1724-34. Epub 2020/09/02 as doi:
531 10.1056/NEJMoa2026116.
- 532 34. Kalkan Yazici M, Koc MM, Cetin NS, Karaaslan E, Okay G, Durdu B, Sumbul B,
533 et al. Discordance between Serum Neutralizing Antibody Titers and the Recovery
534 from COVID-19. J Immunol 2020 Nov 15;205 10:2719-25. Epub 2020/09/27 as
535 doi: 10.4049/jimmunol.2000840.

- 536 35. Bal A, Pozzetto B, Trabaud MA, Escuret V, Rabilloud M, Langlois-Jacques C,
537 Paul A, et al. Evaluation of high-throughput SARS-CoV-2 serological assays in a
538 longitudinal cohort of patients with mild COVID-19: clinical sensitivity, specificity
539 and association with virus neutralization test. Clin Chem 2021 Jan 5. Epub
540 2021/01/06 as doi: 10.1093/clinchem/hvaa336.
- 541 36. Padoan A, Bonfante F, Pagliari M, Bortolami A, Negrini D, Zuin S, Bozzato D, et
542 al. Analytical and clinical performances of five immunoassays for the detection of
543 SARS-CoV-2 antibodies in comparison with neutralization activity. EBioMedicine
544 2020 Dec;62:103101. Epub 2020/11/08 as doi: 10.1016/j.ebiom.2020.103101.
- 545

546 **Figure Legends**

547 **Fig. 1. A)** *The Sensitivity Improved Two-Test (SIT²) algorithm includes sensitivity*
548 *improvement by adapted cut-offs and a subsequent specificity rescue by re-testing all*
549 *samples within the re-testing zone of the screening test by a confirmatory test. B)*
550 *Testing algorithm for SIT² utilizing a screening test on an automated platform*
551 *(ECLIA/Roche, CMIA/Abbott, CLIA/DiaSorin) and a confirmation test, either on one of*
552 *the remaining platforms or tested by means of ELISA (Technozym RBD, NP). C) All test*
553 *results between a reduced cut-off suggested by the literature, and a higher cut-off,*
554 *above which no more false-positives were observed, were subject to confirmation*
555 *testing. **... results between 12.0 and 15.0, which are according to the manufacturer*
556 *considered borderline, were treated as positives; ***... suggested as a cut-off for*
557 *seroprevalence testing; ****... determined by in-house modeling; ¹... see (21); ²... see*
558 *(22); ³... see (23).*

559 **Fig. 2.** *False-positives (FP)/false-negatives (FN) (A) and total error (B) of single tests,*
560 *tests with reduced thresholds according to (21-23), orthogonal testing algorithms (OTAs)*
561 *and the Sensitivity Improved Two-Test (SIT²) algorithm at 5 and 20% estimated*
562 *seroprevalence. Data in (B) were compared by Mann-Whitney tests (unpaired) or*
563 *Wicoxon tests (paired). *... $P < 0.05$; **... $P < 0.01$; ***... $P < 0.001$.*

564 **Fig. 3.** *Sensitivities of single tests, orthogonal testing algorithms (OTAs) and the*
565 *Sensitivity Improved Two-Test (SIT²) algorithm. The dotted line indicates the sensitivity*
566 *of virus neutralization test (VNT).*

567 **Fig. 4.** *Differences in sensitivity and specificity (mean \pm 95% confidence interval) between*
568 *the Sensitivity Improved Two-Test (SIT²) algorithm and standard orthogonal testing*

