Public Health England

Protecting and improving the nation's health

Supporting information for the PHE commercial serology assay evaluations

Contents

Document control

Current version publication date	Owner	Amendments
March 2021	Richard Clayton, Ashley Otter	Addition of Siemens IgG v2
August 2020	Richard Clayton	

Executive summary

This document provides additional information on the approach followed in Public Health England to the evaluation of commercial serological assays to inform a decision on the use of the assays by NHS laboratories for the detection of anti-SARS-CoV-2 antibodies in patient samples.

It is intended to provide background to the work undertaken in the early stages of the pandemic response as reported in the individual evaluation reports published on this site. The paper details how and when the evaluations were carried out and provides greater detail about the sample sets used in the different evaluations.

PHE is grateful for the advice and challenge provided by external collaborators Professor Sheila Bird, MRC Biostatistics Unit at the University of Cambridge and Professor Jon Deeks, Professor of Biostatistics at the University of Birmingham, who informed the content of this paper. PHE colleagues contributing were:

- Abbie Bown
- Jackie Duggan
- Kevin Brown
- Mary Ramsay
- Nick Andrews
- Richard Clayton
- Richard Vipond
- Sharon Peacock
- Stephanie Migchelsen
- Tim Brooks
- Tristana Perez
- Ashley Otter

Why did PHE undertake these rapid evaluations?

PHE was asked by the Department of Health and Social Care in late April to perform a rapid evaluation of numerous serological assays to inform a decision on the use of the assays by NHS laboratories for the detection of anti-SARS-CoV-2 antibodies in patient samples. The primary purpose of these was to evaluate performance against the performance claims of the manufacturers.

These rapid evaluations were completed in a short time span, generally under a week with schedules constrained by availability of equipment and reagents. An initial generic protocol to provide the best analysis possible based around the samples held by PHE was prepared on 28 April and the first 2 assays (Abbott and Roche) were run the following week. The 900-sample panel was reduced for the Roche as only 600 tests were available at that time.

The start dates for review of all the assays examined were determined by the actual availability of the definitive assay as many manufacturers announced their assays before the release date of the production standard product. All tests were performed following the manufacturer's instructions, and representatives of the companies oversaw the work on their product. At the time the protocol was defined no standards or study guidance existed beyond generic approaches. Scientific studies or validations as described by the National Institute for Health and Care Excellence (NICE) did not appear until nearly 2 months after the start of the work.

When were the evaluations carried out?

The table below shows the start dates, end dates and publication dates of the studies undertaken. The EuroImmun evaluation reflects samples analysed over an extended period as the assay was in use for seroprevalence work in PHE.

Table 1. Dates of evaluations

Evaluation Assay	Start date	End date	Date published
Euroimmun Anti-SARS-CoV-2 ELISA IgG	5 April	21 May	19 June
Abbott Anti-SARS-CoV-2 IgG	4 May	7 May	23 May Updated 8 June
Roche Elecsys Anti-SARS-CoV-2	5 May	7 May	23 May Updated 11 June
Ortho Clinical Diagnostics VITROS Immunodiagnostic Products Anti-SARS-CoV-2 IgG	11 May	15 May	29 May
DiaSorin LIAISON SARS-CoV-2 S1/S2 IgG	2 June	10 June	19 June

Evaluation Assay	Start date	End date	Date published
Siemens Atellica-IM Total (COV2T) SARS-CoV-2	3 June	12 June	23 June
Ortho Clinical Diagnostics Vitros Immunodiagnostic Products Anti- SARS-CoV-2 Total	2 June	11 June	23 June
Beckman Coulter Access Anti- SARS-CoV-2 IgG	24 June	6 July	
Siemens Anti-SARS-CoV-2 IgG	6 July	14 July	
Siemens Ant-SARS-CoV-2 IgG v2	5 November	15 December	4 March

What kind of samples were used?

Serum samples were used. All the assays we evaluated recommend the use of serum or plasma.

How sick were the patients that the samples came from?

The aim was to use samples representative of the general population. Most samples came from community cases, very few were admitted to hospital and those that were may have only been admitted for isolation during the containment phase. This group was chosen specifically because they had mild disease and would better reflect cases detected through population-based seroprevalence.

How many samples did you use?

