

1 **A Reagent and Virus Benchmarking Panel for a Uniform Analytical Performance**

2 **Assessment of N Antigen–Based Diagnostic Tests for COVID-19**

3 Allison Golden*, Jason L. Cantera, Lorraine Lillis, Thanh T. Phan, Hannah Slater, Edwin J.

4 Webb, Roger B. Peck, Gonzalo J. Domingo, David S. Boyle

5

6 All authors: PATH; 2201 Westlake Avenue, Suite 200; Seattle, Washington 98121, USA

7 *Corresponding author

8

9 **ABSTRACT**

10 Rapid diagnostic tests (RDTs) that detect antigen indicative of SARS-CoV-2 infection can help
11 in making quick health care decisions and regularly monitoring groups at risk of infection. With
12 many RDT products entering the market, it is important to rapidly evaluate their relative
13 performance. Comparison of clinical evaluation study results is challenged by protocol design
14 variations and study populations. Laboratory assays were developed to quantify nucleocapsid (N)
15 and spike (S) SARS-CoV-2 antigens. Quantification of the two antigens in nasal eluates
16 confirmed higher abundance of N than S antigen. The median concentration of N antigen was 10
17 times greater than S per genome equivalent. The N antigen assay was used in combination with
18 quantitative RT-PCR to qualify a panel composed of recombinant antigens, inactivated virus, and
19 clinical specimen pools. This benchmarking panel was applied to evaluate the analytical
20 performance of the SD Biosensor STANDARD Q COVID-19 Ag test, Abbott Panbio COVID-19
21 Ag Rapid Test, Abbott BinaxNOW COVID-19 Ag test, and the LumiraDx SARS-CoV-2 Ag
22 Test. The four tests displayed different sensitivities toward the different panel members, but all
23 performed best with the clinical specimen pool. The concentration for a 90% probability of
24 detection across the four tests ranged from 21 pg/mL to 102 pg/mL of N antigen in the extracted
25 sample. Benchmarking panels provide a quick way to verify the baseline performance of a
26 diagnostic and enable direct comparison between diagnostic tests.

27

28 **KEY WORDS**

29 SARS-CoV-2, rapid antigen diagnostic test, rapid diagnostic test, COVID-19, screening,
30 nucleocapsid antigen, spike antigen, genome equivalent, PCR

31 **ABBREVIATIONS**

32 BEI, BEI Resources; BSA, bovine serum albumin; CDC, US Centers for Disease Control and
33 Prevention; COVID-19, coronavirus disease 2019; Ct, cycle threshold; *E. coli*, *Escherichia coli*;
34 EUA, emergency use authorization; EUL, emergency use license; GE, genome equivalent; LOD,
35 limit of detection; LLOQ, lower limit of quantification; MSD, Meso Scale Discovery; N,
36 nucleocapsid; NIAID, US National Institute of Allergy and Infectious Diseases; NIH, US
37 National Institutes of Health; PBS, phosphate-buffered saline; qRT-PCR, quantitative reverse
38 transcription polymerase chain reaction; RDT, rapid diagnostic test; RT-PCR, reverse
39 transcription polymerase chain reaction; S, spike; SARS-CoV-2, severe acute respiratory
40 syndrome coronavirus 2; TCID₅₀, 50% tissue culture infective dose; ULOQ, upper limit of
41 quantification.

42 INTRODUCTION

43 Diagnostic tools are essential for surveillance and control of the COVID-19 pandemic.¹ While
44 reverse transcription polymerase chain reaction (RT-PCR) from a nasopharyngeal or nasal swab
45 is the gold standard for confirmation of infection with SARS-CoV-2, the complexity of such
46 tests requires sophisticated laboratory systems, and imposes logistical challenges for its effective
47 use in scenarios requiring either a fast time-to-result or where laboratory systems are not robust.
48 Rapid diagnostic tests (RDTs) designed to detect viral antigens, primarily the nucleocapsid
49 protein (N) antigen, hold promise for testing in settings where RT-PCR cannot be implemented,
50 and as both time and cost-saving measures for frequent testing and entry points.¹⁻³

51 The earliest-to-market RDTs were tested in clinical studies to assess their performance and
52 utility. The results of these clinical studies are informative in terms of clinical performance
53 within the context of the studies conducted, but they also highlight the variability in clinical
54 performance as driven by the study design, target population, and other study-specific factors.⁴⁻⁶

55 There is need for a performance assessment of these tests that is less study specific and can be
56 performed in multiple laboratories, the results of which would enable more direct comparison of
57 performance across different RDTs.⁷ If the results of this assessment can be linked to clinical
58 data, they may also be indicative of clinical performance. Currently, assessments of analytical
59 performance have been expressed primarily through comparison to SARS-CoV-2 viral RNA in
60 terms of cycle thresholds (Ct) or to cultured virus infective units, with increasing correlation to
61 viral copy number quantification to aid in standardization of results.⁷⁻¹³ Complementary to these
62 efforts, this article presents a quantitative open-platform assay for the N and spike (S) antigens, a
63 comparison of genome equivalents (GEs) to N and S antigen concentration from clinical

64 samples, a panel of reagents with which to assess the performance of the RDTs representing
65 multiple sources of target analyte protein, and the results from assessment of four different
66 emergency use–authorized (EUA)/licensed (EUL) COVID-19 rapid antigen diagnostic tests.

67 **MATERIALS AND METHODS**

68 **Materials**

69 Full-length recombinant N antigens expressed in *Escherichia coli* and in HEK293 mammalian
70 cells were purchased from Native Antigen (Kidlington, United Kingdom) and Acro Biosystems
71 (Newark, Delaware, USA), respectively. Recombinant S antigen, in a stabilized trimeric form
72 and expressed in HEK293 mammalian cells, was purchased from Acro Biosystems.

73 The following reagents were obtained through BEI Resources (Manassas, Virginia, USA); and
74 the US National Institute of Allergy and Infectious Diseases (NIAID)/US National Institutes of
75 Health (NIH), contributed by the US Centers for Disease Control and Prevention (CDC): SARS-
76 Related Coronavirus 2, Isolate USA-WA1/2020, Gamma-Irradiated (NR-52287), and Genomic
77 RNA from SARS-Related Coronavirus 2, Isolate USA-WA1/2020 (NR-52285). The following
78 reagents were obtained through BEI Resources and NIAID/NIH: Human Coronavirus 229E (NR-
79 52726), and Human Coronavirus OC43 (NR-52725). NR-52287 used for N antigen
80 determination had assigned values for the concentration of infectious virus of 2.8×10^5 50%
81 tissue culture infective dose (TCID₅₀)/mL and RNA (4.1×10^9 copies/mL), determined prior
82 to inactivation. The viral RNA control for RT-PCR was prepared from USA-WA1/2020 (47.5
83 ng/mL total RNA with an estimated 1.84×10^7 GEs/mL).

84 Using the GenBank sequence of the 419 as a full-length SARS-CoV-2 N protein, GenBank
85 accession number QHO62115.1, with an additional cleavage site and a polyhistidine tag, an
86 extinction coefficient was calculated based on ProtParam¹⁴ of Abs 0.1% (= 1 g/L) of 0.959. This
87 extinction coefficient was used to confirm stock protein concentration of recombinant SARS-
88 CoV-2 antigen that was aliquoted and stored at -80°C.

89 The buffer diluent contained 1X phosphate-buffered saline (10 mM PBS, 2.7 mM potassium
90 chloride, 137 mM sodium chloride pH 7.4) with 1% bovine serum albumin (BSA), 1X PBS +
91 1% BSA. Negative swab pool diluent contains pooled discarded SARS-CoV-2 PCR-negative
92 human nasal swabs eluted into 1X PBS.

93 **Clinical samples**

94 De-identified nasal swab eluates were acquired from the Washington COVID-19 Biorepository
95 (Seattle, Washington, USA) or Boca Biolistics (Pompano Beach, Florida, USA). Nasal swab
96 eluates used in this study were prepared in either 1X PBS or Clinical Transport Medium (Noble
97 Biosciences, Gyeonggi-do, Republic of Korea).

98 **Quantification of SARS-CoV-2 N and S antigens using SARS-CoV-2 antigen**

99 **immunoassays**

100 An immunoassay detecting SARS-CoV-2 N and S antigens was developed using the Meso Scale
101 Discovery (MSD) platform (Meso Scale Diagnostics, Rockville, Maryland, USA), which uses
102 electrochemiluminescence for detection.

103 Antibodies sourced from Sino Biological (Beijing, P.R. China) was used for N antigen detection
104 and an antibody pair from Leinco Technologies (Fenton, Missouri, USA) for S antigen detection.

