



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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Clinical evaluation of commercial automated SARS-CoV-2 immunoassays



Maximilian Kittel^{b,*}, Maria Christina Muth^{a,1}, Ingrid Zahn^a, Heinz-Jürgen Roth^a, Margot Thiaucourt^b, Catharina Gerhards^b, Verena Haselmann^b, Michael Neumaier^b, Peter Findeisen^a

^a MVZ Laboratory Dr. Limbach & Colleagues, Heidelberg, Germany

^b Institute for Clinical Chemistry, Medical Faculty Mannheim of the University of Heidelberg, Mannheim, Germany

ARTICLE INFO

Article history:

Received 9 September 2020

Received in revised form 30 November 2020

Accepted 2 December 2020

Keywords:

Automated SARS-CoV-2 antibody detection

High throughput testing

Pandemic control

Seroprevalence

Seroconversion

Longitudinal monitoring of antibody

development

ABSTRACT

Objective: Numerous immunoassays for detecting antibodies directed against SARS-CoV-2 have been rapidly developed and released. Validations of these have been performed with a limited number of samples. The lack of standardisation might lead to significantly different results. This study compared ten automated assays from six vendors in terms of sensitivity, specificity and reproducibility.

Methods: This study compared ten fully automated immunoassays from the following vendors: Diasorin, Epitope Diagnostics, Euroimmun, Roche, YHLO, and Snibe. The retrospective part of the study included patients with a laboratory-confirmed COVID-19 infection, and controls comprised patients with a suspected infection, in whom the disease was excluded. Furthermore, biobanked sera were taken as negative controls (n = 97). The retrospective part involved four groups: (1) laboratory-confirmed COVID-19 infection (n = 183); (1B) suspected COVID-19 infection (n = 167) without a qRT-PCR result but positive serological results from at least two different assays, and suspected COVID-19 infection due to a positive serological result from the Roche assay (n = 295); (2) biobanked sera obtained from patients before the emergence of SARS-CoV-2 (n = 97) as negative controls; and (2A) probably COVID-19-negative sera with negative serological results from at least two different assays (n = 152).

Results: Overall diagnostic sensitivities were: Euroimmun (IgA) 87%; Epitope Diagnostics (IgG) 83%; YHLO (IgG) 77%; Roche (IgM/IgG) 77%; Euroimmun (IgG) 75%; Diasorin (IgG) 53%; Epitope Diagnostics (IgM) 52%; Snibe (IgG) 47%; YHLO (IgM) 35%; and Snibe (IgM) 26%. Diagnostic specificities were: YHLO (IgG) 100%; Roche, 100%; Snibe (IgM/IgG) 100%; Diasorin (IgG) 97%; Euroimmun (IgG) 94%; YHLO (IgM) 94%; Euroimmun (IgA) 83%.

Conclusion: Assays from different vendors substantially varied in terms of their performance. These findings might facilitate selection of appropriate serological assays.

© 2020 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread rapidly, and the resulting coronavirus disease 2019 (COVID-19) has been declared a public health emergency of international concern by the World Health Organization. Rapid and

accurate diagnosis of the disease is of the utmost importance for subsequent disease management.

While molecular testing with real-time reverse transcriptase polymerase chain reaction (qRT-PCR) has been the primary means of diagnosing acute SARS-CoV-2 infection, serological testing is gaining importance for diagnosing subacute infections or supporting the diagnosis of respiratory insufficiency in cases where the pathogen is no longer detectable in the upper respiratory tract (Cheng et al., 2020; Loeffelholz and Tang, 2020; Patel et al., 2020; Tang et al., 2020; Yan et al., 2020). Specifically, serology can facilitate the diagnosis of SARS-CoV-2 infections when swab specimens were incorrectly collected and the molecular assays might return a false negative (Zhang et al., 2020). The viral load in swabs decreases in the first week of infection, whereas the

* Corresponding author at: Institute for Clinical Chemistry, Medical Faculty Mannheim of the University of Heidelberg, Theodor Kutzer Ufer 1-3, 68167 Mannheim, Germany.

