Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Diagnostic Microbiology and Infectious Disease 102 (2022) 115587

Contents lists available at ScienceDirect

Diagnostic Microbiology and Infectious Disease

journal homepage: www.elsevier.com/locate/diagmicrobio

Evaluation of five widely used serologic assays for antibodies to SARS-CoV-2

Christopher Stocking[a,1], Laura de Miguel[b], Gabriela Suteu[c], Alexander Dressel[d], Andrea Soricelli[e], Martin Roskos[f], Santiago Valor[g], Caren Mutschmann[a,1], Winfried März[h,i,j,1,*]

[a] SYNLAB Analytics and Services Germany GmbH, Berlin, Germany
[b] SYNLAB Global Diagnostics, Alcobendas, Madrid, Spain
[c] SYNLAB Medical Care Center Leverkusen GmbH, Leverkusen, Germany
[d] D^2 Consulting, Mannheim, Germany
[e] IRCCS SDN, Naples Italy
[f] SYNLAB Holding Deutschland GmbH, Augsburg, Germany
[g] SYNLAB International GmbH, München
[h] SYNLAB Academy, Mannheim, Germany
[i] Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz
[j] Medical Clinic V, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

ARTICLE INFO

Article history:
Received 10 November 2020
Revised in revised form 26 August 2021
Accepted 20 October 2021
Available online 29 October 2021

Key words:
COVID-19
SARS-Cov-2
antibodies
serologic testing

ABSTRACT

Reliable diagnostic technologies are pivotal to the fight against COVID-19. While real-time reverse transcription-polymerase chain reaction (rRT-PCR) remains the gold standard, commercial assays for antibodies against (SARS-CoV-2) have emerged. We sought to examine 5 widely used commercial methods. We measured antibodies against SARS-CoV-2 with assays, Abbott-IgG, Roche-IgT (total antibodies, isotype-unspecific), EUROIMMUN-IgG, EUROIMMUN-IgA, DiaSorin-IgG, in 191 serum samples from patients with rRT-PCR proven COVID-19 between days 0 and 47 after the onset of clinical symptoms and in biobank samples collected in 2018. The assays were calibrated using the manufacturers' instructions; results are in multiples of the assay specific cut-offs (Abbott, Roche, EUROIMMUN) or in arbitrary units (AU/mL, DiaSorin). The assays for IgG and IgT have approximately the same sensitivity and specificity for detecting seroconversion which starts at approximately day 3 after symptom onset, sensitivity reached 93% on day 16 and was 100% for each assay on day 20. The assay for IgA antibodies was superior in sensitivity and had a lower specificity than the others. Bivariate non −parametric correlation coefficients ranged between 0.738 and 0.991. Commercial assays for IgG or total antibodies against SARS-CoV-2 are largely equivalent for establishing seroconversion but differ at high antibody titres. Increased sensitivity to detect seroconversion is afforded by including IgA antibodies. Further international efforts to harmonise assays for antibodies against SARS-CoV-2 are urgently needed.

1. Introduction

Coronavirus disease 2019 (COVID-19) has quickly spread worldwide from Wuhan (China) since December 2019 (Zhu et al., 2020). Reliable diagnostic technologies are pivotal to the fight against COVID-19 (Younes et al., 2020). While real-time reverse transcription-polymerase chain reaction (rRT-PCR) testing of respiratory specimens remains the gold standard, a series of commercial assays for antibodies against SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) has been launched (Younes et al., 2020). Their main indications are: to determine the prevalence rate of SARS-CoV-2 infections in epidemiologic surveys and regional outbreaks, to reclassify ambiguous results of rRT-PCR, to establish the diagnosis of COVID-19 in acute infections once rRT-PCR results have become negative (Guo et al., 2020, Kucirka et al., 2020, To et al., 2020, Wikramaratna et al., 2020), to retrospectively infer infection late after onset following the relief of symptoms, and to potentially confirm immunity or the efficacy of vaccines. Available tests rely on the SARS-CoV-2 spike (S) protein and/or the nucleocapsid (N) protein (Ou et al., 2020). Tests detecting IgG, IgA, or IgM isotypes or combinations thereof are available.

The world-wide, explosive propagation of COVID-19 had prompted the emergency development and accelerated approval of tests by manufacturers and authorities, however, only a few robust

* Corresponding author. Tel.: +49 621 43179432; fax: +49 621 43179433
E-mail address: Winfried.Maerz@synlab.com (W. März).
[1] These authors contributed equally

https://doi.org/10.1016/j.diagmicrobio.2021.115587

evaluations of available products have been completed (AdministrationFood and, 2020, Hong KH, 2020). The Cochrane Collaboration concludes that the numbers of individuals contributing data within pertinent studies are small and are usually not based on tracking the same groups of patients over time (Deeks et al., 2020).

We therefore sought to examine 5 widely used, commercial methods for antibodies to SARS CoV-2 in parallel. Most patients selected as part of this study provided at least 2 samples during follow-up.

2. Materials and methods

2.1. Patients and samples

We selectively collected 191 serum samples from patients who had tested positive by real-time reverse transcription-polymerase chain reaction (rRT-PCR) for SARS-CoV-2 and in whom serologic testing had been ordered. The samples came from the Hospital Universitario La Moraleja and the Hospital Virgen del Mar, both located in Madrid, Spain.

We liaised with the responsible physicians to collect information on the onset of clinical signs of COVID-19, symptoms, severity, and course of the disease (Table 1). A timepoint relative to the onset of symptoms was assigned to each sample based upon the clinical information provided by physicians and patients, with day 0 marking symptom onset. Samples were collected from 66 patients, 59 of the patients had samples collected at more than 1 timepoint throughout the course of the infection, and recovery up to day 47.

All results from patient samples were included in this analysis. Analysis of outliers has not been performed nor have any samples been excluded.

The imprecisions of the assays were evaluated using at least 2 levels of quality control material and 1 positive human serum pool, over 6 independent batches with 6 measurements per batch, for a total of 36 measurements for each material. To calculate mean values and standard deviations we used assay readouts after normalizing them against the assay specific cut-off.

Specificities were examined in 262 to 337 frozen samples collected before 2018 for which serology of respiratory infections had been requested.

The ethics board of the Medical Association of Berlin was informed of the study and confirmed in writing that there was no need for further ethical advice, since no additional analyses were performed other than the medically warranted ones.

2.2. Real-time reverse transcription-polymerase chain reaction (rRT-PCR)

rRT-PCR of pharyngeal swaps was performed at SYNLAB Barcelona, Spain, using the VIASURE SARS-Cov-2 Real Time PCR Detection Kit.

