Public Health England

UNIVERSITY OF OXFORD

NHS
Oxford University Hospitals
NHS Foundation Trust

Evaluation of sensitivity and specificity of four commercially available SARS-CoV-2 antibody immunoassays

Public Health England, Porton Down

Nuffield Department of Medicine, University of Oxford

Oxford University Hospitals NHS Foundation Trust

July 2020

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, research, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Published July 2020
PHE publications
gateway number: GW-1386

PHE supports the UN
Sustainable Development Goals

Contents

Executive summary

Public Health England (Porton Down) and a clinical/research team at the University of Oxford and Oxford University Hospitals NHS Foundation Trust were commissioned by the Department of Health and Social Care (DHSC) to evaluate several commercial immunoassays for SARS-CoV-2 antibody detection available on the UK market.
An external appraisal of assay performance is highly desirable in order to determine performance metrics with precision, using a large, well-characterised sample set.

Over a three-week period in May-June 2020, we evaluated four SARS-CoV-2 antibody assays, namely Abbott's SARS-CoV-2 Immunoassay, DiaSorin's LIAISON® SARS-CoV-2 S1/S2 IgG, Roche's Elecsys® Anti-SARS-CoV-2, and Siemens' SARS-CoV-2 Total (COV2T) (referred to as Abbott, DiaSorin, Roche, and Siemens assays hereafter).

Assays were performed in line with the manufacturers' instructions and at the pre-specified thresholds for determining positive vs negative test results. For each test, we calculated sensitivity and specificity, to compare against the UK Medicines and Healthcare products Regulatory Agency (MHRA) Target Product Profile (TPP) for 'enzyme immunoassays' for SARS-CoV-2.

Sensitivity was evaluated on 536 positive samples from unique adult individuals with laboratory-confirmed SARS-CoV-2 infection at ≥20 days post-symptom onset; specificity was evaluated on 994 pre-pandemic (2015-2018) specimens from unique, healthy adult individuals.

Primary results were as follows:

Assay	**Sensitivity [95% CI]**	**Specificity [95% CI]**	**Appraisal against MHRA Target Product Profile (TPP)**
Abbott	92.7 (90.2, 94.8)	99.9 (99.4, 100)	Meets specificity criterion
DiaSorin	95.0 (92.8, 96.7)	98.6 (97.6, 99.2)	Meets specificity criterion
Roche	97.2 (95.4, 98.4)	99.8 (99.3, 100)	Meets specificity criterion
Siemens	98.1 (96.6, 99.1)	99.9 (99.4, 100)	Meets sensitivity and specificity criteria

We also undertook secondary analyses, highlighting that the Roche assay could meet the current MHRA TPP sensitivity criteria with an assay threshold adjustment (eg at a revised assay threshold of ≥0.128 the sensitivity would be 99.4 [95% CI: 98.4, 99.9] with a specificity of 98.1 [95% CI: 97.0, 98.8]).

Further, by optimising assay thresholds to achieve a specificity of ≥98% and extending the sample timeframe specification to ≥30 days post-symptom onset (in lieu of the current MHRA TPP specification of ≥20 days), all four assays would meet the sensitivity criteria.

1. Introduction

1.1. SARS-CoV-2 has emerged as a novel cause of human infection, causing a global pandemic in the first 6 months of 2020, with >8.2 million confirmed cases of infection and 443K deaths (1). Case ascertainment and testing have been critical to controlling the spread of infection, and in developing effective strategies to mitigate the public health and economic impact of this pathogen.

1.2. Laboratory testing for SARS-CoV-2 broadly takes two forms: firstly, direct detection of the presence of virus, by testing respiratory samples with real-time PCR (RT-PCR), and secondly, an assessment of the immunological response to infection, by using serology to determine the presence of antibody (2), and/or neutralisation assays to evaluate the capacity of antibodies to effectively target the virus. Neutralisation assays are time- and resource-intensive tests, and are currently limited in their capacity for rollout. Serological diagnosis has therefore been pursued, with the aim of detecting either specific types of SARS-CoV-2 antibody (IgM, IgG, IgA), or total antibody, supporting several aims:

- at a population level, determining exposure provides insight into spread in communities and healthcare settings, identification of risk groups, and supports tracking and modelling of infection over time
- at an individual level, antibodies are deemed likely to be a correlate of protection against future infection, and may therefore contribute to managing personal risk-assessments
- to support research and development, antibody measurement is a critical tool, particularly in providing quantification of antibody responses in vaccine trials

Antibodies to SARS-CoV-2 typically start to appear >5-7 days after infection (3, 4), and are therefore an unreliable marker for early acute infection. Importantly, it remains unclear what degree of immunity the presence of antibody confers, and how durable this might be.

1.3. To cope with the demand for serological diagnosis, several manufacturers have developed immunoassays that are compatible with current global laboratory infrastructures, including high-throughput analyzers. However, assembling appropriate and large sets of samples to thoroughly test the performance of these assays has been difficult within the very short time frames of assay development and release, and direct comparisons of platforms have been limited.

1.4. The UK Medicines and Healthcare products Regulatory Agency (MHRA) has recently released a 'Target Product Profile for enzyme immunoassays' (5) to support Pillar 3 of the UK testing strategy ("Mass-antibody testing to help determine if people have immunity to coronavirus" (6)), specifying:

- a clinical sensitivity of ≥98% (95% CI: 96-100%) in confirmed SARS-CoV-2 positive cases (defined by RT-PCR) ≥20 days after the appearance of first symptoms;
- a clinical specificity of ≥98% (95% CI: 96-100%) on samples collected >6 months before the first identified cases of SARS-CoV-2 infection

The sensitivity of a test characterises its capacity to identify known positives (ie infected individuals), and the specificity its capacity to identify known negatives (ie uninfected individuals). Of note, there is no clear gold standard against which to evaluate these antibody tests; PCR-positivity is a proxy for the expected presence of antibody, but negative antibody tests in PCR-positive individuals could either represent an issue with antibody test performance, or alternatively be explained by a failure to mount a measurable antibody response (eg in immunocompromise), or through a false-positive PCR test in individuals who have not genuinely had SARS-CoV-2 infection.

1.5. To directly evaluate and compare the sensitivity and specificity of four commercial immunoassays for SARS-CoV-2 antibody (Abbott, DiaSorin, Roche, Siemens; Table S1), we formed a collaboration between Public Health England -Porton Down, Oxford University Hospitals NHS Foundation Trust, and the University of Oxford. Using a large collection of serum/plasma samples from individuals with SARS-CoV-2 infection confirmed by RT-PCR, and a bank of known negative samples collected pre-pandemic, we ran the same samples across all four platforms in a 'head-to-head' evaluation.

2. Methods

2.1. We developed a written protocol for the head-to-head evaluation, which was shared with the four commercial companies and with DHSC at the outset. We provided manufacturers with the opportunity to raise questions and feedback on the protocol. The full data and analysis, including the protocol, will be submitted for publication on a pre-print server (publicly accessible) and in a peer-reviewed journal as soon as possible.

2.2. Samples for testing were collected from adults in the UK, in line with MHRA TPP specifications, namely a ‘known negative’ group of samples collected >6 months prior to the known appearance of SARS-CoV-2 (ie sample collection earlier than July 2019), and a ‘known positive’ group of samples collected from individuals with a previous positive SARS-CoV-2 RT-PCR nose/throat swab, with blood samples taken ≥20 days post-symptom onset. In total, 994 samples from 994 individuals were included in the sensitivity/specificity analyses as the ‘known negative’ cohort, and 536 samples from 536 individuals as the ‘known positive’ cohort. Sample selection and exclusions are shown in Figure 1; characteristics of included samples are presented in Table 1.

