

Evaluation of the accuracy, ease of use and limit of detection of novel, rapid, antigen- detecting point-of-care diagnostics for SARS-CoV-2

L.J. Krüger (1), M. Gaeddert (1), L. Köppel (1), L. E. Brümmer (1), C. Gottschalk (1), I.B. Miranda (2), P. Schnitzler (2), H.G. Kräusslich (2), A.K. Lindner (3.), O. Nikolai (3.), F.P. Mockenhaupt (3.), J. Seybold (4.), V.M. Corman (9,10), C. Drosten (9,10) N.R. Pollock (5.), A.I. Cubas-Atienzar (6.), K. Kontogianni (6.), A. Collins (6), A. H. Wright (6), B. Knorr (7.) A. Welker (7.), M. de Vos (8), J.A. Sacks[#] (8), E.R. Adams[#] (6.), C.M. Denkinger[#] (1) for the study team[&]

1. Division of Clinical Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany

2. Virology, Heidelberg University Hospital, Heidelberg, Germany

3. Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health; Institute of Tropical Medicine and International Health, Berlin, Germany

4. Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health; Medical Directorate, Berlin, Germany

5. Department of Laboratory Medicine, Boston Children's Hospital, Boston, MA, USA

6. Centre for Drugs and Diagnostics Liverpool School of Tropical Medicine, Liverpool, UK

7. Department Public Health Rhein Neckar Region, Germany

8. Foundation for Innovative New Diagnostics, Geneva, Switzerland

9. Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health; Institute of Virology, Berlin, Germany

10. German Centre for Infection Research (DZIF), partner site Charité, Berlin, Germany

[#] Contributed equally

*To whom correspondence should be addressed: Claudia.denkinger@uni-heidelberg.de, Tel.: [06221 56-22999](tel:062215622999)

[&]Members of the study team are detailed at the end of the paper

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract

Background: Reliable point-of-care (POC) diagnostics not requiring laboratory infrastructure could be a game changer in the COVID-19 pandemic, particularly in the Global South. We assessed performance, limit of detection and ease-of-use of three antigen-detecting, rapid POC tests (Ag-RDT) for SARS-CoV-2.

Methods: This prospective, multi-centre diagnostic accuracy study recruited participants suspected to have SARS-CoV-2 in Germany and the UK. Paired nasopharyngeal swabs (NP) or NP and/or oropharyngeal swabs (OP) were collected from participants (one for clinical RT-PCR and one for Ag-RDT). Performance of each of three Ag-RDTs was compared to RT-PCR overall, and according to predefined subcategories e.g. cycle threshold (CT)-value, days from symptoms onset, etc. In addition, limited verification of the analytical limit-of-detection (LOD) was determined. To understand the usability a System Usability Scale (SUS) questionnaire and ease-of-use (EoU) assessment were performed.

Results: Between April 17th and August 25th, 2020, 2417 participants were enrolled, with 70 (3.0%) testing positive by RT-PCR. The best-performing test (SD Biosensor, Inc. STANDARD Q) was 76.6% (95% Confidence Interval (CI) 62.8-86.4) sensitive and 99.3% (CI 98.6-99.6) specific. A sub-analysis showed all samples with RT-PCR CT-values <25 were detectable by STANDARD Q. The test was considered easy-to-use (SUS 86/100) and suitable for POC. Bioeasy and Coris showed specificity of 93.1% (CI 91.0%-94.8%) and 95.8% (CI 93.4%-97.4%), respectively, not meeting the predefined target of ≥98%.

Conclusion: There is large variability in performance of Ag-RDT with SD Biosensor showing promise. Given the usability at POC, this test is likely to have impact despite imperfect sensitivity; however further research and modelling are needed.

Introduction

As of 23rd of September 2020, *SARS-CoV-2* has affected the entire globe with over 30 million confirmed infections worldwide, leading to almost one million fatalities. ¹ Already early in the pandemic, the World Health Organization (WHO) highlighted the critical importance of testing for epidemic control. ² Reverse transcription polymerase chain reaction (RT-PCR) is the current gold standard for the detection of *SARS-CoV-2*. ³ Specialized laboratories for RT-PCR testing are available in high-and low-resource settings. ⁴ However, global shortages of test reagents and limited infrastructure and human resource capacity, have resulted in unavailability of testing or long turnaround times, especially in the Global South. To combat transmission, fast, easy-to-use, accurate, and point-of-care (POC) diagnostic tools are needed on a large scale.

Detection of pathogen-derived antigens on lateral flow assays is a proven concept with excellent performance, e.g. for malaria in blood. ⁵ This method enables rapid, instrument-free, and low-cost POC testing. In respiratory samples, however (e.g. for diagnosis of influenza), performance has been variable. ⁶

Since the onset of the pandemic, numerous companies have been working on bringing a rapid *SARS-CoV-2* antigen-detection test (Ag-RDT) to the market. ⁷ Interest and funding support are substantial ⁸, and the USA, India and Japan, among other countries have approved their use. ⁹⁻¹¹ WHO has recently released interim guidance that the use of Ag-RDTs that meet at least 80% sensitivity and 97% specificity may be considered to diagnose active *SARS-CoV-2* infection “where nucleic acid amplification technology (NAAT) is unavailable or where prolonged turnaround times preclude clinical testing.”. ^{12,13} Few peer-reviewed studies evaluating these novel assays have been published with only one small study, to our knowledge, performed independently of the manufacturer. ^{14,15}

Here we present a manufacturer-independent validation of three novel, antigen-detecting, rapid POC diagnostics for *SARS-CoV-2*. Specifically, we performed a verification of analytical sensitivity, a large multi-centre clinical accuracy study, and an ease-of-use (EoU) assessment. The work was conducted in partnership with the Foundation of Innovative New Diagnostics (FIND) and has informed a WHO recommendation. ¹⁶ FIND is the WHO collaborating centre for Coronavirus disease 2019 (COVID-19) diagnostics.

