# Performance of the Abbott SARS-CoV-2 IgG II Quantitative antibody assay including the new Variants of Concern (VOC 202012/V1 (UK) and VOC 202012/V2 (South Africa)): And first steps towards global harmonization of COVID-19 antibody methods

**Short title:** Quantitative COVID-19 antibody methods

**Authors:** Emma English[1,2], Laura E Cook[2], Isabelle Piec[1], Samir Dervisevic[1,3], William D Fraser[1,2], W. Garry John[1,2]

**Affiliations:** [1]Faculty of Medicine and Health, University of East Anglia, Norfolk, NR4 7TJ, UK, [2]Department of Clinical Biochemistry and Immunology, Norfolk and Norwich University Hospital, Norfolk, NR4 7UY.
[3]Department of Virology, Norfolk and Norwich University Hospital, Norfolk, NR4 7GJ

**Corresponding Author:**

E-Mail: emma.english@uea.ac.uk

Emma English PhD
Faculty of Medicine and Health,
University of East Anglia,
Norwich Research Park
NR4 7TJ
United Kingdom

**Keywords:**

COVID-19, SARS-CoV-2, Antibody assay, serology, evaluation, harmonization, variants

**Abbreviations:**

Evaluation Protocol (EP), CLSI, Clinical & Laboratory Standards Institute (CLSI), Epitope Diagnostics Inc. (EDI), National Institute for Biological Standards and Control (NIBSC), Limit of Detection (LOD), Limit of Blank (LOB), Limit of Quantitation (LOQ), quality control (QC), coefficient of variance (CV), Arbitrary Units (AU/mL), 4-parameter logistic (4PL)

## Abstract

***Background*** In the initial stages of the SARS-CoV-2 (COVID-19) pandemic a plethora of new serology tests were developed and introduced to the global market. Many were not evaluated rigorously and there is a significant lack of concordance in results across methods. To enable meaningful clinical decisions to be made, robustly evaluated, quantitative serology methods are needed. These should be harmonized to a primary reference material, allowing for the comparison of trial data and improved clinical decision making.

***Methods*** A comprehensive evaluation of the new Abbott IgG II anti-SARS-CoV-2 IgG method was undertaken using CLSI based protocols. Two different candidate primary reference materials and verification panels were assessed with a goal to moving towards harmonization.

***Results*** The Abbott IgG II method performs well across a wide range of parameters with excellent imprecision (<3.5%) and is linear throughout the positive range (tested to 38,365 AU/mL). The sensitivity (based on ≥14 day post positive RT-PCR samples) and specificity are 98.3% [90.6-100.0%] and 99.5% [97.1-100%] respectively. The candidate reference materials showed poor correlation across methods with mixed responses noted in methods that use the spike protein versus the nucleocapsid proteins as their binding antigen.

***Conclusions*** The Abbott IgG II anti-SARS-CoV-2 measurement appears to be the first linear method potentially capable of monitoring the immune response to natural infection, including from new emerging variants. The candidate reference materials assessed do not generate uniform results across several methods and further steps are needed to enable the harmonization process.

## Introduction

SARS-CoV-2 has swept the globe at an alarming rate with a reported 90 million cases and 1.9 million deaths by the one year anniversary of the first death of this pandemic (1) and increasing to 119 million cumulative cases and over 2,6 million deaths by the middle of March 2021 (2). In this time there has been an unprecedented global effort to identify new diagnostic tests, treatments and more recently vaccines against the virus and the associated disease, COVID-19.

The development and delivery of a range of vaccines against the virus is underway, with several already approved for use and others in late stage clinical trials (3). Although there are several different approaches to the design of these vaccines a common factor is the use of the spike proteins, in the form of the attenuated whole virus, or portions of the spike protein such as the receptor binding domain or through the use of nucleic acids directed to the synthesis of the spike protein.

In the United Kingdom a national immunization programme has been launched, with a tiered system of invitations to receive the vaccine dependent on risk of a negative outcome from the disease (4). Whilst there is trial data for each of the vaccines in use, there is limited data regarding the quantitative changes in antibody concentrations over time following vaccination. With the introduction of a national immunization programme it will be important to understand the antibody response to immunization in terms of; development, peak concentration and decline over time to assess efficacy of the vaccination delivery. In order to do this, two elements are required; a robust quantitative SARS-CoV-2 IgG method, which is directed against the spike protein and

a commutable standard or reference material to allow comparison of results across different methods and thus different trials or immunization programs (5).