Figure 1. Sample collections and inclusions/exclusions. For de-duplication of samples by individual, the latest sample meeting the MHRA criteria (ie latest sample taken ≥20 days after symptom onset) was included. Table S2 summarises the partial results for the five samples that were of insufficient volume to run across all platforms.

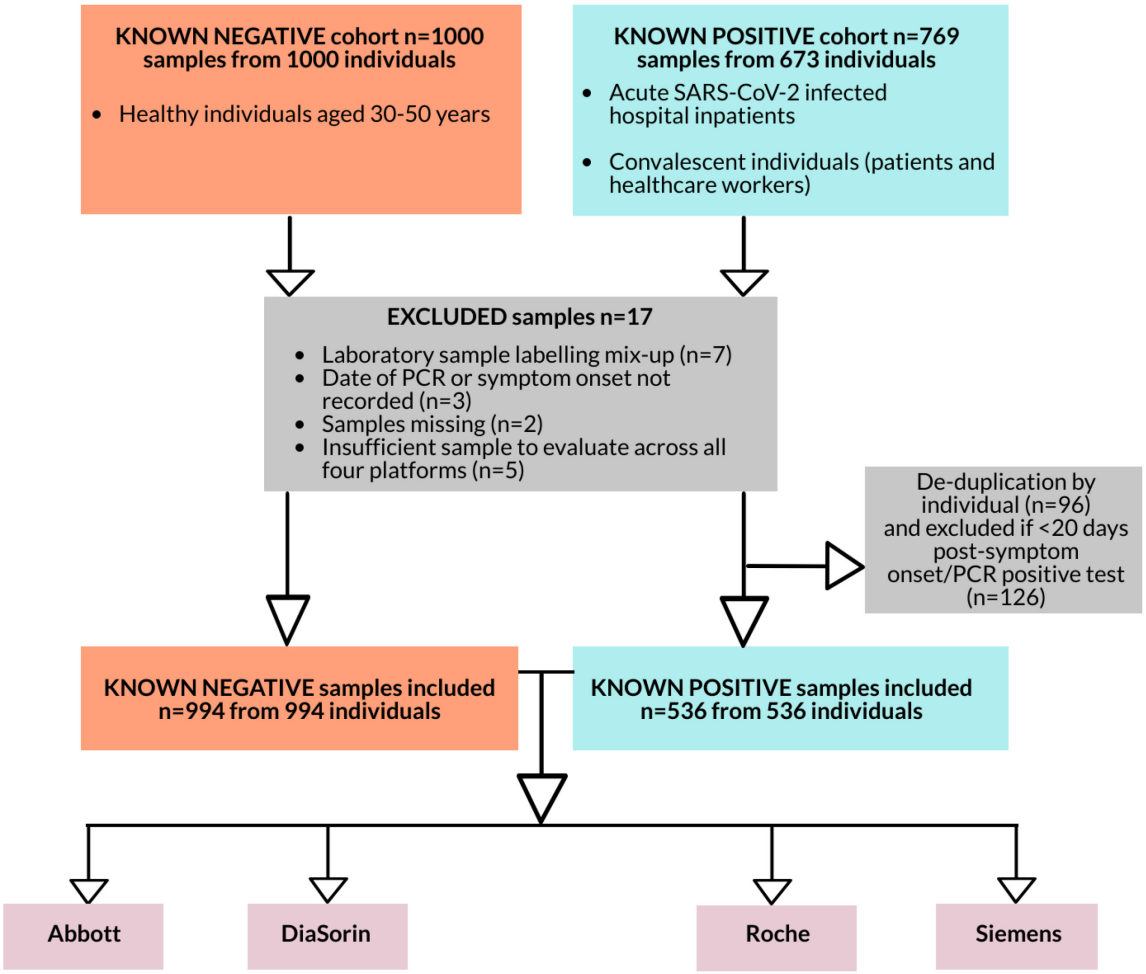


Table 1. Summary of serum samples used for head-to-head analysis of four commercial immunoassay platforms for the detection of SARS-CoV-2 antibodies

Group	Source	Number of samples	Days from symptom onset, median (IQR; min, max; number of samples)	Days from PCR-positive test, median (IQR; min, max; number of samples)
Known negative	Healthy individuals 30-50 years of age, collected between 2015-2018 in Oxfordshire (Oxford BioBank, www.oxfordbiobank.org.uk)	994	n/a	n/a
Known positive	Healthcare workers and patients ≥18 years of age at Oxford University Hospital NHS Foundation Trust, Oxfordshire, UK	158	37 (28-53; 20, 73; n=158)	27 (22-36; 3, 59; n=105)
Known positive	Volunteer plasma donors ≥18 years of age via NHS Blood and Transplant (NHSBT), across the UK	378	All samples ≥28 days post-symptom onset*	44 (40-49; 32, 82; n=378)

* Although specific data on time from symptoms is not available for this group, all donors had to be have been at least 28 days post-symptom onset to be eligible for sampling: see https://www.nhsbt.nhs.uk/plasma-trial/

2.3. Tests were performed in accordance with the manufacturer's instructions by trained laboratory staff in UK Accreditation Service (UKAS) accredited laboratories on appropriate analysers, and with the specified controls and calibrants, at the thresholds set by the manufacturer for testing in the UK (Table S1). The Abbott and DiaSorin tests were performed at the John Radcliffe Hospital Clinical Biochemistry and Microbiology laboratories in Oxford, and the Roche and Siemens tests at PHE Porton Down. For the DiaSorin assay, where the manufacturer specifies additional repeat testing in duplicate in the event that results fall within an equivocal zone (12.0 ≤ x <15.0 AU/mL), we were not able to perform repeat testing due to the limited quantity of sample available; these samples (n=9) were excluded from the final sensitivity/specificity calculations for the DiaSorin assay.

2.4. Data for each assay were collated and checked by an analysis group; the cleaned data were locked prior to analysis. Statistical analyses and data visualisations were performed in R (version 3.6.3). The analysis, results and draft report were reviewed by an external review group (Prof Janet Darbyshire, Emeritus Professor of Epidemiology at University College London, and Prof Sir David Spiegelhalter, Winton Professor of the Public Understanding of Risk, University of Cambridge).