2.3. Determination of antibodies against SARS-CoV-2

We measured antibodies against SARS-CoV-2 with 5 laboratory-based immunoassays under the conditions shown in Table 2. As these methods were not available at a single laboratory site, we distributed aliquots as indicated. Shipment was between 4 and 8°C; the cold chain was not interrupted at any time. All assays were calibrated using the manufacturers' specified procedures and single measurements were performed.

While 4 of the 5 methods measure either IgG or IgA, the Elecsys Anti−SARS-CoV-2 (Roche-IgT) is not isotype-specific for IgG; rather, the package insert states that it is a test for the "qualitative detection of antibodies (including IgG) against" SARS-CoV-2. The EUROIMMUN-IgG and the EUROIMMUN-IgA test may be used simultaneously. We therefore also provisionally combined the 2 tests by assigning an overall "positive" result if one of the 2 tests was "positive" and treating the remaining results as "negative" analyses (EUROIMMUN-IgG/A).

The results of the DiaSorin LIASON SARS-CoV-2 S1/S2 IgG (DiaSorin-IgG) test are in Arbitrary Units per millilitre (AU/mL). For the purpose of this comparison, they have been normalised against the assigned cut-off of 13.5 AU/mL. All other assays provide relative units normalised against the assay specific cut-off. In our analyses, we used those relative units, also for "negative" samples, which then scored below 1.

2.4. Statistical analysis

Imprecision. Imprecision (within-batch, between-batch) was calculated using a 2-factor analysis of variance and expressed as the coefficients of variation (CV). For this purpose, total variance was decomposed into intra-batch and between-batch variance components from which the intra-batch and between-batch standard deviations were calculated (Supplementary Table1).

Bivariate comparisons. Spearman correlation coefficients and the non−parametric regression method of Passing and Bablok (Passing and Bablok, 1983) were performed using MedCalc (Supplementary Table 2). Pearson's correlation coefficients and least squares regressions were not calculated because the CUSUM test (included in *MedCalc*) detected highly significant deviations from linearity for each of the pairwise comparisons. We also produced scatter diagrams along with restricted cubic splines (with knots at the 5th, 27.5th, 50th, 72.5th and 95th percentiles, respectively, of the values of the abscissa variable), whereby we provisionally plotted the SARS-CoV-2 IgG assay from Abbott Diagnostics on the abscissa (Fig. 1A − D). This does not prejudice the interpretation of the data nor indicate that we consider the Abbott Diagnostics assay as a reference method.

In order to examine whether differences between the methods are heterogenous across the concentration range, we generated difference plots according to Bland and Altmann (Bland and Altman, 1986) for descriptive purposes only (Supplementary Fig. 1A−D). We also generated an index for the consistency of the tests (details are provided with Supplementary Fig. 2).

Time course of antibody concentrations. We plotted medians (and interquartile ranges) against the day of blood sampling after the onset (day 0) of clinical symptoms. If less than 10 samples were available for a specific day, samples from several days were pooled together, and the day number displayed then is the mean for the group (Fig. 2).

Sensitivity, specificity. As the concentration of antibodies present are known to rise immediately following the onset of symptoms, the sensitivity of each assay was evaluated relative to the day of onset of symptoms using the same strata as in Fig. 2 (Table 3A), and the cut-offs provided by the manufacturers. For this purpose, we treated "indeterminate" results as "negative."

We provisionally defined overall sensitivity as the cumulative sensitivity on day 16, which likely marks the timepoint at which the window for positive rRT-PCR results closes (Younes et al., 2020).

To scrutinize potential differences between the assays, we performed 2 multivariate logistic regressions using time strata (less than 3 days, 3−7 days, 8−13 days, 14 days or more) as independent, and the test result as the dependent variables. In the first model, we used the 5 assays separately, in the second one, we replaced the 2 EUROIMMUN tests by their combination EUROIMMUN-IgG/A (Table 3B).

We also visualized the time-dependency of the assays' sensitivities by creating monotonic cubic splines with 4 knots at the centres of time quartiles (boundaries 3, 8, and 14 days) which assumed mean values of 0.7 (quartile 1), 4.8 (quartile 2), 10.3 (quartile 3) and 23 (quartile 4) days, respectively, once considering each test individually and once after replacing EUROIMMUN-IgG and EUROIMMUN-IgA by EUROIMMUN-IgG and (Fig. 3B).

Table 1
Clinical characteristics 11of patients included in this evaluation.[a]

Characteristic *N* =	All Patients 66		Survivors 57		Deceased 9	
Mean age ± SD[b] (range), yrs.	68 ± 14	(30–97)	65 ± 12	(30–84)	86 ± 9	(69 – 97)

Gender	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)
Men, *n* (%)	44	66.7	38	66.7	6	66.7
Women, *n* (%)	22	33.3	19	33.3	3	33.3

Comorbidities	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)
Hypertension, *n* (%)	10	15.2	9	15.5	1	11.1
Dyslipoproteinaemia, *n* (%)	7	10.6	6	10.3	1	11.1
Diabetes mellitus, *n* (%)	5	7.6	4	6.9	1	11.1
COPD[c], *n* (%)	2	3	2	3.4	0	0
Ischaemic heart disease	3	4.5	3	5.2	0	0

Respiratory symptoms	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)
Pneumonia, *n* (%)	48	72.7	41	70.7	7	77.8
Shortness of breath, n (%)	32	48.5	28	48.3	4	44.4
Cough, *n* (%)	36	54.5	36	62.1	0	0

Other symptoms	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)	Number (*n*)	Percentage (%)
Fever, *n* (%)	45	68.2	40	69.0	5	55.6
Diarrhoea, *n* (%)	4	6.1	4	6.9	0	0
Musculoskeletal pain, *n* (%)	4	6.1	4	6.9	0	0
Headache, *n* (%)	4	6.1	4	6.9	0	0
Nausea, *n* (%)	3	4.5	3	5.2	0	0
Confusion, disorientation, *n* (%)	2	3	0	0.0	2	22.2

Disease severity		
Mild (no hospital admission)	10	15.2
Moderate (hospital admission without ICU)	42	63.6
Severe (hospital admission with ICU)	14	21.2

[a] Only those characteristics mentioned 3 times or more are displayed.
[b] SD = standard deviation
[c] COPD = chronic obstructive pulmonary disease

Table 2
Assays used for the measurement of antibodies for SARS-CoV-2.