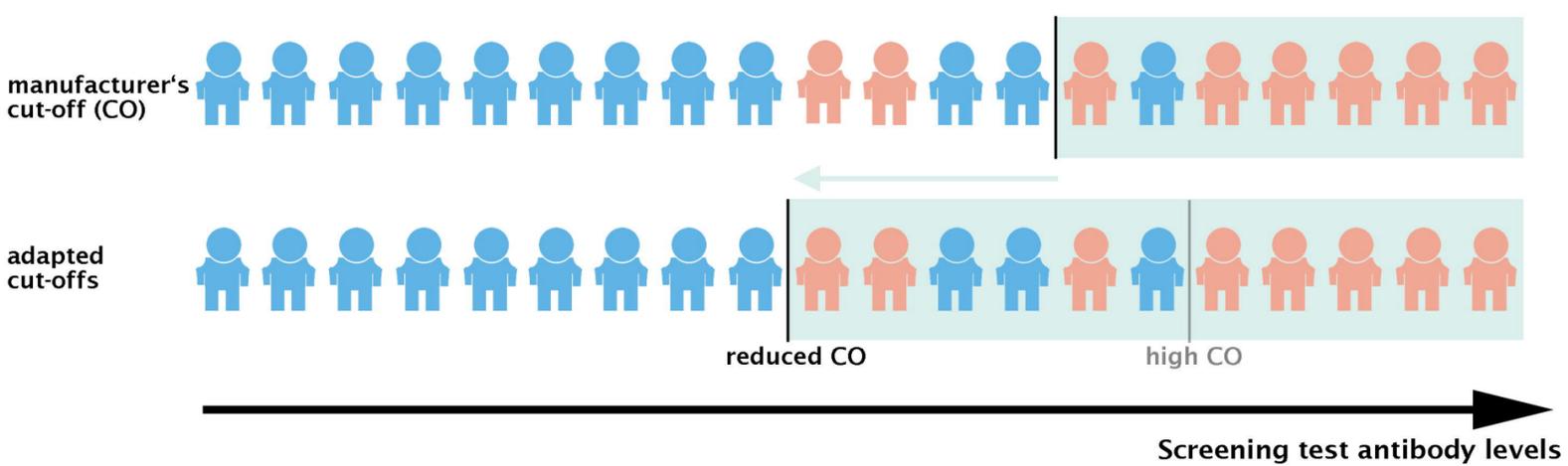
569 *algorithms (OTAs) within the UK validation cohort. *... P<0.05; **...P<0.01;*

570 ****...P<0.001; ****...P<0.0001*

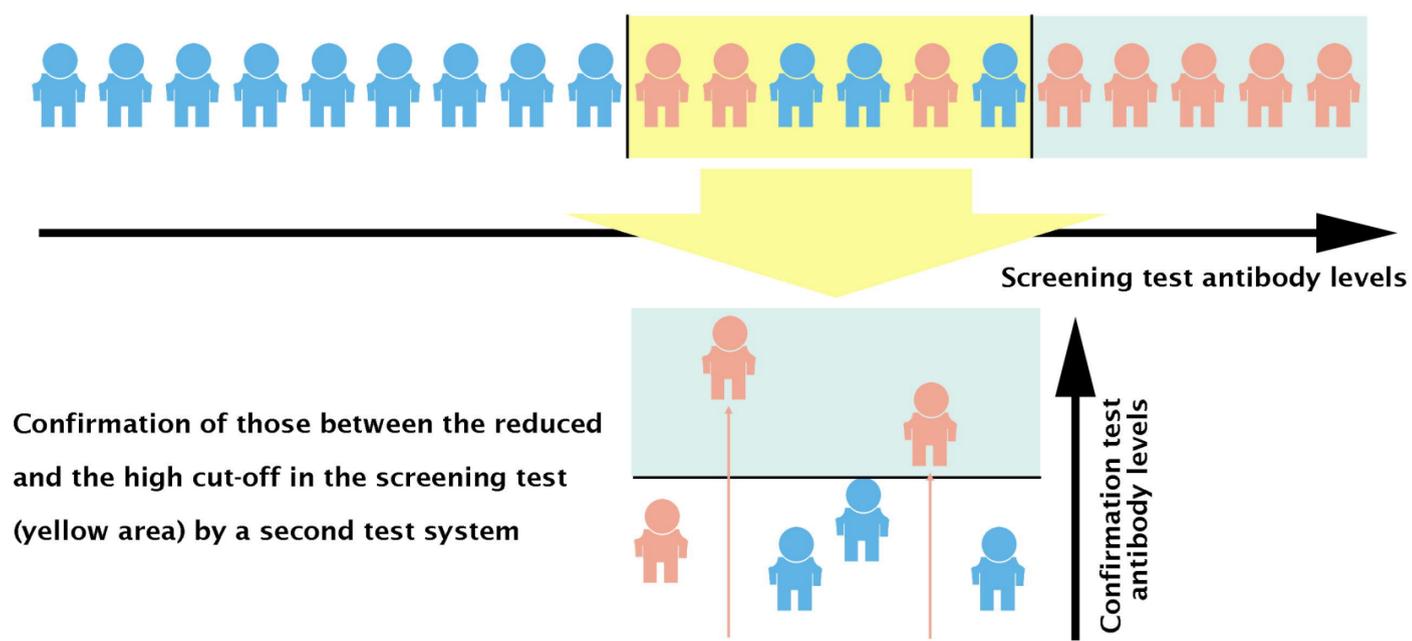
571 **Fig. 5.** *Comparing false-positives (FP), false-negatives (FN), and total error (TE) for two*
572 *selected test systems, **A)** Abbott, **B)** DiaSorin, between different Sensitivity Improved*
573 *Two-Test (SIT²) combinations and the respective single test within the UK validation*
574 *cohort for different estimated seroprevalences.*