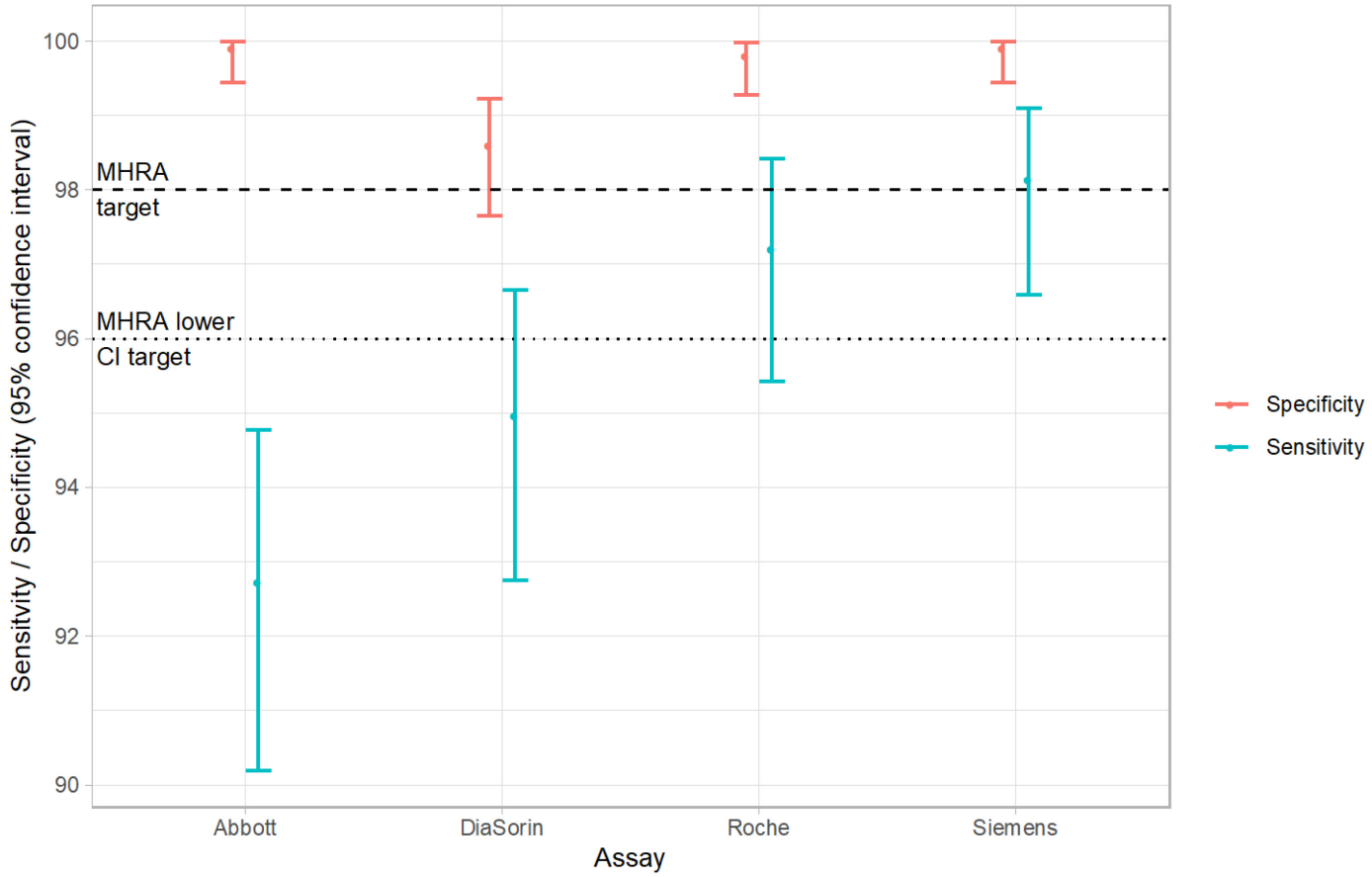
3. Results: Primary analysis

3.1. We calculated sensitivity and specificity and 95% confidence intervals (exact binomial method) for each assay, using the manufacturers' thresholds and in line with the current MHRA TPP criteria. The results are presented in Table 2 and Figure 2. The Siemens assay met the current MHRA TPP criteria for both sensitivity and specificity; the Abbott, DiaSorin and Roche assays met the current MHRA TPP criteria for specificity only.

Table 2. Sensitivity and specificity (95% confidence interval, CI) for each assay using the current MHRA TPP criteria: ≥20 days post-symptom onset in confirmed laboratory cases of SARS-CoV-2 for positive cases, and >6 months prior to the first known COVID-19 cases for negatives. Equivocal results were excluded from the calculation of sensitivity and specificity for the DiaSorin assay (n=9).

Assay	Number PCR-positive	Number detected	Number not detected	*Number equivocal	Sensitivity (95% CI)	Number pre-pandemic controls	Number detected	Number not detected	*Number equivocal	Specificity (95% CI)
Abbott	536	497	39	n/a	92.7 (90.2, 94.8)	994	1	993	n/a	99.9 (99.4, 100)
DiaSorin	536	509	20	7	95.0 (92.8, 96.7)	994	12	980	2	98.6 (97.6, 99.2)
Roche	536	521	15	n/a	97.2 (95.4, 98.4)	994	2	992	n/a	99.8 (99.3, 100)
Siemens	536	526	10	n/a	98.1 (96.6, 99.1)	994	1	993	n/a	99.9 (99.4, 100)

Figure 2. Sensitivity and specificity (95% confidence intervals) plotted for each assay using the current MHRA TPP criteria: ≥20 days post-symptom onset in confirmed laboratory cases of SARS-CoV-2 for positive cases, and >6 months prior to the first known COVID-19 cases for negatives. The MHRA TPP target performance is shown (dashed line) including the required lower bound of the 95% confidence interval (dotted line) for both sensitivity and specificity. Data are presented for 994 known negative samples and 536 known positive samples run on each assay; equivocal results were excluded from the calculation of sensitivity and specificity for the DiaSorin assay (n=9).



4. Results: Secondary analysis

4.1. The distribution of numerical values obtained for each sample from each assay is shown in Figure 3, illustrating the separation of sample values around the thresholds set by the manufacturers.

4.2. We used ROC curves to investigate the performance of each assay (Figure 4). ROC curves evaluate the trade-off between true positive rates (ie assay sensitivity) versus false positive rates (ie 1-specificity) at a given assay threshold. They can therefore be used to illustrate how assay threshold adjustment can impact on whether an assay achieves the MHRA criteria for both sensitivity and specificity. An assay threshold adjustment to ≥0.128 would have resulted in the Roche assay meeting both the current MHRA TPP sensitivity and specificity criteria (revised sensitivity: 99.4 [95% CI: 98.4, 99.9] revised specificity of 98.1 [95% CI: 97.0, 98.8]; Figure 4, Appendix Figure S1, Table S3). Sensitivity and specificity for each assay with revised thresholds chosen to ensure all assays met the target specificity ≥98.0% are shown in Appendix Figure S1.

Figure 3. Distribution of numerical results obtained for each assay, using samples defined according to the current MHRA TPP criteria. Assay thresholds (set by the manufacturers) are shown as dashed lines. For the purposes of plotting values on a log scale, values of zero were set to the lowest non-zero value and results of greater or less than the largest or smallest values were truncated to the largest and smallest values. Data are presented for 994 known negative samples and 536 known positive samples run on each assay.