Methods

Ethic statement

The study protocol was approved by the ethical review committee at the Heidelberg University Hospital for the study sites in Heidelberg and Berlin (Registration number S-180/2020), and by the NHS review board, IRAS number 282147, for the study in Liverpool, UK. Each participant provided written informed consent.

Tests evaluated

Three novel rapid antigen-based diagnostics were assessed:

(a) Bioeasy 2019-nCoV Ag Fluorescence Rapid Test Kit (Time-Resolved Fluorescence).¹⁷ (Shenzhen Bioeasy Biotechnology Co. Ltd., Guangdong Province, China; henceforth called Bioeasy), (b) COVID-19 Ag Respi-Strip¹⁸ (Coris Bioconcept, Gembloux, Belgium; henceforth called Coris) and (c) STANDARD Q COVID-19 Ag Test¹⁹ (SD Biosensor, Inc. Gyeonggi-do, Korea; henceforth called SD Biosensor).

All assays utilize the lateral flow assay principle for the detection of viral antigens. Both Coris and SD Biosensor use colloidal gold conjugated to antibodies to enable a colour change and read-out with the naked eye, while Bioeasy uses fluorescent readout and thus requires the use of a proprietary reader device. Coris is designed as a dipstick, whereas Bioeasy and SD Biosensor use a cassette-based assay format. For Coris and SD Biosensor the manufacturers' instructions for use (IFU) were followed; except for Bioeasy, the IFU was followed except the developer requested for pipettes to be used to transfer adequate quantities of liquid (in the IFU no pipette is needed and a nozzle is provided).

¹⁷ Bioeasy and SD Biosensor include swabs within their test kits for sample collection, while for Coris, a commercially available swab was used: the eSwab system from Copan Diagnostics (Murrieta, CA, USA), which contains 1ml of Amies solution and was approved for use by Coris. A prespecified quantity of proprietary buffer solution was applied to the lateral flow cassette for Bioeasy and SD Biosensor and into the test tube with the dipstick for Coris as per the respective IFU. The readout was done within the recommended time for each Ag-RDT (10 minutes for Bioeasy, 15 minutes for Coris and 15 to 30 minutes for SD Biosensor).

Analytical study

A limited verification of the analytical sensitivity (lower limit of detection) was performed using standardized, quantified viral cultures.

The SARS-CoV-2 isolate REMRQ0001/Human/2020/Liverpool was propagated in Vero E6 cells (C1008; African green monkey kidney cells). Cells were maintained in Dulbecco's minimal essential medium (DMEM) containing 10% foetal bovine serum (FBS) and 0.05 mg/mL gentamycin at 37°C with 5% CO₂ and FBS concentration was reduced to 4% for viral propagation.

Ten-fold serial dilutions of virus stock starting from 1.0 x 10⁶ Plaque forming units (PFU)/ml were made using culture media as a diluent (DMEM + 2% FBS + 0.05 mg/ml gentamycin). Two-fold dilutions were made between 1.0 x 10⁴ to 1.0 x 10². Lateral flow tests were performed in triplicate. For SD Biosensor and Bioeasy, the swab included in each individual kit was immersed into the virus dilutions and spiked (soaked). Subsequently, the test was performed using the spiked swab following the manufacturer's instructions. For Coris, the eSwab from Copan was spiked with serial dilutions of virus and then added to the 1 ml of Amies solution contained in the swab tube. 100µl of each of the sample dilutions were added to the test tube with the dipstick and was read-out as per manufacturer's instructions.

Coris and SD Biosensor tests were read by two operators, each blinded to the result of the other. A third operator read any discrepant tests and final result was agreed (2/3). The Bioeasy result is a digital readout, and the first operator recorded the result, which was checked by the second operator, and printed out directly from the machine.

Clinical diagnostic accuracy study

Study design and participants

The primary objective of the clinical multicentre diagnostic accuracy study was to estimate the sensitivity and specificity of a single antigen test compared to a single PCR test as the reference method.

Participant recruitment and sample collection was conducted at three sites: (1) A drive-in testing station in Heidelberg, Germany; (2) a clinical ambulatory testing facility in Berlin, Germany and (3) secondary care facilities at Liverpool University Hospital Foundation Trust, Royal and Aintree (LUHFT), Liverpool, United Kingdom. Participants screened for the study were adults (age ≥18 years) identified as at risk for SARS CoV-2 infection according to the local department of public health based on exposure to a SARS CoV-2 positive person, suggestive symptoms, or travel to a high risk area. Participants were excluded if they had previously been diagnosed with SARS CoV-2 or if their command of either English or German was insufficient to give informed consent for participation.

Study procedures

Antigen-detecting testing

After obtaining written consent, participants underwent a first swab for testing with RT-PCR, which was taken either as (i) a nasopharyngeal (NP) swab or (ii) oropharyngeal (OP) swab (if clinically contra-indicated to do a NP swab (e.g. because of risk of bleeding)) in Heidelberg or (iii) combined as NP/OP swab in Berlin and Liverpool as per institutional recommendations (with OP done first, followed by NP with the same swab). Subsequently, participants underwent collection of the same type of swab again for Ag-RDT.

Ag-RDTs were performed at the sample collection site in Heidelberg and Berlin. In Liverpool samples were transported on ice to a category 3 facility and then processed. For on-site testing, separate areas were maintained for clean material (without sample) and 'unclean' areas of benches (material with sample) to avoid cross-contamination. Disinfection was performed and gloves exchanged after each sample. Negative control testing with buffer without sample was performed regularly to capture possible contamination. For the Ag-RDTs with visual read-out (Coris and SD Biosensor), the results were interpreted by two operators, each blinded to the result of the other. If a discrepant result was obtained, both operators re-read the result and agreed on a final result. Invalid results were repeated once using the remaining buffer according to the respective IFUs.