105 The capture antibodies were labeled with biotin using the EZ-Link[®] Sulfo-NHS-LC-LC-
106 Biotinylation kit (ThermoFisher Scientific, Waltham, Massachusetts, USA) and the detector
107 antibodies were labeled with SULFO-TAG[™] (GOLD SULFO-TAG NHS-Ester, Meso Scale
108 Diagnostics). Any unbound biotin or SULFO-TAG was removed using desalting columns
109 (Zeba[™], 40k MWCO, ThermoFisher Scientific). The concentrations of antibodies were
110 measured at 280 nm via a spectrophotometer (NanoDrop[™] 2000C, ThermoFisher Scientific),
111 and concentration of detector antibody following labeling was assigned 90% of the concentration
112 prior to desalting. Standards were prepared from recombinant HEK293-expressed full-length
113 SARS-CoV-2 N protein and stabilized trimeric S protein (Acro Biosystems).

114 **Clinical sample testing with SARS-CoV-2 antigen assays**

115 N and S assays were run separately, using 25 μ L per well of the 0.5 μ g/mL biotinylated capture
116 antibody was used to coat a blocked SECTOR small spot streptavidin plate (Meso Scale
117 Diagnostics). Analysis of signal and quantification of unknowns relative to the standard curve
118 were conducted using Meso Scale Diagnostics' Discovery Workbench 4.0 software. For
119 quantification, standards and blank were fit with a 4-parameter log logit fit with 1/y² weighting.
120 The lower limit of detection (LOD) was defined by the software's curve fitting. The lower limit
121 of quantification (LLOQ) was defined by the lowest concentration of standard with signal above
122 the following: the limit of blank plus 10 times the standard deviation of the limit of blank.¹⁵ The
123 upper limit of quantification (ULOQ) was defined by both software and a back-calculated
124 recovery average of 100% \pm 20%. Standard curves spanned 0.128 pg/mL to 50 ng/mL of N
125 antigen,⁷ and 0.128 pg/mL to 1,250 pg/mL of S antigen.

126 The concentration of SARS-CoV-2 N antigen was measured in 405 residual nasal swab eluates,
127 characterized by PCR at CLIA registered clinical laboratories, collected in July through
128 December 2020 in Washington state. Samples were selected across a range of Ct values from
129 original testing, with selection biased toward higher Ct values. Clinical samples either found or
130 anticipated to be over the detection range were diluted either 5-fold or 20-fold to bring them into
131 range, if volume allowed. Replicate well values for positives with a coefficient of variation
132 greater than 20% were repeated.

133 **Molecular testing for SARS-CoV-2**

134 Viral RNA was extracted from samples using the QIAamp[®] Viral RNA Mini Kit (Qiagen,
135 Valencia, California, USA) according to the manufacturer's instructions and eluted in 100 μ L
136 buffer. A quantitative RT-PCR (qRT-PCR) assay to estimate the SARS-CoV-2 GE/mL was
137 developed using the N1 primer set developed by the CDC with primers and probe procured from
138 Integrated DNA Technologies (Coralville, Iowa, USA). Each 20 μ L final reaction volume
139 contained 5 μ L of 4X TaqPath[™] 1-Step RT-qPCR Master Mix (ThermoFisher Scientific), 0.5
140 μ L of probe (5 μ mol/L), 0.5 μ L each of forward and reverse primers (20 μ mol/L), 8.5 μ L of
141 nuclease-free water, and 5 μ L of nucleic acid extract. Amplification was performed on an
142 Applied Biosystems[™] 7300 Real-Time PCR instrument (ThermoFisher Scientific).
143 Thermocycling conditions consisted of 15 minutes at 50°C, 2 minutes at 95°C, and 45 cycles of
144 3 seconds at 95°C and 30 seconds at 55°C. The cutoff for positive samples was less than 40
145 cycles. The median Cts were used to determine the viral GE concentration using a standard curve
146 from SARS-CoV-2 genomic RNA (Isolate USA-WA1/2020).

147 **Benchmarking panel**

148 Commercially sourced, full-length, His-tagged recombinant N protein as described above was
149 used. Radiation-inactivated, cultured SARS-CoV-2 virus (BEI Resources, NR-52287) stocks
150 were thawed and diluted into either buffer or negative swab pool. Serial dilutions were further
151 made into diluent and aliquots were frozen. Clinical nasal swab discards from five different
152 individuals, positive for SARS-CoV-2 by qRT-PCR, were selected and combined. The combined
153 samples were then serially diluted into negative swab pool, aliquoted, and frozen. The aliquots
154 were tested by qRT-PCR to quantify viral GE/mL. A benchmarking panel composed of dilutions
155 of recombinant proteins, inactivated viral lysate, and clinical specimen pool was then defined
156 and applied to all diagnostic tests in this study. The panel members were characterized for N
157 antigen concentration using the N antigen immunoassay as well as qRT-PCR for the clinical
158 specimen dilutions. The benchmarking panel is described in Table 1.

159 **Table 1. Components of the benchmarking panel used in evaluation of rapid diagnostic**
160 **tests to detect nucleocapsid and spike SARS-CoV-2 antigens.**

Category	Antigen source	Diluent/Matrix	Number of dilutions	N antigen concentration range
Recombinant N	HEK293-expressed, His tagged	1X PBS, 1% BSA	10	0.1–50 ng/mL, 5,000 ng/mL
Recombinant N	<i>E. coli</i> -expressed, His tagged	1X PBS, 1% BSA	10	0.1–50 ng/mL, 5,000 ng/mL
Inactivated	Gamma-	1X PBS, 1% BSA	8	0.05–10 ng/mL,

virus	inactivated SARS-CoV-2 (202-WA-1)			representing 10^4 – 10^7 GE/mL
Inactivated virus	Gamma- inactivated SARS-CoV-2 (202-WA-1)	Negative swab pool	8	0.05–10 ng/mL, representing orders of 10^4 – 10^7 GE/mL
Clinical dilutions	PATH biorepository	Negative swab pool	16	0.025–225 ng/mL, representing orders of 10^3 – 10^7 GE/mL
Specificity	OC43 and 229E cultured viral lysates	1X PBS, 1% BSA	1	Not detected; quantity of huCoV-specific N antigen not determined; 5,000 TCID ₅₀ /mL OC43 and 10,000 TCID ₅₀ /mL huCoV229E
Specificity	Diluent controls	Negative swab pool	1	Not detected
Specificity	Diluent controls	1X PBS, 1% BSA	1	Not detected

161 Abbreviations: BEI, BEI Resources; BSA bovine serum albumin; *E. coli*, *Escherichia coli*; GE,
162 genome equivalent; PBS, phosphate-buffered saline; TCID₅₀, 50% tissue culture infective dose.

163 **Rapid diagnostic tests used**

164 Four EUA and/or EUL SARS-CoV-2 RDTs were evaluated using N antigen benchmarking
165 panels: the Abbott BinaxNOW™ COVID-19 Ag card test (EUA), Abbott Panbio™ COVID-19
166 Ag Rapid Test Device (EUL), LumiraDx SARS-CoV-2 Ag Test (EUA), and SD Biosensor
167 STANDARD™ Q COVID-19 Ag Home Test (EUL). Tests were assigned, in no particular order,
168 identification numbers of RDT 1 through RDT 4, for the purpose of this publication, to de-
169 identify specific results.

170 **Evaluation of rapid diagnostic tests with benchmarking panels**

171 Each RDT was run using a minimum of five replicates per panel concentration and type. Panel
172 member aliquots were thawed on ice and mixed gently. A pipetted volume of the panel mixed
173 into the rapid test-specific extraction buffer simulated the extracted swab material. Thereafter,
174 the instructions for the rapid tests were followed and the diluted panel, at its final concentration
175 in the extraction buffer, was added to the test according to instructions. All panel concentrations
176 were run until two levels of decreasing concentrations were negative for all replicates.

177 **Statistical analysis of detection limits using benchmarking data**

178 Two statistical models were developed to determine the relationships between (1) analyte
179 concentration and RDT test line intensity and (2) analyte concentration and probability of RDT
180 positivity. For (1), a sigmoid function was fit to the data of the form:

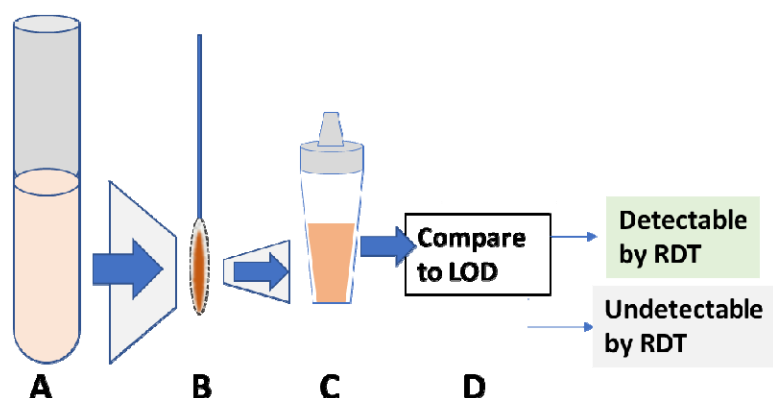
$$\text{test line intensity} = \frac{\alpha}{1 + \exp((m - \log_{10} \text{conc})/s)}$$

181 Where α , m , s are parameters representing the shape of the sigmoid curve and *conc* is the analyte
182 concentration. This model was separately fitted to each data type (clinical dilution, inactivated
183 virus culture, and recombinant), and the step repeated for all four RDTs. For (2), a simple
184 logistic regression model was fitted to the data, where analyte concentration was the independent
185 variable and the RDT result (1 = positive, 0 = negative) is the dependent variable. A categorical
186 factor representing the data type (clinical dilution, inactivated virus culture, recombinant protein)
187 was included as a covariate to allow for data type-specific fitted curves. This step was repeated
188 for all four RDTs. The models described in (1) and (2) were fitted in a Bayesian framework
189 using the R brms package.¹⁶ Noninformative Gaussian priors were used for the parameters and
190 the models run for 5,000 iterations after a burn-in of 2,500 iterations. Convergence of chains was
191 assessed using the R-hat statistic and visual checks. The final fitted lines and surrounding shaded
192 areas represent the median and 95% credible of the expected values of the posterior predictive
193 distributions.