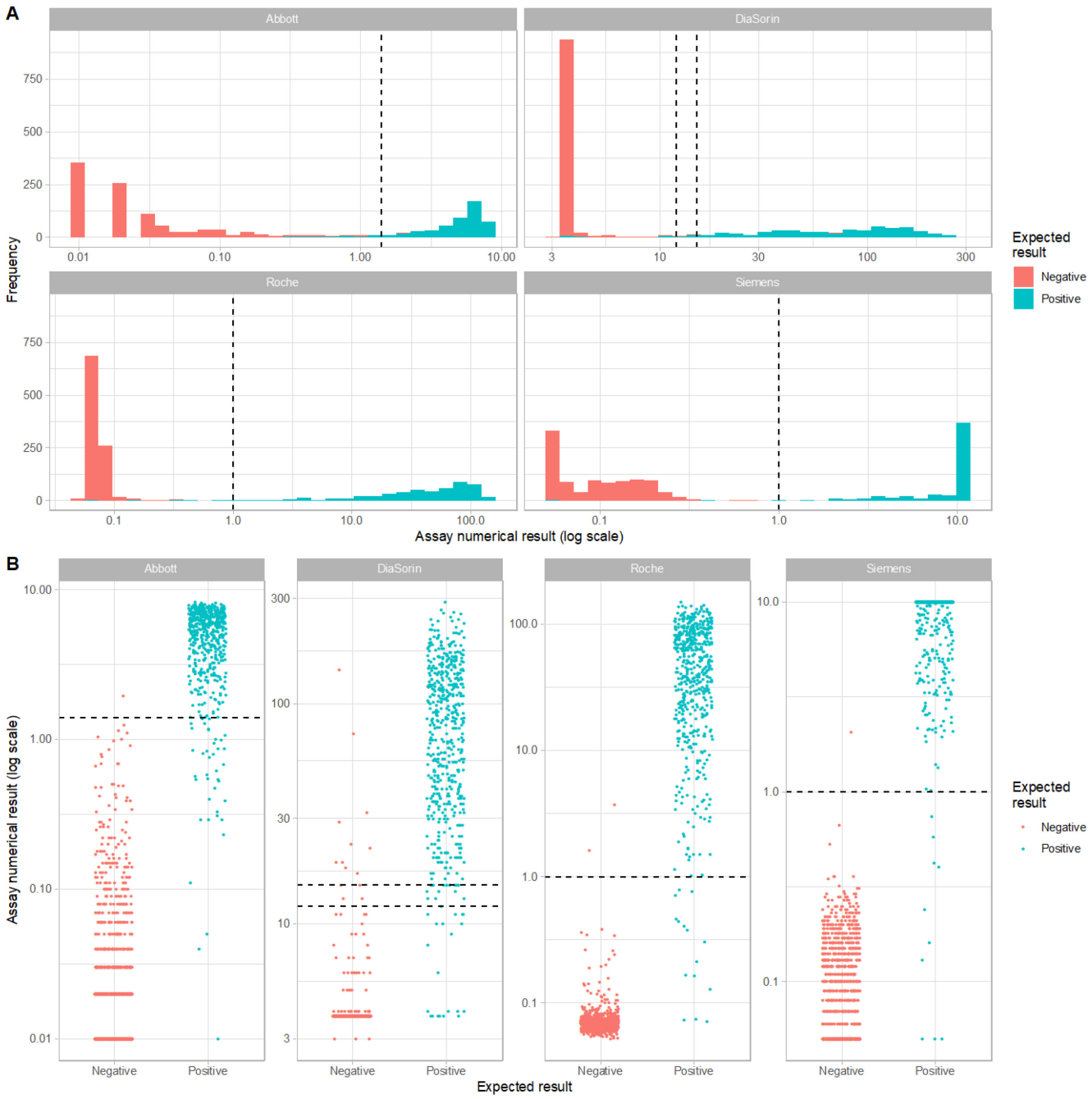
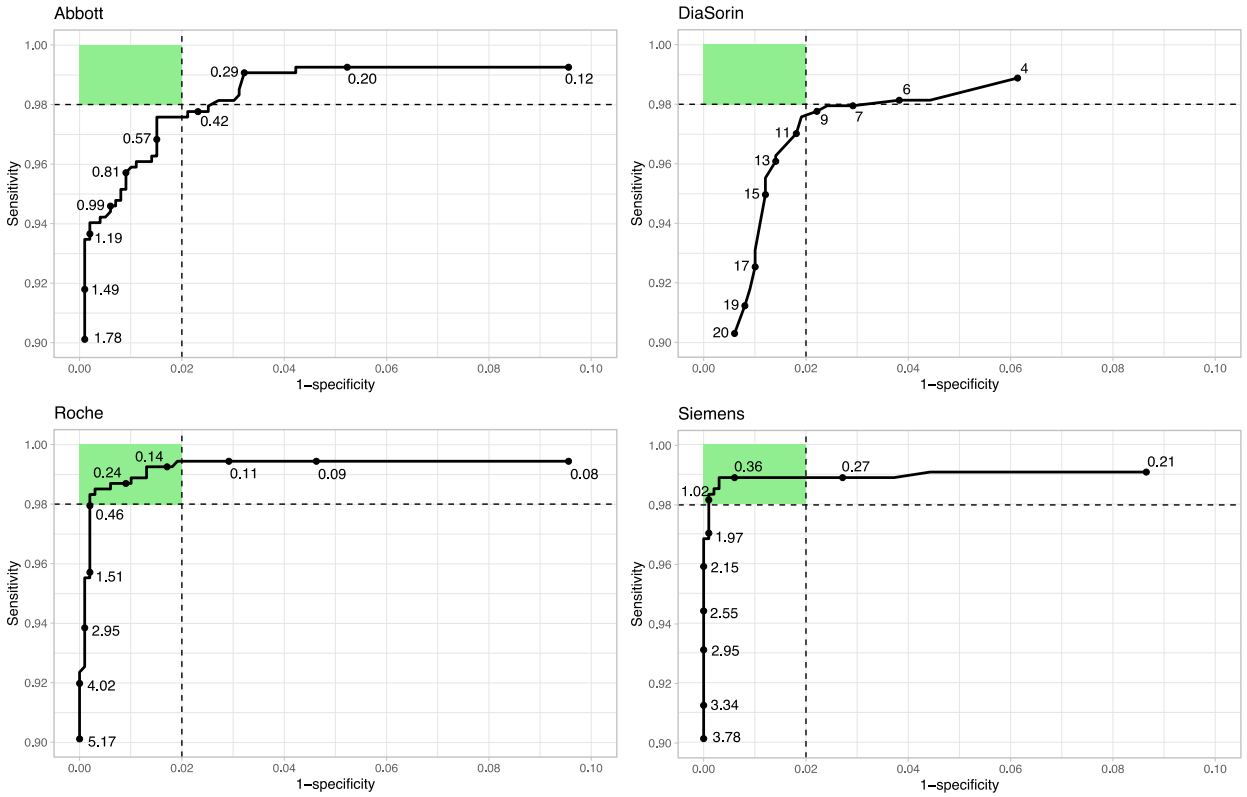


Figure 4. ROC curves for each assay at the current MHRA TPP specification of ≥20 days after the appearance of first symptoms. The green shaded area represents the MHRA TPP sensitivity and specificity of ≥98% and ≥98% respectively. Assay values associated with 10 exemplar points on the ROC curve are shown in each panel. Data are shown based on analysis of 536 samples in the positive category, and 994 pre-pandemic negative samples run on each assay.

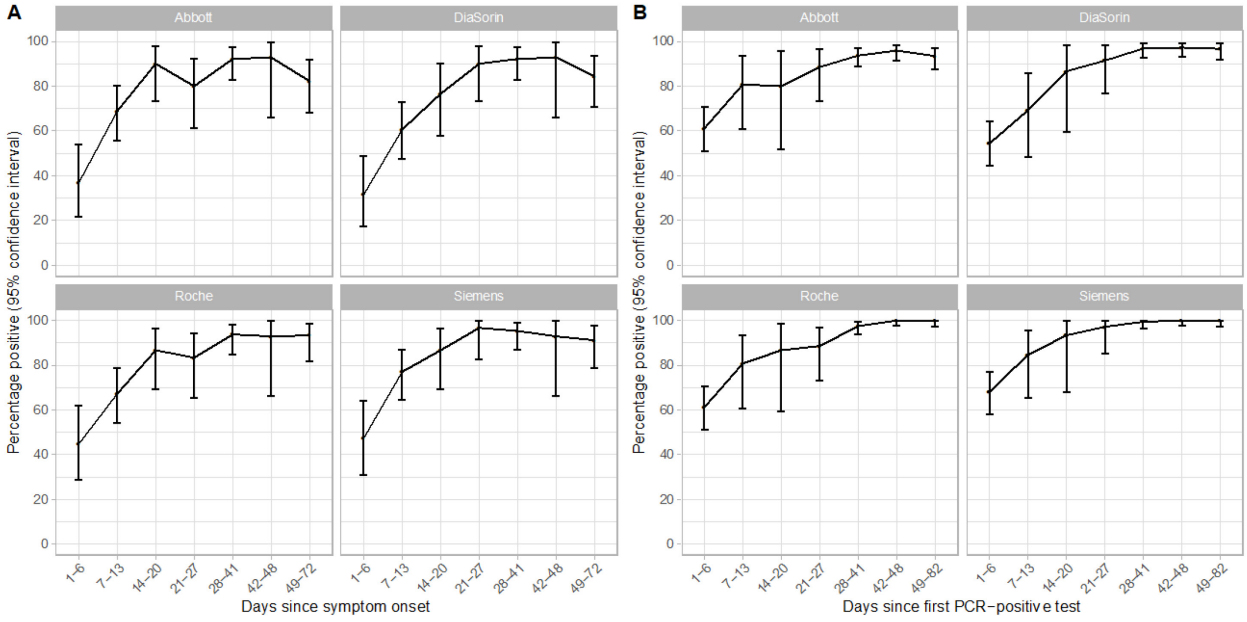


4.3. We also assessed the performance of each assay according to the time-point of sample collection. Based on samples collected ≥30 days after the appearance of first symptoms, the Siemens and Roche assays would meet the current MHRA TPP sensitivity and specificity criteria (Appendix Figure S2). ROC curves are shown for this revised time cutoff (Appendix Figure S3).