E-mail address: maximilian.kittel@umm.de (M. Kittel).

¹ Both authors contributed equally to this work.

concentration of antibodies in serum increases. After approximately 8 days of COVID-19, serological testing is more sensitive than viral nucleic acid detection (Guo et al., 2020; Wolfel et al., 2020). Specifically, antibodies directed against SARS-CoV-2 can be detected in the serum of approximately 40% of COVID-19 patients as early as 7 days after symptom onset, with seroconversion rates rapidly increasing to >90% by Day 14 (Zhao et al., 2020). However, serology devices have just received urgent approval from the National Medical Products Administration in China and from regulatory authorities in various other countries (Farnsworth and Anderson, 2020; Loeffelholz and Tang, 2020). This has resulted in the widespread release of manufacturer-developed laboratory tests that are neither harmonised, standardised, nor thoroughly validated (Farnsworth and Anderson, 2020). Assays vary in terms of their format, the detected antibody class and targeted antigens, and this might lead to poor comparability of results (Okba et al., 2020; Theel et al., 2020). A comprehensive overview of currently available serological test systems can be found on the Johns Hopkins Center for Health Security website (health, 2020).

This study aimed to compare the diagnostic sensitivity and specificity of commercially available SARS-CoV-2 immunoassays. The selection of these tests was based on availability in the current laboratory at the time of testing.

Materials and methods

Serum specimens

Serum tubes (7.5 mL Gel-Monovette; Sarstedt, Nümbrecht, Germany) were used for blood sampling. Specimens were centrifuged at 20 °C for 10 min at 2500 g. Serum was collected and the samples were analysed as soon as possible. If the analyses were delayed, samples were aliquoted and stored between 2–8 °C for a maximum of 3 days. In cases where they had to be stored for longer than 3 days, the serum samples were cryopreserved at –20 °C.

Preanalytical stability

To assess stability of the analytes in serum, five freshly obtained serum samples were analysed with the Roche IgM/IgG assay under the following preanalytical conditions: (1) measurements were performed immediately with the freshly obtained specimens; (2) one aliquot of any specimen was stored at –20 °C prior to delayed measurements; (3) a second aliquot of any specimen was left at room temperature (20–25 °C) for 7 days prior to freezing at –20 °C. All frozen aliquots were thawed and measured independently on Day 7. Measurements were performed with the Roche SARS-CoV-2 assay. These biomaterials were assembled during the period from April to June 2020 inclusive.

Population

Patient serum samples were submitted to the MVZ Labor Dr. Limbach (Heidelberg, Germany) for diagnostic purposes. All leftover serum specimens were anonymised and stored at –20 °C until further use. The following four classes were defined: (1) COVID-19 cases confirmed by qRT-PCR (n = 183); (1B) probably COVID-19-positive sera with no qRT-PCR result but positive serological results from at least two or more different assays (n = 167) and samples with a suspected COVID-19 infection due to a positive serological result from the Roche IgM/IgG assay (n = 295); (2) pre-pandemic sera (n = 97); and (2A) probably COVID-19-negative sera with negative serological results from at least two or more different assays (n = 152). Due to the use of leftover material, not all samples were passed through all platforms. This part of the

study was performed on patient material previously collected for diagnostic purposes (secondary use). The anonymised patient material was used in accordance with the Statement of the Central Ethics Committee of the German Medical Association (Bundesarztekammer, 2003). The subjects for the longitudinal monitoring of antibody development were recruited from the IMMUNITOR study at University Medical Center Mannheim, Medical Faculty Mannheim, University of Heidelberg, Germany (<https://www.immunitor.de>). Recruited subjects with COVID-19 confirmed by qRT-PCR (n = 15) were evaluated. Blood draw was performed at different time points up to 14 weeks after positive qRT-PCR results. Patients' characteristics are shown in Table 1. The study was approved by the Institutional Review Board, and written consent was obtained from each subject before sample collection and analysis.

