

1 **Establishment of an evaluation panel for the decentralized technical evaluation of the**  
2 **sensitivity of 31 rapid detection tests for SARS-CoV-2 diagnostics**

3 Andreas Puyskens<sup>a§</sup>, Eva Krause<sup>a§</sup>, Janine Michel<sup>a§</sup>, Micha Nübling<sup>b</sup>, Heinrich Scheiblaue<sup>b</sup>,  
4 Daniel Bourquain<sup>a</sup>, Marica Grossege<sup>a</sup>, Roman Valusenko<sup>a</sup>, Viktor Corman<sup>c</sup>, Christian  
5 Drosten<sup>c</sup>, Katrin Zwirgmaier<sup>d</sup>, Roman Wölfel<sup>d</sup>, Constanze Lange<sup>e</sup>, Jan Kramer<sup>e</sup>, Johannes  
6 Friesen<sup>f</sup>, Ralf Ignatius<sup>f</sup>, Michael Müller<sup>f</sup>, Jonas Schmidt-Chanasit<sup>g</sup>, Petra Emmerich<sup>g</sup>, Lars  
7 Schaade<sup>a</sup>, Andreas Nitsche<sup>a\*</sup>

8 <sup>a</sup>Robert Koch Institute, Centre for Biological Threats and Special Pathogens – Highly  
9 Pathogenic Viruses (ZBS 1)

10 <sup>b</sup>Testing Laboratory for In-vitro Diagnostic Medical Devices, Paul-Ehrlich-Institute, Paul-  
11 Ehrlich-Straße 51-59, 63225 Langen, Germany

12 <sup>c</sup>Charité – Universitätsmedizin Berlin, Institute of Virology, Rahel-Hirsch-Weg 3, 10117 Berlin,  
13 Germany

14 <sup>d</sup>Bundeswehr Institute of Microbiology, Neuherbergstraße 11, 80937 Munich, Germany

15 <sup>e</sup>LADR Der Laborverbund Dr. Kramer & Kollegen GbR, Lauenburger Straße 67, 21502  
16 Geesthacht, Germany

17 <sup>f</sup>MVZ Labor28 GmbH, Mecklenburgische Str. 28, 14197 Berlin

18 <sup>g</sup>Bernhard Nocht Institute for Tropical Medicine, Arbovirology Department, Bernhard-Nocht-  
19 Straße 74, 20359 Hamburg, Germany

20

21 <sup>§</sup> These authors contributed equally to this work

22

23 \*Corresponding author

24 Prof. Dr. Andreas Nitsche

25 Robert Koch Institute

26 Centre for Biological Threats and Special Pathogens – Highly Pathogenic Viruses

27 Seestraße 10

28 13353 Berlin, Germany

29 Phone: +49 30 187542313

30 E-Mail: NitscheA@rki.de

31

32 **Abstract**

33 Background

34 The detection of SARS-CoV-2 with rapid diagnostic tests has become an important tool to  
35 identify infected people and break infection chains. These rapid diagnostic tests are usually  
36 based on antigen detection in a lateral flow approach.

37 Aims & Methods

38 While for PCR diagnostics the validation of a PCR assay is well established, for antigen tests  
39 e.g. rapid diagnostic tests there is no common validation strategy. Here we present the  
40 establishment of a panel of 50 pooled clinical specimens that cover a SARS-CoV-2  
41 concentration range from approximately  $1.1 \times 10^9$  to 420 genome copies per mL of  
42 specimen. The panel was used to evaluate 31 rapid diagnostic tests in up to 6 laboratories.

43 Results

44 Our results show that there is significant variation in the detection limits and the clinical  
45 sensitivity of different rapid diagnostic tests. We conclude that the best rapid diagnostic  
46 tests can be applied to reliably identify infectious individuals who are presenting with SARS-  
47 CoV-2 loads correlated to  $10^6$  genome copies per mL of specimen. Infected individuals  
48 displaying SARS-CoV-2 genome loads corresponding to less than  $10^6$  genome copies per mL  
49 will be identified by only some rapid diagnostics tests, while many tests miss these viral  
50 loads to a large extent.

51 Conclusions

52 Sensitive RDTs can be applied to identify infectious individuals with high viral loads, but not  
53 to identify infected individuals.

## 54 **Background**

55 PCR-based diagnostics of SARS-CoV-2 is a well-established method with numerous  
56 commercially available kits and in-house assays published over the last months [1]. Although  
57 it is beyond question that in particular real-time PCR provides an unrivalled degree of  
58 analytical sensitivity and reproducibility, there are some obvious limitations [2]. First of all,  
59 PCR-based diagnostics requires a functioning laboratory infrastructure and skilled personnel.  
60 The PCR reaction itself requires only about 2 hours, including pre- and post-analytical steps;  
61 however, time to result in high-throughput mode is about 24 hours or more.

62 The need for faster and simpler approaches to diagnose a SARS-CoV-2 infection is therefore  
63 evident as well as the need for on-site tests and for diagnostics in regions of lower standards  
64 of laboratory infrastructure [3–5]. One promising technology is using lateral flow  
65 immunoassays detecting SARS-CoV-2-specific proteins in respiratory secretions which  
66 operate within less than 30 minutes (Rapid Diagnostic Test: RDT) [6]. The trade-off for a  
67 simple and quick diagnostic test is an often significantly lower analytical sensitivity and  
68 specificity compared to nucleic acid amplification techniques like PCR [7,8].

69 Beside the traditional RDTs whose read-out is performed visually, there are assays that  
70 utilize readers for the identification of positive signals. These readers can provide a better  
71 sensitivity, reproducibility and objectivity, in particular with fluorescence-based-formats;  
72 however, the mobility and parallel testing of many specimens can be negatively affected.

73 Here we describe the decentralized evaluation of the sensitivity of 31 RDTs by up to six  
74 German laboratories investigating an identical panel of 50 pools of clinical specimens. By this  
75 approach, we generated at least two independent results per RDT and hence addressed

76 inter-laboratory variations that have to be considered when RDTs are performed in different  
77 locations by different persons.