RT-PCR testing

The swab (eSwab, Copan) for RT-PCR testing was sent to the referral laboratory in the provided Amies solution. The RT-PCR assays were performed according to routine procedures at the referral laboratory. In Heidelberg, the *SARS CoV-2* assay from TibMolbiol (Berlin, Germany), the Allplex *SARS-CoV-2* Assay from Seegene (Seoul, South Korea) or the Abbott (Illinois, US) RealTime 2019-nCoV assay was performed. In Berlin the Roche Cobas *SARS CoV-2* assay (Pleasanton, CA United States) on the Cobas® 6800 or 8800 system or the *SARS CoV-2* assay from TibMolbiol (Berlin, Germany) was performed. In Liverpool, the Genesig® Real-Time Coronavirus COVID-19 PCR assay (Genesig, UK) was performed. Assays were calibrated with the E-gene assay as described before.^{3,19} Only samples that showed a signal above the threshold in the relevant RT-PCR target regions for each assay were considered to be positive. Staff performing the Ag-RDTs were blinded to results of RT-PCR tests and vice versa.

After completion of testing, the participants in Heidelberg were contacted for an interview focusing on symptoms and co-morbidities via telephone or email according to the indicated preference on site. In Berlin and Liverpool participants were interviewed directly on-site after obtaining the consent (questionnaire available in the supplement material Section (B)).

All data were captured in a dedicated database utilizing Research Electronic Data Capture (REDCap™, www.prorect-redcap.org) hosted at the Heidelberg University Hospital or Liverpool School of Tropical Medicine (LSTM). REDCap™ is a secure, web-based application, which provides audit trails for tracking data manipulation and export procedures. The protocol is provided in the supplement material Section (C).

System Usability Scale and Ease of Use Assessment

To understand the usability of the three Ag-RDTs, a standardized System Usability Scale (SUS)²⁰ questionnaire and a ease-of-use assessment (EoU), specifically developed for the study, were performed. The questionnaire can be found in the supplement material Section (D). The laboratory personnel from each study site (at least 3 operators per test) were invited to complete the surveys separately for each assay. The complete EoU survey can be found in the supplement material Section (D). For the interpretation of the SUS, it can be said that a SUS score above 68 would be considered above average and anything below the score of 68 is below average.²⁰

Statistical Analysis

In the analytical assessment, the limit of detection (LOD) was determined as the lowest concentration for which all three replicates scored positive.

In the clinical accuracy study, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the Ag-RDTs with 95% confidence intervals (CIs) were estimated by comparing the Ag-RDT result to the gold standard, RT-PCR, on the same day from the same participant (statistical analysis plan available upon request). A sub-analysis was performed by CT-value.²¹⁻²³ Furthermore, a conversion of CT-values for RT-PCR tests into viral-load was performed by calibrating the test systems using quantified specific in vitro-transcribed RNA.²⁴

A sample size of 1000 participants per Ag-RDT under evaluation was targeted but given the variable prevalence throughout the pandemic this target was left flexible. Interim analyses were performed at predefined sample sizes (25%, 50%) and an evaluation of an Ag-RDT was stopped if the predefined performance criterion for specificity ($\geq 98\%$, considering the upper bound of the confidence interval) was not met.²⁵ Invalid Ag-RDT results were reported separately, whereas participants with invalid RT-PCR results were excluded from the analysis. Inter-operator variability was assessed for tests read with the naked eye. The analysis was conducted with the statistical software R (R Foundation for Statistical Computing, Vienna, Austria).

For the usability assessment, the SUS score was calculated taking the mean of all answers given for each test. For the EoU assessment, verbal answers were transferred into numerical values. Taking the mean across values, all answers given for each test were summarized into a heat map. Both SUS and Ease of Use assessment were analysed with Microsoft Excel.

Role of the funding source

The study was supported by FIND, Heidelberg University Hospital and Charité – University Hospital internal funds. Pfizer funded the clinical team in Liverpool, UK. FIND provided input on the study design, and data analysis in collaboration with the rest of the study team. The other external funders of the study had no role in study design, data collection, or data analysis. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Analytical sensitivity and exclusivity

SD Biosensor was positive in all replicates at 5.0×10^3 PFU/ml, and only one out of three replicates at 1.0×10^3 PFU/ml, which is approximately ten-fold higher than the supplier-reported LOD of $3.06 \times 10^{2.2}$ TCID₅₀/ml as per ATCC.org.¹⁹ The Bioeasy assay was positive on all three replicates at 5.0×10^3 PFU/ml (approximately 7.14×10^3 TCID₅₀/ml) and only one out of three replicates at 2.5×10^3 PFU/ml (approximately 3.57×10^3 TCID₅₀/ml). Coris was positive on all three replicates at 1.0×10^4 PFU/ml (approximately 1.43×10^4 TCID₅₀/ml), which is two-fold higher than the supplier-reported LOD of 7.14×10^3 TCID₅₀/ml.²⁶ Coris was positive on only one of three replicates at 5.0×10^3 PFU/ml (approximately 7.14×10^3 TCID₅₀/ml).

Clinical diagnostic accuracy study

Between April 17th and August 25th 2020, 3160 participants with presumed SARS-CoV-2 infections were invited to participate in the study. In total, 2417 met inclusion criteria, provided consent and were included in the analysis (Figure 1). Of those, 1263 participants were tested using SD Biosensor from July 20th to July 31st 2020 in Heidelberg (N=334), from June 3rd to July 31st 2020 (N=910) in Berlin, and May 20th to June 12th 2020 (N=19) in Liverpool. Separately, 729 participants were tested using Bioeasy, 506 of whom were enrolled in Heidelberg between April 17th and May 3rd 2020,

and 223 in Berlin between May 14th and June 3rd 2020. 425 enrolled participants were tested using Coris: 286 in Heidelberg from 11th to 25th May 2020; 105 in Berlin from 19th to 25th August 2020; and 34 in Liverpool from May 12th to June 12th 2020. The variance in sample size is explained by the fact that both the Coris and Bioeasy evaluations were aborted after interim analysis indicated that specificity did not meet predefined performance criteria, and after a work-up for reasons for the false-positive results was completed.

