medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

SARS-CoV-2 antibody testing in a UK population: detectable IgG for up to 20 weeks post infection.

Authors: Louise J. Robertson[1], Julie S. Moore[1], Kevin Blighe[1], Kok Yew Ng[2], Nigel Quinn[3], Fergal Jennings[3], Gary Warnock[4], Peter Sharpe[3], Mark Clarke[5], Kathryn Maguire[5], Sharon Rainey[5], Ruth Price[1], William Burns[2], Amanda Kowalczyk[1], Agnes Awuah[1], Sara McNamee[2], Gayle Wallace[6], Steve Segar[1], Connie Chao Shern[7], M. M. Andrew Nesbit[1], James McLaughlin[2*], Tara Moore[1&7*].

[1]Biomedical Sciences Research Institute, Ulster University, Northern Ireland
[2]Nanotechnology and Integrated Bioengineering Centre, Ulster University, Northern Ireland
[3]Clinical Biochemistry Laboratory, Southern Health and Social Care Trust, Northern Ireland
[4]Microbiology Laboratory, Southern Health and Social Care Trust, Northern Ireland
[5]Northern Ireland Blood Transfusion Service, Belfast City Hospital, Northern Ireland
[6]Department of Anaesthesia, Royal Victoria Hospital, Belfast Health and Social Care Trust, Northern Ireland.
[7]Avellino, 1505 Adams Dr, Menlo Park, CA 94025, United States

*Joint corresponding authors- Professor Tara Moore tara.moore@ulster.ac.uk Professor James McLaughlin jad.mclaughlin@ulster.ac.uk

Abstract

Background

The SARS-CoV-2 pandemic necessitated rapid and global responses across all areas of healthcare, including an unprecedented interest in serological immunoassays to detect antibodies to the virus. The dynamics of the immune response to SARS-CoV-2 is still not well understood.

Methods

We measure SARS-CoV-2 antibody levels in plasma samples from 880 people in Northern Ireland by Roche Elecsys Anti-SARS-CoV-2 IgG/IgA/IgM, Abbott SARS-CoV-2 IgG and EuroImmun IgG SARS-CoV-2 ELISA immunoassays to analyse immune dynamics over time. Using these results, we develop a 'pseudo gold standard' reference cohort against which to assess immunoassay performance. We report performance metrics for the UK-RTC AbC-19 rapid lateral flow immunoassay (LFIA)

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

against a characterised panel of 304 positives established using the 'pseudo gold standard' system and 350 negative samples.

Results

We detect persistence of SARS-CoV-2 IgG up to 140 days (20 weeks) post infection, across all three antibody immunoassays, at levels up to 4.4 times the cut-off for a positive result by Roche measurement. Using our 'pseudo gold standard' cohort (n=348 positive, n=510 negative) we determine the sensitivity and specificity of the three commercial immunoassays used (EuroImmun; Sens. 98.9% [97.7-99.7%]; Spec. 99.2% [98.4-99.8%]; Roche; Sens. 99.4% [98.6-100%]; Spec. (96.7% [95.1-98.2%]; Abbott; Sens. 86.8% [83.1-90.2%]; Spec. (99.2% [98.4-99.8%]). The UK-RTC AbC-19 lateral flow immunoassay using shows a sensitivity of 97.70% (95.72%-99.34%) and specificity of 100% (100.00-100.00%).

Conclusions

Through comprehensive analysis of a large cohort of pre-pandemic and pandemic individuals, we show detectable levels of IgG antibodies, lasting up to 140 days, providing insight to immunity levels at later time points. We propose an alternative to RT-PCR positive status as a standard for assessing SARS-CoV-2 antibody assays and show strong performance metrics for the AbC-19 rapid test.

Introduction

The World Health Organization declared a pandemic in March 2020 due to severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), identified late 2019 in Wuhan, China, causing COVID-19 disease[1,2].

A global race ensued to develop diagnostic assays, with the most common being viral

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

RNA detection (RT-qPCR assays), to detect acute infection[3]. RT-qPCR assays are labour and reagent intensive, limited by a short temporal window for positive diagnosis, and exhibit potential for false negative results[4]. Evidence suggests sensitivity of RT-qPCR can be as low as 70%[5]. Lockdown measures and "flattening the curve" strategies meant many infected individuals were instructed to self-isolate and were not offered a diagnostic RT-qPCR, with much of the testing limited to patients admitted to hospital, who perhaps reflect a more severely infected cohort. Consequently, a potentially large number of cases were unconfirmed or undetected[6].

The ability to accurately detect SARS-CoV-2 specific antibodies, which develop after an immune response is evoked, is vital for building biobanks of convalescent sera for treatment, monitoring immune response to infection and assessing responses to vaccination programmes. Knowledge of antibody levels, indicative of prior virus exposure, will play an important role in public health policy and understanding SARS-CoV-2 epidemiology, including determining seroprevalence. Knowledge of the timing for when antibody levels can be measured is essential to allow accurate reporting.

Commercial serology immunoassays are mostly laboratory-based and measure IgG antibody levels in plasma or serum. Alternatively, lateral flow immunoassays (LFIAs), require a finger prick blood sample and can be used at point-of-care (POC) or in the home; particularly important in the context of lockdown enforcement during the pandemic. Currently, a limited number of laboratory-based chemiluminescence immunoassays are approved for use in the UK including the Roche Elecsys Anti-SARS-CoV-2 IgG/IgA/IgM against the SARS-CoV-2 Nucleocapsid antigenic region (Roche Diagnostics, Basel, Switzerland) and the Abbott SARS-CoV-2 IgG assay against the same antigenic region (Abbott Diagnostics, Abbott Park, IL, USA).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

The complexities of the humoral immune response to SARS-CoV-2 is a much-debated topic. The 'sero-silent', those who do not make IgG, may remain asymptomatic or conversely may become particularly unwell and unfortunately die[7]. Others produce IgG against different antigenic regions of the virus such as the spike or nucleocapsid protein. Patients who remain asymptomatic may mount a humoral immune response which is short-lived, with detectable levels of antibody falling after 2 months[8]. This, alongside potentially low sensitivity and lack of RT-PCR test availability across the UK has hindered development of well characterised serology samples known to contain IgG antibodies to SARS-CoV-2.

Herein, we describe use of Roche and Abbott commercial immunoassays, as well as the EuroImmun Anti-SARS-CoV-2 ELISA-IgG against the S1 domain of the spike antigenic protein of SARS-CoV-2 (EuroImmun UK, London, UK) to characterise a large number (880) of pre-pandemic and pandemic COVID-19 blood samples from within Northern Ireland and report on longevity of IgG antibodies detected. Presently, there is no gold standard assay for comparison. Therefore, we describe a 'pseudo gold-standard' against which to evaluate assays. Furthermore, we present results of an independent validation of the UK-RTC AbC-19 POC LFIA against a cohort of 304 known positives according to this 'pseudo gold-standard' system and 350 known negative samples for IgG to SARS-CoV-2.

Methods

Study Design

This study was approved by Ulster University Institutional Ethics committee (REC/20/0043), South Birmingham REC (The PANDEMIC Study IRAS Project

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

ID: 286041Ref 20/WM/0184) and adhered to the Declaration of Helsinki and Good Clinical Practice.

Participant samples

The flow of participant samples is summarised in Figure S1. All participants provided informed consent with no adverse events. An online recruitment strategy was employed, with the study advertised through internal Ulster University email, website and social media. A BBC Newsline feature providing the pandemic study email address also prompted interest from the general population.

The first 800 respondents who expressed interest were provided with an online patient information sheet, consent form and health questionnaire and invited to register to attend a clinic. Exclusion criteria related to blood disorder or contraindication to giving a blood sample. To enrich the cohort for positive samples, further participants were invited if they had previously tested PCR positive or had the distinctive symptom of loss of taste and smell. Blood sampling clinics were held at locations around Northern Ireland between April and July 2020 resulting in collection of 263 10ml EDTA plasma samples from 263 separate study participants. A small cohort (n=19) of anonymised plasma samples were obtained from Avellino, USA. Additional anonymised plasma samples were obtained from Southern Health and Social Care Trust (SHSCT) Biochemistry Laboratory (n=195), and Northern Ireland Blood Transfusion Service (NIBTS, n=184) through convalescent plasma programs.

Pre-pandemic samples (prior to June 2019, n=136) were obtained from Ulster University ethics committee approved studies with ongoing consent and from NIBTS

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

(n= 200, more than 3 years old). Plasma samples were used at no more than 3 freeze-thaw cycles for all analyses reported within this manuscript.

Clinical information

Basic demographic information and data about positive RT-PCR result and time from symptom onset was provided by PANDEMIC study participants through the secure online questionnaire. Anonymised participant samples from USA, SHSCT and NIBTS were provided with age, gender and time since PCR-positive, where a previous test had been carried out.

Laboratory-based immunoassays

Details of laboratory immunoassays are summarised in supplementary methods and Table S1.