Serological testing

The SARS-CoV-2 serology was routinely performed on an iFLASH 1800 using the IgM and IgG assays from Shenzhen YHLO Biotech. In addition, alternative serological assays from six suppliers listed in Table 2 were performed in parallel. All serum antibody tests were performed with fully automated analysers according to the manufacturers' instructions. Titers were calculated and results interpreted according to the manufacturers' protocols.

Reproducibility

Intraday and interday precision were determined for assays from Roche and Shenzhen YHLO Biotech. For the Elecsys Anti-SARS-CoV-2 electrochemiluminescence immunoassay (ECLIA), commercially available controls with defined cut-off interval (COI) values and one human serum pool near the cut-off value were used. Pooled patient samples were used to evaluate the Shenzhen YHLO Biotech assay. Due to unavailability of commercially available controls, the evaluation of the Shenzhen YHLO Biotech test was conducted exclusively by pooling patient material. Repeated measurements for controls and pooled serum samples were performed five times on five consecutive days, in line with the EP15-A3 protocol from the Clinical and Laboratory Standards Institute (CLSI).

Table 1
Characteristics of patients and controls. Descriptive statistics of patients with SARS-CoV-2-positive qRT-PCR results from University Hospital Mannheim.

Sex	(n = 15)
Female	73.3%
Male	33.3%
Age (years)	average: 51.2 (23–64)
Weight (kg)	average 87.4 (70–165)
Symptoms	
Fever	60.0%
Night sweat	46.7%
Diarrhoea	20.0%
Cough	40.0%
Shortness of breath	40.0%
Muscle pain	80.0%
Nausea	33.3%
Anosmia	80.0%
Comorbidities	
Autoimmune deficiency	13.3%
Hypertension	40.0%
Asthma	20.0%
Active cancer disease	0.0%
Renal impairment	0.0%
Hepatic disease	0.0%

Table 2

Overview of suppliers. Chemiluminescence immunoassay (CLIA), enzyme-linked immunosorbent assay (ELISA), nucleocapsid protein (N-protein), spike protein (S-protein).

Supplier (alphabetic order)	Analyser	Method	Antibody	Antigen
DiaSorin	Liaison	CLIA	IgG	S1- and S2 Subunit of S-protein
Epitope Diagnostic	DSX	ELISA	IgG and IgM	N-protein
Euroimmun	WorkStation	ELISA	IgG and IgA	S1-subunit of S-protein
Roche Elecsys	Cobas 801	ECLIA	IgG/M	N-protein
Yhlo Biotech	iFlash 1800	CLIA	IgG and IgM	S- and N-protein
Snibe Diagnostic	Maglumi 800	CLIA	IgG and IgM	S- and N-protein

Statistical analyses

Sensitivity was defined as the proportion of patients correctly identified as having SARS-CoV-2 infections that were confirmed by positive qRT-PCR in respiratory samples. Specificity was defined as the proportion of SARS-CoV-2 immune-naive study participants accurately identified as negative for COVID-19. The clinical accuracies of the ELISA assays were examined by using area under the receiver operating characteristic (AUROC) plots with Abacus 2.0 software (LABanalytics GmbH, Germany). AUROCs were calculated as the fraction ‘correctly identified as positive’ and the fraction ‘falsely identified as positive’ determined according to manufacturers’ cut-off values for positive results. The EP15-A3 model from the CLSI was used to estimate imprecision, reproducibility, linearity, and limit of detection (LOD). These validation steps were carried out using Abacus 2.0 (LABanalytics GmbH, Germany).

For normalisation of data, the Z-score was calculated (i.e. the difference between the measured value and the assigned value corrected for the variability) (Coucke and Soumali, 2017).