## 78 **Material & Methods**

### 79 **Evaluation panel**

80 To enable the systematic and comparable decentralized evaluation of numerous RDTs, a  
81 panel of 50 samples was compiled by pooling in total approximately 500 specimens which  
82 had been collected between March and September 2020 (Panel 1V1). For this purpose, up to  
83 ten respiratory specimens obtained for routine diagnostics with different virus loads were  
84 pooled. Pools were frozen at -80 °C. Real-time PCR [(9)] was applied to determine the RNA  
85 load per pool. In vitro RNA (provided by WHO) as well as the quantitative reference material  
86 provided by INSTAND were used for quantification (<https://www.instand-ev.de>). Finally, the  
87 panel covered a range of SARS-CoV-2 RNA from  $1.1 \times 10^9$  genomes per mL down to 420  
88 genomes per mL. The study obtained ethical approval by the Berliner Ärztekammer (Berlin  
89 Chamber of Physicians, Eth 20/40). When Panel 1V1 was used up, new pools were generated  
90 by diluting the same samples as for Panel 1V1 (except for 4 pools 1–4 that had to be  
91 constituted from new clinical specimens collected between October 2020 and January 2021),  
92 resulting in comparable virus loads as determined by PCR, and the panel was labelled Panel  
93 1V2. Panel 1V1 was compared to Panel 1V2 with RDT #3 and RDT #31 and showed identical  
94 results. Panel 1V1 and Panel 1V2 have been extensively used to evaluate the sensitivity of  
95 122 RDTs as described in the tandem publication by Scheiblaue et al.)

96 Previous studies revealed that a minimal RNA genome copy number of  $10^6$  genome copies  
97 per mL of specimen represents an amount of infectious virus particles which is required for  
98 successful virus propagation in cell culture [9–12]. To correlate the pools to potential

99 infectivity by a specimen, we subjected the pools with  $\geq 10^6$  genome copies per mL,  
100 corresponding to a CT value  $< 25$ , to cell culture. Confirmation of replication-competent  
101 SARS-CoV-2 was achieved by inoculation of VeroE6 cells with the respective pools. Pools  
102 containing infectious SARS-CoV-2 were subsequently titrated on VeroE6 cells. However, even  
103 if pools containing higher amounts of SARS-CoV-2 RNA showed generally higher titres than  
104 those with lower genome numbers, we observed no significant correlation between the  
105 genome load and the titre (data not shown).

106 The specifications of the 50 pools are listed in table 1; pools allowing SARS-CoV-2  
107 propagation are marked in bold.

108

109 Table 1: Characteristics of the 50 pools constituting Panel 1V1 and Panel 1V2.

Panel 1V1			Panel 1V2		
Pool-Nr.	CT/5 $\mu$ L of RNA	RNA subjected to test	Pool-Nr.	CT/5 $\mu$ L of RNA	RNA subjected to test
1	<b>17.55</b>	1.1E+07	1	<b>17.31</b>	1.31E+07
2	<b>20.54</b>	1.4E+06	2	<b>19.08</b>	3.87E+06
3	20.38	1.6E+06	3	<b>19.62</b>	2.67E+06
4	<b>20.98</b>	1.0E+06	4	<b>20.61</b>	1.35E+06
5	<b>20.28</b>	1.7E+06	5	<b>20.60</b>	1.36E+06
6	<b>20.20</b>	1.8E+06	6	21.21	8.96E+05
7	<b>21.71</b>	6.4E+05	7	22.15	4.70E+05

8	21.95	5.4E+05	8	22.32	4.18E+05
9	22.14	4.7E+05	9	23.13	2.39E+05
10	22.88	2.8E+05	10	23.21	2.27E+05
11	22.34	4.1E+05	11	23.13	2.27E+05
12	<b>21.82</b>	5.9E+05	12	22.12	4.79E+05
13	23.32	2.1E+05	13	25.29	5.42E+04
14	24.28	1.1E+05	14	24.97	6.76E+04
15	<b>24.14</b>	1.2E+05	15	24.38	1.01E+05
16	22.55	3.6E+05	16	22.88	2.84E+05
17	24.00	1.3E+05	17	24.81	7.54E+04
18	25.30	5.4E+04	18	28.33	6.71E+03
19	25.50	4.7E+04	19	25.45	2.39E+05
20	26.27	2.8E+04	20	29.46	2.27E+05
21	25.54	4.6E+04	21	25.95	2.27E+05
22	25.87	3.7E+04	22	27.42	4.79E+05
23	<b>24.04</b>	1.3E+05	23	24.45	5.42E+04
24	25.24	5.6E+04	24	25.20	6.76E+04
25	29.70	2.6E+03	25	25.07	1.01E+05
26	25.47	4.8E+04	26	26.32	2.84E+05
27	25.14	6.0E+04	27	26.12	7.54E+04
28	27.14	1.5E+04	28	27.41	6.71E+03
29	27.15	1.5E+04	29	27.34	1.33E+04
30	28.86	4.7E+03	30	27.24	1.42E+04

31	25.27	5.5E+04	31	26.24	1.42E+04
32	26.44	2.5E+04	32	26.64	2.14E+04
33	28.96	4.4E+03	33	28.92	4.47E+03
34	27.89	9.1E+03	34	27.82	9.53E+03
35	27.04	1.6E+04	35	26.66	2.12E+04
36	28.13	7.7E+03	36	27.05	1.62E+04
37	30.54	1.5E+03	37	30.13	1.95E+03
38	28.14	7.6E+03	38	29.36	3.31E+03
39	29.76	2.5E+03	39	30.12	1.96E+03
40	27.65	1.1E+04	40	28.19	7.39E+03
41	30.13	1.9E+03	41	30.14	7.39E+03
42	28.43	6.2E+03	42	29.48	3.04E+03
43	31.05	1.0E+03	43	31.61	7.04E+02
44	29.24	3.6E+03	44	29.51	2.98E+03
45	30.10	2.0E+03	45	31.19	9.40E+02
46	31.54	7.4E+02	46	31.34	8.48E+02
47	35.19	6.0E+01	47	34.55	9.34E+01
48	32.06	5.2E+02	48	31.19	9.40E+02
49	35.22	5.9E+01	49	36.04	3.35E+01
50	36.36	2.7E+01	50	35.83	3.87E+01

110 CT values given in bold indicate that SARS-CoV-2 could be propagated in cell culture.

111

112

113 **RDT selection**

114 RDTs included in this study were selected according to availability at the time of the study

115 (table 2). No technical assumptions were made in the RDT selection process.