This study evaluates several assay performance criteria (such as precision and sensitivity) using recognized and standardized evaluation protocols (EP), for the new Abbott SARS-CoV-2 IgG II Quant method on the Alinity i system (Abbott Diagnostics, Chicago, US) and explores the different materials available that may form the basis of a candidate international reference standard for harmonization programs (6).

## Methods

***Sample collection and storage*** All procedures were performed in accordance with the ethical standards of University of East Anglia for de-identified samples for method development and in concordance with the Helsinki Declaration. Serum samples were collected, anonymized, aliquoted and stored at –80$^{o}$C until analysed. SARS-CoV-2 Positive samples (P) were from patients with PCR confirmed infection (AusDiagnostics platform, Chesham, UK, The Panther, Manchester, UK and The Altona Hamburg, Germany). All RT-PCR assays have dual genome target. Key performance testing including precision, limit of quantitation (LoQ), linearity, and method comparison were assessed per Clinical and Laboratory Standards Institute (CLSI) protocols which ensured standardized testing procedures were used.

***Summary of Abbott assay*** The SARS-CoV-2 IgG II Quant assay is an automated two-step, chemiluminescent microparticle immunoassay (CMIA). It is used for the qualitative and quantitative determination of IgG antibodies to the receptor binding domain (RBD) of the S1 subunit of the spike protein of SARS-CoV-2, in human serum

and plasma on the Alinity i system. The sequence used for the RBD was taken from the WH-Human 1 coronavirus, GenBank accession number MN908947. The analytical measurement interval is stated as 21-40,000 AU/mL and positivity cut-off is ≥50 AU/mL (manufacturer defined).

***CLSI EP-5 and EP-15 imprecision*** Both CLSI EP-5 and EP-15 based protocols were used to evaluate the imprecision of the assay. For the EP-15 study, three different QC levels (Abbott Diagnostics) were used (one negative and two positive) and five replicates were measured twice a day for five days. For the EP-5 study, four patient serum pools (one negative and 3 different positive) were measured in duplicate, twice a day for 20 days.

***CLSI EP-6 linearity*** Dilutions of a high patient sample (mean value 38,365 AU/mL, from triplicate measurement) were made using the Abbott diluent to generate a series of samples with antibody concentrations over 95% of the analytical measurement range of the assay. All samples were measured in triplicate.

***Limit of Quantitation (LOQ) and Limit of Detection (LOD)*** The LOQ was determined by measuring five negative patient pools in quintuplicate, twice a day for two days. The LOQ was estimated as the lowest concentration with a 20% CV (7). As defined in CLSI EP17, LOD is determined by utilising both the measured limit of the blank (LOB) and test replicates of a sample known to contain a low concentration of analyte using the equation $LoD = LoB + 1.645(SD_{\text{low concentration sample}})$ (7) The Abbott diluent was used to determine LOB as it is the diluent used for on-board dilution.

***Cross reactivity samples*** Negative control samples were from healthy patients with no recorded history of infection or immune disorders and collected in 2018, prior to the emergence of COVID-19. Pre-pandemic samples from patients who had a range of confirmed respiratory infections (including Influenza A, B and seasonal coronaviruses were included in the cross-reactivity analysis). Samples from patients positive for thyroid stimulating immunoglobulin (TSI) were analysed to test the non-specific binding of non-SARS-CoV-2 antibodies in the assay. These groups of samples are referred to as N (negative control), CR (cross-reactivity), TSI (patients with thyroid stimulating immunoglobulin); for further details on samples and collection please see reference (8). A total of 334 individual serum samples (143 P, 65 N, 97 CR and 29 TSI) were analysed for SARS-CoV-2 IgG antibodies.

***Specificity and sensitivity analysis*** The quantitative IgG levels were measured at different time points after a confirmatory RT-PCR test for SARS-CoV-2 allowing for analysis at pre and post 14 days from RT-PCR date, along with the cross reactivity samples these were used to determine sensitivity and specificity.