Assay	SARS-CoV-2 IgG	Elecsys Anti–SARS-CoV-2	Anti–SARS-CoV-2 ELISA (IgG)	Anti–SARS-CoV-2 ELISA (IgA)	LIASON SARS-CoV-2 S1/S2 IgG
Manufacturer	Abbott Diagnostics	Roche	EUROIMMUN	EUROIMMUN	DiaSorin
Abbreviation	Abbott-IgG	Roche-IgT	EUROIMMUN-IgG	EUROIMMUN-IgA	DiaSorin-IgG
Order reference number	6R86	09203095190	EI 2606-9601 G	EI 2606-9601 A	311450
Ig Class	IgG	IgT	IgG	IgA	IgG
Antigen	nucleocapsid protein of SARS-CoV-2	nucleocapsid protein of SARS-CoV-2	S1 domain	S1 domain	S1 and S1 domains
Method	automated chemiluminescence microparticle immunoassay	automated electrochemiluminescence immunoassay	automated Enzyme-linked immunosorbent assay (ELISA)	automated Enzyme-linked immunosorbent assay (ELISA)	automated chemiluminescence immunoassay
Analyser	Abbott ARCHITECT i2000SR	Roche Cobas e411	EUROIMMUN Analyzer I	EUROIMMUN Analyzer I	DiaSorin LIASON XL
Site of analysis	Augsburg[a]	Berlin[b]	Augsburg[a]	Augsburg[a]	Leverkusen[c]
Data level claimed[d]	Qualitative	Qualitative	Semi-quantitative	Semi-quantitative	Quantitative
Unit	None	None	None	None	AU/mL[e]
Interpretation[f]					
Negative	<1.4	<1.0	<0.8	<0.8	<12
Indeterminate	N/A[g]	N/A	≥0.8 to <1.1	≥0.8 to <1.1	12-15
Positive	≥1.4	≥1.0	≥1.1	≥1.1	≥15

[a] SYNLAB Medical Care Center Augsburg GmbH, Gubener Straße 39, 86156 Augsburg, Germany
[b] SYNLAB Analytics and Services Germany GmbH, Turmstraße 21, 10559 Berlin, Germany
[c] SYNLAB Medical Care Center Leverkusen GmbH, Paracelsusstraße 13, 51375 Leverkusen Germany
[d] according to the product specifications of the manufacturers
[e] AU/mL = arbitrary units/mL
[f] only positive results used for the calculation of the sensitivity
[g] N/A = not applicable y

Specificity was expressed as the proportion of subjects in the biobank samples with negative test results in the total of samples tested (Table 3C).

Positive predictive values (PPVs), negative predictive values (NPVs). We visualised PPVs and NPVs according to hypothetical disease prevalence rates in the range of 0.001 through 0.05 (Fig. 4).

If not stated otherwise, statistical analysis was performed with MedCalc Version 11.2.0.0. The restricted cubic spline regressions together with their standard errors (regarding the model) were calculated with the aid of the functions *ols* and *rcs* of the library *rms* (R version 4.0.2).

3. Results

3.1. Clinical characteristics of patients

191 samples were collected from 66 patients. The detailed clinical characteristics of these patients are shown in Table 1. Their mean age was 68 years with a range of 30 to 97 years. Among the comorbidities, hypertension, dyslipoproteinaemia, diabetes mellitus, chronic obstructive lung disease, and ischemic heart disease were prevailing. The mean time interval between the onset of symptoms and hospital admission was 5 days. As expected, the most frequent clinical features of COVID-19 were pneumonia, cough, shortness of breath, and fever (Table 1). 57 patients recovered, 9 died. Characteristics of the deceased and the surviving patients are also provided with Table 1.

3.2. Imprecision

Supplement Table 1 shows coefficients of variation (CVs) obtained in control sera at low (negative) and high (positive) antibody titres and in an in-house pool of positive sera. At concentrations below the diagnostic cut-off, total CVs were around 10%, with the exception of the EUROIMMUN-IgA assay showing a CV of 48.7%. At high concentrations, the CVs ranged from 1.7 to 9.3% for Abbott-IgG and EUROIMMUN-IgG and between 2.5 and 6.8% for Roche-IgT and EUROIMMUN-IgA.

3.3. Bivariate correlations

The scatterplots shown in Fig. 1 indicate substantial scattering of the results in all pairwise comparisons. The CUSUM test for linearity indicated significant ($p < 0.01$) deviations from linearity for all comparisons. In the light of the scattering of the data, we dispensed with the calculation of Pearson correlation coefficients and least squares regression lines and visualised the relationships using splines (Fig. 1). Spearman non–parametric bivariate correlation coefficients were calculated for descriptive purposes and ranged between 0.738 and 0.861 (Supplementary Table 2, upper diagonal). Slopes and intercepts of linear regression lines according to the Passing and Bablok are shown in the lower diagonal of Supplementary Table 2. Comparing all other methods to the Abbott Diagnostics test reveals slopes below 1.0 and intercepts above zero. These numbers are likely owing to the marked deviations from linearity and should therefore be interpreted with caution. Plotting Roche-IgT versus Abbott-IgG shows that above concentrations of 4, Roche scores approximately threefold compared to Abbott. In contrast, only slightly higher scores are seen with the EUROIMMUN-IgG compared with Abbott; starting at concentrations of about 8, the spline of EUROIMMUN-IgG versus Abbott-IgG flattens with no further increase of EUROIMMUN-IgG. This behaviour is even more pronounced for the EUROIMMUN-IgA which in tendency may decrease while Abbott-IgG exceeds the concentration of 8. For descriptive purposes, we provide Bland, and Altmann difference plots in comparison to Abbott IgG as Supplementary Fig. 1A-D. Supplementary Fig. 2 shows that mean consistency of the 5 tests (concordant results of all 5 tests) was high at the concentration around 1 and at