4.4. The sensitivity and specificity of each assay in samples obtained ≥30 days after symptom onset, using revised assay thresholds chosen to ensure all assays had specificity ≥98.0% are shown (Appendix Figure S4), demonstrating that with these revisions, all platforms would meet the TPP sensitivity and specificity criteria. Data are shown based on analysis of 490 samples assigned as positive, and 994 pre-pandemic negative samples run on each assay.

4.5. Antibody responses rose over the first 3-4 weeks from symptom onset (Figure 5). Antibody responses were sustained up to 73 days post symptom onset and up to 82 days post a positive PCR result.

Figure 5. Percentage of tests from SARS-CoV-2 RT-PCR-positive individuals positive over time by serology platform. Samples from <20 days from symptom onset, excluded from the main analysis, are included here. Panel A shows the percentage by time since symptom onset and panel B the percentage by the time since the individual's first positive RT-PCR test.



4.5. For 157 individuals on whom disease severity data were available, there was no evidence of a difference in antibody titre by disease severity (asymptomatic, mild, severe, critical/death) for any of the assays evaluated (Figure S5).

4.6. Finally, we undertook an evaluation of discordance between platforms (Table S4). Seven of the samples classified as 'known positive' based on MHRA criteria tested antibody-negative across all assays, highlighting the difficulties with the lack of a gold standard for these types of evaluations.

Although it is possible that this reflects a detection issue across all four assays, it is also conceivable that there was genuinely no antibody in these samples ie that these represented individuals who genuinely failed to mount an antibody response; that they had a false-positive RT-PCR test for SARS-CoV-2 and had not been infected; or that there was biological interferent in the sample affecting all assays.

5. Conclusions

5.1. Based on the exact current MHRA TPP criteria for evaluating sensitivity and specificity, all assays tested met the specificity requirement, but only the Siemens met both the sensitivity and specificity requirements (Table 1).

5.2. Secondary analyses demonstrate that assay thresholds could be re-evaluated to refine sensitivity/specificity trade-offs, and assay performance is optimised for samples taken ≥30 days after symptom onset.

6. Appendix

6.1. The appendix contains supplementary material to support the results of secondary analyses not included in the main text.

Figure **S1. Sensitivity and specificity (95% confidence intervals) plotted for each assay using the current MHRA TPP criteria with alternative assay thresholds to keep specificity ≥98%.** For each assay the lowest threshold that kept specificity ≥98% was chosen (Abbott = 0.49, DiaSorin = 10, Roche = 0.128, Siemens = 0.29). The MHRA TPP target performance is shown (dashed line) including the required lower bound of the 95% confidence interval (dotted line) for both sensitivity and specificity.

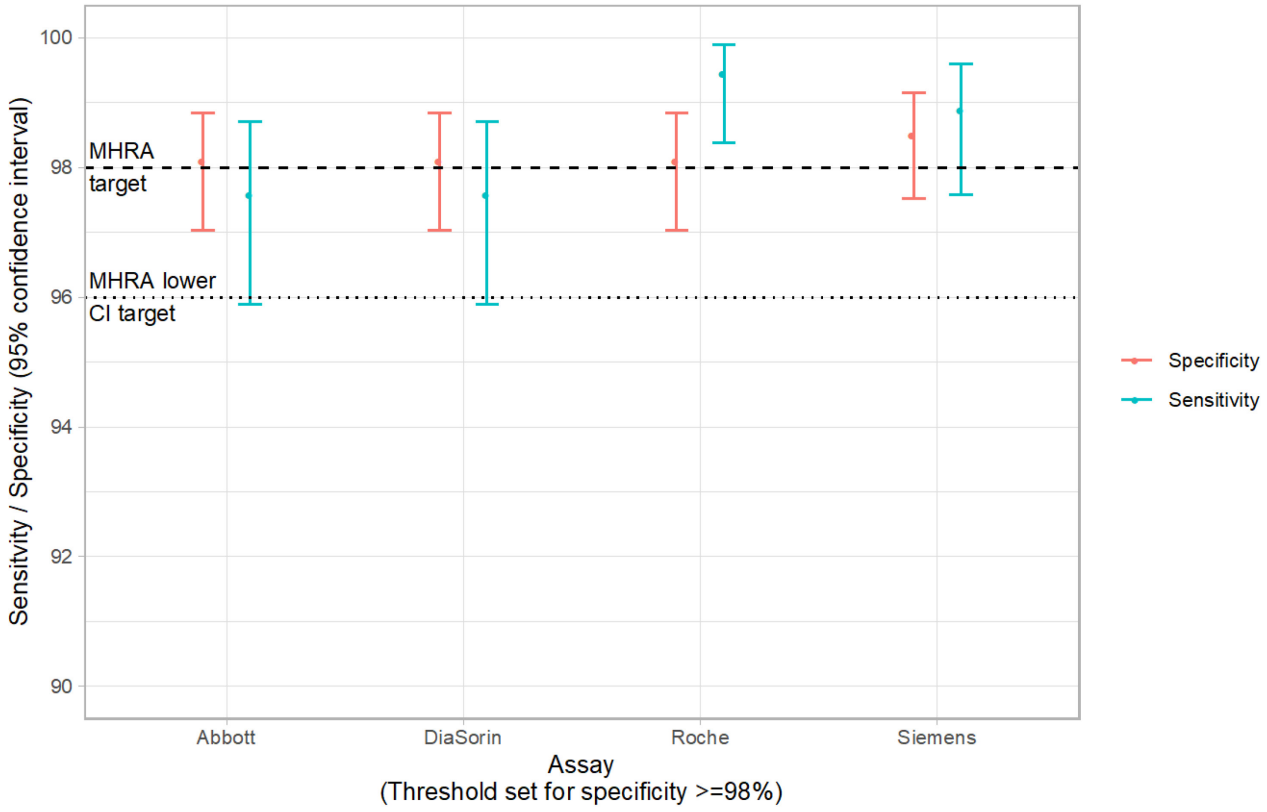


Figure S2. Sensitivity and specificity (95% confidence intervals) plotted for each assay using revised criteria. Sensitivity and specificity were defined using a revised criterion of ≥30 days post-symptom onset in confirmed laboratory cases of SARS-CoV-2 for positive cases. Equivocal results were excluded from the calculation of sensitivity and specificity for the DiaSorin platform (n=9). The MHRA TPP target performance is shown (dashed line) including the required lower bound of the 95% confidence interval (dotted line) for both sensitivity and specificity.