UK-RTC AbC-19 LFIA

UK-RTC AbC-19 POC LFIA testing was conducted at Ulster University according to manufacturer's instructions (details in Table S1). Assays were performed as cohorts, with samples in batches of 10, with one researcher adding 2.5µL of plasma to the assay and a second adding 100µL of buffer immediately following sample addition. After 20 minutes, the strength of each resulting test line was scored from 0-10 according to a visual score card (scored by 3 researchers; Figure S2). A score ≥1 is positive. Details of samples for analytical specificity and sensitivity analysis are available in Supplementary methods.

Statistical analysis

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

As per Daniel WW[9] a minimum sample size based on prevalence can be calculated using the following formula: $n = \frac{Z^2 P(1-P)}{d^2}$, where n = sample size, Z = Z statistic for a chosen level of confidence, P = estimated prevalence, and d = precision. Assuming a prevalence of SARS-CoV-2 of 10% and a precision of 5%, we estimate that the required sample size at 99% confidence (Z = 2.58) to be 240 individuals. If the true prevalence is lower, 5%, the estimated required sample size given a precision of 2.5% is 506 individuals. A minimum sample size of 200 known positives and 200 known negatives was required for validation as per MHRA guidelines for SARS-CoV-2 LFIA antibody immunoassays[10].

Statistical analysis was conducted in in R v 4.0.2[11]. To assess discordance between test results, data was first filtered to include individuals with an Abbott test result in the range ≥0.25 & ≤1.4, with a 2 x 2 contingency table produced that comprised all possible combinations of [concordant|discordant] test results [within|outside of] this range. A p-value was derived via a Pearson χ^2 test after 2000 p-value simulations via the stats package.

To define a gold-standard, a positive result was determined as any individual who was positive by 2 out of 3 immunoassays, while negative was defined as negative by 2 out of 3 immunoassays. EuroImmun Borderline results were excluded. Sensitivity, specificity, and accuracy were then derived via the pROC package, with 95% confidence intervals produced after 2000 bootstraps. ROC analysis was performed via the pROC package. To compare test result (Positive|Negative) to age, a binary logistic regression model was produced with test result as outcome – a p-value was then derived via χ^2 ANOVA. To compare time against test result (encoded continuously), a linear regression was performed. We calculated median per time-

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

period and then converted these to log [base 2] ratios against the positivity cut-off for each assay. All plots were generated via ggplot2 or custom functions using base R[12].

Results

Antibody levels in plasma from 880 individuals were assessed using the three SARS-CoV-2 immunoassays; EuroImmun IgG, Roche Elecsys IgG/IgM/IgA and Abbott Architect IgG (Table S1). This included a negative cohort of 223 pre-pandemic plasma samples collected and stored during 2017 to end of May 2019 to determine assay specificity. Of the 657 participants whose samples were collected post-pandemic, 265 (40.33%) previously tested RT-PCR positive with a range of 7-173 days since diagnosis. A total of 225 participants gave time since self-reported COVID-19 symptoms, with a range of 5-233 days from symptom onset, whilst 198 had no symptom or PCR data available.

Laboratory based antibody immunoassays

A positive result for antibody on one or more of the three laboratory immunoassays was recorded for 385/657 (58.6%) post-pandemic participants. By EuroImmun ELISA, 346 were positive, 20 borderline and 291 were negative. The Roche assay detected 380 positive and 277 negative, whilst Abbott determined 310 positive and 347 negative (Table S2). The median age across all age groups combined was lower for participants testing positive across each of the immunoassays (median [sd] for positive versus negative, respectively: EuroImmun, 41 [13.16] vs 48 [12.95]; Roche, 42 [13.08] vs 48 [13.00]; Abbott, 41 [13.18] vs 47 [13.09]). (Figure S3, $p<0.0001$). When segregated by age group, however, differences were less apparent in certain groups (Figure S4). Excluding the pre-pandemic cohort, this gap reduced but remained statistically

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

significant EuroImmun, 41 [13.18] vs 45 [12.49]; Roche, 42 [13.15] vs 45 [12.49]; Abbott, 41 [13.26] vs 44 [12.63]) (p<0.01) (median [sd] for positive versus negative). Of note, out of 265 individuals with a previous positive RT-PCR result, 14 (5.2%) did not show any detectable antibodies by all three immunoassays, with no association with age, gender or time between test and blood draw (data not shown).

The three commercial laboratory immunoassays provide a ratio value that increases with IgG antibody titre. When correlation between these values is assessed, good overall agreement is observed between the three immunoassays (Figure 1, Figure S5). As highlighted by Rosadas *et al*., we also see significant disagreement in the Abbott 0.25-1.4 range when compared to EuroImmun and Roche (Figure 1a,b; chi-square p-values: EuroImmun vs Abbott, p<0.001; Roche vs Abbott, p<0.001)[13].

Duration of humoral response to SARS-CoV-2

We found IgG antibodies could still be detected in individuals (excluding pre-pandemic) across all three immunoassays used up to week 20 (day 140) (Figure 2). We note a statistically significant decrease in signal with respect to time across each assay (p-value [slope]): EuroImmun, p=0.036 [-0.785]; Roche, p=0.002 [-0.125]; Abbot, p<0.0001 [-3.585]. These remained statistically significant after adjustment for age. Antibody levels (expressed as a ratio of median result per timepoint divided by positivity cut off; Figure 2d) peaked at Week 1-2 for EuroImmun (1.33) and Abbott (1.64), though reached highest levels at Week 8-12 when measured by Roche (5.45). By week 21-24, median score for all tests had dropped below the positivity cut off, though a small number of RT-PCR positive samples remained above the positive cut off at these later timepoints (Figure 2).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Developing a 'pseudo gold-standard'

Use of a previous positive RT-PCR for SARS-CoV-2 viral RNA as a gold standard for the presence of IgG has limitations, therefore we developed a 'pseudo gold-standard' based on laboratory immunoassay antibody results, in which a positive result by two or more immunoassays classes a sample as antibody positive (n=348), whilst a negative result on two or more immunoassays classes a sample as negative (n=510). EuroImmun borderline results were excluded from this analysis (n=22). When assessed against the *pseudo-gold standard,* EuroImmun performed with highest accuracy (Sens. 98.9% [97.7-99.7%]; Spec. 99.2% [98.4-99.8%]; Acc. 99.1% [98.4-99.7%]). Roche showed highest sensitivity and high accuracy (Sens. 99.4% [98.6-100%]; Spec. (96.7% [95.1-98.2%]; Acc. (97.8% [96.7-98.7%]), whilst Abbott performed poorly on sensitivity but had the highest specificity (Sens. 86.8% [83.1-90.2%]); Spec. (99.2% [98.4-99.8%]); Acc. (94.2% [92.7-95.7%]).

Area under the curve indicates best performance by EuroImmun 0.99 (0.984; 0.997), followed by Roche 0.98 (0.972; 0.989), then Abbott 0.93 (0.912; 0.948) (Figure 3).

UK-RTC AbC-19

Using the commercial immunoassays described we established a well characterised serology sample set of 'known positive' and 'known negative' for IgG antibodies to SARS-CoV-2 to evaluate performance metrics for the UK-RTC AbC-19 Rapid POC LFIA.

This LFIA detects IgG antibodies against the spike protein antigen, so we therefore required all samples to be positive by the EuroImmun SARS-CoV-2 IgG ELISA, which likewise detects antibodies against the S1 domain[14]. In line with our 'pseudo gold standard' system, samples were also required to be positive by a second

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

immunoassay (Roche or Abbott). To analyse specificity, we assessed 350 plasma samples from participants classed as 'known negative' on the AbC-19 LFIA. All samples were from individuals confirmed to be negative across all three laboratory assays (Roche, EuroImmun, Abbott). Using these positive n= 304 and negative n=350 antibody cohorts, we determined a sensitivity of 97.70% (95.72%-99.34%) and specificity of 100% (100.00-100.00%) for the AbC-19 LFIA (Table 1).

When used as intended in POC settings, the AbC-19 LFIA provides binary positive/negative results. However, when assessing LFIA in the laboratory, each test line was scored against the scorecard by three independent researchers (0 negative, 1-10 positive; Figure S2). Compared to quantitative outputs from the Abbott, EuroImmun and Roche assays, the AbC-19 LFIA shows strong correlation (Abbott r=0.86 [p<0.001]; EuroImmun r=0.88 [p,0.001]; Roche r=0.83 [p<0.001]; Figure 4, Figure S5).

Analytical specificity and sensitivity of AbC-19 test

We observed no cross-reactivity across samples with known H5N1 influenza, Respiratory syncytial virus, Influenza A, Influenza B, Bordetella Pertussis, Haemophilus Influenzae, Seasonal coronavirus NL63 and 229E on the AbC-19 LFIA. (n=34 samples, n=8 distinct respiratory viruses; Table S3). Against a panel of external reference SARS-CoV-2 serology sample, the AbC-19 LFIA detected antibodies with scores commensurate to the EuroImmun ELISA scores (Figure S6, Table S4).