A

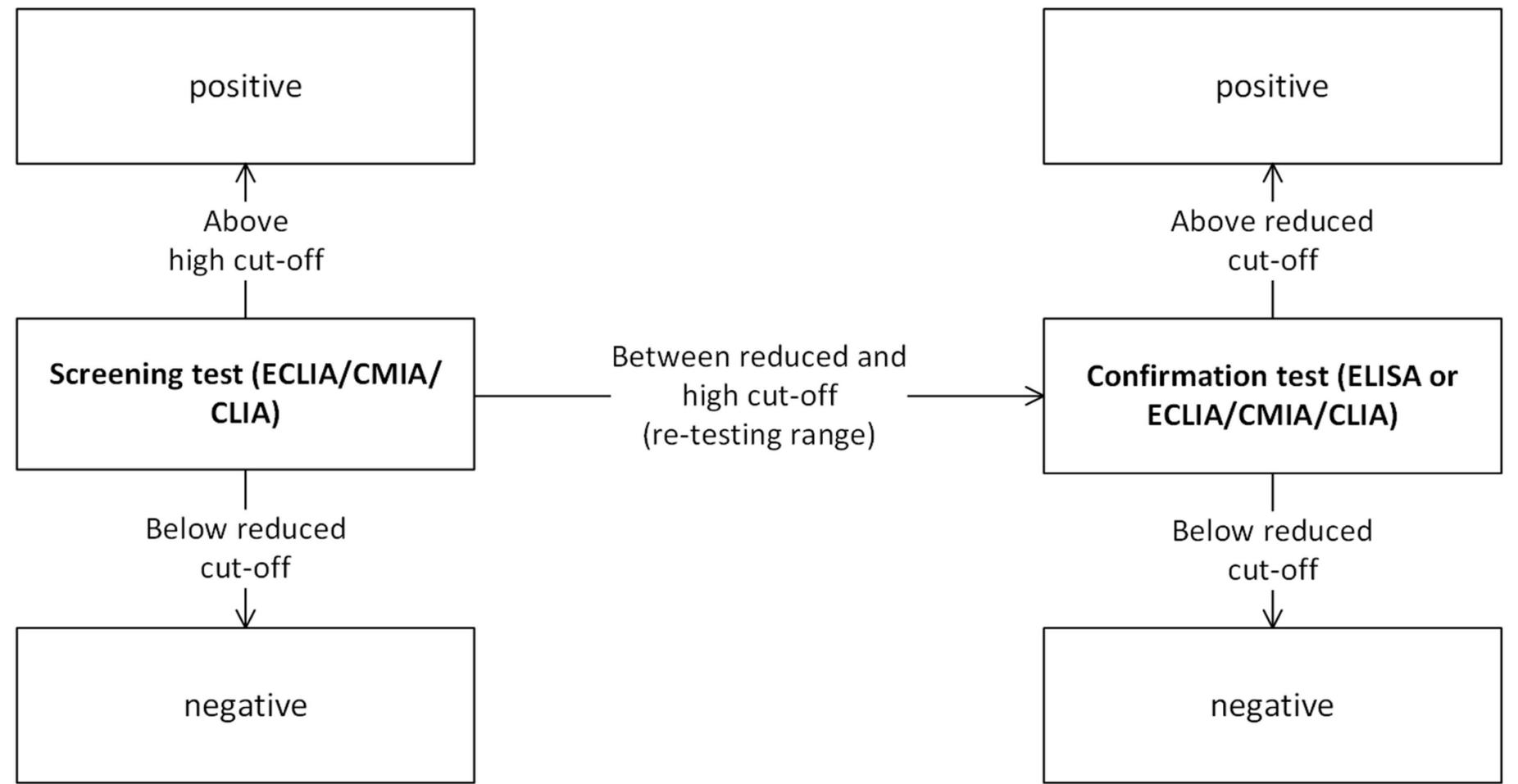
Step 1: Sensitivity Improvement



Step 2: Specificity Rescue



B

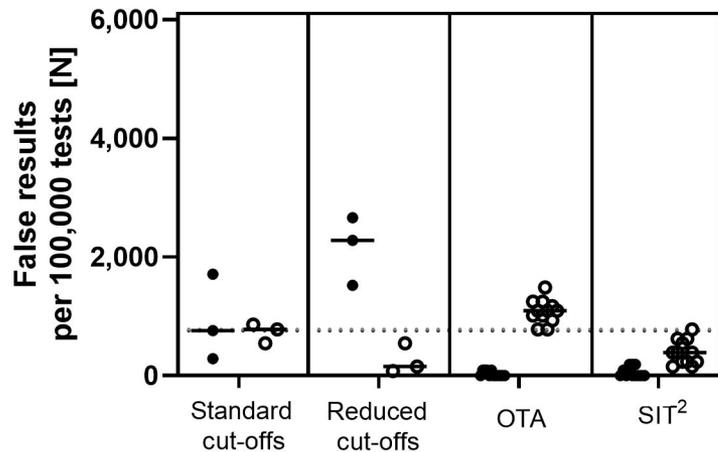


C

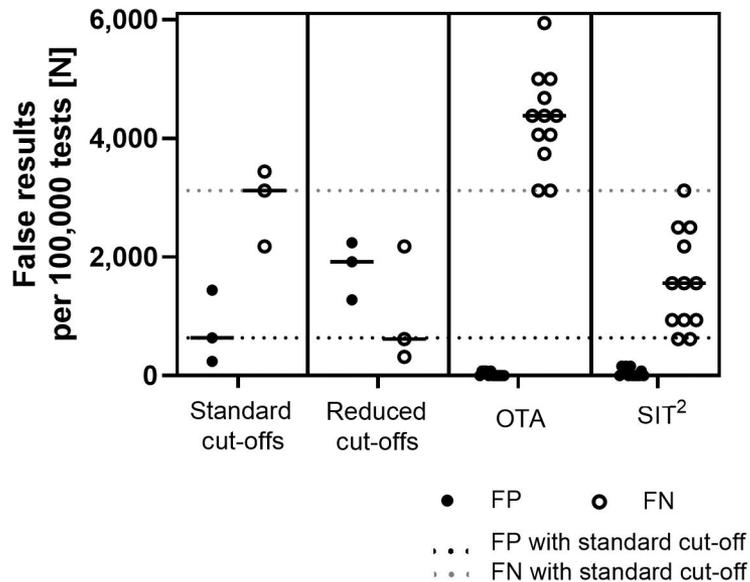
	Roche	Abbott	DiaSorin	Technozym RBD	Technozym NP
Sensitivity	89.1 (79.1-94.6)	84.4 (73.6-91.3)	82.8 (71.8-90.1)	85.9 (75.4-92.4)	78.1 (66.6-86.5)
Specificity	99.7 (99.2-99.9)	99.2 (98.5-99.6)	98.2 (97.3-98.8)	99.2 (98.5-99.6)	99.1 (98.4-99.5)
PPV (5%)	94.6 (84.9-98.2)	84.6 (74.0-91.4)	70.9 (60.2-79.2)	84.9 (74.4-91.6)	82.1 (71.0-89.6)
NPV (5%)	99.4 (98.9-99.7)	99.2 (98.6-99.5)	99.1 (98.4-99.5)	99.3 (98.7-99.6)	98.9 (98.2-99.3)
Cut-off (manufacturer)	1.000 COI	1.40 I	12.0 AU/mL**	8.000 U***	8.000 U***
Reduced Cut-off	>0.165 ¹	≥0.55 ²	≥9.0 ³	>3.000 U****	>3.500 U****
High Cut-off	3.000	3.00	150.0	-	-