Results

Reproducibility

The intra-assay variabilities of the SARS-CoV-2 IgM and IgG tests from Shenzhen YHLO Biotech and Roche were determined five times on five consecutive days, in line with the EP15-A3 protocol from the CLSI. The inter-assay variabilities were determined by a single measurement of a negative and a positive control over 15 days. Figure 1 shows the precision data for Shenzhen YHLO Biotech’s IgM and IgG tests. Reproducibility was high, as coefficients of variation (CV) for all positive specimens were <8%.

Preanalytical stability

Preanalytical stability was high and no statistically significant deviation between results from initial and delayed measurements could be detected. The CV for all measurements was 4.8%, which is within the range of the abovementioned reproducibility.

Linearity

Positive serum specimens with titres at least 10 times above the cut-off value were diluted 1:1 stepwise. The dilution series were performed for the Shenzhen YHLO Biotech IgG assay and for the Roche IgG/IgM assay. Both assays were not linear and saturation was observed for higher concentrations (Supplemental Figure 1). The limit of blank (LOB) for the SARS-CoV-2 IgM test was 0.192 AU/mL, and 0.05 AU/mL for the IgG test. The limit of detection (LOD) for the SARS-CoV-2 IgM test was 0.36 AU/mL, and 0.145 AU/mL for the IgG test. The ECLIA test from Roche yielded a LOD of 0.068 COI and a limit of quantification (LOQ) of 0.082 COI.

Z-scores

The SARS-CoV-2 IgM and IgG assays from Shenzhen YHLO Biotech (cut-off value: 10 AU/mL) and Roche Diagnostics (cut-off value: 1 COI/QE) were compared. The data were standardised using the Z-score calculation with Log (1) as the cut-off value (Coucke and Soumali, 2017). QRT-PCR-confirmed COVID-19 sera (n = 59) and pre-pandemic control sera (n = 50) were used. The Shenzhen YHLO Biotech IgG test correctly identified more COVID-19 cases as positive than the Roche IgM/IgG test (Figure 2). However, in subsequent analyses (see Section ‘Comparison of the two superior tests’), samples were identified that were positive with the Roche IgM/IgG test but negative with the Shenzhen YHLO Biotech IgG test. The Shenzhen YHLO Biotech IgM assay showed the lowest performance. The control sera were correctly identified as negative by both IgG tests. It is worth noting the large interquartile distance of the Roche IgM/IgG assay among the COVID-19-positive sera, and the particularly small interquartile distance among the negative sera.

Sensitivity

Figure 3 shows the sensitivity of SARS-CoV-2 antibody tests from six manufacturers. Serum specimens from qRT-PCR-confirmed patients (Class 1) were analysed. The most sensitive SARS-CoV-2 antibody test was the Euroimmun IgA test with 87%, followed by the Epitope Diagnostics IgG test with 83%, the

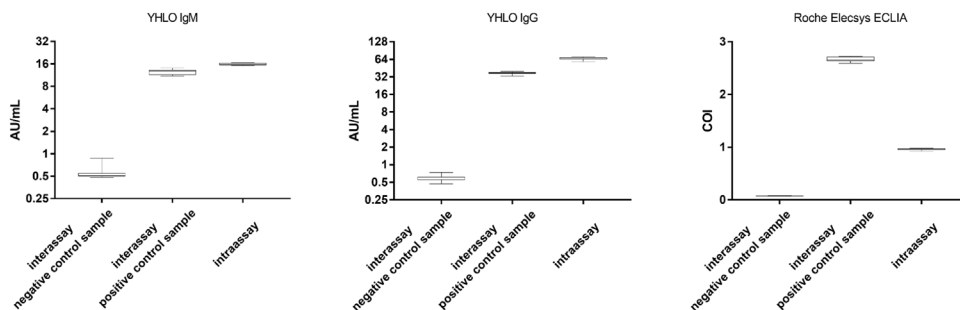


Figure 1. Precision box-whiskers plot of the inter-assay and intra-assay variations in the IgM and IgG SARS-CoV-2 assay from Shenzhen YHLO Biotech and Roche Elecsys Anti-SARS-CoV-2 ECLIA.