Discussion

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Serological antibody immunoassays are an important tool in helping combat the SARS-CoV-2 pandemic. One difficulty faced in validation of antibody diagnostic assays has been access to samples with known SARS-CoV-2 antibody status. As previously described, there is no clear gold standard for reference against which to assess SARS-CoV-2 immunoassays. A positive RT-PCR test has been used previously as a reference standard, although limited by a high rate of false negatives, failure in some cases to develop IgG antibodies (sero-silence or lack of antibody against the same antigenic component of the virus as the immunoassay uses as a capture antigen), or the lack of RT-PCR testing availability early in the pandemic[3,5,15]. Self-assessment of symptoms for COVID-19 disease is a poor indicator of previous infection, even among healthcare workers[16]. Asymptomatic individuals may be unaware of infection and others may harbour pre-existing immunity or elucidate a T cell response.

Our results show strong correlation between all three immunoassays, with shortcomings in the Abbott system output 0.25-1.4 range, as described previously, suggesting an overestimated positive cut-off (Figure 1)[13]. Our detection of antibodies 140 days after RT PCR positive status (20 weeks, and beyond in a small number of samples) indicates persistence IgG antibodies to both the spike protein and nucleocapsid protein, despite typical patterns of antibody decay after acute viral antigenic exposure being rapid[17]. Where others have reported SARS-CoV-2 antibodies decline at 90 days, we also noted a statistically significant decline over time but levels remain detectable at 140 days (Figure 2). We note that IgG levels reach their peak as late as Week 8-12 (Roche ratio 5.45 times threshold cut-off)[18], though this may be an artefact of lower number of participants at earlier timepoints

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

(Figure 2d). Longitudinal studies on SARS-CoV convalescent patients suggests that detectable IgG can still be present as long as 2 years after infection[20]. Further studies are needed on large cohorts with sequential antibody immunoassays performed on symptomatic and non-symptomatic individuals as well as those with mild and severe COVID-19. This is vital to inform vaccine durability, so-called 'immune passports' and in the definition of a protective threshold for anti-SARS-CoV-2 antibodies. This broader detection window supports the use of these immunoassays for seroprevalence screening[19].

To assess sensitivity and specificity, we developed a 'pseudo gold-standard' against which to analyse assays, which does not rely on a single test as reference. A similar approach was used in a recent seroprevalence study in Iceland, whereby two positive antibody results were required to determine a participant positive[15]. The sensitivity and specificity we observed for these laboratory immunoassays did not differ greatly from Public Health England evaluations, though EuroImmun performed with higher sensitivity (98.9%) than the reported 72% and Abbott with lower (86.8%) than the reported 92.7%[21–23].

Our evaluation of performance metrics for the rapid POC finger-prick AbC-19 LFIA gave 97.7% sensitivity and 100% specificity. Strong correlation was observed in quantitative score between results on all immunoassays with the highest observed between EuroImmun and AbC-19 LFIA (Figure S5). This is to be expected, given both the AbC-19 LFIA and EuroImmun ELISA detect antibodies against spike protein. For the assessment of immunity to prior natural infection as well as to immunisation, it is important to note IgG antibodies against SARS-CoV-2 spike protein detected by

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

laboratory-based EuroImmun ELISA and AbC-19 POC LFIA are known to correlate with neutralizing antibodies, which may confer future immunity [24,25].

The 'pseudo gold-standard' system may artificially raise the threshold for positive sample inclusion, by requiring a positive result by two immunoassays. However, similar issues have been raised when using previous RT-PCR result or definitive COVID-19 symptoms as inclusion criteria given these will likely skew a cohort towards more severe disease, which may increase immune response[5]. In the absence of a clear gold standard test, our 'pseudo-gold standard' system relies on no single test and instead takes an average of three. We observed a high number of positive participants in our cohorts (n=348; 40.9%), a reflection of biases in the recruitment methods, enriching for positive samples.

The validation of the AbC-19 LFIA reported excellent performance metrics and we note it uses plasma from venous blood samples, as opposed to the use of a finger prick blood sample. A matrix study comparing the result from finger prick blood versus plasma is under investigation by this PANDEMIC Study team @Ulster University. When this LFIA was used on our cohort, a number of the positive results scored low, 1/10 (using the score card under laboratory conditions) with a faint test band, visible to a trained laboratory scientist but perhaps difficult to identify as positive by the general public conducting a self-test (Figure S6). This faint line may be reflective of the longer time from infection. If this AbC-19 LFIA is to be rolled out as a home-testing kit, it is important to determine if members of the public observe the same results as observed in the laboratory. Again, this is the aim of an ongoing usability study by the PANDEMIC study team @Ulster University.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Lastly, it is important to consider prevalence when interpreting an assays' sensitivity and specificity as in a low prevalence scenario, as even slightly lowered performance metrics (such as the MHRA required 98%) can result in large numbers of false negative and false positive results[10,26]. Given the high sensitivity of the AbC-19 LFIA observed in this study, a false positive result is unlikely, however false negatives may occur. For prevalence studies, this may underestimate true prevalence (though could be corrected statistically). At an individual level, a false negative result may cause anxiety or over-caution- particularly if we progress through this pandemic and become more informed about SARS-CoV-2 immunity. A positive antibody test may be linked to prevention of future infection following COVID-19 illness or post-vaccination.

Acknowledgements: We are extremely grateful to all the people of Northern Ireland who took part in this study and gave blood during the pandemic. We are indebted to the phlebotomists- Geraldine Horrigan, David Hunter and Pamela Taylor who conducted all the blood draws whilst ensuring the highest possible level of safety to the participants. We are also extremely grateful to Kingsbridge Private Hospital for sponsorship and providing everything needed for blood collection including the clinical space. We acknowledge Dr Tony Byrne for use of the SafeWater laboratory and Professor Gareth Davison for laboratory space and equipment during the pandemic within a locked down University.

Declaration of interests:

At the time of this study TM and JML acted as advisors to CIGA HealthCare, an industrial partner in the UK Rapid Test Consortium. No personal financial reward or

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

renumeration was received for this advisory role. At the time of submission of this manuscript TM and JML no longer held these advisory positions.

All other authors have no potential conflict of interest to report.

Funding statement:

Funding for this study was obtained from UK-RTC to cover all laboratory costs incurred.

References

1 World Health Organisation. Rolling updates on coronavirus disease (COVID-19). 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen (accessed Aug 11, 2020).

2 Lu R, Zhao X, Li J, *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 2020; **395**. DOI:10.1016/S0140-6736(20)30251-8.

3 Petherick A. Developing antibody tests for SARS-CoV-2. *Lancet* 2020; **395**: 1101–2.

4 Winichakoon P, Chaiwarith R, Liwsrisakun C, *et al.* Negative nasopharyngeal and oropharyngeal swabs do not rule out COVID-19. J. Clin. Microbiol. 2020; **58**. DOI:10.1128/JCM.00297-20.

5 Watson J, Richter A, Deeks J. Testing for SARS-CoV-2 antibodies. *BMJ* 2020; **370**. DOI:10.1136/bmj.m3325.

6 Black JRM, Bailey C, Przewrocka J, Dijkstra KK, Swanton C. COVID-19: the case for health-care worker screening to prevent hospital transmission. *Lancet*

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

(London, England) 2020; **395**: 1418–20.

7 Buss LF, Prete CA, Abrahim CMM, *et al.* COVID-19 herd immunity in the Brazilian Amazon. *medRxiv* 2020; : 2020.09.16.20194787.

8 Long Q-X, Tang X-J, Shi Q-L, *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* 2020; **26**: 1200–4.

9 Daniel WW. Biostatistics : a foundation for analysis in the health sciences, 7th Editio. New York: John Wiley & Sons, Ltd, 1999.

10 Medicines and Healthcare product Regulatory Agency. Target product profile: antibody tests to help determine if people have immunity to SARS-CoV-2. 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/881162/Target_Product_Profile_antibody_tests_to_help_determine_if_people_have_immunity_to_SARS-CoV-2_Version_2.pdf (accessed April 24, 2020).

11 The R Development Core Team. R: A language and environment for statistical computing. ISBN 3-900051-07-0. 2017.

12 Wickham H. ggplot2 Elegant Graphics for Data Analysis. 2016 DOI:10.1007/978-0-387-98141-3.

13 Rosadas C, Randell P, Khan M, McClure MO, Tedder RS. Testing for responses to the wrong SARS-CoV-2 antigen? *Lancet (London, England)* 2020; **396**: e23.

14 UK-RTC and Abingdon Health. Charting the course to a post-COVID world. 2020.

15 Gudbjartsson DF, Norddahl GL, Melsted P, *et al.* Humoral Immune Response to SARS-CoV-2 in Iceland. *N Engl J Med* 2020; published online Sept 1.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

DOI:10.1056/NEJMoa2026116.