Positive samples

The aim was to have a panel of around 100 positive samples for each evaluation. At the beginning of the evaluation process in late April 2020, very few serum samples were available to us as we are not an organisation that treats patients. See below for a description of how samples were sourced and selected.

Negative samples

A larger panel of negative samples was used. The target sample size was 400 samples where possible.

Confounder samples

Table 2. Total samples used for each evaluation

Type	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensG v1	SiemensG v2
Positive	96	93*	93*	93*	100	100	100	100	100	115
Negative	395	399	387	399	399	399	399	399	399	348
Confounders	354	100	85	100	100	100	100	100	100	152

*100 samples were initially selected for the panel however 7 were excluded post analysis as these were found to be PCR negative

Where were samples sourced from?

Positive samples

The analysis was carried out on samples accessed using our relationships with hospital laboratories and professional bodies. The convalescent samples were mainly identified from the First Few 100 surveillance study (the first few hundred UK cases of PCR-confirmed SARS-CoV-2) and through the Royal College of General Practitioners (RCGP) surveillance study – samples were submitted by GPs in the community. You can read more about the FF100.

Samples were selected where the volume was sufficient to cover multiple assays. Most samples were between 200 to 400 µl in volume. Of the FF100 collection, a maximum of 82 samples had sufficient volume to use. FF100 samples were available in both locations where evaluations were completed and a common sample set was selected. However, there were 5 samples run in Colindale where

There was insufficient volume in Porton Down. There were 8 samples run in Porton Down where there was insufficient volume in Colindale. An additional set of 14 samples were obtained from cases admitted to the Royal Free Hospital (RFH) but these samples had very limited demographic information. As only time since admission to hospital was recorded (not time since onset of illness) these were replaced as new short onset samples became available. Other sources included Basingstoke Hospital and the Porton Down laboratory.

Table 3. Origin of positive samples used in evaluations

Source	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensG v1	Siemens G v2
FF100	82	79	79	79	79	70	70	70	70	72
RFH	14	14	14	14	14	0	0	0	0	17
Basingstoke	0	0	0	0	5	26	26	26	26	23
Porton*	0	0	0	0	2	4	4	4	4	3
Total	96	93	93	93	100	100	100	100	100	115

*= received in RIPL or staff samples

Negative samples

These were all collected from historical samples held by PHE's Sero-epidemiology Unit (SEU) in Manchester or from existing reference panels held on-site at Porton and Colindale. All negative samples were collected before December 2019.

Supporting information for the PHE commercial serology assay evaluations

Confounder samples

Confounder negative samples from the Sero-Epidemiology Unit (SEU), Manchester were tested including; samples positive for seasonal coronavirus, and a second group who tested positive for rheumatoid factor, Epstein-Barr virus (EBV), cytomegalovirus (CMV) or varicella- zoster virus (VZV). Individual reports describe the precise mix of confounder samples used.

How many patients were these samples from?

The intention was to use individual patient samples. However, some of the positive samples provided to us in the FF100 collection were repeat samples, in that 6 patients provided samples early and later following the onset of symptoms. In early evaluations, these were used to achieve the target totals

Were all the assays evaluated using the same samples?

We were not able to evaluate all the commercial assays using the same set of samples, as there was a limited volume of each sample. Where possible, samples were replaced from the same original source and matched as closely as possible.

The tables below show the number of samples used across all the evaluations to-date: Abbott, Euroimmun, Roche, Ortho IgG, Siemens Total, Ortho Total, Diasorin, Beckman and Siemens IgG (v1 and v2). The table identifies samples that were in common pairwise between evaluations, although these were not contemporaneous.

Positive samples

When a positive sample needed to be replaced, the substituted samples were matched for date of onset of symptoms to date of collection to the best of our ability. See the following question for a breakdown of demographic and onset of illness data.

We were able to evaluate Abbott, Euroimmun, Roche and Ortho IgG assays using broadly the same set of samples, as shown below: 88 of the samples used to evaluate the Abbott assay were used to evaluate Euroimmun, Roche and Ortho IgG as well as an additional 5 samples.

We were able to use 82 of the original samples for the Siemens Total evaluation, which verlaps with 87 of the 93 used on the Euroimmun, Roche and Ortho IgG assays. Following this evaluation, many samples were fully consumed. Only 58 samples from the original samples remained and we introduced 42 new samples into the panel for the Ortho Total and Diasorin evaluations.