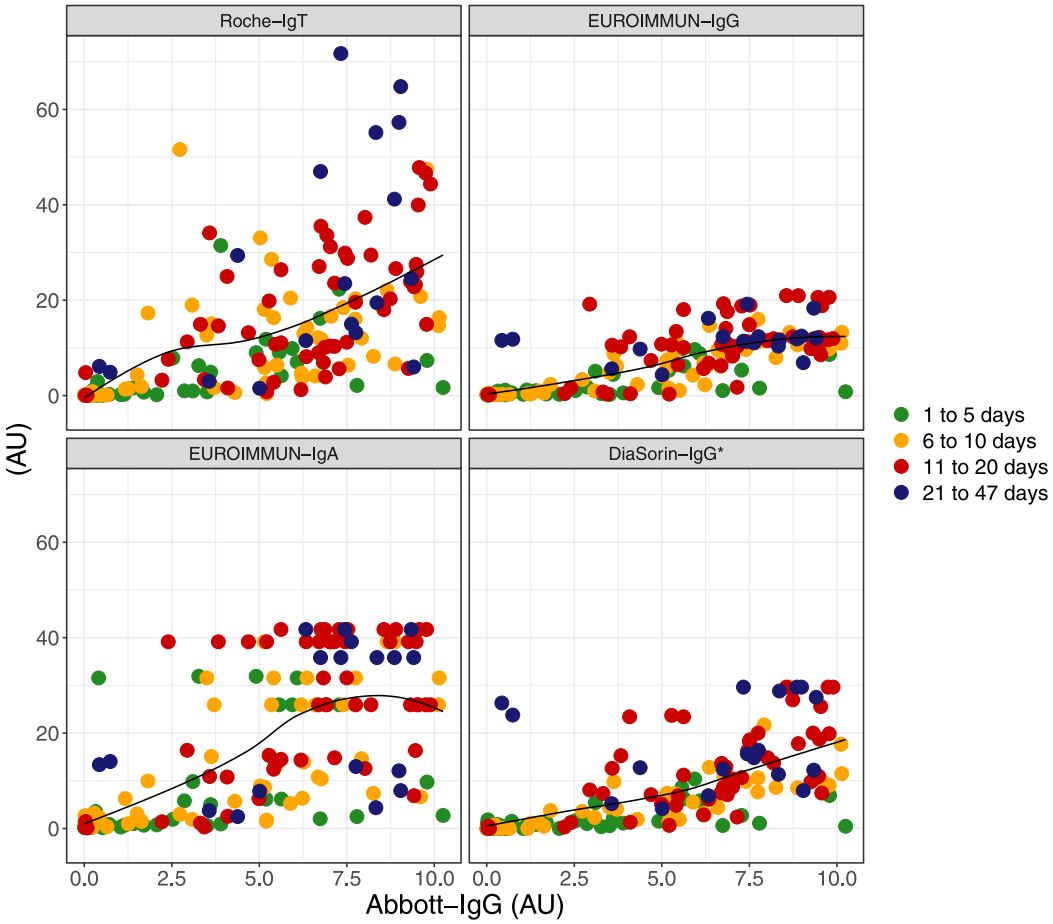

Fig. 1. Scatterplots of relative concentrations of antibodies against SARS-CoV-2 produced by 5 commercial assays. *Top left panel:* Roche-IgT *versus* Abbott-IgG; *top right panel:* EUROIMMUN-IgG *versus* Abbott IgG; *bottom left panel:* EUROIMMUN-IgA *versus* Abbott-IgG; *bottom right panel D:* DiaSorin-IgG *versus* Abbott-IgG. The colours of the data points indicate the time of blood sampling after the onset of clinical symptoms of COVID-19. *Green:* 1 to 5 days; *orange:* 6 to 10 days; *red:* 11 to 20 days; *blue:* 21 to 47 days. The solid line represents a spline with knots at the 5th, 27.5th, 50th, 72.5th and 95th percentiles of the values on the abscissa variable. AU: arbitrary units, multiples of the cut-off (Color version of the figure is available online.)

day 1, decreased to approximately 40 and 80%, respectively, at the concentration of 2 and at day 2, and then increased to approach 100% at a concentration of 5 and day 6, respectively.

3.4. Time course of antibody concentrations

The median concentrations obtained with Abbott-IgG, Roche-IgT, EUROIMMUN-IgG and EUROIMMUN-IgA started to rise between days 3 and 5 post onset of symptoms, the median concentrations of the DiaSorin-IgG started to increase between days 5 and 6 (Fig. 2). All tests plateaued at day 20 after the onset of symptoms. After this time-point, Abbott-IgG did not show any change, Roche-IgT showed some fluctuation, EUROIMMUN-IgG a slight decline, EUROIMMUN-IgA a strong decline and DiaSorin-IgG a further increase between day 35 and 45, but this may be due to a smaller number of measurements available for day 45 ($n = 5$, Fig. 2).

3.5. Sensitivity and specificity

We analysed the sensitivities of each of the tests in relation to the day post onset of symptoms. As displayed in Table 3A, the cumulative sensitivity increased along the time axis. Until about day 12 the EUROIMMUN-IgA appeared as the most sensitive, whereas DiaSorin-IgG appeared the least sensitive assay.

Sensitivity was then 93% across all tests on day 16 post onset of clinical symptoms. One patient became negative after testing positive for 1 assay. This patient with moderate clinical symptoms had a negative titre on the Abbott-IgG assay at day 37 and day 45 after 4 positive results from day 10 to 17, however, the maximum titre was only 3.83. All other assays remained positive.

Multivariate logistic regression accounting for the time after symptom onset suggests that the EUROIMMUN-IgA is approximately 2-fold more sensitive than the other tests, with no material differences between them (Table 3B). Also, the combined use of the EUROIMMUN-IgA and the EUROIMMUN-IgG was more sensitive than the 3 remaining tests. This was confirmed by visualization of the assays' sensitivities as a function of the time after symptom onset using monotonous cubic spline interpolation (Fig. 3): The EUROIMMUN-IgA and the EUROIMMUN-IgG/A had a higher sensitivity than the other tests up to 10 days after the onset of symptoms while later the assays' sensitivities converged.

We used biobank samples collected before 2018 to examine specificity. The numbers of samples used for each test and results are displayed with Table 5B. The specificities were 1.00 for the Abbott-IgG and the Roche-IgT test; they ranged from 0.953 through 0.989 to 0.996 for EUROIMMUN-IgA, DiaSorin-IgG and EUROIMMUN-IgG, respectively.

3.6. Positive predictive values (PPVs), negative predictive values (NPVs)

Fig. 4 shows the PPVs and NPVs according to assumed prevalence rates in the populations between 0 and 0.05 calculated using the