16 Mulchandani R, Taylor-Phillips S, Jones H, *et al.* Self assessment overestimates historical COVID-19 disease relative to sensitive serological assays: cross sectional study in UK key workers. *medRxiv* 2020; : 2020.08.19.20178186.

17 Andraud M, Lejeune O, Musoro JZ, Ogunjimi B, Beutels P, Hens N. Living on Three Time Scales: The Dynamics of Plasma Cell and Antibody Populations Illustrated for Hepatitis A Virus. *PLoS Comput Biol* 2012; **8**: e1002418.

18 Ibarrondo FJ, Fulcher JA, Goodman-Meza D, *et al.* Rapid Decay of Anti–SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N Engl J Med* 2020; **383**: 1085–7.

19 Deeks JJ, Dinnes J, Takwoingi Y, *et al.* Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst Rev* 2020; published online June 25. DOI:10.1002/14651858.CD013652.

20 Wu L-P, Wang N-C, Chang Y-H, *et al.* Duration of antibody responses after severe acute respiratory syndrome. *Emerg Infect Dis* 2007; **13**: 1562–4.

21 Public Health England. Evaluation of the Abbott SARS-CoV-2 IgG for the detection of anti-SARSCoV-2 antibodies. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/890566/Evaluation_of_Abbott_SARS_CoV_2_IgG_PHE.pdf.

22 Public Health England. Evaluation of the Euroimmun Anti-SARS-CoV-2 ELISA (IgG) serology assay for the detection of anti-SARS-CoV-2 antibodies. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/893433/Evaluation_of_Euroimmun_SARS_CoV_2_ELISA_

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

IgG__1_.pdf.

23 Public Health England. Evaluation of Roche Elecsys AntiSARS-CoV-2 serology assay for the detection of anti-SARS-CoV-2 antibodies. 2020.

24 Iyer AS, Jones FK, Nodoushania A, *et al.* Dynamics and significance of the antibody response to SARS-CoV-2 infection. *medRxiv* 2020; : 2020.07.18.20155374.

25 Addetia A, Crawford KHD, Dingens A, *et al.* Neutralizing antibodies correlate with protection from SARS-CoV-2 in humans during a fishery vessel outbreak with high attack rate. *medRxiv* 2020; : 2020.08.13.20173161.

26 SPI-B Scientific Pandemic Influenza Group on Behaviours. SPI-B: Pre-empting possible negative behavioural responses to COVID-19 antibody testing, 13 April 2020. 2020 https://www.gov.uk/government/publications/spi-b-pre-empting-possible-negative-behavioural-responses-to-covid-19-antibody-testing-13-april-2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table 1: UK-RTC AbC-19 LFIA performance metrics.

Test	**AbC-19**
Total Negative	350
True Negative	350
False Positive	0
Total Positive	304
True Positive	297
False Negative	7
Sensitivity % (95 CI)	97.70% (95.72%-99.34%)
Specificity % (95 CI)	100.00% (100.00%-100.00%)
Positive predictive value	100.00%
Negative predictive value	98.04% (96.42%-99.43%)
Accuracy % (95 CI)	98.93% (98.01%-99.69%)

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Figures

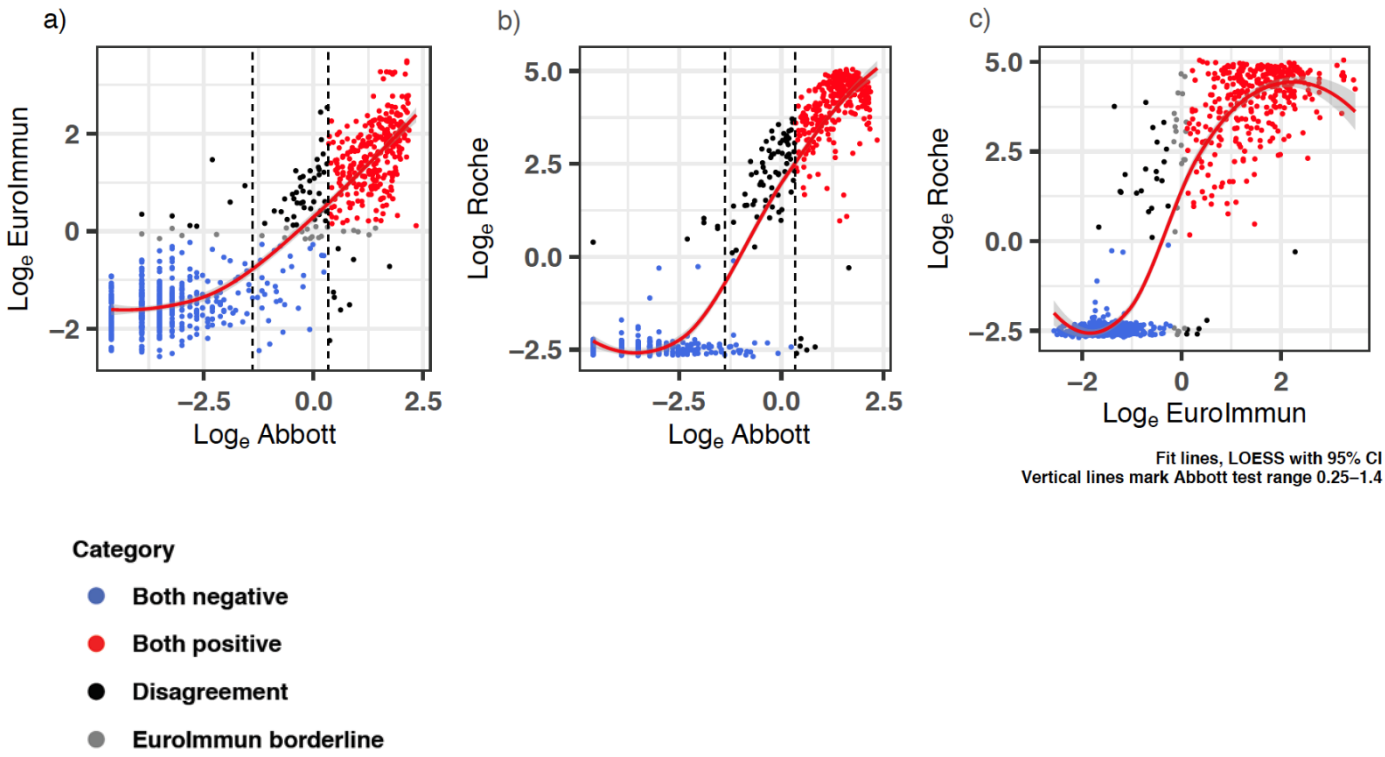
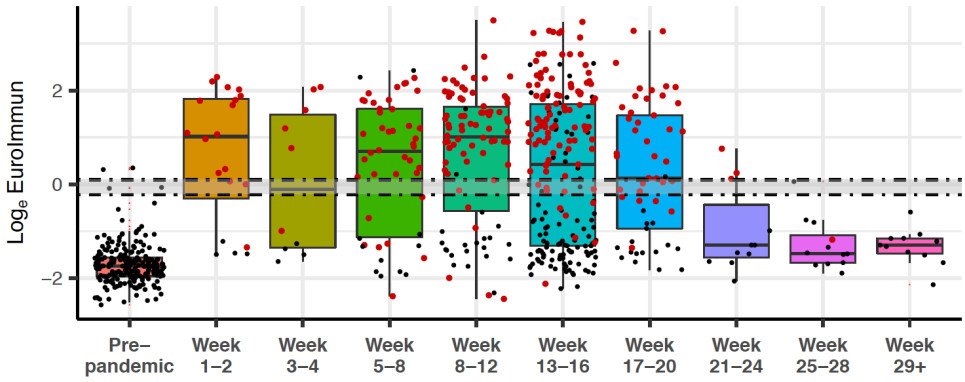


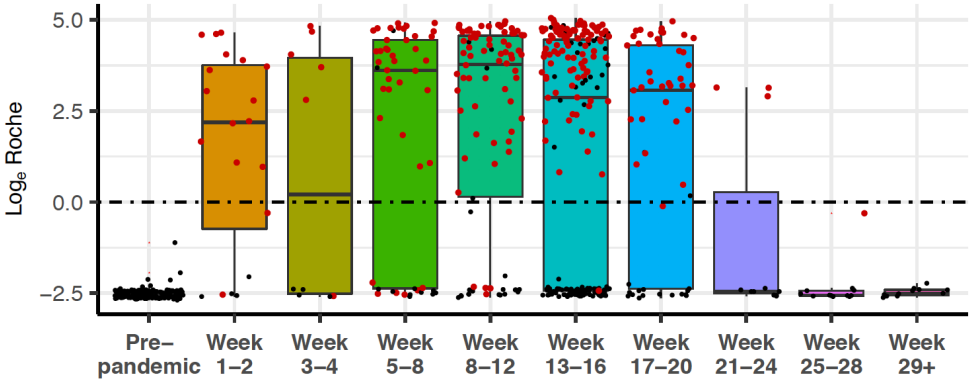
Figure 1: Two-way correlation scatter plots comparing a) EuroImmun b) Abbott and c) Roche immunoassays. Pearson χ^2 test was used to assess correlations. The results for each test were log transformed to ensure results follow a normal distribution. Negative agreement shown as blue dots, red dots show positive agreement for the two immunoassays, whilst black dots show disagreement and grey dots as the EuroImmun borderline results. Vertical lines mark the Abbott test range 0.25-1.4. n=880. The graphs show positive correlations between all immunoassays evaluated, with the fewest disagreement of results between the Log of Roche and the Log of EuroImmun. Fit lines LOESS, with 95% confidence interval shaded.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