194 **Simulation of detection of clinical samples by rapid diagnostic tests**

195 The detection limits derived from benchmarking data were used to simulate the detection of
196 clinical samples that had been characterized for SARS-CoV-2 N antigen concentration by MSD
197 assay. For each clinical sample, the quantity of N antigen was assumed to be concentrated into
198 the source swab, which had been extracted and diluted into 3 mL transport medium. The quantity
199 of antigen was then divided by the manufacturer-designated volume of rapid test extraction
200 buffer to simulate the final concentration of N antigen that would be added to the test. Finally,
201 this final concentration was compared to detection limits generated in analysis of the
202 benchmarking data to determine whether the sample would be designated as detectable (Figure
203 1).

204 **Figure 1.** Representation of simulated clinical sample detection by rapid diagnostic test (RDT).
205 **A.** Clinical samples consisting of swab eluate in 3 mL transport medium; N antigen
206 concentration measured. **B.** Total quantity of antigen assumed to be on swab. **C.** Total quantity
207 of antigen is diluted in extraction buffer of RDT to calculate final concentration of N antigen
208 added to RDT. **D.** Final concentration compared to 90% probability lower limit of detection from
209 benchmarking to determine whether detectable by RDT.



210

211 Abbreviations: LOD, limit of detection; RDT, rapid diagnostic test.

212 RESULTS

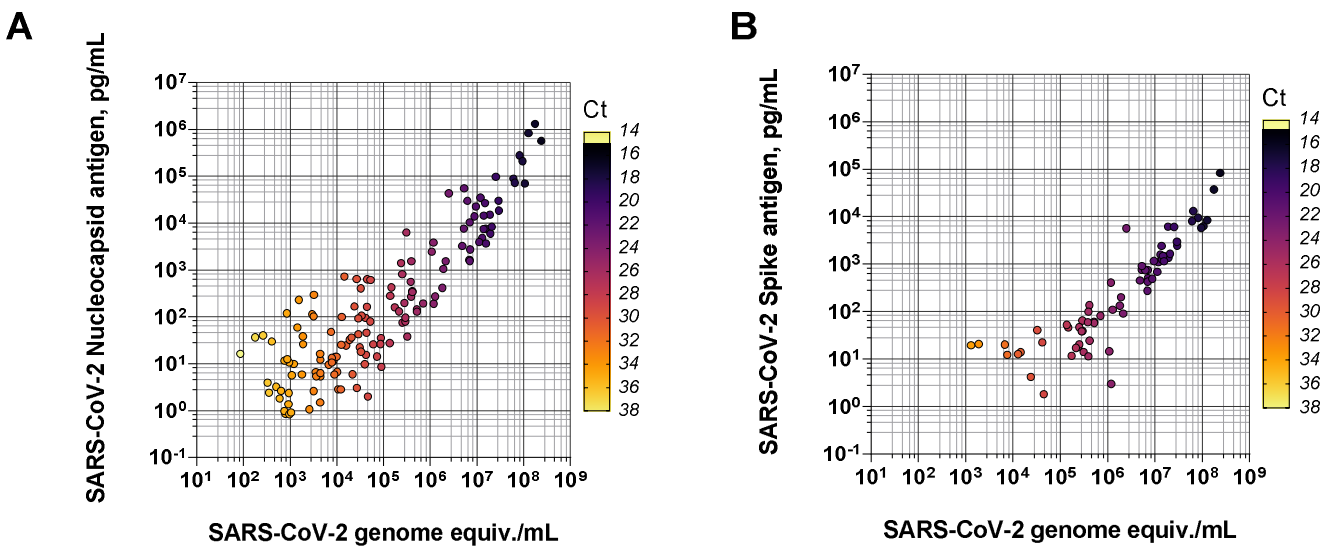
213 Analytical performance for the SARS-COV-2 N and S proteins

214 The SARS-CoV-2 N antigen immunoassay on the MSD platform had an average LOD of 0.45
215 pg/mL (0.25 to 0.93 pg/mL range), an LLOQ of 3.2 pg/mL, and a ULOQ of 50 ng/mL. The
216 SARS-CoV-2 S antigen immunoassay had an average LOD of 6.2 pg/mL (2.1 to 9.0 pg/mL
217 range), an LLOQ of 80 pg/mL, and a ULOQ of 250 ng/mL. Both assays were nonreactive or
218 below detection limits for panel diluents, transport media, and common human coronavirus
219 lysates OC43 and 229E.

220 **Distribution of N and S antigens in clinical samples**

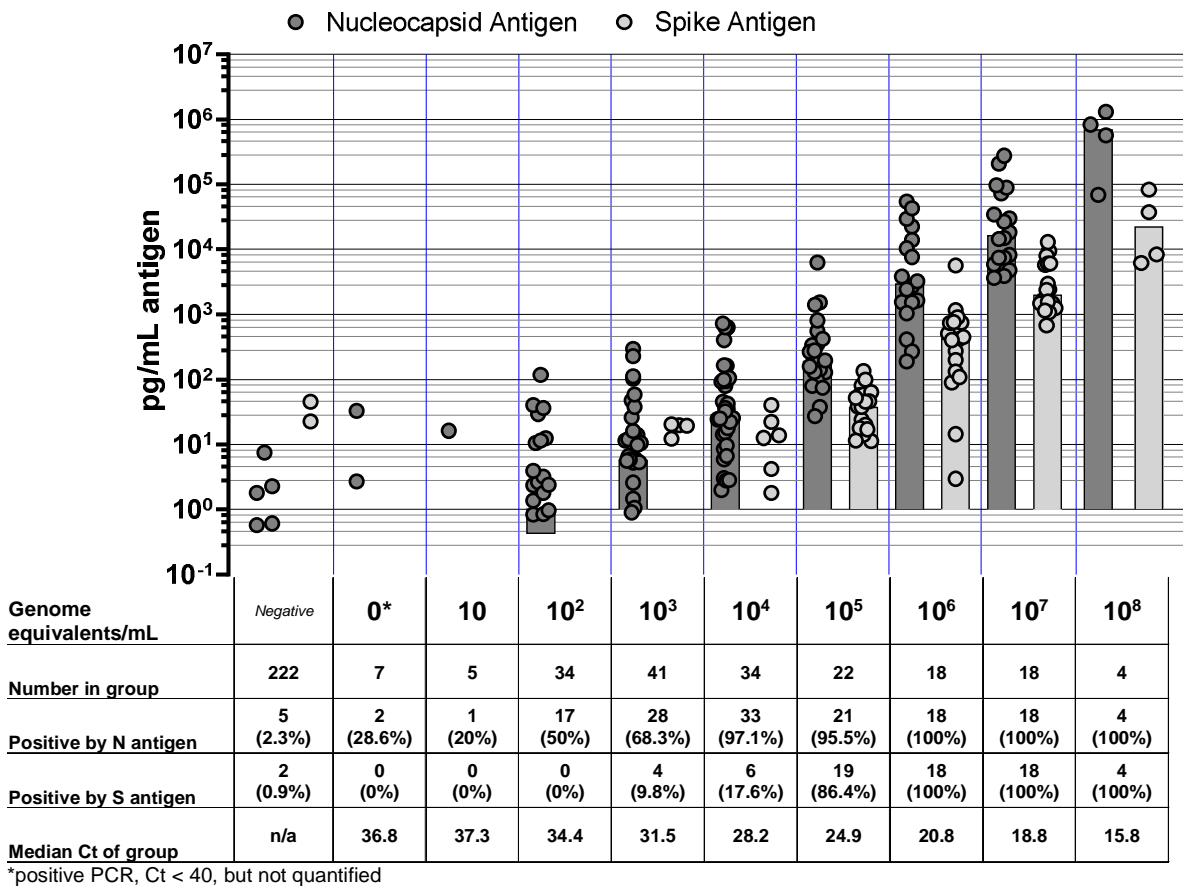
221 All 200 presumed positive samples and 100 of the 205 negative samples were analyzed by qRT-
222 PCR for SARS-COV-2. Here, 182 samples were confirmed positive by qRT-PCR and 99 of the
223 negatives were confirmed negative, with one repeatedly testing positive by qRT-PCR. The
224 remaining negatives were assigned as qRT-PCR negative based on previously assigned qRT-
225 PCR results, resulting in 183 PCR-positive samples and 222 PCR-negative samples. N and S
226 antigen quantification was conducted on all 405 specimens. A positive correlation was found
227 between the N antigen concentration and GEs (Figure 2A and 2B). The N antigen assays showed
228 a high percent positive agreement ($> 95\%$) for specimens containing 10^4 or more GEs/mL
229 SARS-CoV-2 (Figure 3). The percent positive agreement progressively dropped with decreasing
230 concentrations of GEs. For the S antigen, the percent positive agreement was 86.4% at 10^5
231 GEs/mL, with a sharp drop in agreement for lower concentrations. For all samples for which N
232 antigen was within the LOQ of the assay ($n = 120$), the mean and median per GE N antigen
233 observed were 12.7 fg and 1.5 fg N antigen (range of 0.1 to 204.2 fg/GE). For concentrations of
234 SARS-CoV-2 GEs less than 10^4 GEs/mL, the quantity of N antigen varied more widely for
235 positive samples within the LOQ, with a trend toward highest N antigen per GE. For all samples
236 for which S antigen was within the LOQ of the assay ($n = 41$), the mean and median per GE S
237 antigen observed were 0.2 fg and 0.1 fg S antigen (range of 0.04 to 2.3 fg/GE).