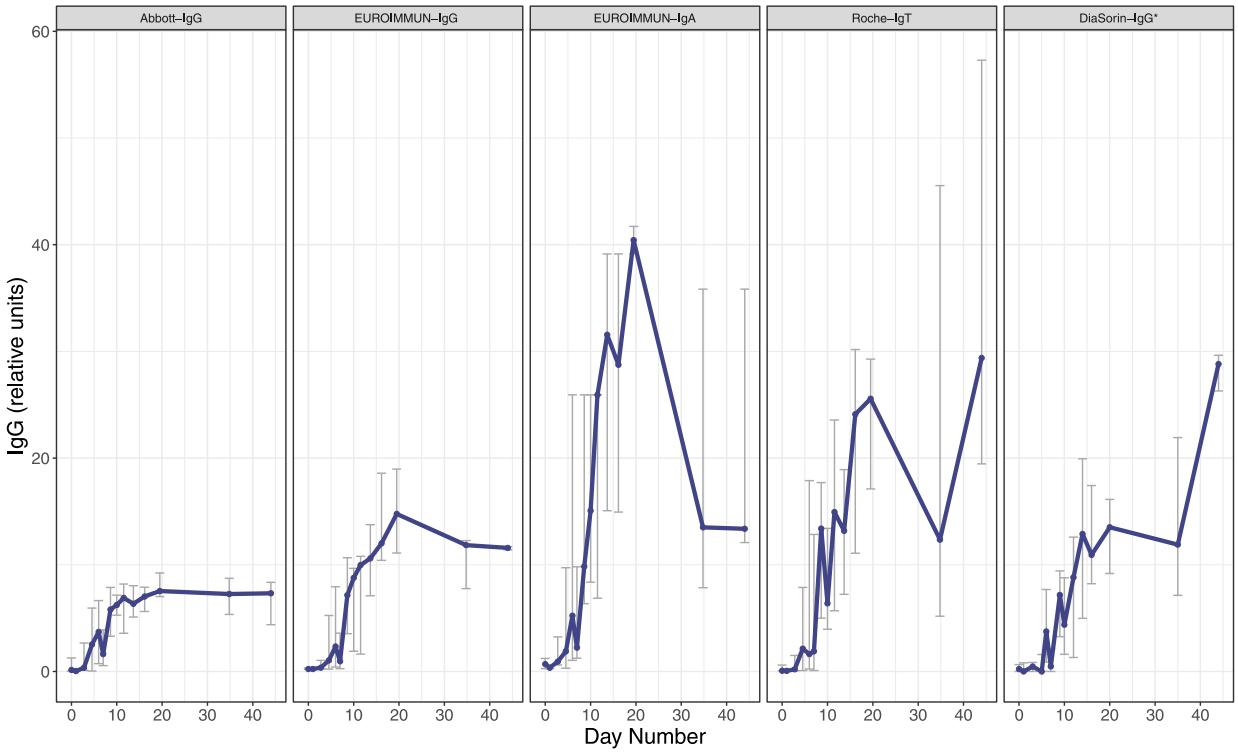

Fig. 2. Time course of antibody concentrations. Medians (and interquartile ranges) plotted against the day of blood sampling after day 0 using the same day groups as in Table 5A. *First panel:* Abbott-IgG; *second panel:* EUROIMMUN-IgG; *third panel:* EUROIMMUN-IgA; *fourth panel:* Roche-IgT; *fifth panel: DiaSorin*-IgG.

sensitivities at day 16 displayed in Table 3A. At an assumed prevalence rate of 0.003 for Germany, the PPV and NPV of the Abbott-IgG and the Roche-IgT were 1.0. Due to the low prevalence rate assumed, the PPVs of the EUROIMMUN-IgA, EUROIMMUN-IgG and the DiaSorin-IgG were 0.41, 0.06 and 0.20, respectively. The NPVs of all tests were 1.00 at the prevalence rate of 0.003.

4. Discussion

This is a systematic comparison of contemporary commercial assays to measure antibodies against SARS-CoV-2. We made the following key observations: The median antibody concentrations started to increase between days 3 and 5 post onset of symptoms, with the exception of the DiaSorin-IgG which started to increase between days 5 and 6. Until about day 12 the sensitivity of EUROIMMUN-IgA was highest. Combined use of the EUROIMMUN-IgG and EUROIMMUN-IgA offered the highest sensitivity to detect seroconversion early after infection. Sensitivity was 93% across all tests on day 16 and reached 100% on day 20 post onset of clinical symptoms.

At concentrations way above the decision range and/or 20 days after onset of symptoms our comparison reveals substantial random scattering of results. We need to recognize, however, that only results of the DiaSorin-IgG are claimed to be quantitative by the manufacturer, while EUROIMMUN provides semi-quantitative and Abbott and Roche qualitative results. Anyway, the vast scattering of results at concentrations way above the decision ranges suggests that the methods may not be interchangeable for the long-term follow-up of SARS-CoV-2 antibody titres, at least as long quantitative tests will become widely available.

Only few studies have compared isotype-specific (IgM, IgG, and IgA) sensitivities of antibodies for SARS-CoV-2 (Guo et al., 2020, (Xiao et al., 2021). Here, seroconversion started at day 3 after the onset of clinical symptoms (15%–38% of samples positive), reached 50% on day 5 (42–54) and 93% on day 16. On day 20 seroconversion was complete for all assays (Table 5). This is in line with other observations (Ai et al., 2020, Gudbjartsson et al., 2020, Kabesch et al., 2020, Okba et al., 2020, Tan et al., 2020).

In line with previous work (Okba et al., 2020), the EUROIMMUN-IgA a higher sensitivity as the IgG- and IgT- based tests. There was no material difference in sensitivity between the Roche-IgT and those tests capturing IgG only. This is in agreement with the observation that there is only a tiny, if any, time shift between the emergence of specific IgM and IgG in the plasma upon SARS-CoV-2 exposure (Xiao et al., 2020), as it has also been shown for SARS and MERS. The obvious difference in the detection of early seropositivity between EUROIMMUN-IgA and the other tests may have clinical relevance if one accounts for the documented limitations of rRT-PCR. While rRT-PCR is widely accepted as the gold standard in the diagnostics of acute COVID19 infections it has been argued that the overall sensitivity of PCR is limited to 80% and can be improved by combination with antibody testing (Zhao et al., 2020). This may be of special importance if the clinical situation is suggestive of COVID-19 while PCR results are negative.

While all tests were highly consistent for negative samples obtained early in the course of the infection, there was some inconsistency around the decision cut-offs, and the days of seroconversion (Supplementary Fig. 2). The reason for this may lie in the comparatively low reproducibility at low concentrations within the so called “gray zone.” It is hence obvious that single results close to the assay cut-offs need to be interpreted cautiously. Also, and consequentially, it might be prudent to reproduce such results by duplicates, in an independent sample obtained a few days later to prove titre changes, or to verify them by another assay, and/or a confirmatory test like Western blotting once this approach becomes widely available.

Titres produced by Abbott-IgG stayed constant after 20 days, while there was a fluctuation of titres obtained with Roche-IgT (decrease first, increase thereafter), which might relate to an isotype shift. Similarly, EUROIMMUN-IgG showed a moderate and the

Table 3A
Rates of positive results (sensitivity) of 5 commercial assays for SARS-CoV-2 antibodies[a]. Anti−SARS-CoV-2 ELISA (IgG) EUROIMMUN and Anti−SARS-CoV-2 ELISA (IgA) EUROIMMUN were not considered separately. Rather, the combination of both tests was evaluated.