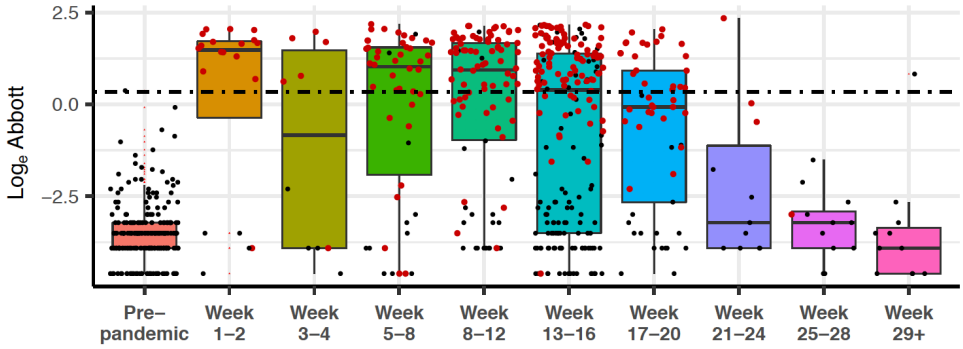
a)



b)



c)



d)

	Ratio Antibody level : assay positivity cut-off									
	Week									
	Pre-2020	1-2	3-4	5-8	8-12	13-16	17-20	21-24	25-28	29+
EuroImmun	-2.65	1.33	0.2	0.87	1.32	0.47	0.04	-2.01	-2.26	-2.01
Roche	-3.64	3.16	3.05	5.21	5.45	4.14	4.42	-3.54	-3.69	-3.61
Abbott	-5.54	1.64	-0.51	0.99	0.86	0.08	-0.59	-5.13	-5.13	-6.13
Sample number (n=)	223	20	10	50	90	202	53	11	12	11

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Figure 2: SARS-CoV-2 antibody levels by (a) EuroImmun, (b) Roche, and (c) Abbott, relative to weeks since first reported symptoms or positive PCR result (where data available). RT-PCR positive individuals are denoted by red dots, while individuals with time since symptom data are denoted in black. Dashed lines delineate $\log_e$ equivalent of positivity threshold (EuroImmun 1.1, Roche 1.0, Abbott 1.4) for each test, and the negativity threshold for EuroImmun (0.8; borderline result between the two lines). Black bars indicate median, within IQR (interquartile range) boxes for EuroImmun/Roche/Abbott value. Red triangles indicate outliers, based on 1.5* IQR (interquartile range). **(d)** Antibody level ratios for assays over time show varying peaks levels depending on test. Calculated by first establishing the median per time period, then calculating log2 ratio for each period versus each respective assay positivity cut-off.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

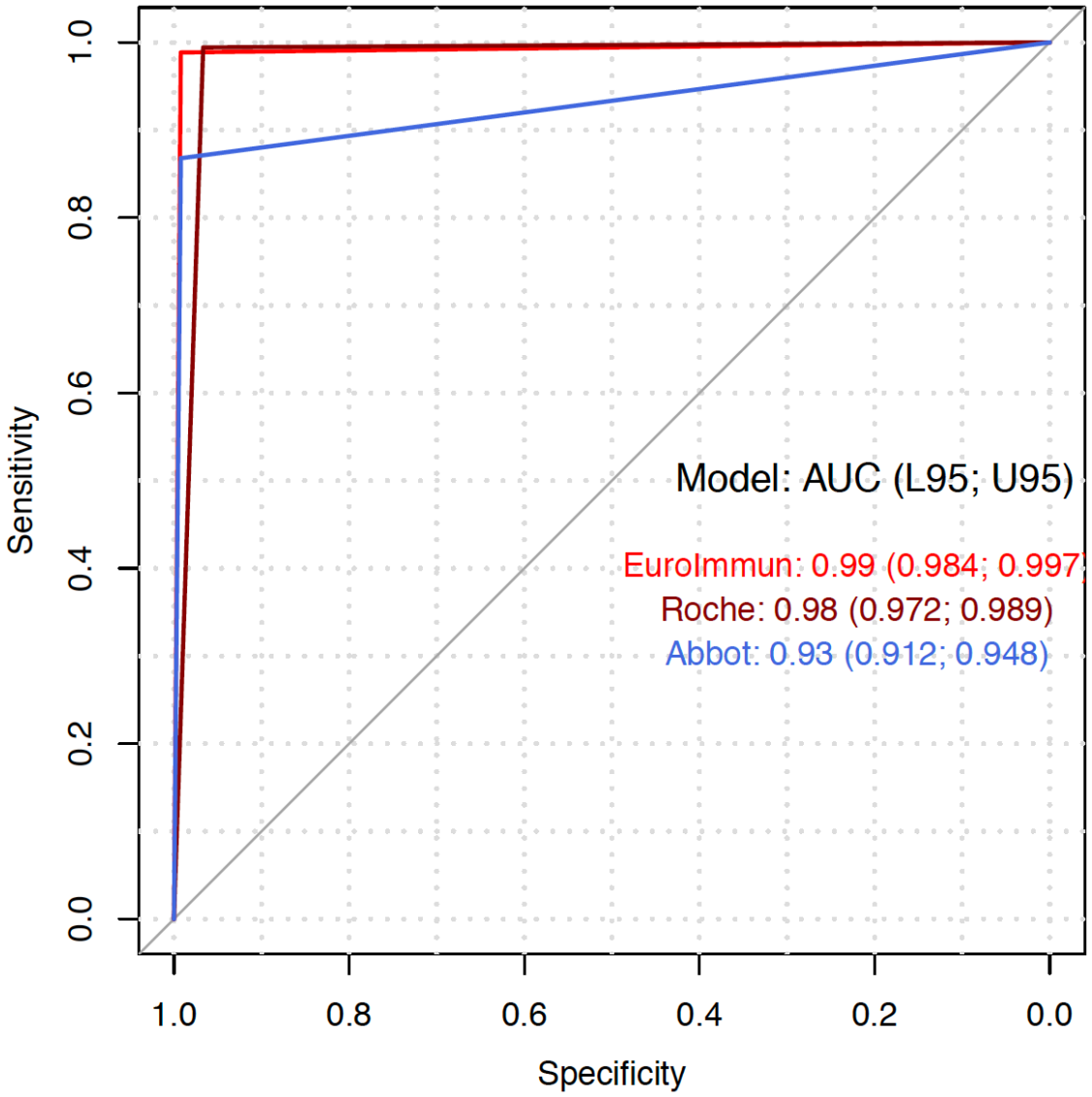


Figure 3: Receiver operating curve for Abbott, EuroImmun and Roche antibody immunoassays. To evaluate the performance and accuracy of each test binary logistic regression was carried out and the area under the curve (AUC) analysis was carried out to include lower and upper 95% confidence limits for each test; EuroImmun (red), Roche (maroon), Abbott (blue). The AUC for each test is close to 1, indicating the immunoassays perform with a high level of accuracy. Diagonal grey line is the reference line (AUC=0.5).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20201509.this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