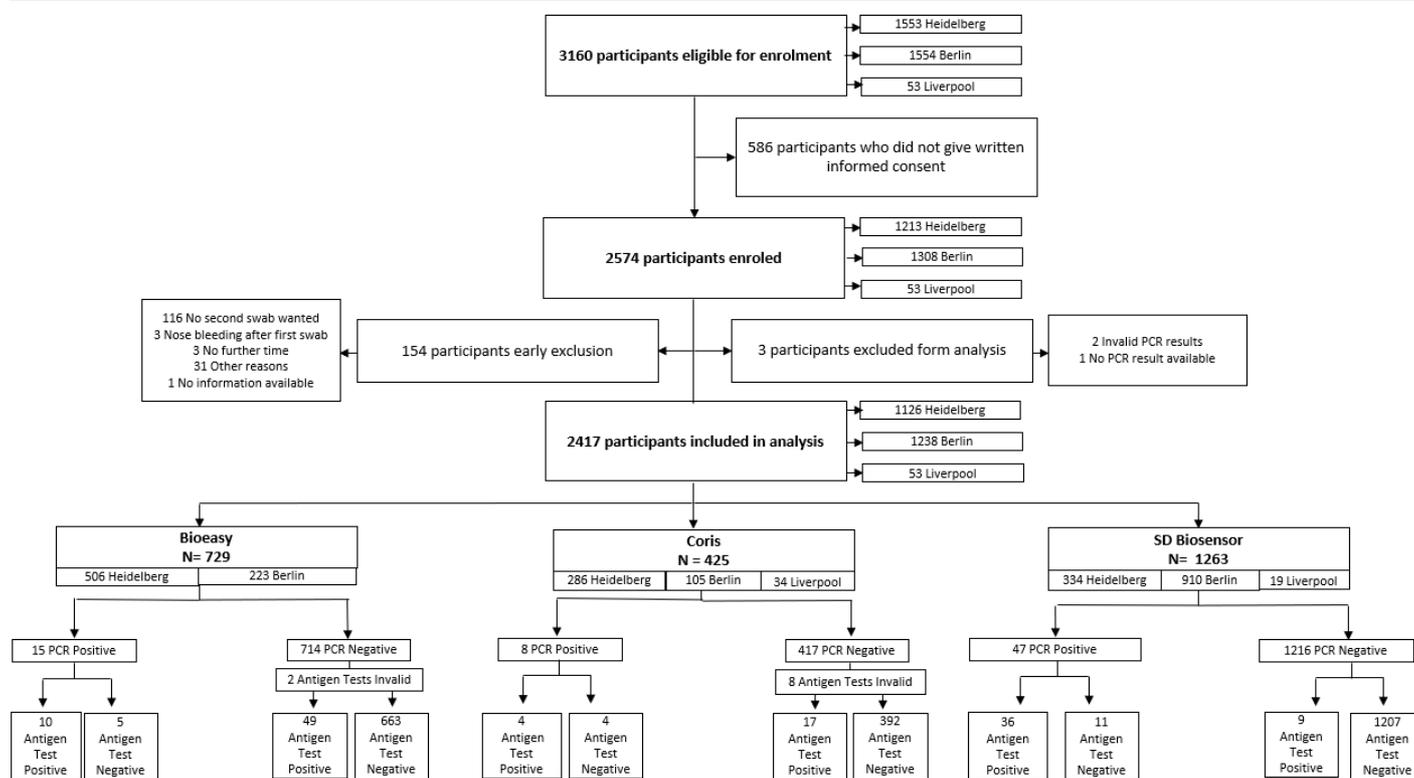


Figure 1. Study Flow

We present clinical and demographic characteristics overall and by Ag-RDT in Table 1 (Supplement Tables 2, 3 and 4 show study characteristics by study site). The average age of participants presenting for testing was 40.4 years (Standard Deviation [SD] 14.3) with 52.8% being female and 35.9% having comorbidities. On the day of testing, 80.7% of participants had one or more symptoms consistent with COVID-19 (questionnaire in the supplement material Section (B)), with most (1868 (79.3%)) reporting being only mildly ill and only 6 (0.3%) stating that they are completely bedridden due to illness (a more detailed list of symptoms and comorbidities is provided in the supplement Table 5). Average duration of symptoms at the time of swab collection/testing was 5.0 days (SD 9.6 days). Of the 2407 analysed participants, 70 (2.9%) tested positive by RT-PCR, of which 30 (42.9%) had CT values of <25. A description of symptoms for all participants tested positive by RT-PCR is provided in supplement Table 5.