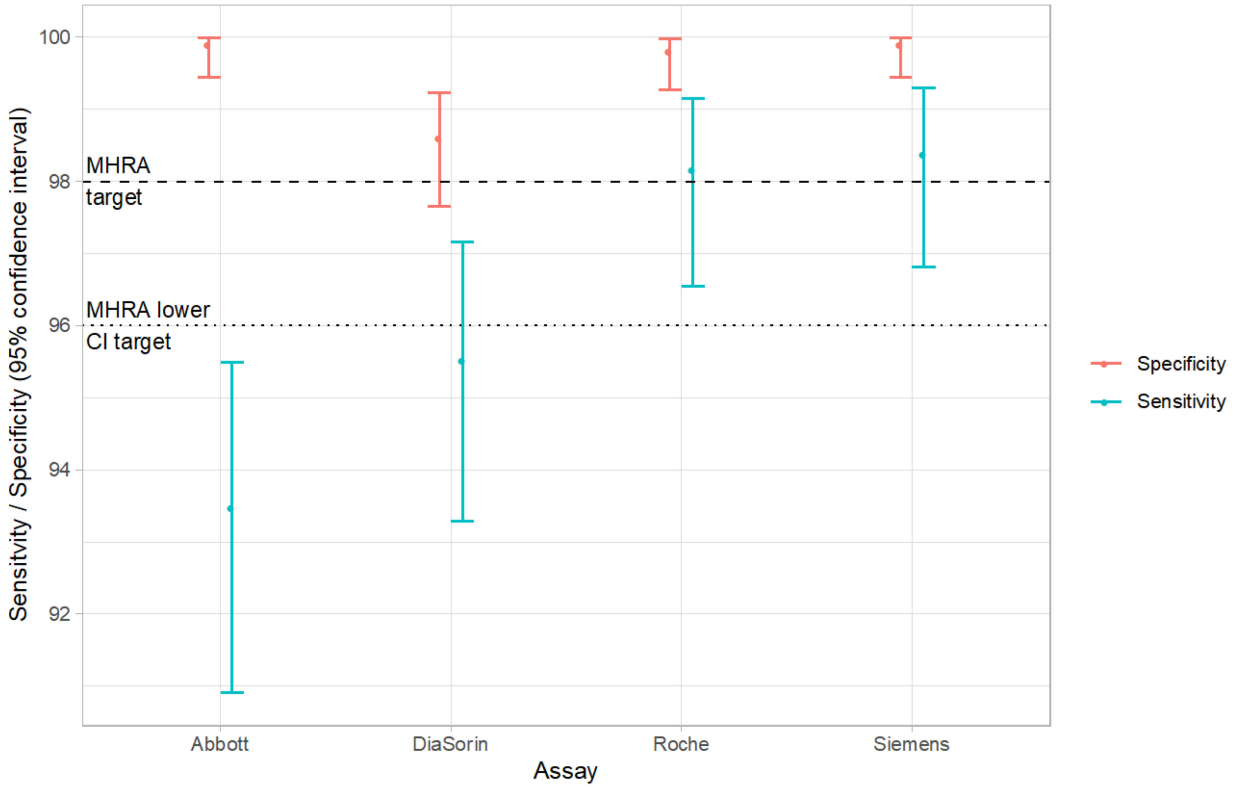


Figure S3. ROC curves for each assay at the current MHRA TPP specification revised ≥30 days after the appearance of first symptoms. The green shaded area represents the MHRA TPP sensitivity and specificity of ≥98% and ≥98% respectively. Assay values associated with 10 exemplar points on the ROC curve are shown in each panel. Data are shown based on analysis of 490 samples in the positive category, and 994 pre-pandemic negative samples run on each assay.

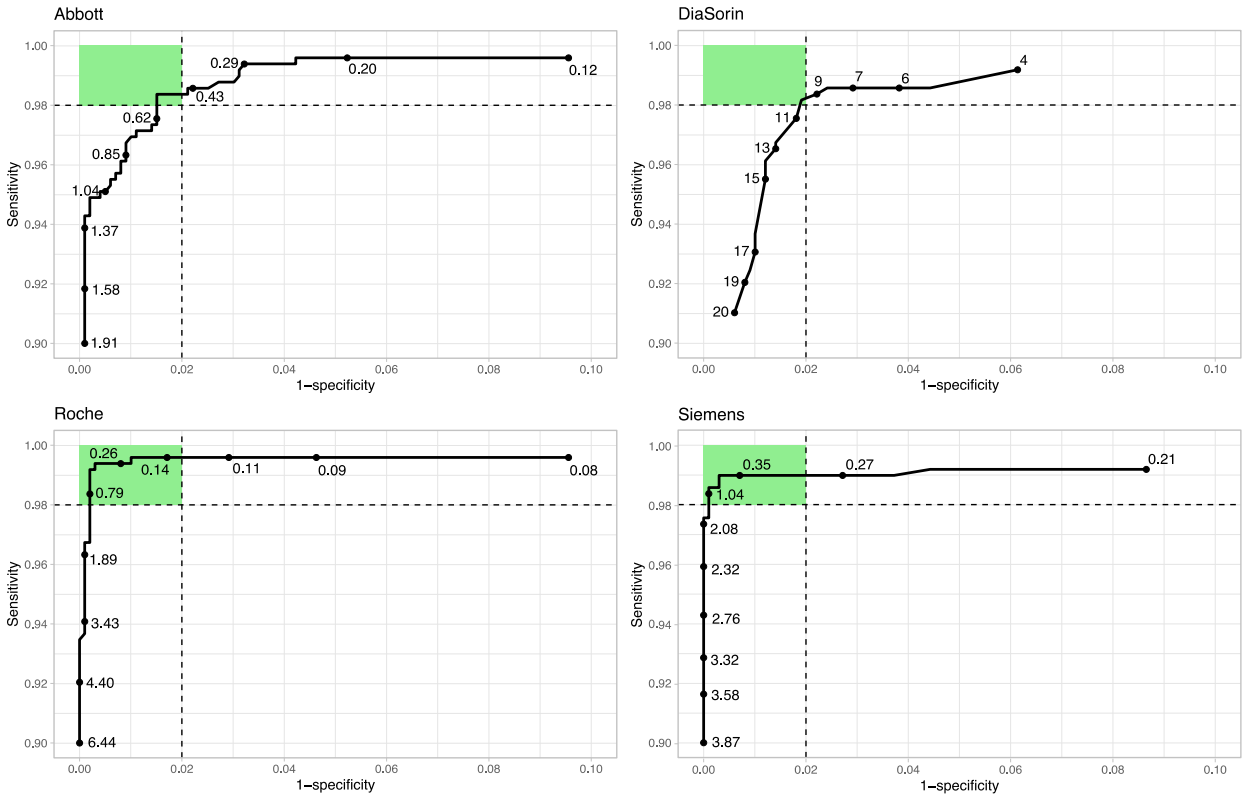


Figure S4. Sensitivity and specificity (95% confidence intervals) plotted for each assay using the current MHRA TPP criteria with alternative assay thresholds to keep specificity ≥98% and revised criteria to show samples ≥30 days after the appearance of first symptoms. For each assay the lowest threshold that kept specificity ≥98% was chosen (Abbott = 0.49, DiaSorin = 10, Roche = 0.128, Siemens = 0.29). The MHRA TPP target performance is shown (dashed line) including the required lower bound of the 95% confidence interval (dotted line) for both sensitivity and specificity.