116

117 Table 2: RDTs evaluated in this study

#	Name	Manufacturer/distributor	# of evaluating laboratories
Classical Point-of-Care tests			
1	SARS-CoV-2 Rapid Antigen Test	SD BIOSENSOR, distr. by Roche Diagnostics GmbH	6
2	Panbio™ COVID-19 Ag Rapid Test Device (Nasopharyngeal)	Abbott	5
3	RIDA® QUICK SARS-CoV-2 Antigen	R-Biopharm AG	3
4	Cleartest	servoprax GmbH	2
5	Dedicio COVID-19 Ag plus Test	Jiangsu Changfeng Medical Industry Co., Ltd., distr. by nal von minden GmbH	3
6	NADAL® COVID-19 Ag Test	nal von minden GmbH	3
7	BIOCREDIT COVID-19 Ag	RapiGEN Inc., distr. by Weko Pharma	3
8	STANDARD™ Q COVID-19 Ag Test	SD BIOSENSOR	2
9	BIOSYNEX COVID-19 Ag BSS	Biosynex Swiss SA	2
10	MEDsan® SARS-CoV-2 Antigen Rapid Test	MEDsan GmbH	2

11	NowCheck® COVID-19 Ag Test	BIONOTE, distr. by concile	3
12	COVID-19 Rapid Ag Test Cassette	Zhejiang Orient Gene Biotech Co., Ltd	2
13	ESPLINE® SARS-CoV-2	Fujirebio Inc.	2
14	COVID-19 Ag Respi-Strip	Coris BioConcept	2
15	Healgen Coronavirus Ag Rapid Test Cassette (Swab)	Healgen Scientific LLC	2
16	Keul-o-test®	Günter Keul GmbH	2
17	Acro Rapid Test COVID-19 Antigen Rapid Test (Nasopharyngeal Swab)	Acro Biotech. Inc.	2
18	COVID-19 Antigen Schnelltest (Nasen-Rachenabstrich)	MEXACARE GmbH	2
19	COVID-19 Antigen Rapid Test Kit	Beijing Beier Bioengineering Co., Ltd.	2
20	mö-screen Testkit Corona Antigen Nasenabstrich	möLab GmbH	2
21	Medicovid SARS-CoV-2 AG Antigen Schnelltest	Xiamen Biotime Biotechnology Co., Ltd., distr. by MEDICE Arzneimittel Pütter GmbH & Co. KG	2
22	GENEDIA COVID-19 Ag	Green Cross Medical Science Corp.	2
23	COVID-19 Antigen Test Kit	Guangdong Wesail Biotech Co., Ltd., distr. by Bio-Gram Diagnostics GmbH	2
24	COVID-19 Antigen Schnelltest (Colloidal Gold)	Joinstar Biomedical Technology Co., Ltd., distr. by care impuls Vertriebs GmbH	2
25	blnk COVID-19 Antigen Rapid Test (Nasopharyngeal Swab)	Hangzhou Realy Tech Co. Ltd., distr. by TREKSTOR GmbH	2
26	Novel Coronavirus (COVID-19)	Hangzhou Laihe Biotech Co., Ltd.,	2

	Antigen Test Kit	distr. by Lissner Qi GmbH	
27	COVID-19 Antigen Rapid Test Strip	Koch Biotechnology (Beijing) Co., Ltd.	2
	RDTs requiring a read-out device		
28*	STANDARD™ F COVID-19 Ag FIA	SD Biosensor	4
29*	Sofia SARS Antigen FIA	Quidel Corporation	3
30*	ScheBo® SARS-CoV-2 Quick™ Antigen	Schebo Biotech AG	3
31*	Wantai SARS-CoV-2 Antigen Rapid Test (FIA)	Beijing Wantai Biological Pharmacy	2

118 RDTs marked with \* require a test-specific device and are not to be judged visually with the  
 119 device.

120

121

## 122 RDT procedure

123 In general, RDTs were used as recommended by the manufacturer’s instructions. 50 µL of  
 124 each pool were either subjected to the RDT using all buffers in volumes as recommended by  
 125 the manufacturer, or the swab included in the RDT kit was used to absorb the 50 µL of each  
 126 pool and then subjected to the RDT procedure as recommended by the manufacturer.  
 127 Results obtained by visual examination of the test device by different laboratories were  
 128 categorized as “positive” or “negative” and subjected to statistical analysis by using the  
 129 GraphPad Prism software as indicated in the results section. Results were only accepted  
 130 when the control band was positive, which was the case in more than 99% of the tests run.

## 131 Results

## 132 **Evaluation panel establishment**

133 Prior to distribution of the panel we compared the detectability of specimens before and  
134 after freezing at -40°C with two RDTs (#2, #3) and observed no significant differences in the  
135 virus concentration range close to the detection limit (data not shown). Furthermore, to test  
136 whether pooling and freezing had an impact on the detectability of specimens, we compared  
137 up to 44 fresh clinical specimens, representing a CT value range from 20 to 35 ( $1.8 \times 10^7$  to  
138  $7.0 \times 10^3$  genome copies per mL), with the 32 pools of Panel 1V1 covering the same range  
139 with in total 10 RDTs (#2, #3, #7, #8, #10, #11, #14, #21, #28, #31). In none of the RDTs  
140 tested we observed a significant discrepancy between the detectability of fresh specimens  
141 and pools with a similar CT value (data summarized in supplemental figure 1).

142

143 Supplemental figure 1: Comparison of 10 randomly selected RDTs with pools from Panel 1V1  
144 ( $n \leq 32$ ) and fresh clinical specimens ( $n \leq 44$ ), covering a genome load from approximately  $10^7$   
145 to  $7 \times 10^3$  genome copies per mL. None of the RDTs showed a significant difference in the  
146 detectability of pools in comparison to fresh specimens.