A concordance analysis was undertaken comparing the Abbott quantitative method with 3 other SARS-CoV-2 IgG immunoassay methods: 1) Epitope Diagnostics Inc. (EDI, San Diego, CA, USA) performed using the Agility ELISA automate (Dynex Technologies, Chantilly, VA, USA), 2) Abbott Diagnostics (qualitative method) (Maidenhead, UK) on the Alinity™ i analyser and 3) DiaSorin (Dartford, UK) on the Liaison XL analyser.

***New variant samples*** As the pandemic progresses new variants of the virus emerge raising concern that the mutations in these variants may render immunoassays ineffective, as the antigenic changes that arise may no longer represent the antigenic regions of the reagents in the assay. The main SARS CoV-2 lineage circulating in the autumn of 2020 in the UK was B.1.177 (the Spanish lineage). However, since January 2021 the 'UK variant' (VOC 202012/01 or B1.1.7) has become the predominant virus in the UK (9). Furthermore, by the beginning of March 2021 there have been in total 266 confirmed and probable cases of 'South Africa' variant (VOC 202012/02 (B.1.351). These new variants have several mutations in different parts of SARS CoV-2 genome with some within the Receptor Binding Domain increasing the virus transmissibility. The RBD of the 'UK' (VOC 202012/01 or B1.1.7) strain, which is the predominant virus in the UK, contains mutation N501Y in the RBD domain of the Spike protein amongst 15 other mutations in the other genome areas. This variant has since January 2021 acquired another RBD mutation, E484K, in addition to the variant defining mutations which resulted in designation as VOC 202102/02 (B1.1.7 cluster with E484K). E484K is currently the mutation with most evidence of causing antigenic change (10). The RBD of 'South Africa' (VOC 202012/02 or B.1.351) variant contains RBD K417N mutation in subsets of isolates in addition to the E484K and N501Y RBD mutations.

Viruses containing the above mutations are not very similar to the predominant virus (B.1.177 ) in circulation in the summer and autumn of 2020.

We analysed samples from patients proven to have the VOC 202012/v-1 (UK) strain which is now a predominant virus in the UK, as well as one imported cases of the VOC 202012/v-2 (South Africa) strain.

***CLSI EP-9 (trueness)*** Although a small number of assays are marketed as quantitative methods there are no standardized reporting units making method comparisons difficult. It was considered inappropriate to evaluate trueness using a standard CLSI EP-9 protocol. See section below for further discussion.

***Identifying potential standards and reference materials –*** In order to progress the harmonization of SARS-CoV-2 Ig immunoassay methods, a certified reference material (CRM) is needed. Several candidate standard materials were evaluated on each of the four methods described above. Three different materials were obtained from the National Institute for Biological Standards and Control (NIBSC) and assayed. These consisted of:

- A CE marked verification panel of 37 samples (NIBSC code 20/B770). Each sample consisted of 0.3 mL of human plasma containing the bacteria growth inhibitor Bronidox® at 0.05% (w/v). Twenty-three samples were convalescent plasma packs known to be anti-SARS-CoV-2 positive and the remaining 14 were detailed as negative.
- A CE marked 'working standard' (NIBSC code: 20/162) intended for use as a Diagnostic Calibrant to monitor the sensitivity of assays. The standard consisted of convalescent plasma positive for anti-SARS-CoV-2 antibodies pooled from three different donors. Frozen liquid (0.3 mL) was supplied. 'The material has been assigned an arbitrary unitage of 1000U.' A series of dilutions were made to assess the linearity at the positive cut off value.
- A CE marked 'Quality Control 1' (NIBSC code: 20/B764) intended for internal quality control use for immunoassays that detect SARS-CoV-2 antibodies. The

material was supplied as a ready to use reagent of plasma positive for anti-SARS-CoV-2 antibodies, derived from two different donors and diluted in defibrinated convalescent plasma, preserved with Bronidox® at 0.05% (w/v).

A panel of heat inactivated 'reference materials' with 5 positive samples and 1 negative sample from Technopath, were assayed (Technopath, Tipperary, Ireland). The manufacturer information suggests the samples are a series of pre-diluted samples from a positive stock. The primary material, the diluent used, and the heat inactivation process were not described.