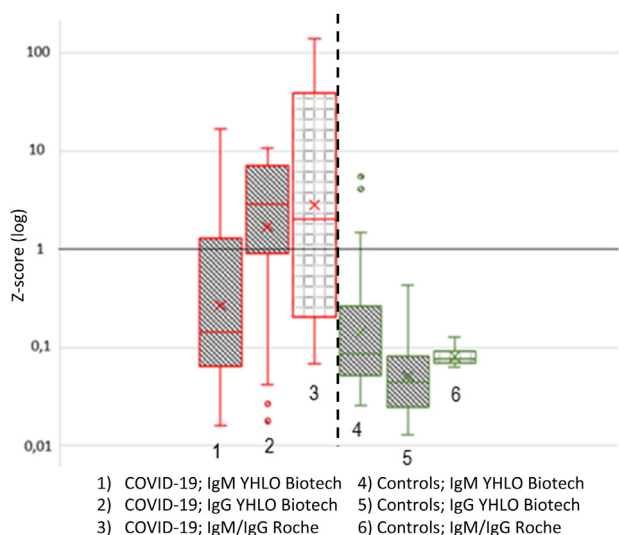


Figure 2. Z-Score comparison. Comparison between the IgM and IgG SARS-CoV-2 assay from Shenzhen YHLO Biotech and the IgM/IgG assay by Roche Diagnostics. Data were transformed to Z-scores, as described in the materials and methods section, and the cut-off values were uniformly set to log (1). Measurement of qRT-PCR-confirmed COVID-19 cases (n = 59) and pre-pandemic negative controls (n = 50).

Shenzhen YHLO Biotech IgG test with 77%, the Roche IgM/IgG test with 77%, the Euroimmun IgG test with 75%, the Diasorin IgG test with 53%, the Epitope Diagnostics IgM test with 52%, the Snibe IgG Test with 47%, the Shenzhen YHLO Biotech IgM test with 35%, and the Snibe IgM test with 26%. However, if only sera collected at least 10 days after positive RT-PCR were included in the sensitivity calculation, the sensitivity changed as follows: Shenzhen YHLO Biotech IgG 89%; Euroimmun IgG 88%; Roche IgM/IgG 86%; Euroimmun IgA 85%; Epitope Diagnostics IgG 78%; Snibe IgG 58%; Diasorin IgG 58%; Epitope Diagnostics IgM 44%; Shenzhen YHLO Biotech IgM 33%; and Snibe IgM 26%.

Specificity

Serum specimens of pre-pandemic patients (Class 2) were analysed. Specificity of 100% was achieved with the Shenzhen YHLO Biotech IgG assay, Roche IgM/IgG assay and Snibe IgM and IgG assays. Diasorin had a specificity of 97%, Euroimmun IgG had 94%, Shenzhen YHLO Biotech IgM had 94%, and Euroimmun IgA

had 83%. Due to their low sensitivity performance and the limited amount of sample material, the assays from Epitope Diagnostics were not further tested for specificity.

AUROC analysis

The diagnostic power of all IAs listed in Table 2 was determined by area under the receiver operating characteristic (AUROC) analyses. AUROC values ranged from 0.97 (Shenzhen YHLO Biotech IgG) to 0.66 (Snibe IgM) and are shown in Figure 4. The IgG assay from Shenzhen YHLO Biotech and the Roche IgM/IgG assay had the highest AUROC values of 0.97 and 0.96, respectively. The AUROC values of all tested IAs are displayed in Table 3.

Concordance analysis

Identical specimens of proven (Class 1) and probable (Class 2) COVID-19 patients, as well as controls from pre-pandemic patients (Class 2) and SARS-CoV-2-negative patients (Class 2A) were tested in parallel with all IAs listed in Table 1. The highest concordance was observed between the Shenzhen YHLO IgG assay and the Roche IgM/IgG assay, with 83%, 94%, 100%, and 100% for classes 1, 1B, 2, and 2A, respectively (Supplementary Figure 2).