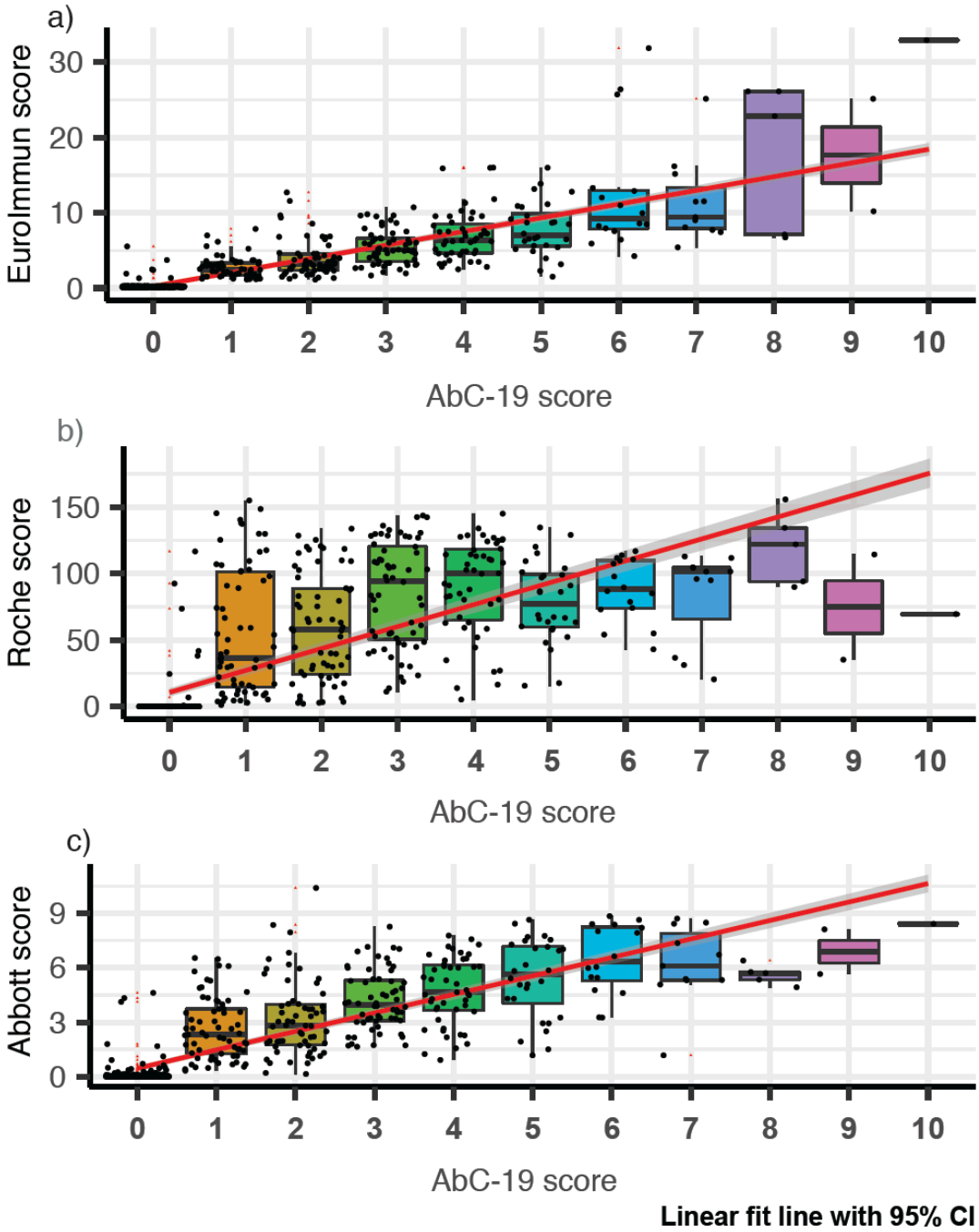


Figure 4: AbC-19 correlation to a) EuroImmun b) Roche and c) Abbott scores. Box plots overlaid on scatter plot, comparing AbC-19 test scores to EuroImmun, Roche and Abbott quantitative antibody values. Red linear line of best fit with 95% confidence interval shaded in grey. Black bars indicate median, within IQR (interquartile range) boxes for EuroImmun/Roche/Abbott value. Red triangles indicate outliers, based on 1.5* IQR (interquartile range).

Supplementary Materials

Supplementary Methods

Laboratory-based immunoassays

Researchers were blinded to other test results when processing these assays.

EuroImmun Anti-SARS-CoV-2 ELISA-IgG (EuroImmun, EI 2606-9601 G) was carried out according to manufacturer's instructions. Optical density (OD) at 450nm and reference OD at 620nm was read on BMG Labtech Fluostar Omega spectrophotometer (BMG Labtech). Ratios were calculated by dividing absorbance of the clinical sample by the absorbance of EuroImmun calibrator, with a score of < 0.8 determined negative, ≥ 0.8 to <1.1 borderline and ≥ 1.1 positive. For a portion of samples provided by NIBTS, EuroImmun IgG assay data was provided to researchers by NIBTS.

Roche Elecsys immunoassay (Roche Diagnostics, kit 09203079190) was carried out according to manufacturer's instructions on the Roche cobas e601 (C6000 line) or e801 (C8000 line) analysers. The analyser automatically calculates the cut-off based on the measurement of ACOV2 Cal1 (negative) and ACOV2 Cal2 (positive). The result of a sample is given either as reactive or non-reactive as well as in the form of a cut-off index (COI; signal sample/cut-off). A score of <1.0 is determined negative, while a score ≥ 1.0 is positive.

Abbott Architect SARS-CoV-2 immunoassay was carried out according to manufacturer's instructions on the Abbott Architect i2000SR analyser (Abbott,

kit 18115FN00, calibrator kit 17412FN00, Control kit 17531FN00). The external control is entered into a Quality Monitor programme and must be within 3 standard deviations of the mean (cumulative; External control NIBSC QCRSARSCoV-2QC1 Lot 20/B764-01). Results are reported by dividing the sample result by the calibrator result. The result unit for the SARS-CoV-2 IgG assay is Index (Sample/Calibrator). A ratio of < 1.4 is determined negative and ≥ 1.4 is determined positive.

Analytical specificity and sensitivity assessment

Four virology samples (H5N1 influenza serology 7/150, RSV serology 16/284, Influenza B 9/222 and Bordetella Pertusis 89/530) were obtained from NIBSC (National Institute for Biological Standards, Herts, UK). An additional 30 serology samples from known virus infections were a kind gift from SugenTech, Soeul, Korea. 15 of these virology samples were obtained from Trina (Trina Bioreactives AG, Switzerland) from 5 different individuals per virus (Influenza A IgG, Influenza B IgG and RSV IgG). A further 15 of these virology samples were obtained from AbBaltris, Kent from 5 different individuals per virus (Haemophilus Influenza IgG, Seasonal Coronavirus NL63 and 229E Seasonal Coronavirus). All these serology samples alongside a panel of 6 external standard research reagents (Table S4; NIBSC; Cat: 20/118 and 20/130) were assessed on the AbC-19 LFIA to confirm analytical specificity and sensitivity.

Table S1: Summary specifications for SARS-CoV-2 immunoassays investigated.

Immunoassay	Principle	Antigen Target	Assay Time (min)	Antibody Detected	Measurement	Result	Calibration	Evaluation of results	Results
EuroImmun ELISA	Enzyme-linked immunosorbent assay (enzyme-HRP)	S1 domain of the spike protein	120	IgG	Photometric measurement of the color intensity using wavelength of 450 nm and a reference wavelength between 620 nm and 650 nm	OD (Optical density)	One Positive calibrator	OD of clinical sample/OD of calibrator	< 0.8 Negative, ≥ 0.8 to <1.1 Borderline, ≥ 1.1 Positive
Roche Elecsys immunoassay	Electro-chemiluminescence	Nucleocapsid	18	IgG, IgA and IgM	Application of a voltage to the electrode then induces chemiluminescent emission which is measured by a photomultiplier	RLU (Relative Light Intensity)	One Positive calibrator and one Negative calibrator	The analyzer automatically calculates the cut-off based on the measurement of ACOV2 Cal1 (negative) and ACOV2 Cal2 (positive). The result of a sample is given either as reactive or non-reactive as well as in the form of a cut-off index (COI; signal sample/cut-off).	< 1.0 Negative, ≥ 1.0 Positive
Abbott Architect SARS-CoV-2	Chemiluminescent microparticle immunoassay	Nucleocapsid	30	IgG	The resulting chemiluminescent reaction is measured as a relative light unit (RLU).	RLU (Relative Light Intensity)	One Positive calibrator	Results are reported by dividing the sample result by the calibrator result (mean of 3 calibrators). The default result unit for the SARS-CoV-2 IgG assay is Index (S/C).	< 1.4 Negative, ≥ 1.4 Positive
AbC-19	Rapid Point of Care Lateral Flow Immunoassay	Full length Spike protein	20	IgG	The colour intensity of the test line is analysed using the reference score card.	Binary	The presence of a control line indicates the test is valid.	A result is positive if there is both a test line and a control line, whilst a result is negative if only the control line is present.	Using the reference score card; Positive scores ≥1 Negative scores=0

Table S2: Post-pandemic participant laboratory-based SARS-CoV-2 antibody result. Breakdown of individual immunoassay results or result by one or more test.

Test	Positive (%)	Borderline (%)	Negative (%)
Abbott	310/657 (47.2%)	n/a	347/657 (62.8%)
EuroImmun	346/657 (52.7%)	20/657 (3.2%)	291/657 (44.4%)
Roche	380/657 (57.8%)	n/a	277/657 (42.2%)
One or more test	385/657 (58.6%)	3/657 (0.45%)	269/657 (40.9%)

Table S3: Analytical specificity analysis on the AbC-19 LFIA LFIAs were assessed using 34 serum samples with known other respiratory viruses, negative results for all suggests analytical specificity for SARS_CoV_2 IgG.