Assay/D[b]	*N*[c]	SARS-CoV-2 IgG Abbott Diagnostics (Abbott-IgG)	Anti−SARS-CoV-2 Roche (Roche-IgT)	Anti−SARS-CoV-2 ELISA (IgG) EUROIMMUN (EUROIMMUN-IgG)	Anti−SARS-CoV-2 ELISA (IgA) EUROIMMUN (EUROIMMUN-IgA)	Anti−SARS-CoV-2 ELISA (IgA) and (IgG) EUROIMMUN combined[c] (EUROIMMUN-IgA/G)	LIASON SARS-CoV-2 S1/S2 IgG[d] (DiaSorin-IgG)
0	16	4 (25)	3 (19)	2 (13)	4 (25)	4 (25)	3 (19)
1	13	1 (8)	0 (0)	0 (0)	1 (8)	1 (8)	2 (15)
3	26	9 (35)	8 (31)	6 (23)	10 (38)	12 (46)	4 (15)
5	13 (12)	7 (54)	7 (54)	6 (46)	7 (54)	7 (54)	5 (42)
6	11	7 (64)	7 (64)	6 (55)	7 (64)	7 (64)	8 (73)
7	11 (10)	7 (64)	6 (55)	4 (36)	8 (73)	8 (73)	5 (50)
9	18	14 (78)	15 (83)	14 (78)	16 (89)	16 (89)	14 (78)
10	11 (10)	10 (91)	10 (91)	9 (82)	11 (100)	11 (100)	9 (90)
12	18 (17)	16 (94)	17 (100)	13 (76)	17 (100)	17 (100)	14 (88)
14	15 (14)	15 (100)	14 (93)	13 (87)	14 (93)	14 (93)	13 (93)
16	14 (13)	13 (93)	13 (93)	13 (93)	13 (93)	13 (93)	12 (92)
20	10	10 (100)	10 (100)	10 (100)	10 (100)	10 (100)	10 (100)
35	10	9 (90)	10 (100)	10 (100)	10 (100)	10 (100)	10 (100)
44	5	4 (80)	5 (100)	5 (100)	5 (100)	5 (100)	5 (100)

[a] The cells contain absolute numbers and percentages in brackets.
[b] If less than 10 samples were available for a specific day, samples from several days were combined. The displayed day number is then the mean for the group.
[c] The Anti−SARS-CoV-2 ELISA (IgG) EUROIMMUN and Anti−SARS-CoV-2 ELISA (IgA) EUROIMMUN were not considered separately. Rather, the combined test was considered positive if any of the tests was positive.
[d] There are 6 less samples for LIASON SARS-CoV-2 S1/S2 IgG, the number of samples for this test are in brackets in the second column.

Table 3B
Predictors of positive antigen test amongst SARS-CoV-2 positive samples in a multivariate model.

Assay	OR (95% CI)
Model 1 (each test individually)	
Abbott-IgG	0.64 (0.50, 0.81)
Roche-IgT	0.62 (0.49, 0.78)
DiaSorin-IgG	0.51 (0.39, 0.66)
EUROIMMUN-IgG	0.41 (0.32, 0.52)
EUROIMMUN-IgA	1.0 (reference)
Time (after onset of symptoms) quartile	
1st (less than 3 d)	1.0 (reference)
2nd (3 to 7 d)	3.26 (1.85, 6.64)
3rd (8 to 13 d)	27.4 (14.60, 68.66)
4th (14 d and more)	78.49 (38.12, 385.47)
Model 2 (EUROIMMIUN assays replaced by EUROIMMUN-IgA/G)	OR (95% CI)
Abbott-IgG	0.60 (0.47, 0.77)
Roche-IgT	0.58 (0.36, 0.53)
DiaSorin-IgG	0.48 (0.36, 0.63)
EUROIMMUN-IgA/G	1.0 (reference)
Time (after onset of symptoms) quartile	
1st (less than 3 d)	1.0 (reference)
2nd (3 to 7 d)	2.98 (1.70, 5.88)
3rd (8 to 13 d)	26.72 (14.14, 71.72)
4th (14 d and more)	69.36 (34.31, 325.77)

Table 3C
Specificity of 5 commercial assays for SARS-CoV-2 antibodies.

	SARS-CoV-2 IgG Abbott Diagnostics (Abbott-IgG)	Anti−SARS-CoV-2 Roche (Roche-IgT)	Anti−SARS-CoV-2 ELISA (IgG) EUROIMMUN (EUROIMMUN-IgG)	Anti−SARS-CoV-2 ELISA (IgA) EUROIMMUN (EUROIMMUN-IgA)	LIASON SARS-CoV-2 S1/S2 IgG[4] (DiaSorin-IgG)
N =	276	337	276	276	262
Positive, *n*	0	0	1	13	3
Specificity	1.000	1.000	0.996	0.953	0.989
95% confidence interval	0.987−1.000	0.989−1.000	0.980-1.000	0.921−0.975	0.967- 0.998

EUROIMMUN-IgA and a sharp decline of signals beyond day 20, in agreement with Tan et al. (Tan et al., 2020). These data of course need to be interpreted with a sense of proportion, since they rely on the small number of samples available beyond day 20. However, they illustrate the current uncertainty as to whether, and how long antibodies against SARS-CoV-2 would circulate in the serum at detectable concentrations late after an infection. In any case, the findings suggest that the most reliable timepoint to test for SARS-CoV-2 exposure