***Statistics*** Calculations were performed using SPSS Statistics (IBM) 25.0.0.1. or GraphPad Prism version 8.0 (GraphPad Software, Inc., USA). Cohen's Kappa tests was used to determine the concordance between the assays. Analysis of CLSI EP-15 was performed using the software EP evaluator (Data Innovations, Build 11.3.0.23). Throughout the tables, figures, and legends, the following terminology is used to show statistical significance: *P<0.05; **P<0.01 and ***P<0.001.

## Results

***CLSI EP-5 and EP-15 imprecision*** Supplemental Table 1 shows the performance of a negative and two positive QC samples using the CLSI EP-15 based protocol. The mean value of the negative sample was very low (3.4 AU/mL) and outside of the analytical measurement range, hence the high CV at that level, however the SD value is low (SD 0.7) inferring good performance. The imprecision of the positive QC material was low (3.0% and 3.3% total CV).

The imprecision data based on CLSI EP-5 is presented in Table 1. Again, this shows a high CV in the negative samples (mean 2.7 AU/mL, SD 1.74) but the more relevant total CV in the each of the three positive samples remains below 3.5% CV.

***CLSI EP-6 linearity*** CLSI states for EP-6 that goals for linearity should be derived from goals for bias and should be less than or equal to these goals. Figure 1 shows the linearity of a diluted sample over the working range of the assay, it is linear up to 38,365 AU/mL as tested (manufacturer claim is 40,000 AU/mL). This was the highest patient sample value available measured neat that was under 40,000 AU/mL and was acceptable as this high sample was within 5% of the upper limit. Supplemental Table 2 details the % difference from the target values for each dilution. Linear fitting was performed and showed a slope of 1.004 (95%CI 0.9923-1.017) with a r2=0.9992. The model was tested and returned a p $P<0.001$, indicating non deviation from the linearity.

***Limit of Quantitation (LOQ) and Limit of Detection (LOD)*** The means and the CVs (%) of the samples for the LOQ were 3.8 AU/mL (59.3%), 18.1 AU/mL (13.6%), 30.4 AU/mL (5.4%), 36.9 AU/mL (3.4%) and 52.2 AU/mL (3.9%). The LOQ was estimated as the lowest concentration with a 20% CV. Using a 4-PL curve-fit; the LOQ was calculated at 15.4 AU/mL. As defined in CLSI EP17, LOD is determined by utilising both the measured LOB (0.1 AU/mL) (lower limit of blank) and test replicates of a sample known to contain a low concentration of analyte. The LOD was determined as 4.3 AU/mL using the Abbott diluent.

***Cross reactivity samples*** Patient samples (n=97) from people with respiratory infections, collected in 2018 and 2019 prior to the COVID-19 pandemic were analysed. The results ranged from 1.1 AU/mL to 48.3 AU/mL, with a mean value of 6.9 AU/mL. All but one samples were reported as negative, but one was also close to the cut-off of 50 AU/mL (manufacturer defined). One sample had a value of 140.5 AU/mL but was negative on the qualitative IgG assay and the IgM assay. In addition to these CR, 29 TSI samples were also analysed and all returned negative values from 0.0-29.0 AU/mL. Data is summarised in Table 2.

***Specificity and sensitivity*** Table 2 details the sensitivity and specificity of the method with analysis of SARS-CoV-2 RT-PCR positive samples at all time points and at >14 days post confirmatory test. The data shows that the method has a sensitivity of 91.6% in all time points and 98.3% at >14days, and a specificity of 99.4%.

The analysis of assay concordance revealed a mixed pattern of agreement with the highest between the Abbott quantitative method (IgG II) and the Abbott qualitative method with a Cohen's Kappa of 0.965 and agreement of 98.4% and the poorest between the DiaSorin and other methods (Cohen's Kappa of 0.930, agreement of 96.7% between the Abbott quantitative and the DiaSorin methods, see figure 2.

***New variant samples*** The results clearly show that the Abbott IgG II method detects the original strain of SARS-CoV-2 as well as two new Variants of Concern, the VOC 202012/v-1 (UK) strain and the VOC 202012/v-2 (South Africa) strain. Figure 3 shows a time course for a subset of 4 different patients charting the increase in antibody levels post confirmatory RT-PCR.