Comparison of the two superior tests

Due to the necessity of higher throughput, the initially established assay from Shenzhen YHLO Biotech was substituted by the IgM/IgG assay from Roche Diagnostics. Routine samples with positive results from the Roche IgM/IgG assay (n = 295) were subsequently measured with the Shenzhen YHLO Biotech assays. The Roche IgM/IgG assay showed 91.86% (271 of 295) agreement with the Shenzhen YHLO Biotech IgG assay. However, in 8.14% (24 of 295) of all cases, only the Roche IgM/IgG assay was positive and the Shenzhen YHLO Biotech IgG assay was negative. The median values (standard deviations, SD) of these 24 divergent results were 3.3 (SD: 3.35) for the Roche assay (cut-off: 1.0 COI) and 5.2 (SD: 3.3) for the Shenzhen YHLO Biotech IgG (cut-off: 10 AU/mL).

Longitudinal monitoring

Sera were analysed with the IgM and IgG tests from Shenzhen YHLO Biotech. The kinetics of the immune response were modelled with COVID-19-positive sera confirmed by qRT-PCR (n = 15), and blood draw was performed at various time points up to 14 weeks

Manufacturer	Measurand	Sensitivity (%) days after positive RT-PCR													
		Total		- 6-0		1-7		8-14		15-21		22-28		≥ 29	
		n	%	n	%	n	%	n	%	n	%	n	%	n	%
Snibe Co.	IgM	10/38	26%	2/5	40%	0/10	0%	3/10	30%	3/9	33%	0/1	0%	0/1	0%
	IgG	18/38	47%	3/5	60%	0/10	0%	7/10	70%	5/9	56%	0/1	0%	1/1	100%
Yhlo Biotech	IgM	49/142	35%	10/17	59%	6/25	24%	7/17	41%	7/25	28%	7/19	37%	11/32	31%
	IgG	130/168	77%	14/18	78%	8/25	32%	17/18	94%	23/26	88%	26/29	90%	33/39	85%
Euroimmun AG	IgA	53/61	87%	8/9	89%	6/6	100%	10/13	77%	12/13	92%	10/11	91%	5/6	83%
	IgG	48/64	75%	6/9	67%	1/6	17%	13/13	100%	9/13	69%	11/11	100%	5/6	83%
Epitope Diagnostics	IgM	15/29	52%	5/7	71%	0/1	0%	3/6	50%	2/9	22%	3/4	75%		
	IgG	24/29	83%	6/7	86%	1/1	100%	6/6	100%	6/9	67%	3/4	75%		
DiaSorin	IgG	29/55	53%	5/6	83%	0/7	0%	4/7	57%	3/8	38%	7/9	78%	6/11	55%
Roche Diagnostics	IgM/IgG	68/86	77%	4/6	67%	3/10	30%	3/3	100%	7/13	54%	12/15	80%	36/36	100%

Figure 3. Diagnostic sensitivity. Comparison of serological results for all immunoassays listed in Table 2: Serum specimens of qRT-PCR-confirmed patients (Class 1) were analysed. Blood from patients with infection was drawn on various days after positive qRT-PCR results, as shown in the figure. The dates on which the qRT-PCRs were performed are unknown for 13 samples.