147

## 148 **RDT analytical sensitivity**

149 The panel was shipped on dry ice to the participating laboratories. Prior to use, the panel  
150 was thawed, mixed and aliquoted, and identical aliquots were used immediately to allow  
151 maximum comparability between laboratories and test days. Figure 1 summarizes the  
152 workflow as recommended to the participating laboratories. Assessment of results was  
153 performed visually by experienced laboratory personnel.

154

155 Figure 1: Recommendation for panel usage to guarantee maximum comparability between  
156 different laboratories and points in time of validation.

157

158 Figure 2 comprises the results obtained for the sensitivity of 27 RDTs that could be analysed  
159 visually (RDT #1–27) and the results for RDTs that needed a device for read-out (RDT #28–  
160 31). Based on binary logistic regression of merged results obtained from different  
161 laboratories, the 50% probability to detect a certain genome load was calculated as a marker  
162 for the detection limit of an RDT. Six out of 31 RDTs showed a 50% detection probability for  
163 genome loads higher than  $10^6$  genome copies per mL. Out of 31 RDTs, 24 had a 50%  
164 detection probability of less than  $10^6$  genome copies per mL, while 15 had a 50% detection  
165 probability of less than  $10^5$  genome copies per mL. The most sensitive RDTs detected  
166 approximately 75,000 genome copies per mL with a probability of 90%, while the least  
167 sensitive RDT showed a 90% detection probability for  $2.3 \times 10^7$  genomes per mL.

168 Figure 2 also shows that using the RDT-specific swabs to absorb the pool material prior to  
169 testing can lead to a loss of analytical sensitivity of approximately the factor 10 to 50,  
170 although some of the RDTs showed only small differences between direct application of the  
171 pool and swab usage. This indicates a varying efficiency of the absorption and release  
172 characteristics for SARS-CoV-2 particles from the swabs.

173

174 Figure 2: Analytical sensitivity expressed by the 50% detection probability of 31 RDTs. RDTs  
175 28 to 31 require an additional device for reading out the result. Binary logistic regression was

176 applied to calculate the analytical sensitivity from the 50 samples included in the panel.  
177 Open blue bars represent the results when virus-containing pools were directly subjected to  
178 the RDT without using a swab. Open green squares represent the corresponding results  
179 when using a swab. Crossed squares indicate that no results were determined with swabs,  
180 while green crosses show the one positive result from a pool with the corresponding virus  
181 load to CT 17.55, pool #1.

182

### 183 **RDT clinical sensitivity**

184 To determine the clinical sensitivity of an RDT, results of different laboratories were merged  
185 and categorized according to the genome load, e.g. CT value obtained for each pool by real-  
186 time PCR [8a]:  $CT < 25$  ( $10^6$  genome copies per mL),  $25 < CT < 30$  and  $CT > 30$ . Sensitivities of each  
187 RDT to identify pools correctly were calculated for each CT category.

188 According to previous studies,  $CT < 25$  corresponds to an increased probability of a specimen  
189 to successfully propagate SARS-CoV-2 in cell culture [9–11]. This was confirmed for 9/18  
190 pools with CT values  $< 25$  for Panel 1V1 and 5/17 pools for Panel 1V2. Pools with a  $CT > 30$   
191 were highly unlikely to contain virus amounts high enough to grow in cell culture. Specimens  
192 with CT values between 25 and 30 very rarely propagated virus in cell culture.

193 Figure 3 summarizes the sensitivities for the 31 RDTs evaluated. Blue circles represent  
194 results when no swab was used and green circles those when swabs were used. Figure 3A  
195 shows the sensitivity for the whole number of 50 SARS-CoV-2-positive pools of the panel.  
196 Results are further categorized according to the different virus loads represented by a pool.  
197 Figure 3B covers pools with  $CT < 25$  ( $n=18$  Panel 1V1,  $n=17$  Panel 1V2, potentially infectious).  
198 Figure 3C shows pools with  $CT > 25$  ( $n=32$  Panel 1V1,  $n=33$  Panel 1V2) and figure 3D pools

199 with a CT value between 25 and 30, which is the range where the RDTs showed significant  
200 differences (n=23 for both panels). The number of laboratories contributing to a result is  
201 given in table 1.

202 The data reveal that for virus loads higher than  $10^6$  genome copies per mL (CT<25) the  
203 sensitivity of 26/31 RDTs was higher than 80%, indicating that these RDTs would potentially  
204 identify infectious specimens with a probability of 80% (figure 3A). For the latter RDTs, the  
205 proficiency to detect these pools that contain culturable SARS-CoV-2 was even better with  
206 values up to 100% (data not shown). For a virus load  $<10^6$  genome copies per mL (CT>25),  
207 none of the evaluated tests could surpass a sensitivity of 80% (figure 3B). In the range  
208 between CT 25 to CT 30, 10/31 RDTs reached a sensitivity of 80% and higher, even if five  
209 further RDTs showed sensitivities close to 80% (figure 3D). Finally, when the sensitivity to  
210 detect all 50 pools of the panel was determined, only 4/31 RDTs passed a sensitivity of 80%  
211 or higher, with two of the four RDTs requiring a detection device for read-out. However,  
212 10/31 RDTs showed an overall sensitivity higher than 70% (figure 3A). Using a swab to  
213 reduce the volume of the pool material in the RDT test procedure can lead to reduced  
214 sensitivity due to loss of virus particles in the absorption and release process. Hence, as  
215 described for the analytical sensitivity based on the RNA detection limit, clinical sensitivities  
216 were lower for most RDTs when a swab was used.

217

218 Figure 3: Clinical sensitivities of 31 RDTs as determined by two to six laboratories using 50  
219 pools from evaluation Panels 1V1 and 1V2. Blue circles indicate direct application of the pool  
220 to the test buffer and green circles indicate results using the RDT-specific swab. Each circle

221 represents the results of one laboratory (symbols may cover each other). RDTs 28 to 31  
222 require an additional device for reading out the result.

223

224 Even if most of the RDTs were analysed in two independent laboratories only, RDTs 1–5 and  
225 28–30 were evaluated 3–6 times with or without using swabs. As shown in figures 2 and 3,  
226 there can be significant variability for some tests, which is most likely due to the more  
227 subjective interpretation of a positive signal; however, most results were very similar even  
228 when generated in different laboratories on different days.