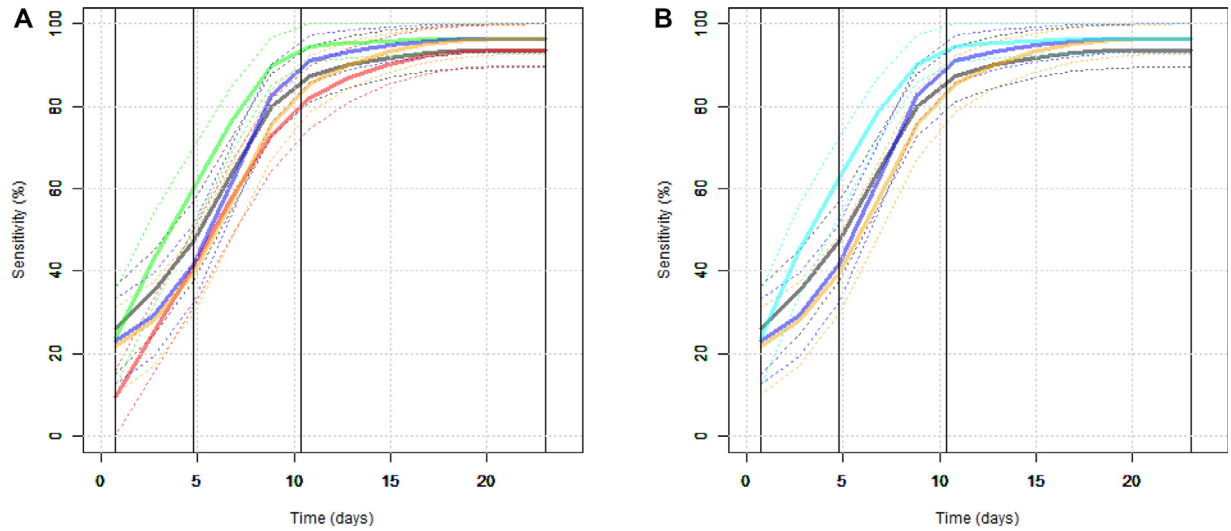

Fig. 3. Time course of the sensitivities of 5 commercial assays for antibodies for SARS-CoV2. The abscissa shows the day of blood sampling after symptom onset, the ordinate the respective sensitivities (with lower and higher 95% CI bounds in dotted lines) modelled as monotonous cubic spline with nodes at the means of day quartiles (less than 3 days, 3 days through 7 days, 8 days through 13 days, more than 14 days, highlighted by black vertical lines). *Panel A: black:* Abbott-IgG; *red*: EUROIMMUN-IgG; *green*: EUROIMMUN-IgA; *blue*: Roche-IgT; *orange*: Diasorin-IgG; *Panel B: black:* Abbott-IgG; *blue*: Roche-IgT; *orange*: Diasorin-IgG; *cyan*: Euroimmun-IgG and Euroimmun-IgA combined (Euroimmun-IgG/A) (Color version of the figure is available online.)

would be no earlier than day 20 after onset of symptoms (World Health Organization, 2020), while the possibility exists that samples become negative thereafter. It is completely consistent with this estimate that we attained 100 % sensitivity at day 20.

A recent validation study including lateral flow tests and ELISAs reported specificities between 0.84 and 1.0 (Whitman et al., 2020). As assessed using frozen biobank samples collected from 2018 and earlier, we found 2 tests producing specificities of 1.00 (Abbott-IgG and Roche IgT). This corresponds to the specificity reported previously for the Roche-IgT (0.998 obtained by analysing 5272 serum samples) (https://www.fda.gov/media/137605/download). The specificities of 2 other tests came close to unity (EUROIMMUN-IgG and DiaSorin-IgG), whereas the specificity of the EUROIMMUN-IgA was materially lower (0.95). While more true negative samples would be required to draw definite conclusions in regard to the specificities, even a very low rate of false positives would translate into a considerable restriction of the predictive value of a positive test in a low prevalence setting as it is encountered with COVID-19 (Fig. 4). If applied to epidemiologic studies, even very low false positive rates would hence produce an artefactual overestimation of the disease prevalence rate paired with an underestimation of the case fatality. Tentatively, and until more data for test specificity become available, the Abbott-IgG, and the Roche-IgT might in this respect have some advantages over other tests in epidemiologic surveys. This may, however, not apply to any clinical situation in which the pre-test probability is higher than in the general population.

4.1. Limitations

We studied symptomatic patients only and cannot infer on the strength of the immune response in patients without or with few symptoms. In severe COVID-19 cases seroconversion may occur early (Okba et al., 2020) while mild or asymptomatic infections may even lack seroconversion (Gudbjartsson et al., 2020, Yongchen et al., 2020, Zhou et al., 2013). Approximately, seroconversion was seen in 50% of the samples at day 5 after symptom onset. It is a limitation of the

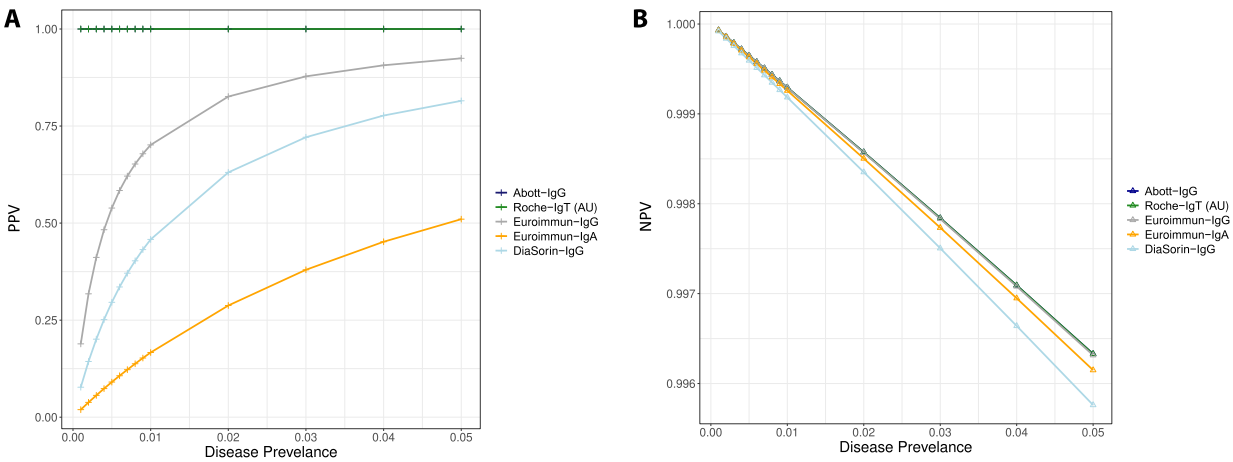

Fig. 4. Predictive values of positive tests (PPV, *panel A*) and predictive values of negative tests (NPV, *panel B*) of 5 commercial assays for antibodies against SARS-CoV-2 in relation to disease prevalence rates between 0.001 and 0.05. *Dark blue:* Abbott-IgG; *green:* Roche IgT; *gray:* EUROIMMUN-IgG; *orange:* EUROIMMUN-IgG; *light blue:* DiaSorin. Note that the lines for Abbott-IgG and Roche-IgT and for Abbott-IgG, Roche-IgT and EUROIMMUN-IgG in *panels A and B*, respectively, are coinciding (Color version of the figure is available online.)

current work, that the timepoint of symptom onset was obtained retrospectively, so that evidently the estimate of the time to seroconversion depends on the precision of the information conveyed by patients to treating physicians. Also, as a consequence of the limited clinical information and the small number of severe COVID-19 cases, we were not able to reliably correlate the kinetics of antibody concentrations with disease severity (Huang et al., 2020, Okba et al., 2020, Tan et al., 2020)

We also have no knowledge on the long-term persistence of antibodies beyond the time frame of our study which extends to 48 days only after the onset of clinical symptoms. It thus remains open whether any of the tests would be suitable to monitor the immune status in persons presenting later after an infection and/or with mild or no symptoms (Gudbjartsson et al., 2020).

It is also a limitation that lateral flow tests, legions of which have been launched for on-site testing, have not been included into the current comparison.

Finally, neutralization tests (Haveri et al., 2020, van der Heide, 2020, Wolfel et al., 2020, Wu et al., 2020) are not currently available to our laboratories. It hence remains outside the scope of this article whether or to what extent the assays examined would reflect resilience against subsequent SARS-CoV-2 infections.