238 **Figure 2. A.** Correlation between N antigen concentration (pg/mL) and N genome equivalents
239 (copies/mL). Mean cycle threshold (Ct) values are color coded to provide an indication of the
240 corresponding Ct values. **B.** Correlation between S antigen concentration (pg/mL) and N genome
241 equivalents (copies/mL). Mean Ct values are color coded to provide an indication of the
242 corresponding Ct values.



243
244

245 **Figure 3.** Antigen concentration over a range of SARS-CoV-2 genome equivalents detected by
246 Meso Scale Discovery antigen quantification. Characteristics of groups of viral genome
247 equivalent concentrations are shown in the table below the graph. Bars indicate the median of the
248 group of antigen-positive points shown.



249
250 Abbreviations: Ct, cycle threshold; N, nucleocapsid; PCR, polymerase chain reaction; S, spike.

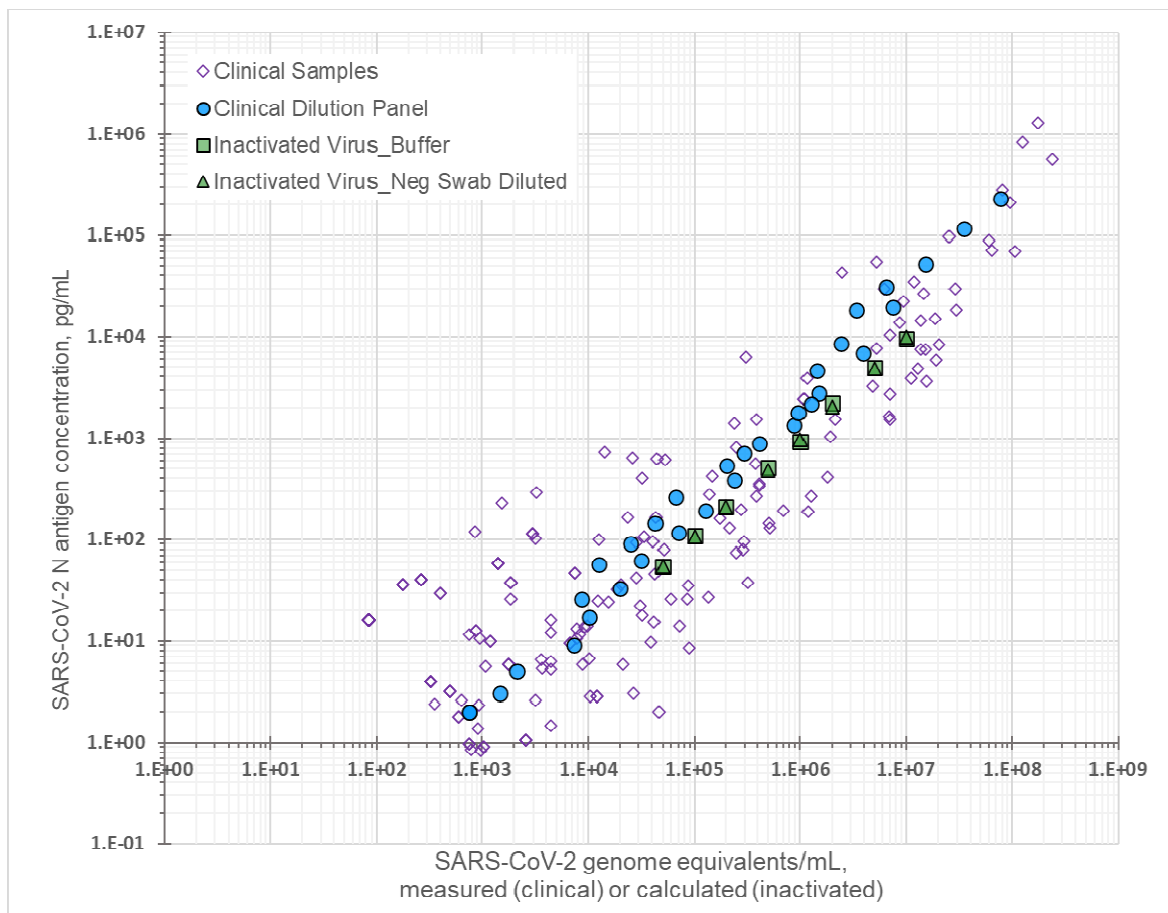
251 Benchmarking panels

252 Benchmarking panels were prepared to span concentrations of N antigen corresponding to LODs
253 expected in rapid tests following dilution into the extraction buffer. As most RDTs only detect
254 the N antigen, S antigen analysis was not included in the panel. HEK293-expressed and *E. coli*-
255 expressed recombinant proteins were prepared in both buffer and negative swab pool dilution
256 matrices, and linear fit of measured concentration as compared to target concentration was
257 greater than 0.98 over the range of the panels of 0.2 ng/mL to 50 ng/mL.

258 **Quantification of inactivated virus and clinical dilutions, compared to genome equivalents**

259 Though all panels were prepared from a single lot of inactivated SARS-CV-2 (BEI Resources),
260 comparisons between two lots showed that for each lot, the per-calculated GE concentration of N
261 antigen had a median of 1.0 fg/GE (range of 0.51 to 1.1 fg/GE) over the dilutions. In contrast,
262 when compared to TCID₅₀, the median amount of N antigen per TCID₅₀ in the dilutions was
263 5,820 fg/TCID₅₀ for BEI Resources lot 70033322 and 668 fg/TCID₅₀ for lot 70035888. Both
264 irradiated virus and clinical pool samples behaved similarly, even in terms of the per GE
265 concentration of N antigen (Figure 4). For the clinical specimen pool dilution series, the median
266 per GE quantity of N antigen was 2.4 fg/GE (range of 1.2 to 5.3 fg/GE).

267 **Figure 4.** The relationship between N antigen concentration and genome equivalents for panels
268 of diluted clinical specimens and inactivated virus. Blue circles indicate the clinical pool dilution
269 series. Green squares and triangles indicate BEI Resources irradiated virus lot 7003588 in buffer
270 and negative swab pool, respectively.