SAMPLE	Number of samples	Number of AbC-19 Positive results	Number of AbC-19 Negative results
H5N1 Influenza (NIBSC 7/150)	1	0	1
RSV (NIBSC 16/284)	1	0	1
Influenza B (NIBSC 9/222)	1	0	1
Bordetella Pertussis (NIBSC 89/530)	1	0	1
Influenza A	5	0	5
Influenza B	5	0	5
Respiratory syncytial virus	5	0	5
Haemophilus Influenzae	5	0	5
Seasonal coronavirus NL63	5	0	5
Seasonal coronavirus 229E	5	0	5

Table S4: AbC-19 LFIA results with NIBSC external reference samples

NIBSC standard serology samples were provided with a data sheet indicating the SARS-CoV-2 antibody levels. We measured SARS-CoV-2 antibody levels in these samples and obtained similar results with the EuroImmun IgG ELISA in our laboratory.

NIBSC #	AbC-19 LFIA result	Ulster University lab result	NIBSC provided antibody data				
		EuroImmun IgG (S1 domain)	EuroImmun IgG (S1 domain)	EuroImmun IgA	In-house IgG S1	In-house IgG N	In-house IgG sSpike
20/120	pos (10)	pos (8.39)	pos (8.59)	pos (10.1)	5580	3417	2693
20/122	pos (7)	pos (3.49)	pos (3.47)	pos (1.1)	3202	2425	1488
20/124	pos (1)	pos (1.56)	pos (1.62)	pos (1.84)	1636	3296	118
20/126	pos (1)	neg (0.60)	neg (0.64)	pos (1.63)	1181	995	8
20/128	neg (0)	neg (0.23)	neg (0.21)	neg (0.02)	<50	<50	<50
20/130	pos (8)	pos (6.96)	pos (7.77)	pos (9.74)	5388	17197	2707

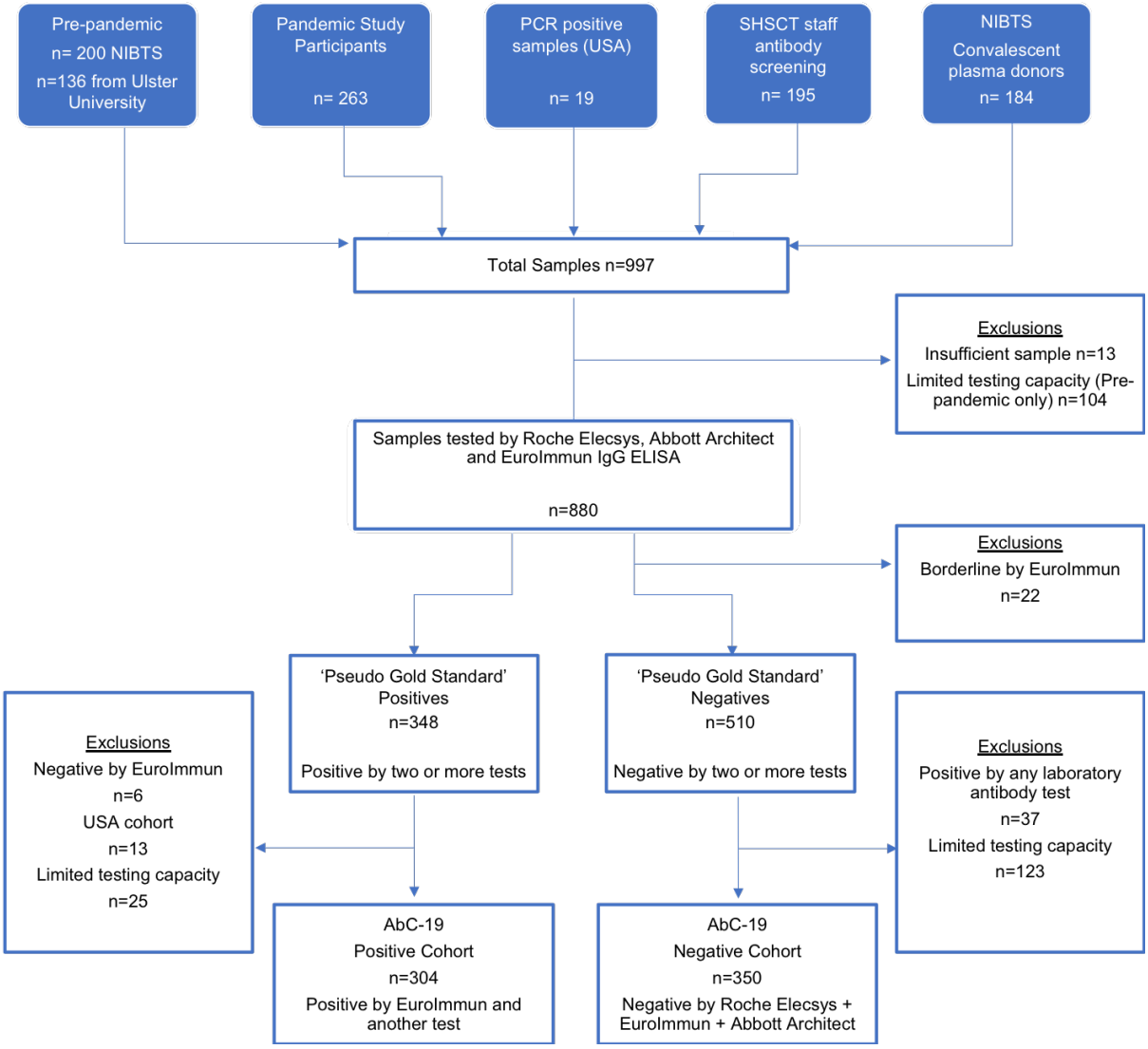


Figure S1: Flow of participant plasma samples through the study.

All available samples from participants within each cohort, and the included and excluded samples at all stages. Freeze thaw cycles were closely monitored for all sample aliquots.

Figure S2: Visual Score card for quantitative interpretation of AbC-19 LFIA test bands. A scale of 0 (not pictured, negative-no test line visible) to 10 (positive-strongest test line). Any LFIA scoring 1 or above was classified as positive.

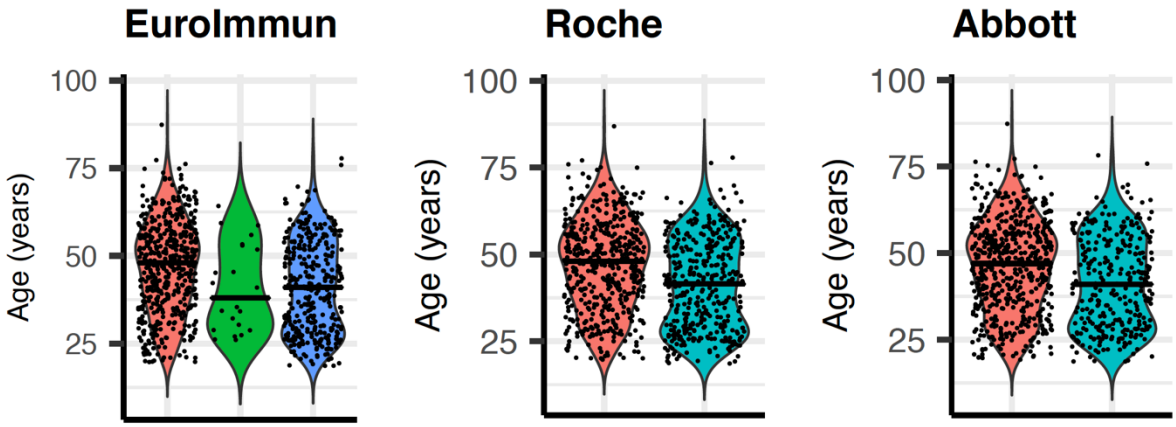


Figure S3: Age violin plots overlaid with scatter for samples included in correlation analysis (where age data available) n=880.

The above graphs allow comparison of the distributions and probability density of ages for EuroImmun, Roche and Abbott immunoassays. Wider areas of the violin plot represent high probability density, whilst narrow areas represent low probability density. Horizontal bar indicates median age. The red violin plots represent the negative results, the green violin plot represent the borderline results and the blue/turquoise violin plots represent the positive results.