Table 1. Study population characteristics

		Overall	SD Biosensor	Coris	Bioeasy
Age in years Information available on N=2405	Mean (SD)	40·4 (14·3)	37·6 (12·7)	44·9 (15·4)	42·7 (14·9)
Gender Information available on N=2361*	Female N(%)	1246 (52·8%)	630 (50·2%)	248 (60·3%)	368 (52·8%)
BMI >25 Information available on N=2290	N(%)	884 (38·6%)	451 (36·9%)	181 (47·6%)	252 (36·6%)
Comorbidities Information available on N=2407	N(%)	863 (35·9%)	361 (28·6%)	198 (47·5%)	304 (41·8%)
Symptoms on Testing Day[#] Information available on N=2355	N(%)	1901 (80·7%)	1054 (84·4%)	283 (68·9%)	564 (81·2%)
Duration of symptoms from day of testing in days[#] Information available on N=1830	Mean (SD)	5·0 (9·6)	3·7 (5·6)	6·2 (14·0)	7·0 (12·2)
Previous test negative Information available on N=1928	Yes N(%)	236 (12·2%)	125 (12·5%)	38 (12·6%)	73 (11·7%)
	None N(%)	1692 (87·8%)	878 (87·5%)	264 (87·4%)	550 (88·3%)
RT-PCR positives	N	70	47	8	15

* Excluding one diverse gender; # Differences between subpopulations are not statistically significant

The best performing test among the three was SD Biosensor with a sensitivity of 76·6% (36 out of 47 positives detected; CI 62·8-86·4) and specificity of 99·3% (CI 98·6-99·6) (Table 2). The sensitivity was 100% (CI 82·4-100) for the 18 samples with CT-values <25, while for those with CT ≥25, only 18 out of 29 scored positive (62·1%; CI 44·0-77·3) (Table 3). Corresponding viral load to CT values of RT-PCR positive results are provided in the supplement Table 6. A sub-analysis of sensitivity of SD Biosensor by sampling strategy suggested that a NP swab alone might be less effective than a combined NP/OP swab, however confidence intervals were overlapping (Table 4). No difference was seen by duration of symptoms (one vs two weeks). A trend was observed for increased sensitivity by disease severity, although again confidence intervals were overlapping (58·8% [36·0-78·4], 85·7% [65·4-95·0], 100% [51·0-100] for normal activity, mild and severe illness, respectively). No invalid results occurred. Two test results were excluded due to handling issues (read-out after time limit). Inter-operator reliability was near perfect with a kappa of 0·99. Overall, the usability was considered very favourable, with a SUS score of 86/100, and the only challenges flagged by the EoU were occasional difficulties when interpreting the test result (i.e. intensity of band; Figure 2).

Bioeasy had a sensitivity of 66·7% (10 out of 15 detected; CI 41·7-84·9) and specificity of 93·1% (49 false-positives; CI 91·0-94·8) (Table 2.). The sensitivity was 88·9% (8/9 detected; CI 56·5-99·4) for samples with CT-values <25, while for

those with CT ≥ 25 only 2 out of 6 tested positive (33.3%; CI 9.7-70.0). Further sub-analyses were not performed for sensitivity because of the limited number of positive samples. Two invalid results occurred in 729 tests (0.3%). Overall, the usability was considered favourably with a SUS score of 73/100. The EoU identified a high number of test execution steps (including precision pipetting) which resulted in challenges when performing multiple tests at the same time possibly hindering the test's wide-spread use (Figure 2).

Coris showed a sensitivity of 50% (4 out of 8; CI 21.5-78.5) and specificity of 95.8% (CI 93.4-97.4). Two out of the three samples with CT < 25 were detected, while for the five specimens with CT ≥ 25 , sensitivity was 40%. Further sub-analyses were not performed for sensitivity because of the limited number of positive samples. The low specificity determined the stop of the study as the upper margin of 95% CI was below 98%. Inter-operator reliability was very good with a kappa of 0.95. Invalid tests occurred on 8 occasions (1.9%) either due to the control line not being visible or due to smearing of the readout window. Four results were excluded due to handling issues (read-out after time limit). Overall, the usability was considered unfavourable with a SUS score of 48/100 and the EoU pointed towards challenges due to inconsistent test result interpretation (often only very faint lines visible) and deficiencies in both the test kit quality and design (Figure 2).

An investigation of the false-positives was performed for all assays by (1) testing negative buffer (without swab) interspersed with patient samples; (2) repeating RT-PCR from Ag-RDT buffer/swab sample for a subset of false-positive Ag-RDT samples (where feasible from available quantity of buffer); (3) having the sampling performed by the same person for Ag-RDT and PCR test (to limit variability in sampling); and (4) contacting participants again and retesting with Ag-RDT and RT-PCR if they had persistent symptoms or COVID-19 cases in their surroundings. None of this work-up yielded any evidence of cross-contamination, limited RT-PCR sensitivity or sampling error, which makes cross-reactivity most likely as the reason for the false-positives (target of cross-reactivity unknown).

Table 2. Antigen test performance overall

Overall	Ag-Test positive/ PCR positive (TP)	Ag-Test negative/PCR positive (FN)	Ag-Test positive/ PCR negative (FP)	Ag-Test negative/PCR negative (TN)	Sensitivity (95% CI)	Specificity (95% CI)
Results overall						

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .

SD Biosensor	1263	36 (2.9%)	11 (0.9%)	9 (0.7%)	1207 (95.6%)	76.6% (62.8-86.4)	99.3% (98.6-99.6)
Bioeasy	727	10 (1.4%)	5 (0.7%)	49 (6.7%)	663 (91.2%)	66.7% (41.7-84.8)	93.1% (91.0-94.8)
Coris	417	4 (1.0%)	4 (1.0%)	17 (4.1%)	392 (94.0%)	50% (21.5-78.5)	95.8% (93.4-97.4)

Table 3. Subanalysis for antigen test performance in different CT-value categories for only RT-PCR positive results

	Overall	Ag-Test positive	Ag-Test negative	Sensitivity (95% CI)
CT value PCR <25				
SD Biosensor	18 (60%)	18 (100%)	0 (0%)	100% (82.4-100)
Bioeasy	9 (30%)	8 (88.9%)	1 (11.1%)	88.9% (56.5-99.4)
Coris	3 (10%)	2 (66.7%)	1 (33.3%)	66.7% (20.8-98.3)
CT value PCR >=25				
SD Biosensor	29 (69%)	18 (62.1%)	11 (37.9%)	62.1% (44.0-77.3)
Bioeasy	8 (19%)	2 (25%)	6 (75%)	33.33% (9.7-70.0)
Coris	5 (12%)	2 (40%)	3 (60%)	40% (11.8-76.9)