272 **Benchmarking of rapid antigen detection tests for SARS-CoV-2**

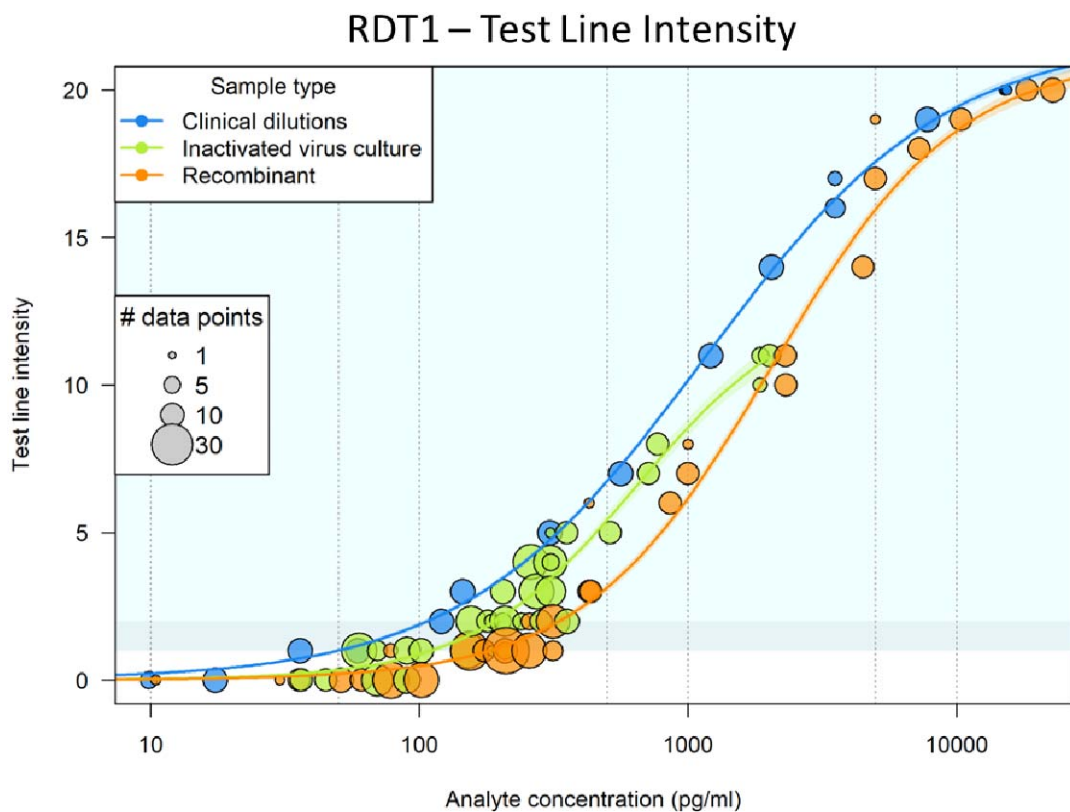
273 Four RDTs were evaluated with the benchmarking panels: SD BIOSENSOR Abbott BinaxNOW
274 COVID-19 Ag card test, Abbott Panbio COVID-19 Ag Rapid Test Device, LumiraDx SARS-
275 CoV-2 Ag Test, and SD Biosensor STANDARD Q COVID-19 Ag Home Test. The tests were
276 anonymized using identification numbers RDT 1 through RDT 4, in no particular order, for
277 comparative presentations of the results.

278 **Reactivity of panel types**

279 RDT line intensity scored either with a score card provided by the manufacturer or a preset
280 universal score card allowed an assessment of the signal-dose response per panel member type.

281 An illustrative example is given in Figure 5 for RDT 1. Recombinant proteins from mammalian
282 and *E. coli*-based expression systems produced a strong dose-response signal on the RDTs and
283 were similarly reactive. Greater reactivity was observed with both inactivated viral culture and
284 diluted clinical positives, based on the final concentration of N antigen. Binary positive/negative
285 results were used to estimate the probability of detection at different antigen concentrations. This
286 was performed for all four tests (Figure 6). All tests performed best against the clinical specimen
287 pool dilutions in terms of 90% probability of detection, and then differentially against the
288 different panel members. RDTs had lower reactivity to the inactivated virus when it was diluted
289 into the buffer diluent versus negative swab pool diluent (Table 2).

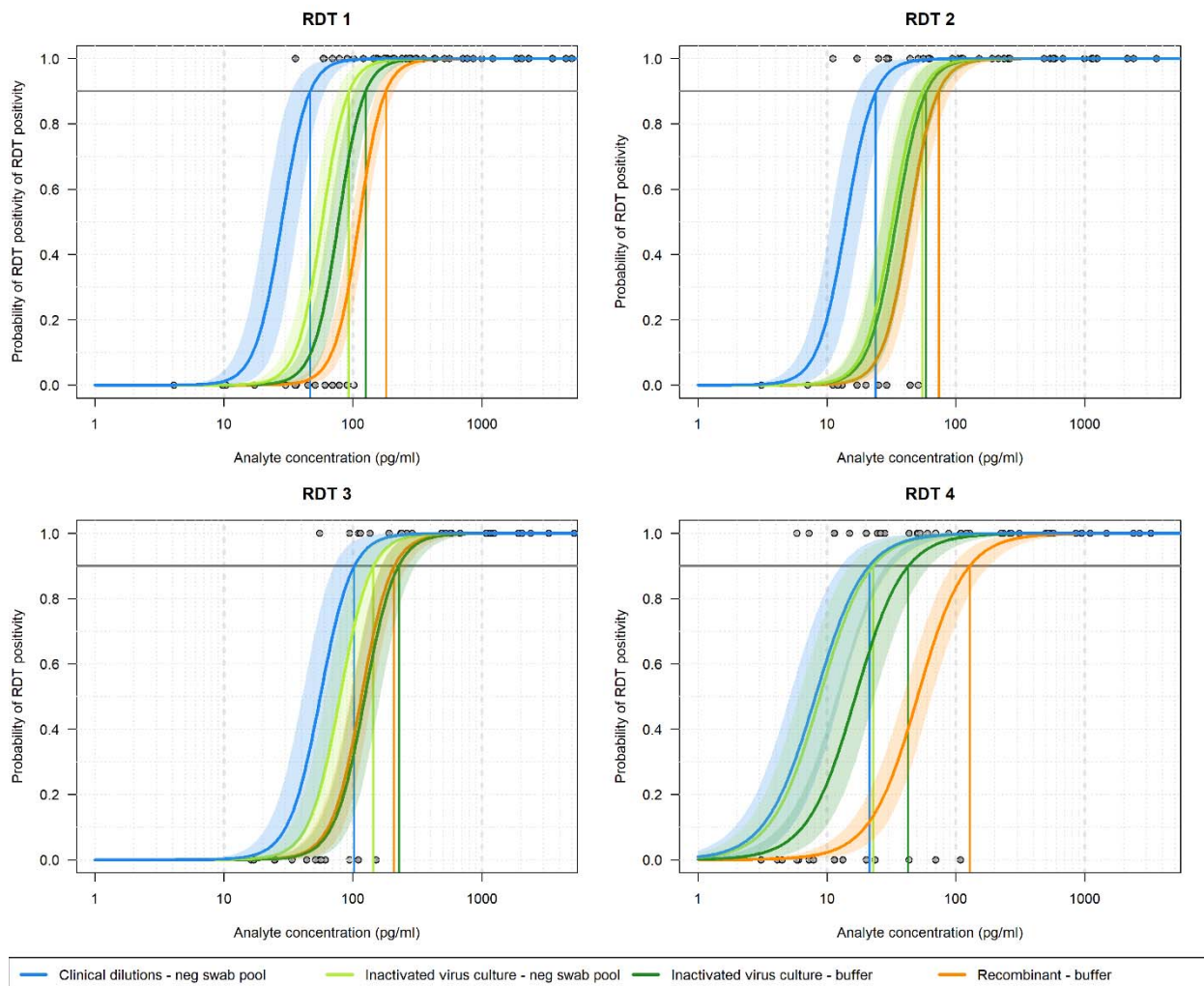
290 **Figure 5.** Illustrative sub-benchmarking panel member analysis for SARS-CoV-2 rapid
291 diagnostic test 1. Circled positions indicate the replicates with a given test line intensity result for
292 the concentration of antigen (analyte) panel added. Each panel subset has been given a different
293 color code: blue for clinical specimen pool, green for inactivated virus, and orange for
294 recombinant antigen. The size of the circles indicates the number of replicates supporting each
295 data point. Test line intensity is shown on a scale of 0 (negative) and 1 through 20, representing
296 from least to most intense visible test line of positive results.



297

298 Abbreviation: RDT, rapid diagnostic test.

299 **Figure 6.** Probability of detection against antigen concentrations for the four antigen detection
300 rapid diagnostic tests per benchmarking panel type and diluent matrix. Probability of detection of
301 positive for the different benchmarking panel member types for each test in each panel (A, B, C,
302 and D). Blue lines indicate clinical specimen dilution; green lines represent inactivated virus; and
303 orange lines represent recombinant protein. Confidence intervals (95%) are indicated in
304 matched-color shading.