## 229 **Discussion**

230 RDTs are promising tools in the diagnostic portfolio of tools for the identification of SARS-  
231 CoV-2-infected individuals [13–15]. Since these tests do not use amplification of the target,  
232 like PCR, their analytical sensitivity is usually limited. Hence, the evaluation of RDTs plays a  
233 major role in defining the proper applications of RDTs. In contrast to PCR, where the  
234 specimen can be inactivated, RDTs should be evaluated with clinical material that contains  
235 native viruses to mirror the diagnostic application as authentically as possible. However, the  
236 systematic comparison of various RDTs in different laboratories at different times makes  
237 larger sample volumes and a good storage stability necessary.

238 Multiple sampling of naso- and/or oropharyngeal swabs is hampered by reproducibility.  
239 Even sampling the same patient with several swabs consecutively will likely lead to a  
240 variation in the virus load per swab which has to be controlled by real-time PCR, changing  
241 the test procedure. Because most of the RDT protocols require the clinical specimen to be  
242 sampled with a swab from which virus has to be eluted in the system-specific buffer, one

243 swab cannot be used more than once without changing the protocol. Therefore, the  
244 decentralized evaluation of various RDTs in different laboratories is obviously difficult. So far,  
245 clinical samples with semi-quantified SARS-CoV-2 concentrations or virus propagated in cell  
246 culture have been used to validate RDTs [16–18]. This can be done for a limited number of  
247 tests in a short period of time but is not suitable when numerous RDTs have to be compared  
248 regularly in different laboratories.

249 Therefore, we established an evaluation panel that was used to determine the analytical and  
250 clinical sensitivity of RDTs providing comparable results. The main basis of each pool were  
251 dry swabs in PBS originally used for PCR diagnostics. Since some of the swabs were  
252 transported to the laboratory in medium, the final concentration of medium in pools was  
253  $\leq 20\%$  for each pool,  $\leq 10\%$  for 38 pools and  $0\%$  for 25 pools. Ten RDTs were randomly  
254 selected to calculate whether the medium content had an impact on the sensitivity. As  
255 shown in supplemental figure 2 for one of the RDTs, binary logistic regression revealed no  
256 difference in the detection probability of pools containing varying amounts of medium.  
257 Although these results are only derived from ten RDTs, we believe that this low medium  
258 content has no influence on the test sensitivity in general, since pools with medium  
259 percentages between  $10\%$  and  $20\%$  were distributed over the whole panel 1V1.

260

261 Supplemental figure 2: Impact of medium contents in the 50 pools included in Panel 1V1.  
262 Binary logistic regression was used to demonstrate that the amount of medium ranging from  
263  $0\%$  to  $20\%$  has no impact on the detection limit of one randomly selected RDT.

264

265 SARS-CoV-2 genome load was determined by real-time PCR in clinical specimens, and those  
266 specimens with similar load were pooled diluted in a background of negative swabs and the  
267 virus load quantified again. The established pools had a volume of 10 mL for both panel 1V1  
268 and 1V2 and covered a genome load from  $1.1 \times 10^9$  genomes per mL down to 400 genomes  
269 per mL, which is the range of typical clinical specimens analysed in our laboratory. Even if  
270 the genome load does not reflect the number of virus particles directly, the RNA copy  
271 number was recently used to estimate the number of virus particles reflecting the infectious  
272 potential of a specimen and can correlate with the N protein concentration in clinical  
273 samples [19].

274 The fact that we used pools of up to 10 clinical specimens facilitates to compensate to some  
275 degree for a potential variation between individual samples, for example varying ratios of  
276 genome copies to the number of viral particles or rather the antigen concentration.

277 We could demonstrate that the pools of the panel showed results comparable to those for  
278 fresh clinical specimens with similar SARS-CoV-2 genome load when selected RDTs were  
279 used, and that freezing at -40 to -80°C did not impact the detectability significantly.

280 Nevertheless, as a trade-off for better reproducibility and comparability, the material used  
281 was not as fresh as in a clinical setting [.

282 Usually, RDTs use swabs that are subjected to the RDT-specific buffer before incubation of  
283 the test membrane is done. Our approach necessarily began with liquid specimens of 50 µL,  
284 which is intended for some RDTs, but not for all of them. However, with the intention to  
285 allow the generation of comparable and reproducible data, we accepted that some of the  
286 buffers were diluted by the fluid of the pools. Adding 50 µl of pools directly to three

287 randomly selected RDTs without additional test-specific buffer led to results indicating that  
288 additional buffers were not required at least in these RDTs.

289 The strategy of using liquid specimens comes with a further benefit, e.g. the option to  
290 cultivate the pools in cell culture, showing the infectivity of pools with a sufficient virus load.  
291 To our knowledge, this is one of the few studies that systematically evaluated several  
292 commercially available Ag RDTs using standardized samples in comparison to qPCR as well as  
293 infectivity data from cell culture [21]. However, since all specimens included in a pool have  
294 been frozen and thawed at least once, the capability to grow in cell culture can be improved  
295 with fresh clinical specimens of a comparable virus amount [22].

296 Variability observed in different laboratories has been described for diseases like malaria  
297 [23] and was significant for some RDTs, but in general highly comparable results were  
298 obtained. Therefore, our results reflect the natural variance that can be expected when  
299 different users apply the RDTs. Over time we observed that users adopted the test and that  
300 the differentiation of weak positive results from negative results became better with  
301 increasing experience. Therefore, it can be assumed that the interpretation of results is  
302 better standardized with RDTs that require a device to read out the signal. Based on the  
303 number of RDTs we have validated, we can confirm that read-out devices can help generate  
304 better reproducible results and reduce the inter-user variance.