Table 4. Subgroup Analysis for SD Biosensor, Inc. STANDARD Q

	Overall	Ag-Test positive/ PCR positive	Ag-Test negative/PCR positive	Ag-Test positive/ PCR negative	Ag-Test negative/PCR negative	Sensitivity (% , 95% CI)	Specificity (% , 95% CI)
Sampling strategy							
NP swab Heidelberg	299 (23·7%)	4	3	6	286	57·1% (25·0-84·1)	97·9% (95·6-99·1)
OP swab Heidelberg	35 (2·8%)	0	0	0	35	NA (no positive PCR results)	100% (90·1-100)
NP/OP swab* Berlin	907 (71·8%)	31	8	3	865	79·5% (64·5-89·2)	99·7% (99·0-99·9)
NP/OP swab Liverpool	19 (1·5%)	0	0	0	19	NA (no positive PCR results)	100% (83·2-100)
Symptom duration – Information available for N=1018[†]							
0-7 days Overall	949 (93·2%)	28	7	7	907	80·0% (64·1-90·0)	99·2% (98·4-99·6)
8-14 days Overall	58 (5·7%)	4	0	0	54	100 % (51·0-100)	100% (93·0-100)
> 14 days Overall	11 (1·1%)	0	0	0	11	NA (no positive PCR results)	100% (74·1-100)
Disease severity in 3. categories – Information available for N=1193							
Normal activity possible Overall	495 (41·5%)	10	7	4	474	58·8% (36·0-78·4)	99·2% (97·9-99·7)
Light restrictions and Able to walk Overall	673 (56·4%)	18	3	5	647	85·7% (65·4-95·0)	99·2% (98·2-99·7)
Limited self-sufficiency and Completely need of care Overall	25 (2·1%)	4	0	0	21	100% (51·0-100)	100% (84·5-100)

* 3 participants in Berlin did not receive the combined swab; [†] No information from N=8 PCR positives out of N=47

Overview & SUS	Bioeasy		On a system usability scale, the test scored 73 out of 100 points. Difficulties could occur mainly due to a high number of test execution steps (including pipetting), long processing time and challenges when performing batch testing.
	Coris		On a system usability scale, the test scored 48 out of 100 points. Difficulties could occur mainly due to inconsistent test result interpretation, quality deficiencies and the number of necessary steps.
	SD Biosensor Q		On a system usability scale, the test scored 86 out of 100 points. Difficulties could occur mainly due to occasional issues with interpreting the test result and complications with storage and stability.

Detailed Ease of Use Assessment

		Bioeasy	Coris	SD Biosensor Q			Bioeasy	Coris	SD Biosensor Q			Bioeasy	Coris	SD Biosensor Q	
Quality of the test's hardware	External paper box of the kit				Ease of test execution	Check expiry date				Ease of result interpretation	Visibility of the control (C) band	NA			
	Assay diluent tube					Remove the test cartridge/device/strip from the pouch					Visibility of the test (T) band	NA			
	Filter cap	NA	NA			Label the test cartridge/device/strip with patient identifier					Read-out from Reader		NA	NA	
	Swab for specimen collection		NA			Label the assay diluent tube with the patient identifier					Interpretation of the test result				
	Test cartridge/device/strip					Open the assay diluent tube by removing the seal					Reader's ease of use		NA	NA	
	Test cartridge/device/strip - packing pouch					Transfer of buffer into diluent tube			NA		Overall satisfaction with the reader		NA	NA	
	Reader		NA	NA		Insert the swab into the tube					Fields of application	Batch testing possible			
	Overall quality of kit components					Swab extraction procedure						Usability in high income countries			
	Size of the cartridge/device/strip					Perform extraction procedure consistently						Range of health care levels (high income countries)			
	Space for labeling on the front					Maintain cleanliness of ancillary devices						Usability in low and middle income (LMI) countries			
	Size of the well to add sample mix		NA			Transferring sample onto device					Range of health care levels (LMI countries)				
	Size of reading window	NA				Transferring exact quantity into the sample well		NA			Storage conditions	Stability of test			
Overall design of the test device				Trouble shooting				Stability of control material							
Test Prep	Instructions for use				Logical sequence of steps			Storage temperature							
	Required training				Overall difficulty of steps			Overall storage conditions							
	Required test runs				Procedure time	Pre-analytic time									
				Analytic time											
				8-hour throughput											

Legend	
NA	Not applicable
	Satisfactory
	Average
	Dissatisfactory

Caption: More explanation on the detailed EoU assessment is available in supplement material section (E).

Figure 2. System Usability Score and Ease of Use assessment results

Discussion

Results of our multi-centre diagnostic accuracy study and analytical assessment show that the SD Biosensor, Inc. STANDARD Q Ag-RDT COVID-19 Ag test can reach excellent sensitivity in high-viral load infections, and good performance in the cohort overall 76·6% (CI 62·8-86·4) with near-perfect specificity (99·3%; CI 98·6-99·6). These results hold promise for the performance of antigen-based diagnosis and the control of the pandemic, particularly in low-resource settings.