A

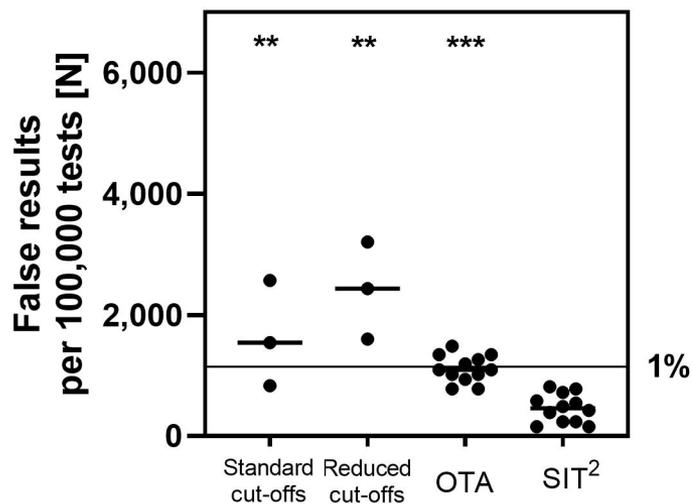
False positives/negatives
(5% seroprevalence estimated)



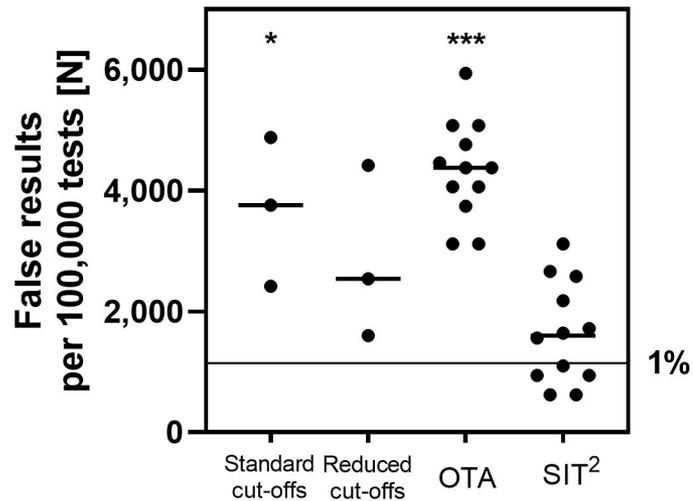
False positives/negatives
(20% seroprevalence estimated)