***Identifying potential standards*** The analysis of the 37 samples from the NIBSC 'verification panel' produced the expected classification of 23 positive and 14 negative samples, with a clear separation between the two groups on the Abbott IgG II assay, see Supplemental Table 3. The range of the positive values was 210 – 9710 AU/mL. The values for the negative samples of 0.7-5.6 AU/mL, were significantly below the ≥50 AU/mL threshold for classification as a positive sample. As expected, these values are markedly different to the values presented for other methods. In order to allow comparison of the results between the different methods and to standardize the results, the individual values were normalized to the highest responding sample for each method which was set to a normalizing value of 1. Thus, the results of the positive samples for each method were divided by the highest value obtained by that method for any of the positive 23 samples. So, if the highest positive values were 50 and 23 for two methods, all positive samples were divided by 50 for the first method and by 23 for the second method. This provided all values as a ratio of the highest value obtained for any sample on an individual method. This should demonstrate if the magnitude of positive response compares between methods, see Supplemental Figure 1. Figure 1 shows that methods which use spike proteins as the assay antigen produced similar results. Methods that used the nucleocapsid antigen are similar. However, there is a lack of agreement between these two method types (spike Ag versus nucleocapsid Ag).

The analysis of the NIBSC 'working standard' (NIBSC code: 20/162) and Quality Control 1 (NIBSC code: 20/B764) samples generated mean values of 14,072 AU/mL and 296.6 AU/mL respectively on the Abbott IgG II method, as a mean of triplicates. This is significantly different to the arbitrary '1000 U' assigned to the working standard.

It should be noted that no volume is detailed in the unit assignment. The dilution of the NIBSC 20/162 working standard proved linear. The slope was 0.9981 (95%CI 0.9836-1.013) with a $r^2$=0.9997. Model test showed no deviation from the linear model with p=0.2500). The % difference from expected is detailed in Supplemental Table 4, the range was 4.6-9.7% difference to a value of 20 AU/mL on dilution.

The Technopath series proved to be linear when measured on the quantitative method, this was expected as the samples represent a dilution series. Although the diluent is not described it does not appear to have had an impact on the linearity of the dilution series. The range of values obtained were 5.6 AU/mL for the negative and 147.5 – 4,098 AU/mL for the positive samples. Figure 4 shows comparison graphs of the values obtained with the Abbott IgG II versus different methods (DiaSorin Liaison XL, EDI, Abbott IgG (qual)). The other methods are clearly calibrated towards the negative/positive threshold and are not linear using this material.

## Discussion

***Quantitative SARS-CoV-2 IgG II method evaluation*** The focus of many method evaluations to date has been on the diagnostic accuracy of the assay; does it correctly identify those with or without antibodies to the SARS-CoV-2 virus? However many publications have neglected to fully address the analytical performance of these methods, which ultimately has a significant impact of the potential clinical utility of these tests. There are over 300 different methods in development or available for purchase that aim to detect SARS-CoV-2 IgG and whilst some have undergone a robust evaluation, many on the market have not (11, 12). This is the first evaluation of

the pre-launch Abbott anti- SARS-CoV-2 IgG assay (IgG II) on the Alinity i system and the data clearly show that it meets many of the expected performance criteria.

The method achieved an excellent precision profile which is well within the ≤ 15% CV often cited as the minimum criteria set by the Food and Drugs Administration (FDA) and the European Medicines Agency (EMA) (13). The method is linear across a wide working range. The sensitivity and specificity are very high at 98.3% [90.6-100%] for samples ≥ 14 days post positive RT-PCR, 91.6% [85.8-95.6%] for all samples and 99.5% [97.1-100%] respectively. Although concordance with other methods varied, it is important to note that this is the first quantitative assay for SARS-CoV-2 IgG and therefore difficult to make direct comparisons, the poorest concordance was with the DiaSorin method, reflecting previously published data (14-16).

It is reassuring to see that the method also identifies antibodies in patients with the two new Variants of Concern (VOC 202012/v-1 (UK) VOC 202012/v-2 (South Africa) strain), and that an increase in antibody levels occurs as the immune response evolves. It is important that any method used to monitor immune response to infection is able to detect antibodies that arise from a variety of emergent variants, otherwise false negative diagnoses may arise.