SARS CoV-2 is a virus with a low dispersion factor, which indicates that a small number of infected individuals are responsible for a large amount of the transmission - the so-called ‘super spreaders’.²⁷ A super-spreader is likely to both have a high viral load, and the occasion to transmit to many people due to aggregation of people in confined spaces.²⁸ With a sensitivity of 100% for the SD Biosensor in infected persons with a high viral load, this Ag-RDT is likely to identify highly contagious individuals, which would more substantially decrease transmission than what would be expected from its imperfect overall sensitivity alone. Furthermore, the rapid turn-around time is likely to result in more rapid isolation of cases and effective contact tracing. The ease-of-use and POC-suitability can enable more frequent testing as compared to RT-PCR and thus, despite its limited sensitivity, the test is expected to have impact on pandemic control based on a recent modelling study.²⁹

Considering the test in low prevalence settings (prevalence 1%), a negative test would have a negative predictive value (NPV) of 99·8%. And even at a prevalence of 10%, 97·4% of negatively tested persons would not be infected. Such an NPV would be considered acceptable outside of high-risk transmission settings (e.g. in a hospital). The positive predictive value (PPV) would be 53% at 1% prevalence, ideally requiring RT-PCR confirmation. At 10% prevalence, the PPV increases to 92%. This suggests that several use-cases could be considered with substantial public health impact, but further implementation research is needed to better detail how the tests can be used to the highest benefit.

Furthermore, stability only up to 30° Celsius and required storage conditions are a limitation for the use in the Global South, and should be considered carefully in such settings.

Our study also shows wide variability in performance between Ag-RDTs with respect to performance and ease-of-use, indicating the importance of manufacturer-independent validations. While our findings largely corroborate the performance reported in the IFU of SD Biosensor, for Coris and Bioeasy, they differed substantially in respect to specificity. For SD Biosensor, the lower overall performance compared to the results in the IFU could relate to the fact that for some patients in our study OP swabs were used. While this is not recommended by the manufacturer, the use of OP swabs when NP is contraindicated, reflects a more clinically realistic scenario. NP/OP swabs were used based on institutional recommendations in Berlin and Liverpool. While more proteases are expected in the saliva that can degrade antigen, there should not be an impact on the sensitivity of Ag-RDTs in our study, as swabs were processed immediately. Another aspect that could explain a lower sensitivity is the population tested in the study. The population tested was likely representative of the spectrum of COVID-19, with less ill patients likely to have lower viral load at diagnosis and more likely Ag-RDT negative. A higher clinical sensitivity would be expected when only hospitalized patients were tested.³⁰ For Bioeasy, our point estimate of specificity falls within the confidence interval of specificity of the only other independently published study, which had few RT-PCR negative cases.¹⁴

Overall, our study had several strengths. The study confirmed the accuracy of the tests by two measures: analytical and clinical accuracy. The population enrolled for testing was representative of the pandemic observed in Europe with a broad spectrum of age group (initially older age group and with ongoing duration of the pandemic an increasingly younger age group presenting for testing) and a spread of clinical presentations (from asymptomatic with high-risk contacts to severely ill). We enrolled in two countries and although enrolment numbers in Liverpool were low, the results confirmed findings in Germany. Due to the wide-spread testing available particularly in Germany, the

population tested is expected to be a representative spectrum of disease. The tests were performed at POC at two sites (Berlin and Heidelberg) thus mimicking the real-world challenges of POC testing. The comprehensive ease-of-use assessment with a standardized SUS-tool and a questionnaire, developed specifically for the study, captured the differences between the tests and highlighted important points for operationalization of the tests.

However, the study also has several limitations. First, the prevalence of positive cases decreased substantially over the course of the study from 10% among tested individuals in early April as documented by the local health department in Heidelberg, Germany, to 3% over the course of the study. In addition, our study assessed a preselected population based on the regulations and criteria from the local health department. Due to the nature of the pandemic, the selection criteria and patient population changed throughout the study as visible from the percentage of patients with symptoms and co-morbidities evaluated with the three Ag-RDTs, although differences between the study populations were not significant. In addition, our study did use OP swabs for a subset of patients as discussed above. Furthermore, the sub-analysis by CT-values has its limitations, as RT-PCR methods varied across sites with different genome targets, PCR instruments and reagents. This might cause misleading comparisons of CT-values between sites.²³ Therefore, we only performed an analysis of dichotomized CT-values and considered <25 as high viral load and ≥ 25 as low viral load. In addition, we translated CT-values into viral load, and determined the variability of the applied *SARS-CoV-2* assays in testing of standardized dilutions of virus to be limited, as shown before.^{21,22}

Conclusion

Our results suggest that, though there is substantial variability between tests, at least one antigen-detection POC test meets WHO targets and is now recommended by the WHO.¹⁶ Given the fast turn-around-time and ease-of-use at POC, as well as the high sensitivity in patients with significant viral load, the test is likely to have impact despite imperfect overall sensitivity; however, further

implementation research, economic evaluation and modelling are needed. With the agreement of Roche and SD Biosensor, manufacturing scale-up and supply-chain solutions should guarantee availability of the test. ³¹

References

1. World Health Organisation. WHO Coronavirus Disease (COVID-19) Dashboard 2020/8/26, 10:41am CEST 2020. <https://covid19.who.int/> (accessed 26.08.20).
2. World Health Organisation. Director-General's opening remarks at the media briefing on COVID-19 – 16 March 2020 16.03.2020 2020 <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020> (accessed 26.08.2020)
3. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA* 2020; **323**(22): 2249-51.
4. Blairon L, Wilmet A, Beukinga I, Tre-Hardy M. Implementation of rapid SARS-CoV-2 antigenic testing in a laboratory without access to molecular methods: Experiences of a general hospital. *J Clin Virol* 2020; **129**: 104472.
5. Lalremruata A, Nguyen TT, McCall MBB, et al. Recombinase Polymerase Amplification and Lateral Flow Assay for Ultrasensitive Detection of Low-Density Plasmodium falciparum Infection from Controlled Human Malaria Infection Studies and Naturally Acquired Infections. *J Clin Microbiol* 2020; **58**(5).
6. Azar MM, Landry ML. Detection of Influenza A and B Viruses and Respiratory Syncytial Virus by Use of Clinical Laboratory Improvement Amendments of 1988 (CLIA)-Waived Point-of-Care Assays: a Paradigm Shift to Molecular Tests. *J Clin Microbiol* 2018; **56**(7).
7. Foundation of Innovative New Diagnostics. FIND SARS-COV-2 Diagnostics Pipeline 2020 <https://www.finddx.org/covid-19/pipeline/> (accessed 30.09.2020).
8. Unitaid. FIND and Unitaid launch call for expression of interest to accelerate availability and manufacturing scale-up of rapid diagnostic tests for COVID-19. Unitaid News-Blog 2020
9. Indian Council of Medical Research. List of rapid antigen kits 2020 https://www.icmr.gov.in/pdf/covid/kits/List_of_rapid_antigen_kits_15092020.pdf (accessed 30.09.2020).