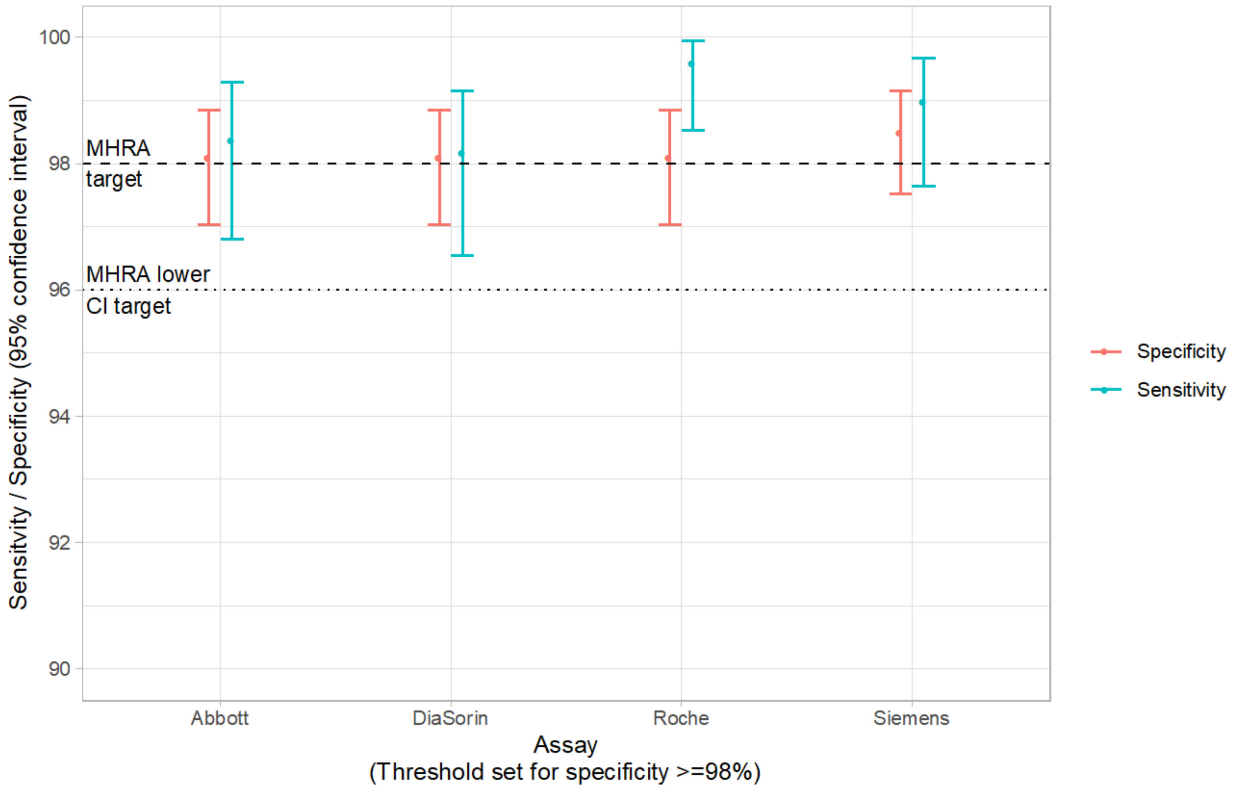


Figure S5. Sensitivity for each assay by disease severity (asymptomatic, mild, severe, critical/death; n=158). Disease severity was defined in line with WHO guidance (7) as follows: asymptomatic = no symptoms; mild = no oxygen requirement; severe = SaO2 ≤93%; critical = respiratory failure requiring intubation; symptom category was assigned on the day of sampling.

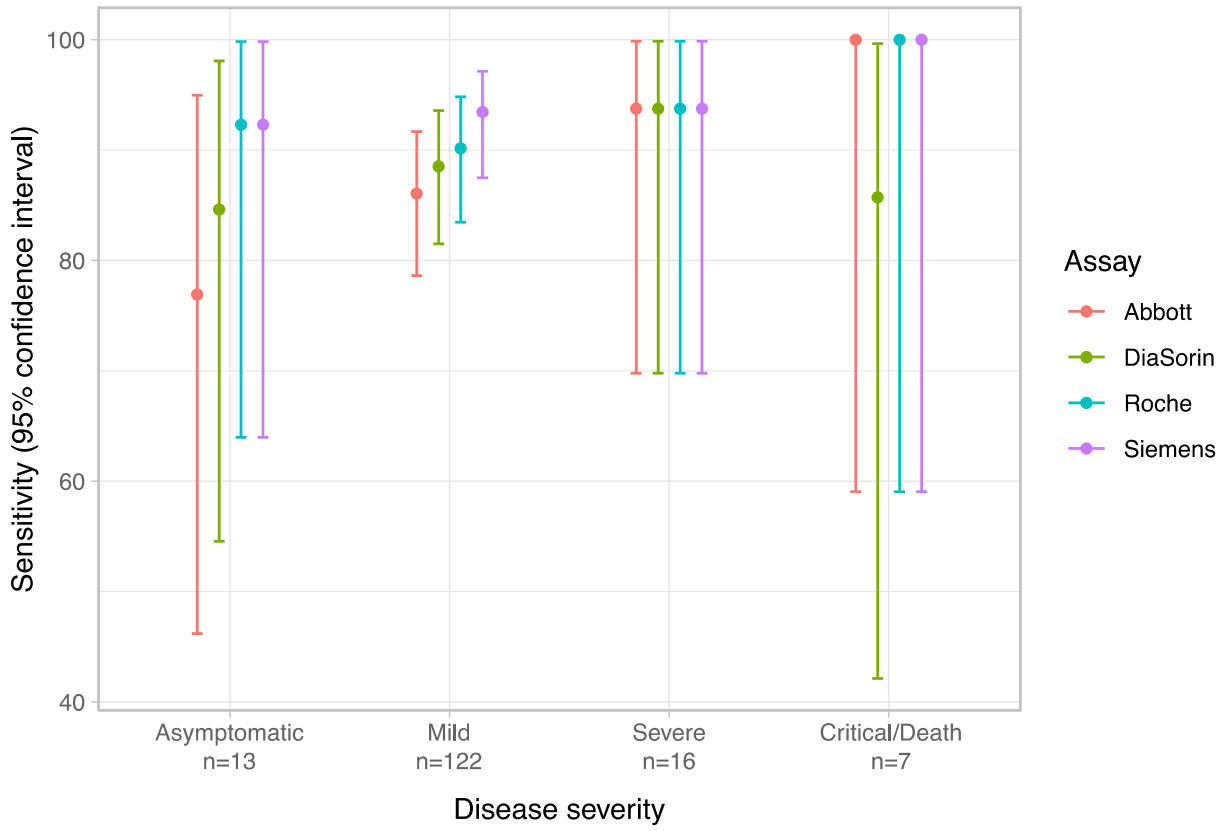


Table S1. Summary of the commercial immunoassays evaluated. Information presented is based on the product literature released by each manufacturer.

Assay and analyser used	**Viral target and antibody type**	**Manufacturers' thresholds**
Abbott SARS-CoV-2 Immunoassay, Architect i2000SR	Nucleocapsid protein, IgG	Negative: <1.4 Positive: ≥1.4
DiaSorin LIAISON® SARS-CoV-2 S1/S2 IgG, LIAISON® XL	Spike protein S1/S2, IgG	Negative: <12.0 AU/mL Equivocal: 12.0 ≤ x <15.0 AU/mL Positive: ≥15.0 AU/mL
Roche Elecsys® Anti-SARS-CoV-2, Cobas e 411	Nucleocapsid protein, Total antibody	Non-reactive: <1.0 Reactive: ≥1.0
Siemens SARS-CoV-2 Total (COV2T), Atellica Solution immunoassay analyzer	Spike protein S1 RBD, Total antibody	Non-reactive: <1.0 Reactive: ≥1.0

Table S2. Partial results for five samples for which there was insufficient sample to run across all four platforms.

Sample barcode	Expected result	Days since symptom onset	Platform	Actual result
900753	Negative	n/a	Abbott	Negative
			DiaSorin	Negative
			Roche	Negative
500379	Positive	40	Abbott	Positive
			DiaSorin	Positive
500380	Positive	41	Abbott	Positive
			DiaSorin	Positive
500381	Positive	41	Abbott	Negative
			DiaSorin	Negative
500384	Positive	42	Abbott	Positive
			DiaSorin	Positive

Table S3. Extended evaluations of sensitivity and specificity (95% CI) for each assay . This includes (i) manufacturer's data for samples taken at least ≥14 days post-symptom onset/PCR test, (ii) adapted MHRA TPP criteria extended to ≥30 days, (iii) per protocol criteria (iv) changing assay thresholds to achieve ≥98% specificity and optimise sensitivity (v) changing assay thresholds to achieve ≥98% specificity and optimise sensitivity, and extending timeframe from symptom onset to sample from ≥20 days to ≥30 days.