**B**

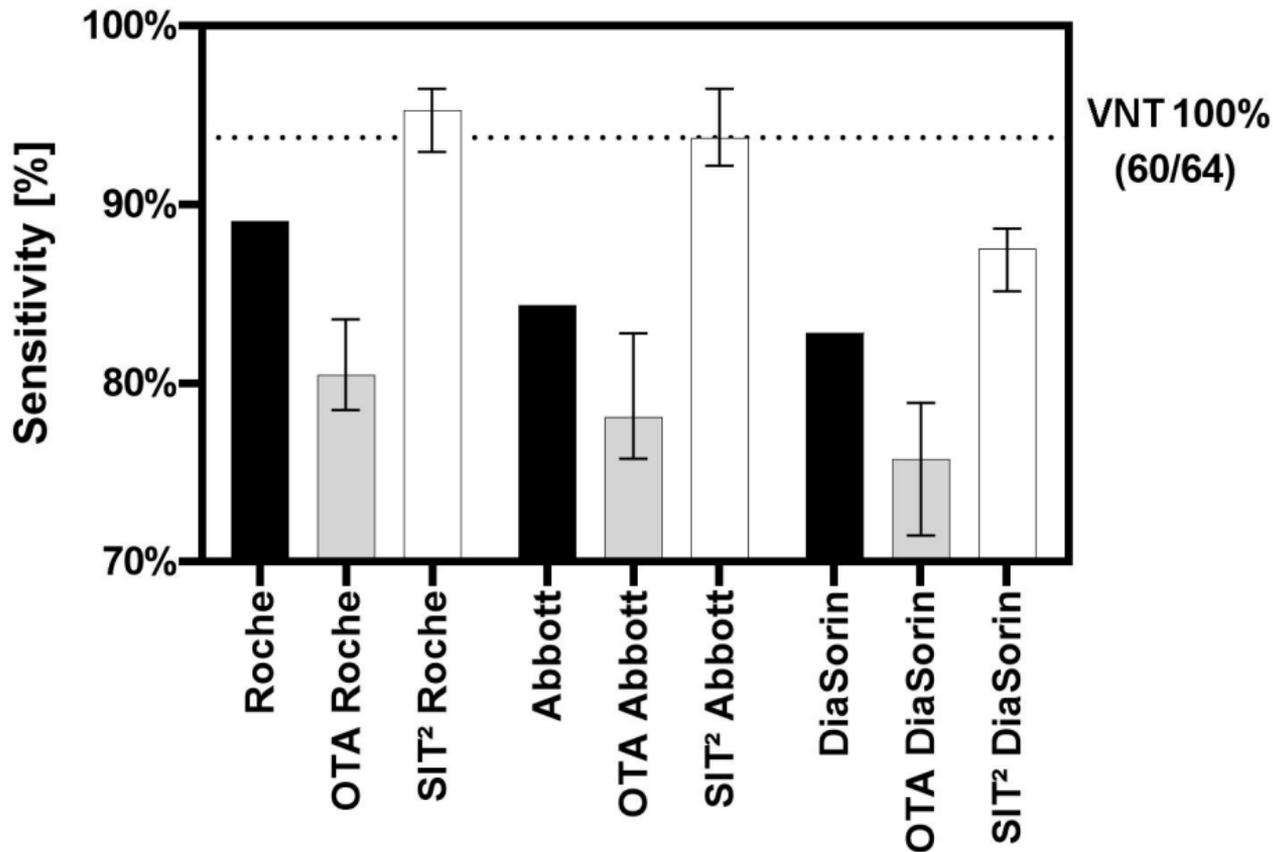
Total error
(5% seroprevalence estimated)



Total error
(20% seroprevalence estimated)