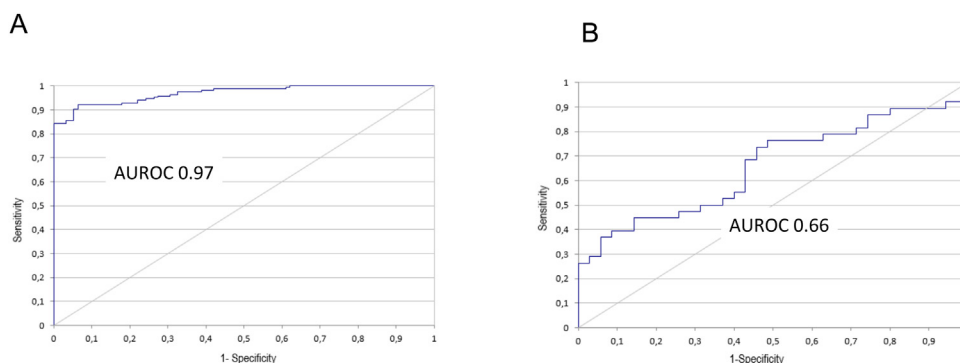


Figure 4. AUROC curves. AUROC graphs for the highest ((A) Shenzhen YHLO Biotech IgG, AUC: 0.97) and lowest ((B) Snibe IgM, AUC: 0.66) values; detailed information can be found in Table 3.

Table 3
AUROC calculations. Diagnostic performance of all immunosorbent assays listed in Table 2, as calculated by AUROC values.

	AUC (CI)	Positive (n)	Negative (n)	Total (n)
YHLO Biotech IgG	0.97 (0.95–0.99)	167	95	262
ROCHE	0.96 (0.92–0.96)	87	50	137
Euroimmun IgA	0.93 (0.88–0.98)	61	35	96
Euroimmun IgG	0.92 (0.86–0.97)	64	35	99
Snibe IgG	0.86 (0.77–0.95)	38	35	73
DiaSorin IgG	0.83 (0.75–0.90)	55	65	120
YHLO Biotech IgM	0.79 (0.74–0.85)	142	95	237
Snibe IgM	0.66 (0.53–0.78)	38	35	73

after positive qRT-PCR results. The kinetics of the antibody response showed high interindividual variability and were grouped into four categories: (I) decreasing, (II) increasing, (III) steady, and (IV) variable concentrations. Most interestingly, antibody formation could not be detected for IgM or IgG, even 9 weeks after positive qRT-PCR (Figure 5).

Discussion

Serology testing for SARS-CoV-2 is increasingly becoming an interesting way of better quantifying the number of COVID-19 infections. Specifically, patients who have a subclinical infection or are perhaps even asymptomatic often lack results from molecular

diagnostic assays. Accordingly, the combination of molecular and serological assays can greatly improve the diagnostic efficacy (Krammer and Simon, 2020; Liu et al., 2020; Lou et al., 2020; Winter and Hegde, 2020). Serological testing is an integral component of a testing algorithm that has been proposed by the Infectious Diseases Society of America (Lu et al., 2020). However, due to their rapid introduction, the available SARS-CoV-2 assays have not been thoroughly validated. Furthermore, different immunoassays vary in terms of format, detected antibody class, targeted antigen, and acceptable specimen types (Bryant et al., 2020; Okba et al., 2020; Theel et al., 2020). Therefore, results from different assays are not necessarily congruent, as the respective diagnostic specificities and sensitivities vary (Younes et al., 2020).

Several method comparisons have been published over the last month. Some authors compared a limited number of tests comprising not more than two immunoassays (Kohmer et al., 2020b; Montesinos et al., 2020; Nicol et al., 2020). In contrast, Lassaunière et al. (2020) evaluated nine commercial SARS-CoV-2 immunoassays; however, the majority of these were point-of-care tests and only two vendors (Wantai and Euroimmun) offered automated test formats. Plebani et al. (2020) compared the reliability of three chemiluminescent immunoassays (CLIA) and two enzyme-linked immunosorbent assays (ELISA). However, the focus of their work was the definition of optimised thresholds that might improve the negative likelihood ratio and reduce the false positives that can hamper any screening approach in an