305

306 Abbreviation: RDT, rapid diagnostic test.

307 **Table 2. Ninety percent probability of detection (95% confidence interval) of N antigen**

308 **final concentration for benchmarking panel categories.**

	Concentration of N antigen (final concentration added to test) with 90% probability of detection, pg/mL				
Panel category ►	Clinical positive	Inactivated virus	Inactivated virus	Recombinant protein (<i>E. coli</i> and	Specificity and diluent controls

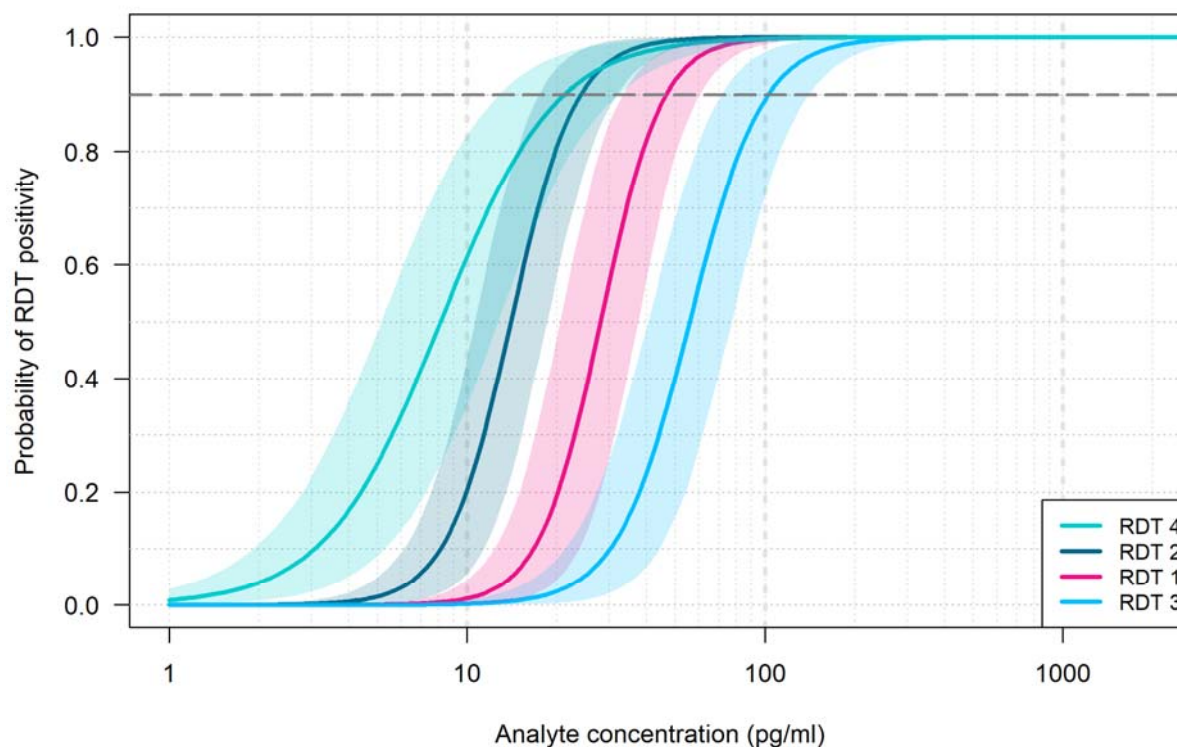
RDT number ▼	Diluted in negative swab pool	Diluted in negative swab pool	Diluted in buffer	mammalian expressed) Diluted in buffer	
RDT 1	46.774 (33.113– 60.256)	93.325 (75.858– 114.815)	125.893 (102.329– 151.356)	181.97 (154.882– 208.93)	Not detected
RDT 2	23.988 (17.378– 31.623)	54.954 (40.738– 72.444)	58.884 (42.658– 79.433)	74.131 (58.884– 91.201)	Not detected
RDT 3	102.329 (70.795– 141.254)	144.544 (104.713– 194.984)	229.087 (165.959– 301.995)	208.93 (165.959– 263.027)	Not detected
RDT 4	21.38 (12.883– 32.359)	22.909 (15.136– 33.884)	42.658 (28.84– 63.096)	128.825 (93.325– 181.97)	Not detected

309 Abbreviations: *E. coli*, *Escherichia coli*; RDT, rapid diagnostic test.

310 Comparative benchmarking results

311 The clinical specimen pool dilution panels were plotted for all four tests, for comparison of
312 analytical performance of the RDTs (Figure 7). The modeled 90% probabilities of detection were
313 found to be 47, 24, 102, and 21 pg/mL of final concentration of N antigen added to test, for
314 RDTs 1 through 4 respectively (Table 2).

315 **Figure 7.** Probability of test positivity versus final N antigen (analyte) concentration added to
316 test for clinical positive dilutions. The four curves indicate the probability of test positivity for
317 each rapid antigen test product. The shaded lines indicate 95% confidence intervals.



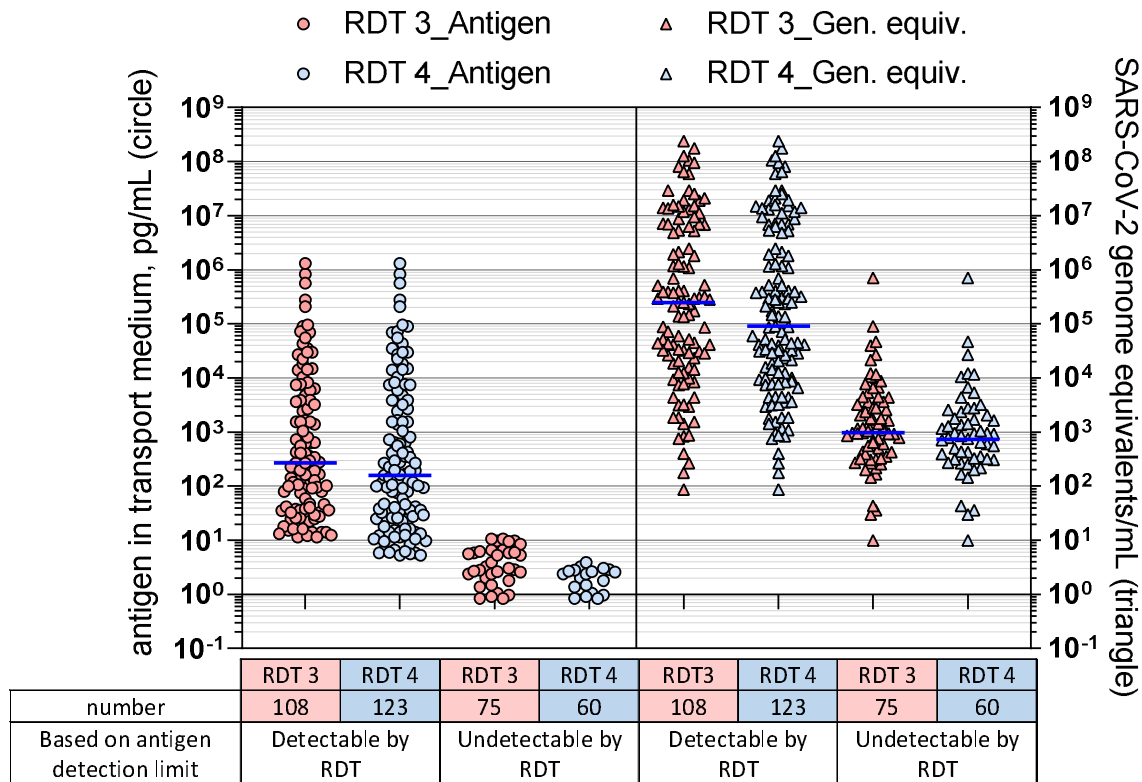
318
319 Abbreviation: RDT, rapid diagnostic test.

320 **Simulating clinical performance**

321 One limitation of benchmarking is that the results are interpretable as a concentration inherently
322 dependent upon assay configuration and the input volumes of the analyte. Interpretation of the
323 final concentration of incoming analyte, as diluted into the assay extraction buffer, allowed
324 normalization across RDTs but created the challenge of direct comparison to qRT-PCR values. If
325 an identical swab were diluted into transport media (typically around 3 mL) and into RDT
326 extraction buffer (typically around 300 μ L), the volume difference could create a 10-fold

327 disparity in analyte concentration from the swab. Figure 8 models this potential dilutional gain
328 when using RDTs to predict what may be observed in a paired swab sampling. Limits of
329 detection determined by the 90% cutoff were compared against the clinical sample set antigen
330 concentrations, with the assumption that extracted material present in the PCR would instead be
331 fully present in the extraction buffer for the rapid test, thus simulating a paired swab experiment.
332 The detection limits based on the concentration of N antigen added to the test were found to
333 affect the number of samples predicted to be detectable.

334 **Figure 8.** The predicted rapid diagnostic test (RDT) positivity from clinical samples based on
335 their antigen detection limit. Results are for RDT 3 (pink shading) and RDT 4 (blue shading),
336 shown as antigen concentration versus positivity (circle) or viral genome equivalent/mL versus
337 positivity (triangle).



338
339 Abbreviation: RDT , rapid diagnostic test.

340 RDTs 3 and 4, which had the highest and lowest analytical detection limits, respectively, based
341 on final concentration, were compared in this simulation. Overlap in N antigen–detectable and –
342 undetectable concentrations of GEs (Figure 8) was observed due to the spread of antigen
343 concentration per GE relationship, but the median GEs/mL between the detectable and
344 undetectable differed by about 2 orders of magnitude for all RDTs.