***Identification of candidate primary reference material*** A robust approach to harmonization of serology methods is essential in order to understand the ongoing impact of both natural infection and vaccination on the immunity of the population to SARS-CoV-2. This study and our previous data (15) have shown a significant disparity in the performance of different commercial methods in terms of linearity (Figure 4),

345 units of measurement (Figure 4) and even sensitivity and specificity (Figure 2, Table
346 2 and reference (15)).

347 This study evaluated candidate reference materials for the harmonization of anti-
348 SARS-CoV-2 antibody methods. Much hope has been placed on the rapid introduction
349 of vaccines against this virus but many questions on their efficacy remain unanswered.
350 Questions such as; are two doses of the vaccine needed? What time interval is most
351 effective? What is the magnitude and duration of the immune response? What level of
352 antibodies in circulation are needed to continue to afford protection? All these
353 questions require multiple, large scale and multi-site studies to answer, which in return
354 need robust and consistent serology measurements. Studies in children who have
355 been vaccinated for Rubella virus show that approximately 9% are seronegative after
356 the first dose decreasing to <1% after the second, clearly indicating the value of the
357 second dose (17, 18). Antibody levels present in a population are higher when due to
358 naturally occurring immunity or post vaccination boosts from virus exposure than
359 levels that arise through immunization alone, meaning the expected values for Rubella
360 antibodies have decreased as immunization programmes have widened their reach
361 (18), this may be mirrored over time with the SARS-CoV-2 antibody levels in the
362 general population meaning any derived target values for positive serology and the
363 limits of quantitation of methods will need to adapt over time.

364 In order to achieve the goal of harmonized serology testing for anti-SARS-CoV-2
365 antibody methods the principles of metrology must be applied (19-21). Key
366 components of a system of traceability include a defined measurand, a primary
367 reference material (preferably approved by a certifying authority such as IFCC, ISO,
368 ICM, NIST etc), a higher order measurement system or reference method procedure
369 and a known calibration hierarchy. Metrological traceability is the property of a

measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty (20). Whilst this has been achieved for analytes such as HbA1c there has been less success with serology testing for viruses such as Rubella and to date very limited attempts for anti-SARS-CoV-2 antibody methods (18, 22, 23).

Hurdles to overcome include the availability of a reference measurement procedure and a primary reference material. The immune response to an antigen challenge is heterogenous and therefore defining the measurand is difficult. Three components comprising the system (or matrix such as plasma), the component (the anti-SARS CoV-2 IgG) and a measurement quantity such as the biological response or biological activity, together form the measurand of interest. It is expected that a primary reference material for such complex analytes will undergo state of the art purification steps with identification of class and subclass of immunoglobulin and some type of functional assessment of biological activity. International Units per millilitre (IU/mL) should be used and the reference material should be commutable across methods (6, 24). Once a primary reference material has been defined all future reference materials should refer to this material rather than the previous batch.

This standardization process has proven difficult for serology methods, with Rubella virus IgG methods an example of how poor agreement is between some methods (25). This has the potential to lead to misinterpretation of results, sometimes causing adverse clinical outcome. Factors which influenced the lack of standardization include: the use of an immunoglobulin preparation from human serum, with limited purification steps, which are not described. The effect on biological activity of the preparation, lyophilization and subsequent reconstitution was not assessed. Guidance on appropriate diluents was not provided. Similarly, these are the same conditions under

which the current available references materials for SARS-CoV-2 antibody methods are prepared and the same questions around performance are being raised.

Our data shows that the current anti-SARS-CoV-2 antibody methods do not compare well in terms of units of measurement, linearity, magnitude of response and relative response in different patient samples. Those methods which detect antibodies directed against the spike protein appear to have greater concordance with each other than those that detect the nucleocapsid. Some of this may be due to the calibration of the methods, with only the Abbott IgG II quantitative method being linear so far. The materials currently available as candidate primary reference material show considerable variation across methods and the preparation and performance of these materials is poorly described. Without steps to improve the quality of these reference materials, including a description of the antibody populations within the reference material (i.e. predominantly nucleocapsid or spike protein recognition), the scientific community is likely to encounter similar pitfalls to those that previous attempts to harmonize serology methods have experienced.