4.2. *Directions for future research*

Due to the threatening spread of COVID-19 serologic tests for COVID-19 have been approved and launched rapidly. Hence, extensive evaluations were skipped for comprehensible reasons. In this respect, the current report represents a snapshot of the *status quo* awaiting confirmation, and extension. In the light of the still low prevalence rate of COVID-19 in many countries, our preliminary data on the specificity of the assays evaluated will need further scrutiny in diverse environments. Further, the long-term course of antibody concentrations will have to be clarified, and research into simple assays specifically indicating protection from SARS-CoV-2 infections or immunity after vaccination is warranted (Whitman et al., 2020). In this regard, surrogate neutralisation assays avoiding the work with viable viruses might be of particular interest (Grzelak L and 10.1101/2020.04.21.20068858).

The scattering of results in our bivariate method comparisons suggests that there is need for reference materials covering the entire concentration range of antibody concentrations and multiple point calibration may be needed to eliminate such differences.

Sufficiently large and challengeable validation studies will be needed which will be facilitated by the formation of international research networking involving the major stakeholders like manufacturers, laboratory service providers, and clinical centres.

Declaration of competing interests

All authors except AD are employed with SYNLAB International GmbH or its local subsidiaries. There are no other known conflicts of interest associated with this publication. There was no financial support to SYNLAB or any of its employees from the manufacturers of the assays used in this evaluation and there has been no other financial support for this work that could have influenced its outcome.

Authors' contributions

CM and SV designed the study; UH, LDM, and GS collected data; CS performed statistical analysis and surveyed laboratory analyses; GS surveyed laboratory analyses; WM wrote the manuscript; all authors validated, reviewed, and edited the manuscript.

Acknowledgments

The authors thank the staff at the Hospital La Moraleja and Virgen del Mar laboratories, especially to the Technical Directors, Drs González and Louzao, for their involnt in the collection and follow up of patients and the staff at the SYNLAB laboratories involved in this study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.diagmicrobio.2021.115587.

References

Administration, Food and Drug. Policy for coronavirus disease-2019 tests during the public health emergency (Revised); 2020. Available at: https://wwwfdagov/media/135659/download. Accessed November 10, 2020.

Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology 2020 200642.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307–10.

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.

Grzelak LTS, Planchais C, Demeret C, Huon C, Guivel F, et al. SARS-CoV-2 serological analysis of COVID-19 hospitalized patients, pauci576, doi:10.1101/2020.04.21.20068858.

Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. N Engl J Med 2020;382(24):2302–15.

Guo L, Ren L, Yang S, Xiao M, Chang YF, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). Clin Infect Dis 2020;71(15):778–85.

Haveri A, Smura T, Kuivanen S, Osterlund P, Hepojoki J, Ikonen N, et al. Serological and molecular findings during SARS-CoV-2 infection: the first case study in Finland, January to February 2020. Euro Surveill 2020;25(11):16–21.

Hong KH LS, Kim TS, Huh HJ, Lee J, Kim SY, et al. Guidelines for laboratory diagnosis of coronavirus disease 2019 (COVID-19) in Korea. Lab Med 2020;40:351–60 doi: 103343/alm2020405351.

Huang AT, Garcia-Carreras B, Hitchings MDT, Yang B, Katzelnick LC, Rattigan SM, et al. A systematic review of antibody mediated immunity to coronaviruses: antibody kinetics, correlates of protection, and association of antibody responses with severity of disease. Nat Commun 2020;11(1):4704.

Kabesch M, Roth S, Brandstetter S, Hausler S, Juraschko E, Weigl M, et al. Successful containment of Covid-19 outbreak in a large maternity and perinatal center while continuing clinical service. Pediatr Allergy Immunol 2020;31(5):560–4.

Kucirka L, Lauer S, Laeyendecker O, Boon D, Lessler J. Variation in false negative rate of RT-PCR based SARS-CoV-2 tests by time since exposure. Preprint at medRxiv 2020 doi: 101101/2020040720051474 2020.

Okba NMA, Muller MA, Li W, Wang C, GeurtsvanKessel CH, Corman VM, et al. Severe acute respiratory syndrome coronavirus 2-specific antibody responses in coronavirus disease patients. Emerg Infect Dis 2020;26(7):1478–88.

Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nature communications 2020;11(1):1620.

Passing H. Bablok. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. J Clin Chem Clin Biochem 1983;21(11):709–20.

Tan W, Lu Y, Zhang S, Wang J, Dan Y, Tan Z, et al. Antibody responses in patients with COVID-19. doi: 101101/2020032420042382 2020.

To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis 2020;20(5):565–74.

van der Heide V. Neutralizing antibody response in mild COVID-19. Nat Rev Immunol 2020;20(6):352.

Whitman JD, Hiatt J, Mowery CT, Shy BR, Yu R, Yamamoto TN, et al. Test performance evaluation of SARS-CoV-2 serological assays. medRxiv 2020 2020.04.25.20074856 https://doi.org/10.1101/2020.04.25.20074856.

Wikramaratna P, Paton RS, Ghafari M, Lourenco J. Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR. Preprint at medRxiv 2020 doi: 101101/2020040520053355 2020.

Wolfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Muller MA, et al. Virological assessment of hospitalized patients with COVID-2019. Nature 2020;581(7809):465–9.

World Health Organization. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases Ig. Publication Server of WHO. Interim guidance 2020 WHO/COVID-19/laboratory/2020.5.

Wu F, Wang A, Liu M, Wang Q, Chen J, Xia S, et al. Preprint at medRxiv; doi: 10.1101/2020.03.30.20047365. 2020.
Xiao AT, Gao C, Zhang S. Profile of specific antibodies to SARS-CoV-2: the first report. J Infect 2020;81(1):147–78.
Xiao T, Wang Y, Yuan J, Ye H, Wei L, Liao X, et al. Early viral clearance and antibody kinetics of COVID-19 among asymptomatic carriers. Front Med (Lausanne) 2021;8:595773.
Yongchen Z, Shen H, Wang X, Shi X, Li Y, Yan J, et al. Different longitudinal patterns of nucleic acid and serology testing results based on disease severity of COVID-19 patients. Emerg Microbes Infect 2020;9(1):833–6.
Younes N, Al-Sadeq DW, Al-Jighefee H, Younes S, Al-Jamal O, Daas HI, et al. Challenges in laboratory diagnosis of the novel coronavirus SARS-CoV-2. Viruses 2020;12(6):582.
Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients with novel coronavirus disease 2019. Clin Infect Dis 2020;71(16):2027–34.
Zhou W, Wang W, Wang H, Lu R, Tan W. First infection by all four non-severe acute respiratory syndrome human coronaviruses takes place during childhood. BMC Infect Dis 2013;13:433.
Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020;382(8):727–33.