### 305 **Influence of swabs on the sensitivity of RDTs**

306 The common RDT test starts with the sampling that results in a swab containing material  
307 from the mucosa of the naso- or oropharynx. Then the swab is put into the RDT-specific  
308 buffer and is subjected to the test membrane. Applying liquid evaluation specimens with  
309 known virus amounts therefore does not consider the impact of the swab on the result. The

310 swab has to absorb liquids from the mucosa potentially containing SARS-CoV-2 or can scrape  
311 cellular material containing virus; probably it is a combination of both. Beside the problem  
312 that the swab does not quantitatively absorb the specimen, virus proteins can stick to the  
313 swab and will not be subjected to the RDT, reducing the analytical sensitivity. While some of  
314 the RDTs do not suffer significantly from using a swab prior to testing (#13, #28, #30), most  
315 of the tests lose sensitivity approximately by the factor 10 to 20. This does not mean that the  
316 respective test device is inferior to the previously mentioned tests, but rather that the swab  
317 is not efficient in absorbing and releasing SARS-CoV-2 from a liquid specimen. However,  
318 speculating that these swabs will come with the same drawbacks when used in clinical  
319 sampling, the loss of sensitivity can also occur. In a patient carrying for example  $10^6$  genome  
320 copies in the nasopharynx, with RNA load used as surrogate for viral particles, the sampling  
321 on the mucosa bears the risk to obtain a false negative result in the RDT due to considerable  
322 loss of antigen in the swab. Further investigations of the efficiency of virus absorption and  
323 release from a swab will help interpret the risk of false negatives by sampling. Besides,  
324 further specimens like saliva are under investigation for their use in RDTs.

325 Finally, our study did not address the specificity of an RDT and cannot assess the risk of false  
326 positive results. However, recent studies show that high specificity is reached by most of the  
327 RDTs evaluated [24]. In addition to SARS-CoV-2-positive specimens, the next version of the  
328 evaluation panel will also include specimens containing SARS-CoV-2-negative specimens  
329 positive for further respiratory viruses.

### 330 **Conclusion**

331 The sensitivity of the 31 RDTs evaluated in this study varied significantly and depended  
332 largely on the virus load in the respective specimen. While four RDTs showed a sensitivity of

333 >80% over the whole range of virus loads investigated, 26 RDTs had a sensitivity of >80% for  
334 potentially infectious specimens, indicating that sensitive RDTs can be used to identify  
335 contagious individuals in various settings, but not to identify infected individuals with lower  
336 virus loads. Our results are in agreement with several other studies not using a standardized  
337 evaluation panel [25], indicating the applicability of the described panel for RDT evaluation.  
338 The minimal performance characteristics for an RDT have been recently discussed to be at  
339 least 80% for symptomatic patients [13]. Considering varying virus loads during the time  
340 course of infection in an individual and in between individuals, the sensitivity of an RDT  
341 should therefore better be attributed to a certain virus load instead of the time after onset  
342 of symptoms or a qualitative PCR result.

#### 343 **Acknowledgements**

344 The authors are grateful to Ursula Erikli for copy-editing. This work was funded by the  
345 Ministry of Health, Germany (Maßnahmepaket 1&2).

#### 346 **Authors contributions**

347 Andreas Puyskens, Eva Krause, Janine Michel and Marica Grossegeesse designed the study,  
348 established the evaluation panel and performed the RDTs with Roman Valusenko. Daniel  
349 Bourquain performed cell culture experiments. Micha Nübling and Heinrich Scheiblauer  
350 characterized the panel and evaluated several RDTs. Viktor Corman, Christian Drosten, Katrin  
351 Zwirgmaier, Roman Wölfel, Constanze Lange, Jan Kramer, Johannes Friesen, Ralf Ignatius,  
352 Michael Müller, Jonas Schmidt-Chanasit, and Petra Emmerich provided data for varying  
353 numbers of RDTs with the evaluation panel. Lars Schaade and Andreas Nitsche  
354 conceptualized the study, analysed the data and wrote the manuscript.

355 References

- 356 1. Nalla AK, Casto AM, Huang M-LW, Perchetti GA, Sampoleo R, Shrestha L, et al.  
357 Comparative Performance of SARS-CoV-2 Detection Assays using Seven Different  
358 Primer/Probe Sets and One Assay Kit. *J Clin Microbiol*. 2020;58(6).
- 359 2. Mackay IM, Arden KE, Nitsche A. Real-time Fluorescent PCR Techniques to Study  
360 Microbial-Host Interactions. Vol. 34, *Methods in Microbiology*. 2004. p. 255–330.
- 361 3. Peeling RW, McNerney R. Emerging technologies in point-of-care molecular  
362 diagnostics for resource-limited settings. Vol. 14, *Expert Review of Molecular*  
363 *Diagnostics*. Expert Reviews Ltd.; 2014. p. 525–34.
- 364 4. Strömer A, Rose R, Schäfer M, Schön F, Vollersen A, Lorentz T, et al. Performance of a  
365 Point-of-Care Test for the Rapid Detection of SARS-CoV-2 Antigen. *Microorganisms*  
366 [Internet]. 2020 Dec 28 [cited 2021 Apr 23];9(1):1–11. Available from:  
367 <http://www.ncbi.nlm.nih.gov/pubmed/33379279>
- 368 5. Jacobs J, Kühne V, Lunguya O, Affolabi D, Hardy L, Vandenberg O. Implementing  
369 COVID-19 (SARS-CoV-2) Rapid Diagnostic Tests in Sub-Saharan Africa: A Review. Vol.  
370 7:557797, *Frontiers in Medicine*. Frontiers Media S.A.; 2020. p. 557.
- 371 6. Kozel TR, Burnham-Marusich AR. Point-of-care testing for infectious diseases: Past,  
372 present, and future. Vol. 55, *Journal of Clinical Microbiology*. American Society for  
373 *Microbiology*; 2017. p. 2313–20.
- 374 7. Berzosa P, De Lucio A, Romay-Barja M, Herrador Z, González V, García L, et al.  
375 Comparison of three diagnostic methods (microscopy, RDT, and PCR) for the detection  
376 of malaria parasites in representative samples from Equatorial Guinea 11 *Medical and*