***Conclusions*** It is clear to see that there is a long road ahead to achieve harmonization of anti-SARS-CoV-2 antibody methods and urgent action is needed to ensure that manufacturers and regulatory bodies work synergistically towards the goal of harmonization.

The Abbott IgG II method performed well in this evaluation and is the only method tested that shows linearity over a wide concentration range and potential external calibration materials. It is suitable for future studies investigating the clinical response to natural infection which are urgently needed.

## Acknowledgments

The authors would like to thank Myra Del Rosario and Martyn Hammond for technical advice. The authors also wish to acknowledge Abbott Diagnostics (Chicago, US) for the supply of the reagents required to undertake this study.

**Table 1:** EP-5 (20 day imprecision data) generated using pool patient material at four levels, one negative and three positives.

| | Negative patient pool | | Positive patient pool 1 | | Positive patient pool 2 | | Positive patient pool 3 | |
|---|---|---|---|---|---|---|---|---|
| | Mean 2.7 AU/mL | | Mean 71 AU/mL | | Mean 283 AU/mL | | Mean 2428 AU/mL | |
| | %CV | SD | %CV | SD | %CV | SD | %CV | SD |
| **Total** | **64** | 1.74 | **2.9** | 2.09 | **3.3** | 9.3 | **3.4** | 83.4 |

**Table 2**: Summary table of the sensitivity, specificity and cross-reactivity samples of Abbott IgG II quantitative anti SARS-CoV-2 IgG method.

| | Total samples tested | SARS-CoV-2 IgG positive | SARS-CoV-2 IgG negative | Result (95% CI) |
|---|---|---|---|---|
| SARS-CoV-2 positive samples (all time points) | 143 | 131 | 12 | 91.6 (85.8-95.6) |
| SARS-CoV-2 positive samples (>14 days, time points) | 57 | 56 | 1 | 98.3 (90.6-100.0) |
| Pre-COVID-19 controls (N) | 65 | 0 | 65 | 100.0 (94.5-100.0) |
| Other respiratory infections (CR) | 97 | 1 | 96 | 99.0 (94.5-100.0) |
| Thyroid stimulating immunoglobulin (TSI) | 29 | 0 | 29 | 100.0 (88.1-100.0) |
| Controls (N, CR, TSI) | 191 | 1 | 190 | 99.4 (97.1-100.0) |