Only 42 samples from the original panel remained for the Beckman Coulter and Siemens IgG evaluations. Due to the Siemens IgG v2 evaluation conducted at a later date, a different panel to many of the other evaluations was used but still had shared samples to other evaluations.

Table 4. Overlap of Positive samples across evaluations

POSITIVES	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensG v1	SiemensG v2
Abbott	96*	88	88	88	82	58	58	42	42	50
Euroimmun		93*	93	93	87	62	62	44	44	48
Roche			93*	93	87	62	62	44	44	48
OrthoG				93*	87	62	62	44	44	48
SiemensT					100	73	73	51	51	53
OrthoT						100	100	78	78	58
DiaSorin							100	78	78	58
Beckman								100	100	57
SiemensG v1									100	57
SiemensG v2										115
	*Note 4 or 7 sample results were removed post-testing									

Negative samples

Table 5. Overlap of Negative samples used across evaluations

NEGATIVES	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensG	SiemensG 2.0
Abbott	395	395	383	395	282	85	85	85	74	94
Euroimmun		399	387	399	285	86	86	86	75	94
Roche e411			387	387	278	84	84	84	74	92
OrthoG				399	285	86	86	86	75	94
SiemensT					399	196	196	196	179	109
OrthoT						399	399	399	374	227
DiaSorin							399	399	374	227
Beckman								399	374	227
SiemensG v1									399	232
SiemensG v2										348

Confounder samples

All substitutes were matched according to confounder disease.

Table 6. Overlap of Confounder samples used across evaluations

CONFOUNDERS	Abbott	Euroimmun	Roche	OrthoG	Siemens	OrthoT	DiaSorin	Beckman	SiemensG	SiemensG 2.0
Abbott	354	50	35	50	50	50	50	42	43	94
Euroimmun		100	85	100	96	73	73	61	59	54
Roche			85	85	81	58	58	46	44	44
OrthoG				100	96	73	73	61	59	54
SiemensT					100	71	71	63	62	58
OrthoT						100	100	84	81	59
DiaSorin							100	84	81	59
Beckman								100	96	65
SiemensG v1									100	68
SiemensG v2										152

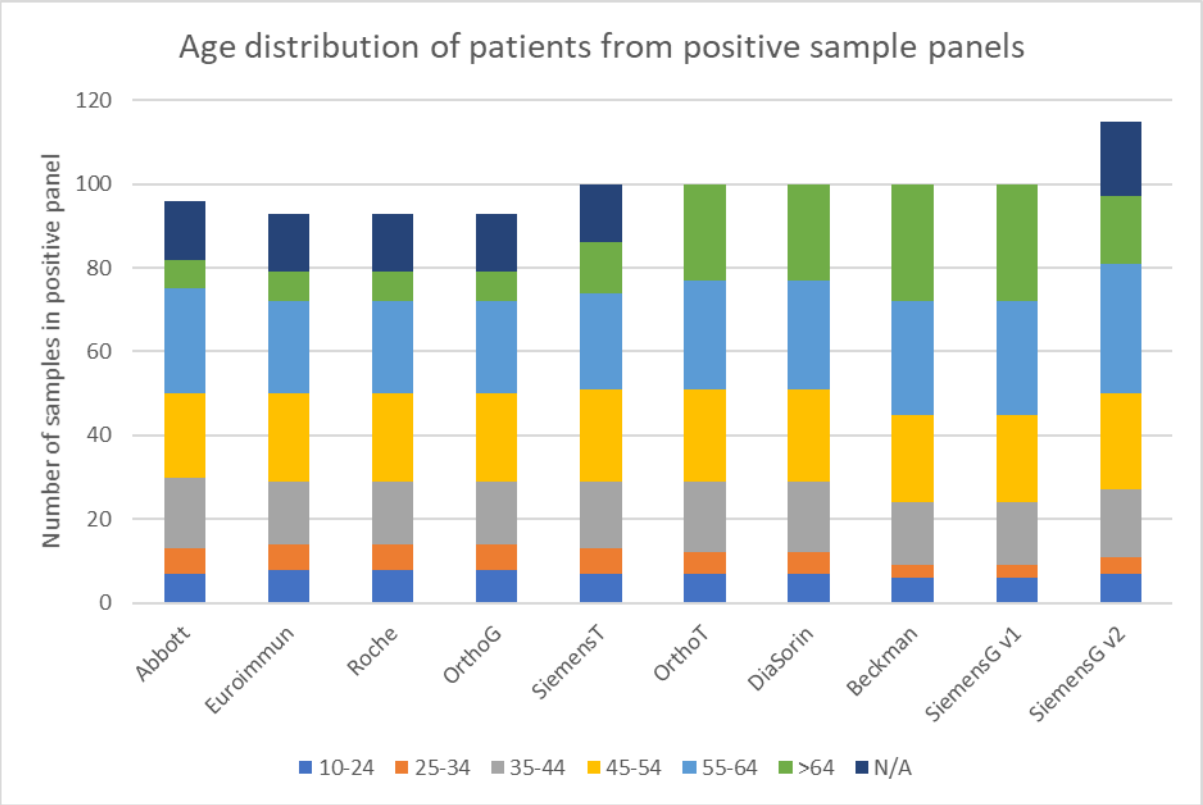
Can you describe the patients that the positive samples came from?