- 377 Health Sciences 1108 Medical Microbiology. Malar J [Internet]. 2018 Sep 17 [cited  
378 2021 Apr 23];17(1):333. Available from:  
379 <https://malariajournal.biomedcentral.com/articles/10.1186/s12936-018-2481-4>
- 380 8. <https://www.finddx.org/sarscov2-eval-antigen/>. A from: FIND evaluation of SARS-  
381 CoV-2 antigen (Ag) detecting tests - FIND.
- 382 9. Janine Michel, Markus Neumann, Eva Krause, Thomas Rinner, Therese Muzeniek,  
383 Marica Grossegeesse, Georg Hille, Franziska Schwarz, Andreas Puyskens, Sophie  
384 Förster, Barbara Biere, Daniel Bourquain, Cristina Domingo, Annika Brinkmann, Lars  
385 Schaade, Livia Schr AN. Resource-efficient internally controlled in-house real-time PCR  
386 detection of SARS-CoV-2. Virol J. 2021;Forthcomin.
- 387 10. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al.  
388 Virological assessment of hospitalized patients with COVID-2019. Nature. 2020  
389 May;581(7809):465–9.
- 390 11. Kim M-C, Cui C, Shin K-R, Bae J-Y, Kweon O-J, Lee M-K, et al. Duration of Culturable  
391 SARS-CoV-2 in Hospitalized Patients with Covid-19. N Engl J Med. 2021;348(7):671–3.
- 392 12. van Kampen JJA, van de Vijver DAMC, Fraaij PLA, Haagmans BL, Lamers MM, Okba N,  
393 et al. Duration and key determinants of infectious virus shedding in hospitalized  
394 patients with coronavirus disease-2019 (COVID-19). Nat Commun. 2021 Dec  
395 1;12(1):267.
- 396 13. La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, et al. Viral RNA  
397 load as determined by cell culture as a management tool for discharge of SARS-CoV-2  
398 patients from infectious disease wards. Eur J Clin Microbiol Infect Dis. 2020;

- 399 14. WHO. COVID-19 Target product profiles for priority diagnostics to support response to  
400 the COVID-19 pandemic v.1.0 [Internet]. 2020 [cited 2021 Apr 22]. Available from:  
401 [https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-](https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1)  
402 [priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1](https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1)
- 403 15. Berger A, Nsoga MTN, Perez-Rodriguez FJ, Aad YA, Sattonnet-Roche P, Gayet-Ageron  
404 A, et al. Diagnostic accuracy of two commercial SARSCoV- 2 antigen-detecting rapid  
405 tests at the point of care in community-based testing centers. PLoS One. 2021 Mar  
406 1;16(3 March 2021).
- 407 16. Mboumba Bouassa RS, Veyer D, Péré H, Bélec L. Analytical performances of the point-  
408 of-care SIENNA™ COVID-19 Antigen Rapid Test for the detection of SARS-CoV-2  
409 nucleocapsid protein in nasopharyngeal swabs: A prospective evaluation during the  
410 COVID-19 second wave in France. Int J Infect Dis. 2021 May 1;106:8–12.
- 411 17. Olearo F, Nörz D, Heinrich F, Sutter JP, Roedel K, Schultze A, et al. Handling and  
412 accuracy of four rapid antigen tests for the diagnosis of SARS-CoV-2 compared to RT-  
413 qPCR. J Clin Virol. 2021 Apr 1;137(April104782).
- 414 18. Corman VM, Haage VC, Bleicker T, Schmidt ML, Mühlemann B, Zuchowski M, et al.  
415 Comparison of seven commercial SARS-CoV-2 rapid point-of-care antigen tests: a  
416 single-centre laboratory evaluation study. The Lancet Microbe. 2021 Apr;
- 417 19. Jääskeläinen AE, Ahava MJ, Jokela P, Szirovicza L, Pohjala S, Vapalahti O, et al.  
418 Evaluation of three rapid lateral flow antigen detection tests for the diagnosis of SARS-  
419 CoV-2 infection. J Clin Virol. 2021 Apr 1;137.
- 420 20. Pollock NR, Savage TJ, Wardell H, Lee R, Mathew A, Stengelin M, et al. Correlation of

- 421 SARS-CoV-2 nucleocapsid antigen and RNA concentrations in nasopharyngeal samples  
422 from children and adults using an ultrasensitive and quantitative antigen assay.  
423 medRxiv. medRxiv; 2020. p. 2020.11.10.20227371.
- 424 21. Kohmer N, Toptan T, Pallas C, Karaca O, Pfeiffer A, Westhaus S, et al. The Comparative  
425 Clinical Performance of Four SARS-CoV-2 Rapid Antigen Tests and Their Correlation to  
426 Infectivity In Vitro. *J Clin Med*. 2021 Jan 17;10(2):328.
- 427 22. Basile K, McPhie K, Carter I, Alderson S, Rahman H, Donovan L, et al. Cell-based  
428 culture of SARS-CoV-2 informs infectivity and safe de-isolation assessments during  
429 COVID-19. *Clin Infect Dis*. 2020 Oct 24;
- 430 23. Boyce MR, Menya D, Turner EL, Laktabai J, Prudhomme-O'Meara W. Evaluation of  
431 malaria rapid diagnostic test (RDT) use by community health workers: A longitudinal  
432 study in western Kenya. *Malar J*. 2018 May 18;17(1).
- 433 24. Baro B, Rodo P, Ouchi D, Bordoy AE, Saya Amaro EN, Salsench S V., et al. Performance  
434 characteristics of five antigen-detecting rapid diagnostic test (Ag-RDT) for SARS-CoV-2  
435 asymptomatic infection: a head-to-head benchmark comparison. *J Infect* [Internet].  
436 2021 Apr [cited 2021 Apr 23]; Available from:  
437 <https://linkinghub.elsevier.com/retrieve/pii/S0163445321001912>
- 438 25. Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Rapid, point-of-  
439 care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection.  
440 *Cochrane Database Syst Rev*. 2020 Aug 26;2020(8).
- 441