345 **DISCUSSION**

346 An open platform assay was developed and described to quantify the N and S antigens in SARS-
347 CoV-2–infected clinical specimens on the MSD platform. The assays showed good performance
348 against RT-PCR–confirmed cases and a panel of negative specimens. Quantification of both the

349 N and S antigens in specimens with associated viral load values showed a positive but not perfect
350 correlation. As anticipated, a higher N antigen concentration was observed per GE in comparison
351 to S antigen concentration, which has also been shown in plasma.^{17,18} These results support the
352 focus on N antigen for RDTs and overall correlation of antigen concentration with genome copy
353 number^{9,19,20}.

354 The antigen assay combined with the qRT-PCR was used to pedigree a panel of reagents
355 designed to benchmark N antigen RDTs. The benchmarking panel consisted of a dilution series
356 of recombinant N antigens, expressed in both prokaryotic and eukaryotic expression systems,
357 two sources of inactivated virus, and a clinical specimen pool. Four widely used tests, either
358 cleared by EUL or EUA, were evaluated against the panel: SD Biosensor STANDARD Q
359 COVID-19 Ag Home Test, Abbott COVID-19 Ag Rapid Test Device, Abbott BinaxNOW
360 COVID-19 Ag card test, and LumiraDx SARS-CoV-2 Ag test.

361 *E. coli*-expressed recombinant N antigen behaved similarly to N antigen expressed in
362 mammalian cells on the N antigen assay. Inactivated virus contained N antigen concentration per
363 viral GEs within the same range as observed in clinical samples. Comparison of the relationship
364 between N antigen concentration and GEs versus TCID₅₀ across two lots from BEI Resources
365 suggests that GEs are more reflective of the anticipated N antigen concentration.

366 The benchmarking results for the four tests highlight a range of reactivities against the different
367 panel components. Consistently, all tests showed improved LOD to the clinical specimens over
368 the other N antigen sources (inactivated virus and recombinant antigen). RDT 4 showed an
369 improved response to negative swab diluent over buffer and highlights the need for
370 comprehensive and test-specific data with a variety of materials before conclusions can be made.

371 Understanding the reactivity against inactivated virus and recombinant protein is valuable, as
372 these can be readily expressed, noninfectious sources of N antigen for emerging virus strains for
373 which there may be a concern for sequence-dependent false negativity.²¹⁻²⁴

374 The 90% probability of detection, a proxy for LOD, for the clinical specimen pool ranged from
375 20 pg/mL to 100 pg/mL across the four tests. While the 90% probability of detection limit model
376 is not a standardized method, it allows a continuous detection response function to be fit to
377 standardized panels, rather than requiring custom dilutions for each test.²⁵ The simulated paired
378 swab results generated by applying analytical detection limits to clinical samples was in
379 alignment with observed clinical performance evaluations, indicating that RDTs detect a high
380 percentage of infections with viral loads associated with Ct values less than 30.^{26,27}

381 While the analytical performance measured through benchmarking may be strongly indicative of
382 the clinical performance of the tests, it cannot be correlated directly to final performance. Some
383 of the factors that will influence the final performance are (1) relative efficiencies for absorption
384 and release of N antigen by the manufacturer's specific nasal swab and elution buffer; (2) the
385 dilution factor and original specimen equivalents loaded on to the test after all processing is
386 conducted; and (3) differing sensitivities of the test to the circulating SARS-CoV-2 strains in the
387 population being sampled. Although RDTs showed the best analytical performance against the
388 diluted clinical specimens pool, inclusion of recombinant protein sources can readily incorporate
389 N antigens into the benchmarking panel with nonsynonymous mutations that may alter the
390 analytical performance of the diagnostic test.^{21,22,28}

391 **CONCLUSION**

392 The benchmarking panel allowed rapid assessment of the analytical performance by
393 manufacturers and third parties in a manner that could be directly compared across RDTs for
394 SARS-CoV-2. Full characterization of both molecular and protein analytes allowed for
395 comparison of results across testing platforms. The benchmarking results were complementary to
396 efforts to produce international standards and to support clinical evaluations.

397 **ACKNOWLEDGMENTS**

398 The authors would like to acknowledge funding by the Bill & Melinda Gates Foundation
399 (<https://www.gatesfoundation.org/>) via grant INV-016821. The funder did not have any role in
400 the study design, data collection and analysis, decision to publish, or preparation of the
401 manuscript. Laboratories providing samples used in this work, accessed through the Washington
402 COVID-19 Biorepository, were The Everett Clinic (part of Optum), FidaLab, Northwest
403 Pathology, Bloodworks Northwest, the Washington State Public Health Laboratories, and the
404 University of Washington School of Medicine. Rapid diagnostic tests used in this work were
405 either purchased using grant funding from the Bill & Melinda Gates Foundation, or donated by
406 partners in support of research via grant from Abbott Laboratories and agreements from
407 LumiraDx and the Washington State Department of Health. Donors did not have any additional
408 role in the study design, data collection and analysis, decision to publish, or preparation of the
409 manuscript. Finally the authors would like to thank Brook Alemayehu and Terri Scott for
410 editorial support with the manuscript.

411 **REFERENCES**

- 412 1. World Health Organization. Interim guidance: Antigen-detection in the diagnosis of SARS-
413 CoV-2 infection. [https://www.who.int/publications-detail-redirect/antigen-detection-in-the-](https://www.who.int/publications-detail-redirect/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays)
414 [diagnosis-of-sars-cov-2infection-using-rapid-immunoassays](https://www.who.int/publications-detail-redirect/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays). Published Online: October 6,
415 2021.
- 416 2. Pollock, N. R.; Berlin, D.; Smole, S. C.; Madoff, L. C.; Brown, C.; Henderson, K.; Larsen,
417 E.; Hay, J.; Gabriel, S.; Gawande, A. A.; Lennon, N. J. Implementation of SARS-CoV2
418 screening in K–12 schools using in-school pooled molecular testing and deconvolution by
419 rapid antigen test. *J. Clin. Microbiol.* **2021**, *59*, e01123-21.
- 420 3. Revollo, B.; Blanco, I.; Soler, P.; Toro, J.; Izquierdo-Useros, J.; Puid, J.; Puid, X.; Navarro-
421 Pérez, V.; Casañ, C.; Ruiz, L.; Perez-Zsolt, D.; Videla, S.; Clotet, B.; Llibre, J. M. Same-
422 day SARS-CoV-2 antigen test screening in an indoor mass-gathering live music event: a
423 randomised controlled trial. *Lancet Infect. Dis.* **2021**, *21*, 1365-1372.
- 424 4. Jimenez, A.; Rees-Channer, R. R.; Perera, R.; Perera, R.; Gamboa, D.; Chiodini, P. L.;
425 Gonzalez, I. J.; Mayor, A.; Ding, X. C. Analytical sensitivity of current best-in-class
426 malaria rapid diagnostic tests. *Malar. J.* **2017**, *16*.
- 427 5. Brümmer, L. E.; Katzenschlager, S.; Gaeddert, M.; Erdmann, C.; Schmitz, S.; Bota, M.;
428 Grilli, M.; Larmann, J.; Weigand, M. A.; Pollock, N. R.; Macé, A.; Carmona, S.; Ongarello,
429 S.; Sacks, J. A.; Denking, C. M. Accuracy of novel antigen rapid diagnostics for SARS-
430 CoV-2: a living systematic review and meta-analysis. *PLOS Med.* **2021**, *18*, e1003735.
- 431 6. Dinnes, J.; Deeks, J. J.; Berhane, S.; Ditttrich, S.; Emperador, D.; Takwoingi, Y.;
432 Cunningham, J.; Beese, S.; Dretzke, J.; di Ruffano, L. F.; Harris, I. M.; Price, M. J.; Taylor-
433 Phillips, S.; Hooft, L.; Mg Leeftang, M.; Spijker, R.; Van den Bruel, A.; Cochrane COVID-

- 434 19 Diagnostic Test Accuracy Group. Rapid, point-of-care antigen and molecular-based tests
435 for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst. Rev.* **2021**, *3*, Cd013705.
436 DOI: 10.1002/14651858.CD013705.pub2.
- 437 7. Corman, V. M.; Haage, V. C.; Bleicker, T.; Schmidt, M. L.; Mühlemann, B.; Zuchowski,
438 M.; Jo, W. K.; Tscheak, P.; Möncke-Buchner, E.; Müller, M. A.; Krumbholz, A.; Drexler,
439 J. F.; Drosten, C. Comparison of seven commercial SARS-CoV-2 rapid point-of-care
440 antigen tests: a single-centre laboratory evaluation study. *Lancet Microbe.* **2021**, *2*, e311-
441 e319.
- 442 8. Schroeder, L. F.; Bachman, M. A.; Idoni, A.; Gegenheimer-Holmes, J.; Kronick, S. L.;
443 Valdez, R.; Lephart, P. R. Predicting direct-specimen SARS-CoV-2 assay performance
444 using residual patient samples. *J. Appl. Lab. Med.* **2021**, *7*, 661-673.
- 445 9. Denzler, A.; Jacobs, M. L.; Witte, V.; Schnitzler, P.; Denking, C. M.; Knop, M. Rapid
446 comparative evaluation of SARS-CoV-2 rapid point-of-care antigen tests. *Infection*. [Online
447 early access]. <https://doi.org/10.1101/2021.07.29.21261314>. Published Online: April 9,
448 2022.
- 449 10. Bekliz, M.; Adea, K.; Essaidi-Laziosi, M.; Sacks, J. A.; Escadafa, C.; Kaiser, L.; Ecker, I.
450 SARS-CoV-2 antigen-detecting rapid tests for the delta variant. *Lancet Microbe.* **2022**, *3*,
451 e90.
- 452 11. Kontogianni, K.; Cubas-Atienzar, A. I.; Wooding, D.; Buist, K.; Thompson, C. R.;
453 Williams, C. T.; Baldwin, L.; Escadafal, C.; Sacks, J. A.; Adams, E. R.; Edwards, T. Lateral
454 flow antigen tests can sensitively detect live cultured virus of the SARS-CoV-2 B.1.1.7
455 lineage. *J. Infect.* **2021**, *83*, e1-e4.