(i)

Manufacturers' sensitivity and specificity evaluations				
Assay	Number of PCR-positive cases	Sensitivity (95% CI)	Number pre-pandemic controls	Specificity (95% CI)
Abbott	88 (≥14 days post-symptom onset)	96.77% (90.86, 99.33)	1070	99.63% (99.05, 99.90)
DiaSorin	14 (≥15 days from RT-PCR-positive test)	97.56% (87.40, 99.57)	1090	98.5% (97.6, 99.2)
Roche	29 (≥14 days from RT-PCR-positive test)	100% (88.1, 100)	5272	99.81% (99.65, 99.91)
Siemens	42 (≥14 days from RT-PCR-positive test)	100.00% (91.59, 100.00)	1091	99.82 (99.34, 99.98)

(ii)

Extending timeframe from symptom onset to sample from ≥20 days to ≥30 days										
Assay	Number PCR-positive	Number detected	Number not detected	Number equivocal	Sensitivity (95% CI)	Number pre-pandemic controls	Number detected	Number not detected	Number equivocal	Specificity (95% CI)
Abbott	490	458	32	n/a	93.5 (90.9, 95.5)	994	1	993	n/a	99.9 (99.4, 100)
DiaSorin	490	468	16	6	95.5 (93.3, 97.2)	994	12	980	2	98.6 (97.6, 99.2)
Roche	490	481	9	n/a	98.2 (96.5, 99.2)	994	2	992	n/a	99.8 (99.3, 100)
Siemens	490	482	8	n/a	98.4 (96.8, 99.3)	994	1	993	n/a	99.9 (99.4, 100)

(iii)

Per protocol evaluation with timeframe from symptom onset set at ≥14 days (NB protocol draft predated the publishing of the MHRA's TPP)										
Assay	Number PCR-positive	Number detected	Number not detected	Number equivocal	Sensitivity (95% CI)	Number pre-pandemic controls	Number detected	Number not detected	Number equivocal	Specificity (95% CI)
Abbott	561	520	41	n/a	92.7 (90.2, 94.7)	994	1	993	n/a	99.9 (99.4, 100)
DiaSorin	561	529	24	8	94.3 (92.0, 96.1)	994	12	980	2	98.6 (97.6, 99.2)
Roche	561	543	18	n/a	96.8 (95.0, 98.1)	994	2	992	n/a	99.8 (99.3, 100)
Siemens	561	548	13	n/a	97.7 (96.1, 98.8)	994	1	993	n/a	99.9 (99.4, 100)

(iv)

Adopting changed assay thresholds to achieve ≥98% specificity and optimise sensitivity										
Assay	Number PCR-positive	Number detected	Number not detected	Number equivocal	Sensitivity (95% CI)	Number pre-pandemic controls	Number detected	Number not detected	Number equivocal	Specificity (95% CI)
Abbott	536	523	13	n/a	97.6 (95.9, 98.7)	994	19	975	n/a	98.1 (97.0, 98.8)
DiaSorin	536	523	13	n/a	97.6 (95.9, 98.7)	994	19	975	n/a	98.1 (97.0, 98.8)
Roche	536	533	3	n/a	99.4 (98.4, 99.9)	994	19	975	n/a	98.1 (97.0, 98.8)
Siemens	536	530	6	n/a	98.9 (97.6, 99.6)	994	15	979	n/a	98.5 (97.5, 99.2)

(v)

Adopting changed assay thresholds to achieve ≥98% specificity and optimise sensitivity and extending timeframe from symptom onset to sample from ≥20 days to ≥30 days										
Assay	Number PCR-positive	Number detected	Number not detected	Number equivocal	Sensitivity (95% CI)	Number pre-pandemic controls	Number detected	Number not detected	Number equivocal	Specificity (95% CI)
Abbott	490	482	8	n/a	98.4 (96.8, 99.3)	994	19	975	n/a	98.1 (97.0, 98.8)
DiaSorin	490	481	9	n/a	98.2 (96.5, 99.2)	994	19	975	n/a	98.1 (97.0, 98.8)
Roche	490	488	2	n/a	99.6 (98.5, 100)	994	19	975	n/a	98.1 (97.0, 98.8)
Siemens	490	485	5	n/a	99.0 (97.6, 99.7)	994	15	979	n/a	98.5 (97.5, 99.2)

Table S4. Summary of concordance**/discordance of results between assays.** "+" denotes a positive result, "-" a negative result and "+/-" an equivocal result (the latter relevant for the DiaSorin assay only).

Abbott result	**DiaSorin result**	**Roche result**	**Siemens result**	**Known negative**	**Known positive**	**Total**
-	-	-	-	976	7	983
-	-	-	+	1	2	3
-	-	+	-	2	2	4
-	-	+	+	0	2	2
-	+	-	-	12	0	12
-	+	-	+	0	6	6
-	+	+	+	0	16	16
-	+/-	-	-	2	0	2
-	+/-	+	+	0	4	4
+	-	-	-	1	0	1
+	-	+	-	0	1	1
+	-	+	+	0	6	6
+	+	+	+	0	487	487
+	+/-	+	+	0	3	3
TOTAL				**994**	**536**	**1530**

7. References

1. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. COVID-19 Dashboard. https://coronavirus.jhu.edu/map.html. Accessed: 17/Jun/2020.\
2. Amanat F, Stadlbauer D, Strohmeier S et al. 2020. A serological assay to detect SARS-CoV-2 seroconversion in humans. Nat Med doi:10.1038/s41591-020-0913-5.
3. Zhao J, Yuan Q, Wang H et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis doi:10.1093/cid/ciaa344.
4. Adams ER, Ainsworth M, Anand R et al. Antibody testing for COVID-19: A report from the National COVID Scientific Advisory Panel [version 1; peer review: awaiting peer review]. Wellcome Open Res 2020, 5:139 (https://doi.org/10.12688/wellcomeopenres.15927.1)
5. UK Medicines and Healthcare products Regulatory Agency (MHRA). Target Product Profile: enzyme Immunoassay (EIA) Antibody tests to help determine if people have antibodies to SARS-CoV-2. https://www.gov.uk/government/publications/how-tests-and-testing-kits-for-coronavirus-covid-19-work/target-product-profile-enzyme-immunoassay-eia-antibody-tests-to-help-determine-if-people-have-antibodies-to-sars-cov-2. Last updated 12/Jun/2020. Last accessed: 17/Jun/2020.
6. Department of Health and Social Care (DHSC). Coronavirus (Covid-19): Scaling up our testing programmes. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/878121/coronavirus-covid-19-testing-strategy.pdf. Last accessed: 17/Jun/2020.
7. World Health Organisation (WHO). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf. Last accessed: 17/Jun/2020.

8. Supporting corporate information

8.1 PHE has a comprehensive system for managing conflicts of interest. The manufacturers provided to PHE free of charge the test kits required for the head-to-head evaluation. They also provided to PHE on a short-term, cost-free loan basis any required proprietary equipment and software.

8.2 In line with its Conflict of Interest Policy, PHE works in co-operation with the pharmaceutical, biotechnology, vaccine, diagnostic, and other healthcare-related industries in order to facilitate the development of products and technologies beneficial to health. In this context, PHE maintains active relationships with a broad range of companies in the UK and internationally.

8.3 The manufacturers were made aware of the results of this evaluation prior to publication but had no editorial rights over the content of the report except to ensure factual accuracy. No such comments were received.

8.4 PHE is not a regulatory body and does not issue accreditation of any testing laboratories or provide approvals, validations or endorsements of any particular products including any SARS-CoV-2 diagnostic assay.

8.5 PHE's name and logo are proprietary to PHE and cannot be used for the purpose of commercial promotion of any particular product.