## References


1. Worldometers.info. COVID-19 Coronavirus Pandemic Dover, Delaware, U.S.A.2021 [updated 10 January, 2021; cited 2021. Available from: https://www.worldometers.info/coronavirus/.
2. Organization WH. Weekly epidemiological update on COVID-19 – 16 March 2021. 2021.
3. Krammer F. SARS-CoV-2 vaccines in development. Nature. 2020;586(7830):516-27.
4. Care DoHaS. UK COVID-19 vaccines delivery plan. London2021.
5. Khoury DS, Wheatley AK, Ramuta MD, Reynaldi A, Cromer D, Subbarao K, et al. Measuring immunity to SARS-CoV-2 infection: comparing assays and animal models. Nat Rev Immunol. 2020;20(12):727-38.
6. Thienpont LM, Van Uytfanghe K, De Leenheer AP. Reference measurement systems in clinical chemistry. Clin Chim Acta. 2002;323(1-2):73-87.
7. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. Clin Biochem Rev. 2008;29 Suppl 1:S49-52.
8. Piec I, English E, Thomas MA, Dervisevic S, Fraser WD, John WG. Performance of SARS-CoV-2 serology tests: Are they good enough? PLoS One. 2021;16(2):e0245914.
9. England PH. Investigation of SARS-CoV-2 variants of concern in England, Technical briefing 6 Public Health England; 2021.
10. England PH. SARS-CoV-2 variants of concern and variants under investigation in England Technical briefing 7. 2021.
11. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database Syst Rev. 2020;6:CD013652.
12. Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui LP, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ. 2020;370:m2516.
13. Kaza M, Karazniewicz-Lada M, Kosicka K, Siemiatkowska A, Rudzki PJ. Bioanalytical method validation: new FDA guidance vs. EMA guideline. Better or worse? J Pharm Biomed Anal. 2019;165:381-5.
14. Boukli N, Le Mene M, Schnuriger A, Cuervo NS, Laroche C, Morand-Joubert L, et al. High Incidence of False-Positive Results in Patients with Acute Infections Other than COVID-19 by the Liaison SARS-CoV-2 Commercial Chemiluminescent Microparticle Immunoassay for Detection of IgG Anti-SARS-CoV-2 Antibodies. J Clin Microbiol. 2020;58(11).
15. Piec I, English E, Thomas MA, Dervisevic S, Fraser WD, John WG. Performance of SARS-CoV-2 Serology tests: Are they good enough? medRxiv. 2020:2020.11.13.20229625.
16. Trabaud MA, Icard V, Milon MP, Bal A, Lina B, Escuret V. Comparison of eight commercial, high-throughput, automated or ELISA assays detecting SARS-CoV-2 IgG or total antibody. J Clin Virol. 2020;132:104613.
17. LeBaron CW, Forghani B, Matter L, Reef SE, Beck C, Bi D, et al. Persistence of rubella antibodies after 2 doses of measles-mumps-rubella vaccine. J Infect Dis. 2009;200(6):888-99.
18. Dimech W, Grangeot-Keros L, Vauloup-Fellous C. Standardization of Assays That Detect Anti-Rubella Virus IgG Antibodies. Clin Microbiol Rev. 2016;29(1):163-74.
19. White GH. Metrological traceability in clinical biochemistry. Ann Clin Biochem. 2011;48(Pt 5):393-409.
20. BIPM I, IFCC, ILAC, IUPAC, IUPAP, ISO, OIML. The international vocabulary of metrology — basic and general concepts and associated terms (VIM). 3rd ed: JCGM; 2012.
21. De Bièvre P. The 2012 International Vocabulary of Metrology: "VIM". Accreditation and Quality Assurance. 2012;17(2):231-2.

22. English E, Lenters-Westra E. HbA1c method performance: The great success story of global standardization. Crit Rev Clin Lab Sci. 2018;55(6):408-19.
23. Hoelzel W, Miedema K. Development of a reference system for the international standardization of HbA1c/glycohemoglobin determinations. J Int Fed Clin Chem. 1996;8(2):62-4, 6-7.
24. [ISO] IOfS. ISO 15194:2009 In vitro diagnostic medical devices — Measurement of quantities in samples of biological origin — Requirements for certified reference materials and the content of supporting documentation. 2009.
25. Dimech W, Panagiotopoulos L, Francis B, Laven N, Marler J, Dickeson D, et al. Evaluation of eight anti-rubella virus immunoglobulin g immunoassays that report results in international units per milliliter. J Clin Microbiol. 2008;46(6):1955-60.

## Figure Legends:

**Figure 1:** Linearity of method over the complete working range of the Abbott IgG II assay using a range of dilutions of a high positive (mean 38,365 AU/mL) in the Abbott diluent. Dash-dot grey line indicates the identity line. The black dotted line represents the 95% likelihood asymmetrical CI of the slope.

**Figure 2**: Cohen's Kappa concordance analysis of the assays and overall (all samples included) agreement of results given as %. Equivocal results were considered negative

**Figure 3:** Representative examples of the quantitative immune response in three different variants of the SARS-CoV-2 virus, including the 'UK' and 'South Africa' variants. The days post-PCR do not necessarily correlate to the day of onset of symptoms or the day of hospitalization.

**Figure 4:** Comparison graphs of the values obtained for the Technopath positive panel with different methods A) Abbott IgG II versus DiaSorin Liaison XL, B) Abbott IgG II versus EDI : C) Abbott IgG II Quantitative (S) versus Abbott IgG Qualitative (R),. Only Abbott quantitative assay showed linearity ($r^2$=0.9984) and was plotted against (A) DiaSorin, Quadratic ($r^2$=0.9988), (B) EDI, 4PL ($r^2$=0.9574) and (C) Abbott Qualitative, 4PL ($r^2$=0.9946).

**Figure 5:** Dilution of NIBSC working standard 20/162 using the Abbott Diluent. Dash-dot grey line indicate the identity line. The black dotted line represents the 95% likelihood asymmetrical CI of the slope