- 456 12. Bekliz, M.; Adea, K.; Essaidi-Laziosi, M; Sacks, J. A.; Escadafal, C.; Kaiser, L.; Eckerle, I.
457 SARS-CoV-2 rapid diagnostic tests for emerging variants. *Lancet Microbe*. **2021**, 2, e351.
- 458 13. Stanley, S.; Hamel, D. J.; Wolf, I. D.; Riedel, S.; Dutta, S.; Contreras, E.; Callahan, C. J.;
459 Cheng, A.; Arnaout, R.; Kirby, J. E.; Kanki, P. J. Limit of detection for rapid antigen testing
460 of the SARS-CoV-2 Omicron and Delta variants of concern using live-virus culture. *J. Clin.*
461 *Microbiol.* [Online early access]. <https://doi.org/10.1128/jcm.00140-22>. Published Online:
462 April 20, 2022.
- 463 14. ExPASy ProtParam tool. <https://web.expasy.org/protparam/>. Accessed May 19, 2021.
- 464 15. Armbruster, D. A.; Pry, T. Limit of blank, limit of detection and limit of quantitation. *Clin.*
465 *Biochem. Rev.* **2008**, 29, S49-S52.
- 466 16. Bürkner, P. C. brms: an R package for Bayesian multilevel models using stan. *J. Stat.*
467 *Softw.* **2017**, 80, 1-28.
- 468 17. Sigal, G. B.; Novak, T.; Mathew, A. Measurement of SARS-CoV-2 antigens in plasma of
469 pediatric patients with acute COVID-19 or MIS-C using an ultrasensitive and quantitative
470 immunoassay. *MedRxiv Prepr. Serv. Health Sci.* [Online early access]. DOI:
471 10.1101/2021.12.08.21267502. Published Online: December 9, 2021.
- 472 18. Hingrat, Q. L.; Visseaux, B.; Laouenan, C.; Tubiana, S.; Bouadma, L.; Yazdanpanah, Y.;
473 Duval, X.; Burdet, C.; Ichou, H.; Damond, F.; Bertine, M.; Benmalek, N.; Choquet, C.;
474 Timsit, J-F.; Ghosn, J.; Charpentier, C.; Descamps, D.; Houhou-Fidouh, N. Detection of
475 SARS-CoV-2 N-antigen in blood during acute COVID-19 provides a sensitive new marker
476 and new testing alternatives. *Clin. Microbiol. Infect.* **2021**, 27, 789.e1-789.e5.
- 477 19. Pollock, N. R.; Savage, T. J.; Wardell, H; Lee, R. A.; Mathew, A.; Stengelin, M.; Sigal, G.
478 B. Correlation of SARS-CoV-2 nucleocapsid antigen and RNA concentrations in

- 479 nasopharyngeal samples from children and adults using an ultrasensitive and quantitative
480 antigen assay. *J. Clin. Microbiol.* **2021**, *59*.
- 481 20. Wang, H.; Hogan, C. A.; Verghese, M.; Solis, D.; Sibai, M.; Huang, C.; Zehnder, J.; Sahoo,
482 M. K.; Pinsky, B. A. Ultra-sensitive severe acute respiratory syndrome coronavirus 2
483 (SARS-CoV-2) antigen detection for the diagnosis of coronavirus disease 2019 (COVID-
484 19) in upper respiratory samples. *Clin. Infect. Dis.* **2021**, *73*, 2326-2328.
- 485 21. Bourassa, L.; Perchetti, G. A.; Phung, Q.; Lin, M. J.; Mills, M. G.; Roychoudhury, P.;
486 Harmon, K. G.; Reed, J. C.; Greninger, A. L. A SARS-CoV-2 nucleocapsid variant that
487 affects antigen test performance. *J. Clin. Virol.* **2021**, *41*, 104900.
- 488 22. Jian, M-J.; Chung, H-Y.; Chang, C-K.; Lin, J-C.; Yeh, K-M.; Chen, C-W.; Lin, D-Y.;
489 Chang, F-Y.; Hung, K-S.; Perng, C-L.; Shang, H-S. SARS-CoV-2 variants with T135I
490 nucleocapsid mutations may affect antigen test performance. *Int. J. Infect. Dis.* **2022**, *114*,
491 112-114.
- 492 23. Barrera-Avalos, C.; Luraschi, R.; Vallejos-Vidal, E.; Mella-Torres, A.; Hernández, F.;
493 Figueroa, M.; Rioseco, C.; Valdés, D.; Imarai, M.; Acuña-Castillo, C.; Reyes-López, F. E.;
494 Sandino, A. M. The rapid antigen detection test for SARS-CoV-2 underestimates the
495 identification of COVID-19 positive cases and compromises the diagnosis of the SARS-
496 CoV-2 (K417N/T, E484K, and N501Y) variants. *Front. Public Health.* **2022**, *9*, 780801.
- 497 24. Bayart, J. L.; Degosserie, J.; Favresse, J.; Gillot, C.; Didembourg, M.; Phanio Djokoto, H.;
498 Verbelen, V.; Roussel, G.; Maschietto, C.; Mullier, F.; Dogné, J-M.; Douxfils, J. Analytical
499 sensitivity of six SARS-CoV-2 rapid antigen tests for Omicron versus Delta variant.
500 *Viruses.* **2022**, *14*, 654.

- 501 25. Clinical and Laboratory Standards Institute (CLSI). *EP12-A2: User Protocol for Evaluation*
502 *of Qualitative Test Performance; Approved Guideline*, 2nd ed.; CLSI: Wayne,
503 Pennsylvania, USA, 2008.
- 504 26. McKay, S. L.; Tobolowsky, F. A.; Moritz, E. D.; Hatfield, K. M.; Bhatnagar, A.; LaVoie S.
505 P.; Jackson, D. A.; Lecy, K. D.; Bryant-Genevier, J.; Campbell, D.; Freeman, B.; Gilbert, S.
506 E. Folster; J. M.; Medrzycki, M.; Shewmaker, P. L.; Bankamp, B.; Radford, K. W.;
507 Anderson, R.; Bowen, M. D.; Negley, J.; Reddy, S. C.; Jernigan, J. A.; Brown, A. C.;
508 McDonald, L. C.; Kutty, P. K.; CDC Infection Prevention and Control Team and the CDC
509 COVID-19 Surge Laboratory Group. Performance evaluation of serial SARS-CoV-2 rapid
510 antigen testing during a nursing home outbreak. *Ann. Intern. Med.* **2021**, *174*, 945-951.
- 511 27. Begum, M. N.; Jubair, M.; Nahar, K.; Rahman, S.; Talha, M.; Sarker, M. S.; Nasir Uddin,
512 A. K. M.; Khaled, S.; Uddin, M. S.; Li, Z.; Ke, T.; Rahman, M. Z.; Rahman, M. Factors
513 influencing the performance of rapid SARS-CoV-2 antigen tests under field condition. *J.*
514 *Clin. Lab. Anal.* **2022**, *36*, e24203.
- 515 28. Jungnick, S.; Hobmaier, B.; Mautner, L.; Hoyos, M.; Haase, M.; Baiker, A.; Lahne, H.;
516 Eberle, U.; Wimmer, C.; Hepner, S.; Sprenger, A.; Berger, C.; Dangel, A.; Ippisch, S.;
517 Hahner, S.; Wildner, M.; Liebl, B.; Ackermann, N.; Sing, A.; Fingerle, V. In vitro rapid
518 antigen test performance with the SARS-CoV-2 variants of concern B.1.1.7 (Alpha),
519 B.1.351 (Beta), P.1 (Gamma), and B.1.617.2 (Delta). *Microorganisms.* **2021**, *9*, 1967.
- 520