10. U.S. Food and Drug Administration. Emergency Use Authorization. 2020
<https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization> (accessed 30.09.2020).
11. Pharmaceutical and Medical Device Agency. PMDA's Efforts to Combat COVID-19 2020
<https://www.pmda.go.jp/english/about-pmda/0002.html> (accessed 30.09.2020).
12. World Health Organisation. Advice on the use of point-of-care immunodiagnostic tests for COVID-19. 08.04.2020 2020 <https://www.who.int/news-room/commentaries/detail/advice-on-the-use-of-point-of-care-immunodiagnostic-tests-for-covid-19> (accessed 26.08.2020)
13. World Health Organisation. Antigen-detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays 2020 https://apps.who.int/iris/bitstream/handle/10665/334253/WHO-2019-nCoV-Antigen_Detection-2020.1-eng.pdf?sequence=1&isAllowed=y (accessed 30.09.2020).
14. Porte L, Legarraga P, Vollrath V, et al. Evaluation of a novel antigen-based rapid detection test for the diagnosis of SARS-CoV-2 in respiratory samples. *Int J Infect Dis* 2020; **99**: 328-33.
15. Mertens P, De Vos N, Martiny D, et al. Development and Potential Usefulness of the COVID-19 Ag Respi-Strip Diagnostic Assay in a Pandemic Context. *Front Med (Lausanne)* 2020; **7**: 225.
16. World Health Organisation. Global partnership to make available 120 million affordable, quality COVID-19 rapid tests for low- and middle-income countries. 2020
<https://www.who.int/news-room/detail/28-09-2020-global-partnership-to-make-available-120-million-affordable-quality-covid-19-rapid-tests-for-low--and-middle-income-countries> (accessed 29.09.2020).
17. BioEasy. Diagnostic Kit for 2019-Novel Coronavirus (2019-nCoV) Ag Test (Fluorescence Immunochromatographic Assay) 2020 <http://en.bioeasy.com/?p=1038> (accessed 26.08.2020)
18. Coris BioConcept. COVID-19 Ag Respi-Strip Detection test of SARS-CoV-2. 2020
<https://v3.globalcube.net/clients/beldico/content/medias/products/microbiology/sars-covid-19-detection-test-leaflet.pdf> (accessed 30.09.2020).

19. SD Biosensor. COVID-19 Ag STANDARDTM Q COVID-19 Ag Test 2020
<http://sdbiosensor.com/xe/product/7672> (accessed 30.09.2020).
20. Bangor A KP, Miller JT. . An Empirical Evaluation of the System Usability Scale. *Intl Journal of Human-Computer Interaction* 2008; **24**: 574-94. .
21. Matheussen V, Corman VM, Donoso Mantke O, et al. International external quality assessment for SARS-CoV-2 molecular detection and survey on clinical laboratory preparedness during the COVID-19 pandemic, April/May 2020. *Euro Surveill* 2020; **25**(27).
22. INSTAND Gesellschaft zur Förderung der Qualitätssicherung in medizinischen Laboratorien e.V. Kommentar zum Extra Ringversuch Gruppe 340 Virusgenom-Nachweis-SARS-CoV-2. 2020.
<https://www.instand-ev.de/System/rv-files/340%20DE%20SARS-CoV-2%20Genom%20April%202020%2020200502j.pdf> (accessed 30.09.2020).
23. Corman VM, Rasche A, Baronti C, et al. Assay optimization for molecular detection of Zika virus. *Bull World Health Organ* 2016; **94**(12): 880-92.
24. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020; **25**(3).
25. World Health Organisation. COVID-19 Target product profiles for priority diagnostics to support response to the COVID-19 pandemic v.0.1. 2020
<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> (accessed 23.09.2020).
26. ATCC. Converting TCID₅₀ to plaque forming units (PFU) Is it possible to determine from the TCID₅₀ how many plaque forming units to expect? 2020 https://www.lgcstandards-atcc.org/Global/FAQs/4/8/Converting_TCID50_to_plaque_forming_units_PFU-124.aspx?geo_country=gb (accessed 23.09.2020 2020).
27. Endo A, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res* 2020; **5**: 67.

28. Kupferschmidt K. Why do some COVID-19 patients infect many others, whereas most don't spread the virus at all? 2020 <https://www.sciencemag.org/news/2020/05/why-do-some-covid-19-patients-infect-many-others-whereas-most-don-t-spread-virus-all> (accessed 02.09.2020 2020).
29. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States. *JAMA Netw Open* 2020; **3**(7): e2016818.
30. Pujadas E, Chaudhry F, McBride R, et al. SARS-CoV-2 viral load predicts COVID-19 mortality. *Lancet Respir Med* 2020; **8**(9): e70.
31. Ltd FH-LR. Roche to launch SARS-CoV-2 Rapid Antigen Test in countries accepting CE mark, allowing fast triage decisions at point of care. 29.09.2020 2020.
<https://www.roche.com/media/releases/med-cor-2020-09-01b.htm> (accessed 29.09.2020 2020).