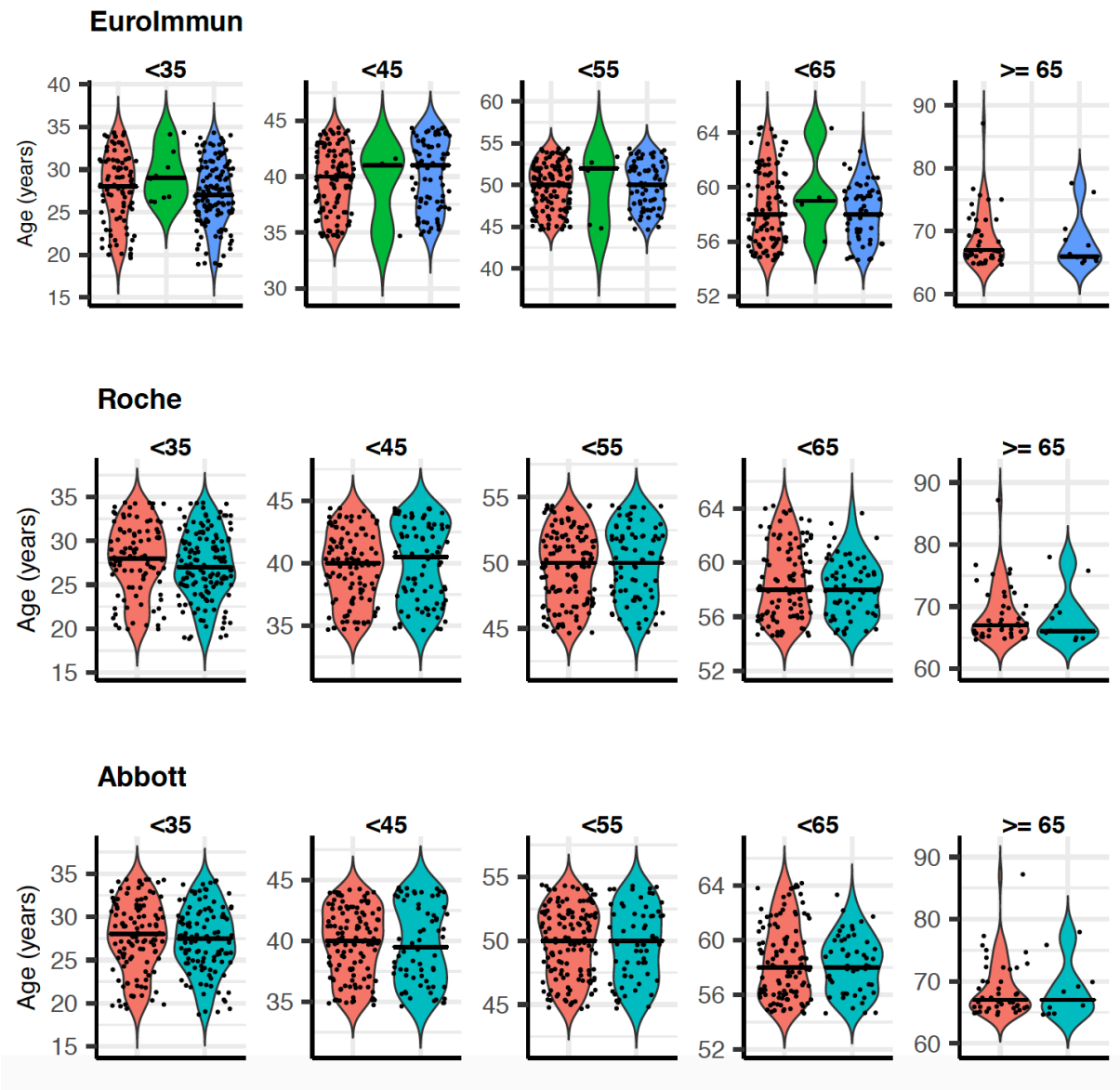


Figure S4: Age violin plots separated into age groups (where age data available) for samples included in correlation analysis.

The above figure presents graphs for each immunoassay (EuroImmun, Roche and Abbott) with the corresponding age groups <35 years, <45 years, <55 years, <65 years and >= 65 years. The red violin plots represent the negative results, the green violin plot represents the borderline EuroImmun results, and the blue/turquoise violin plots represent the positive results (n=848).

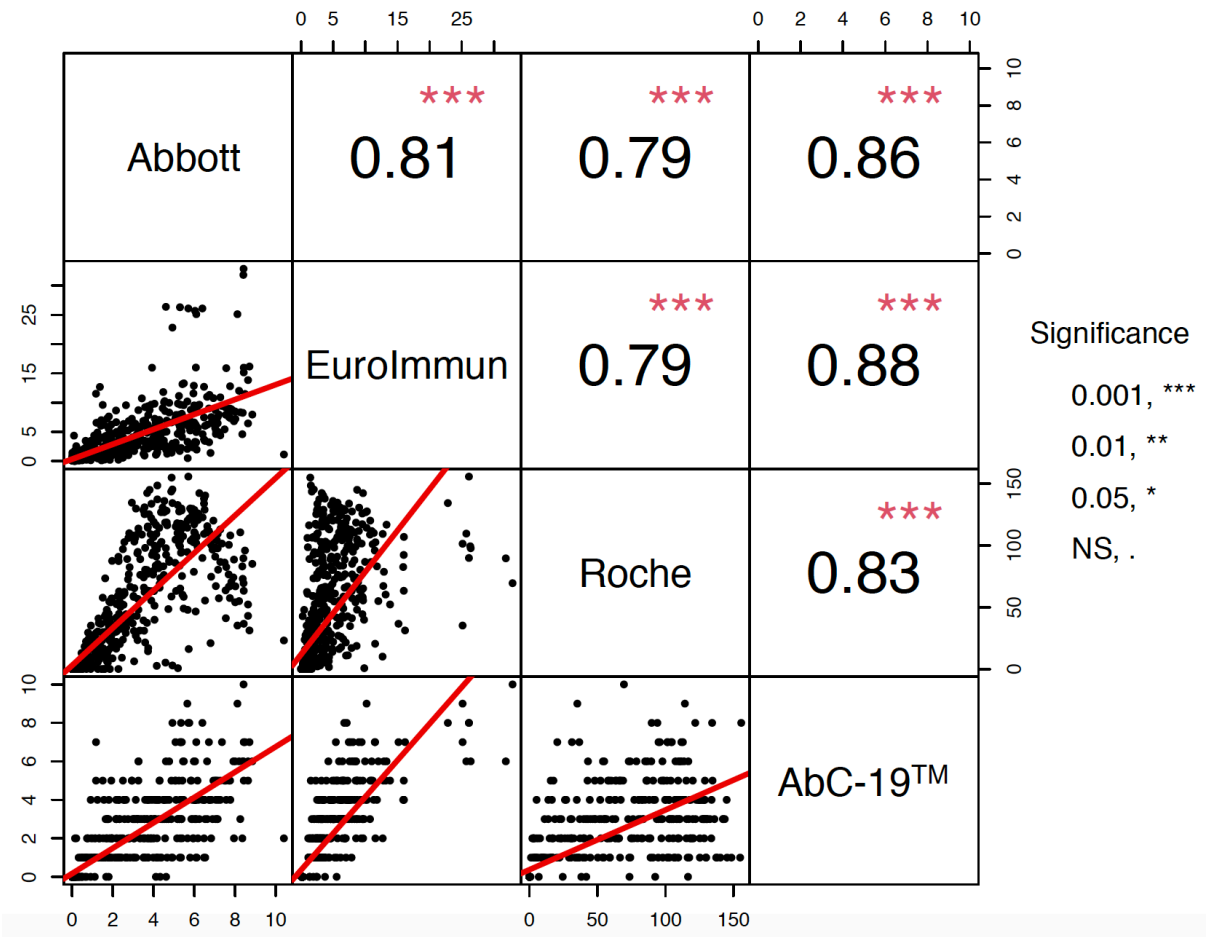


Figure S5: **Correlation matrix between Abbott, EuroImmun, Roche and AbC-19 quantitative output values for SARS-CoV-2 antibody levels.** Strong correlations are observed between all immunoassays. The level of significance was set at $p<0.05$. All immunoassays were significantly correlated $p<0.001$.

Figure S6: NIBSC external reference serology standards and known respiratory virus serology samples.

The scorecard score 0-10 was annotated on test cassette beneath sample ID when agreed by three independent experienced researchers. All LFIAs had a visible control line.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	3
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5/6
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5/6
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	7
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	4
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	7, supp table 1
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	Supp methods, supp table 1
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	Supp methods
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	7/8
	15	How indeterminate index test or reference standard results were handled	7, Supp Fig1
	16	How missing data on the index test and reference standard were handled	Supp Fig 1
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7
	18	Intended sample size and how it was determined	7
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Supp Fig 1
	20	Baseline demographic and clinical characteristics of participants	8/9
	21a	Distribution of severity of disease in those with the target condition	9
	21b	Distribution of alternative diagnoses in those without the target condition	9
	22	Time interval and any clinical interventions between index test and reference standard	5
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Figure 4, Supp Fig 5
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	10/11
	25	Any adverse events from performing the index test or the reference standard	5
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	14/15
	27	Implications for practice, including the intended use and clinical role of the index test	15
OTHER INFORMATION			
	28	Registration number and name of registry	N/a
	29	Where the full study protocol can be accessed	Not yet available
	30	Sources of funding and other support; role of funders	14

CC BY SA

CC BY SA