We do not have demographic information for the 14 positive samples from RFH. The remaining 82 samples used to evaluate the Abbott assay were from 76 patients.

Table 7. Age demographics for positive samples

Age band	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensGv1	SiemensGv2
10-24	7	8	8	8	7	7	7	6	6	7
25-34	6	6	6	6	6	5	5	3	3	4
35-44	17	15	15	15	16	17	17	15	15	16
45-54	20	21	21	21	22	22	22	21	21	23
55-64	25	22	22	22	23	26	26	27	27	31
>64	7	7	7	7	12	23	23	28	28	16
N/A	14	14	14	14	14	0	0	0	0	18
Total	96	93	93	93	100	100	100	100	100	115

Figure 1. Age distribution of patients from positive sample panels



The time since onset of symptoms is provided in each evaluation report. For some of the samples, we only had time since hospital admission, giving an artificially low time. For later evaluations, we were able to replace these and so have more accurate time since symptom onset.

Table 8. Time since onset of disease for positive samples

Interval	Abbott	Euroimmun	Roche	OrthoG	SiemensT	OrthoT	DiaSorin	Beckman	SiemensG v1	SiemensG v2
<= 10	14	14	14	14	14	11	11	11	11	24
11 to 20	5	4	4	4	7	12	12	12	12	15
21 to 30	35	35	35	35	37	37	37	37	37	45
31 to 40	32	30	30	30	32	31	31	31	31	14
41 to 50	10	10	10	10	10	9	9	9	9	10
From 14 days	82	79	79	79	85	85	85	85	85	80
From 21 days	77	75	75	75	79	77	77	77	77	69
All	96	93	93	93	100	100	100	100	100	115

Could you provide the sensitivity data in a different format, say by weekly- increments, rather than 10 days?

Yes, please contact: ripl@phe.gov.uk

Were the samples blinded when they were tested?

Yes, during testing, samples were only identified by their barcode and the scientists conducting the tests were unaware if a sample was positive, negative, or a confounder. Samples were tested on more than one run. The testers had barcodes; but the database was not shared and could not be accessed by the testers.

Evaluations were carried out individually with no cross checking of results between patients or evaluations.

Was any data excluded?

No results were excluded as uninterpretable or borderline. Only exclusions were where no result was generated due to technical issues.

Who carried out the testing?

The majority of the evaluations were undertaken by skilled research scientists in PHE Porton Down laboratory. The Abbott evaluation was undertaken by scientists in the PHE Colindale laboratory. Oversight was provided by a clinical scientist.

Were sera frozen at any point?

Samples which were sourced from the SEU in Manchester were stored at -80 °C prior to shipment. They were thawed at room temperature before aliquoting and shipping to the testing sites.

Can results be compared between evaluations?

Direct comparison between assays was not the purpose of these evaluations.
A head-to-head comparison study of 4 assays was conducted independently and results are published here.

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

Webite: www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Prepared by: Jackie Duggan, Rare and Imported Pathogens Laboratory, PHE Porton Down.

For queries relating to this document, please contact: Tim Brooks, Clinical Services Director, Rare and Imported Pathogens Laboratory, PHE Porton Down.
Email: tim.brooks@phe.gov.uk

Published: March 2021
PHE gateway number: GW-2000

PHE supports the UN Sustainable Development Goals