Figure 1

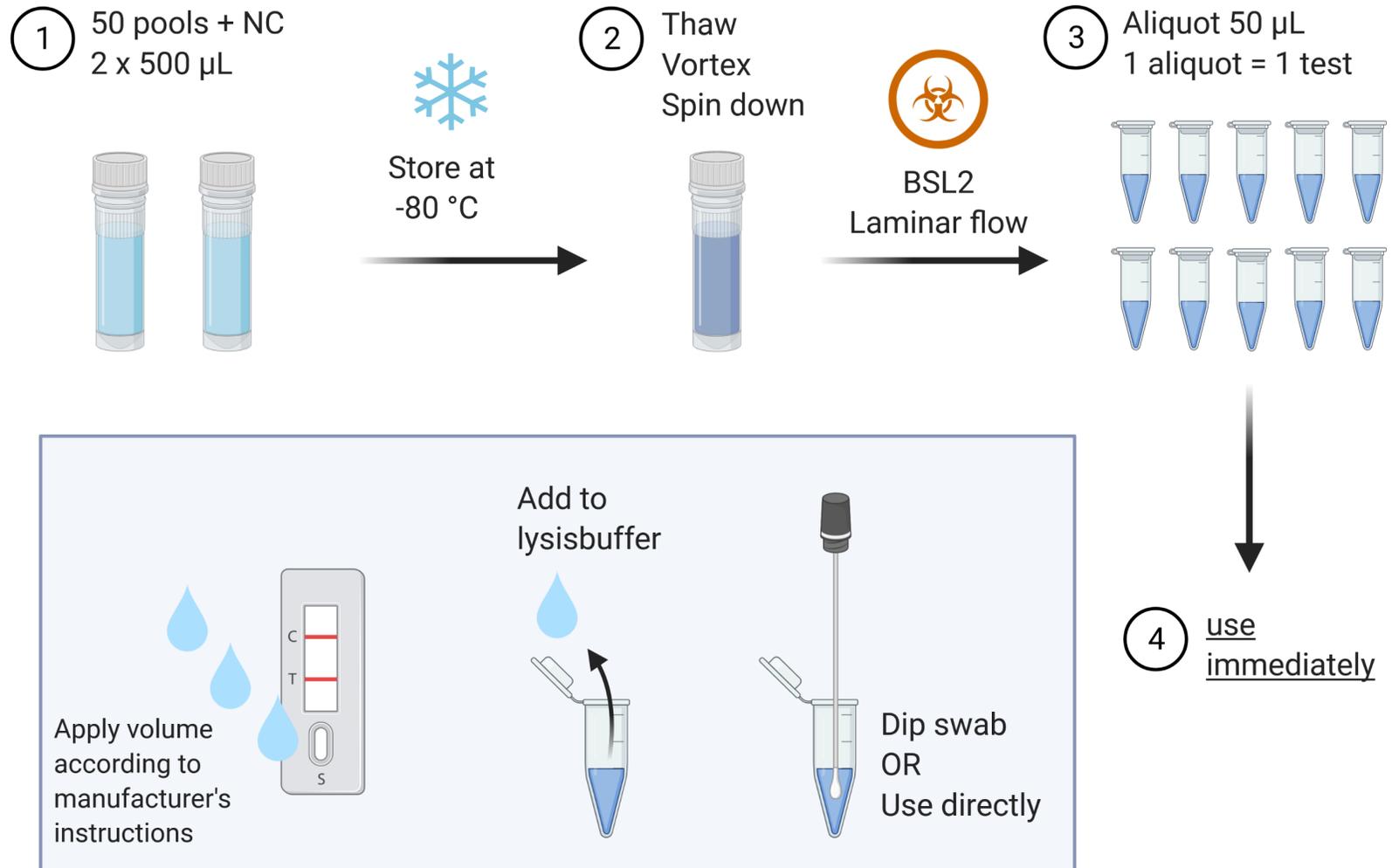


Figure 2

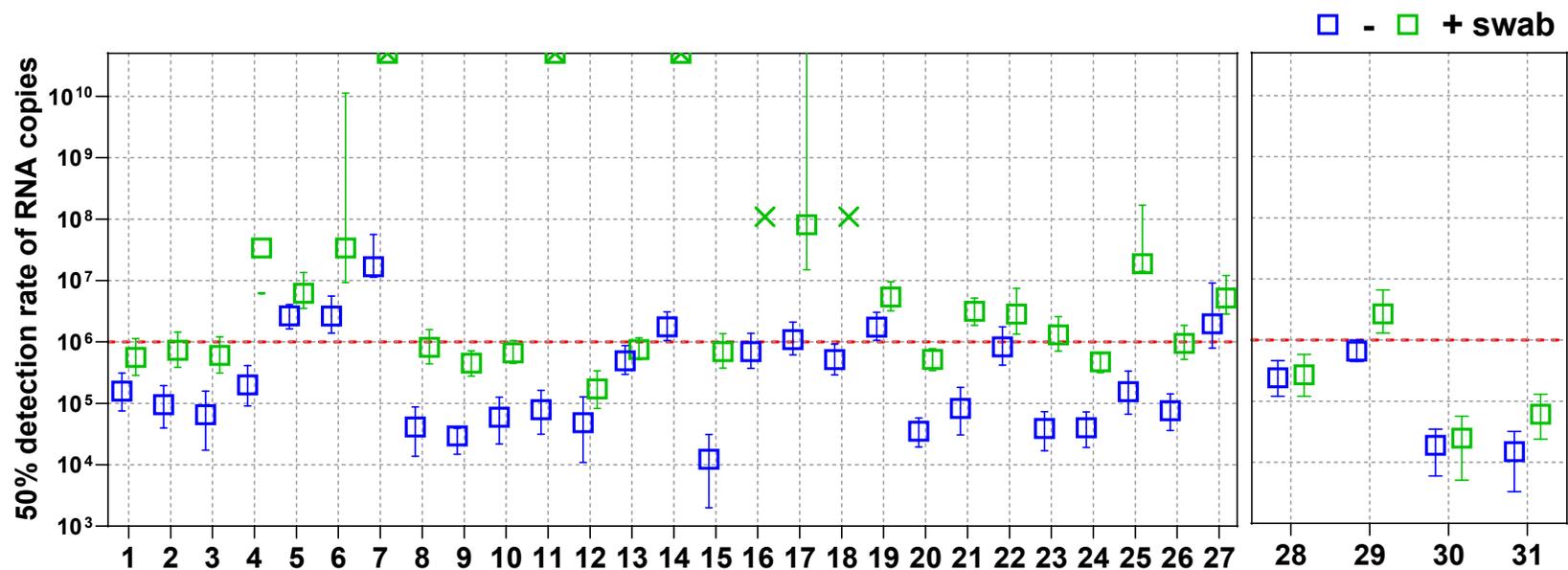


Figure 3