Sensitivity with standard testing, OTA and SIT²



Screening Test

Roche

Abbott

DiaSorin

Siemens

Oxford

Confirmation Test

Roche

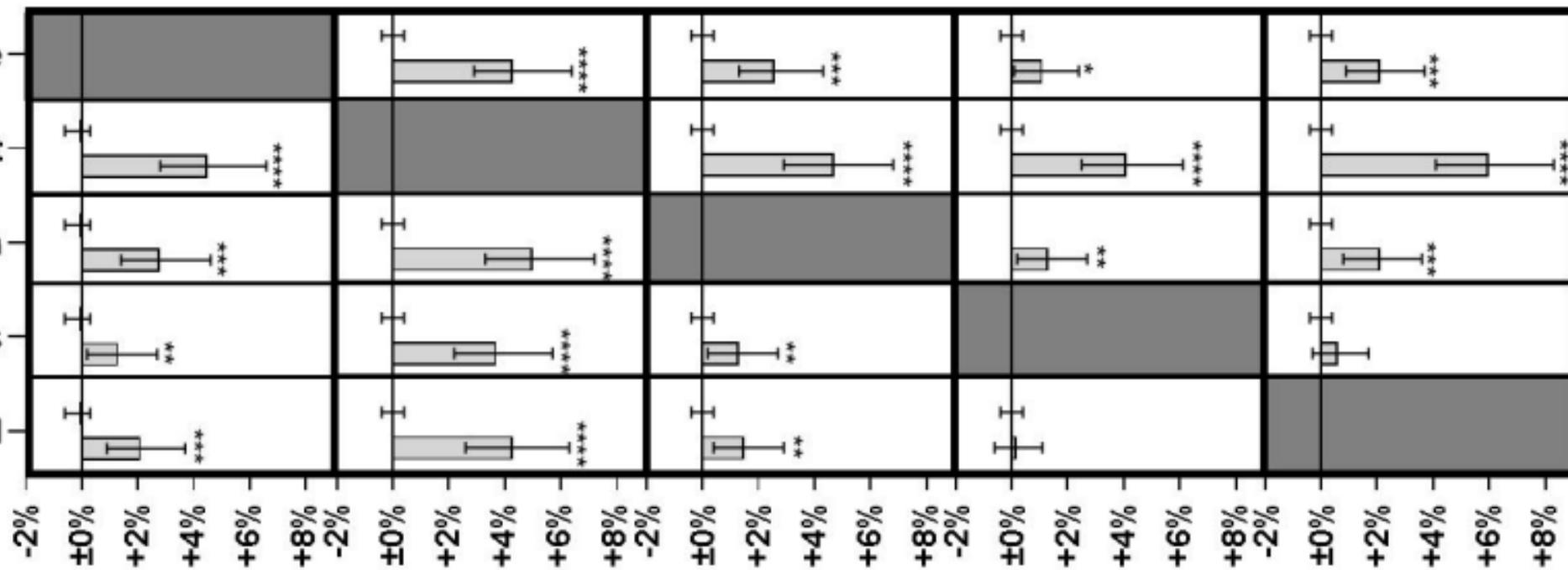
Abbott

DiaSorin

Siemens

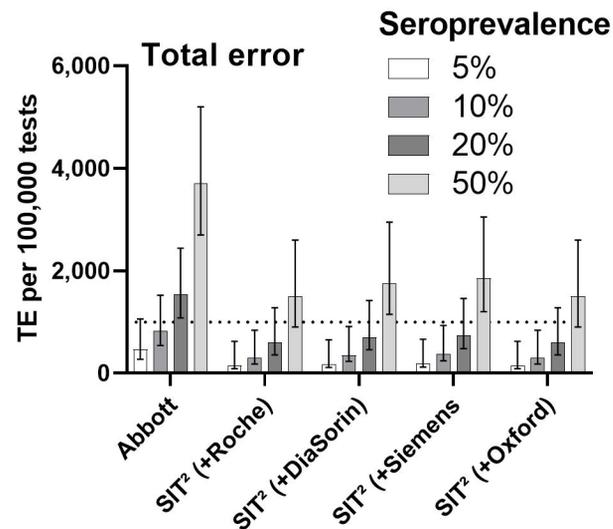
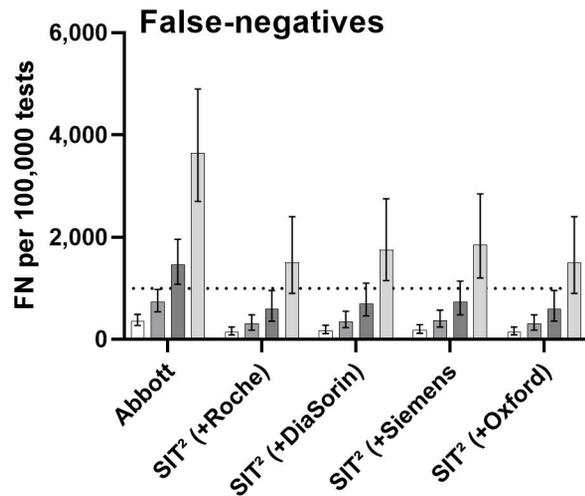
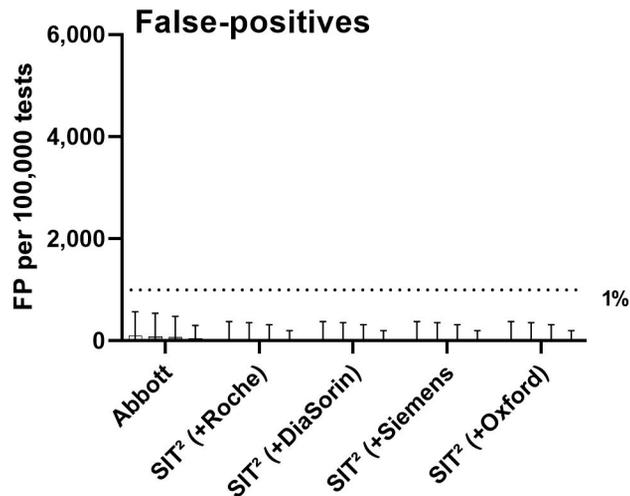
Oxford

Specificity
Sensitivity



$\Delta SIT^2 - OTA$

A Abbott



B DiaSorin