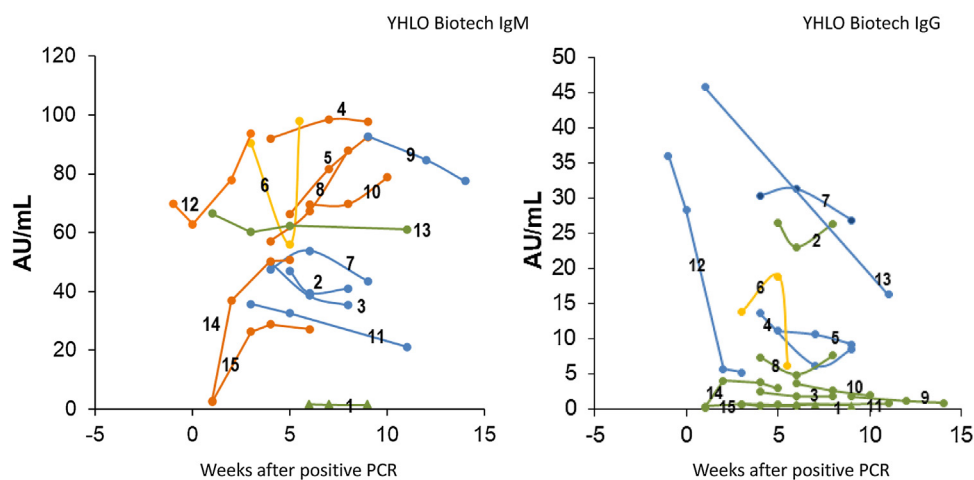


Figure 5. Longitudinal monitoring. Kinetics of antibody concentrations were analysed for 15 patients with qRT-PCR-confirmed COVID-19 infections. Multiple blood samplings were performed for up to 14 weeks after qRT-PCR. The cut-off values for the Shenzhen YHLO Biotech assays (10 AU/mL) are indicated with horizontal lines. The courses of antibody concentrations were grouped into four categories: (I) decreasing (blue), (II) increasing (orange), (III) steady (green), and (IV) variable (yellow). Most interestingly, antibody formation could not be detected for IgM or IgG in one case, even 9 weeks after a positive qRT-PCR result (triangle symbols, Patient 1).

asymptomatic population (Plebani et al., 2020). A recent study from Weidner et al. (2020) focused on a quantitative correlation between different commercial assays, including automated platforms from Euroimmun, Wantai, Roche, and Diasorin. However, the study only included sera from SARS-CoV-2 convalescent patients and no COVID-19-negative controls. In a further study, Kohmer et al. (2020a) compared the performance of six automated immunoassays (from Abbott, Roche, Diasorin, Virclia, Euroimmun, and Virotech). However, the analyses were restricted to IgG and total antibodies, and only limited numbers of COVID-19 patients (n = 45) and non-COVID-19 patients (n = 37) were analysed. In the most recent study, Kruttgen et al. (2020) compared four commercial assays. Again, the study's shortcoming was the limited number of sera from patients who tested positive or negative by SARS-CoV-2 PCR (n = 75). The issue of small sample size has been highlighted by a systematic review and meta-analysis (Lisboa Bastos et al., 2020).

This study presented data from 10 immunoassays comprising IgA, IgM, IgG, and combined IgG/IgM tests from six vendors: Diasorin, Epitope Diagnostics, Euroimmun, Roche, Shenzhen YHLO Biotech, and Snibe. The study included over 900 serum specimens (n = 904) from patients with proven (n = 193) and probable (n = 167) COVID-19 infection, and controls from pre-pandemic patients (n = 97) and probable SARS-CoV-2-negative patients (n = 152). Specimens with SARS-CoV-2-positive results from the Roche IgG/IgM test (n = 295) were re-analysed with the Shenzhen YHLO Biotech IgG assay. As expected, assay performance was markedly different, with AUROC values ranging from 0.66 (Snibe IgM) to 0.97 (Shenzhen YHLO Biotech IgG). The best concordance was observed between the Shenzhen YHLO Biotech IgG and the Roche IgG/IgM tests.

The comparison of the Shenzhen YHLO Biotech IgG assay and the Roche IgG/IgM assay revealed a small subset of specimens that were positive in one of the two assays. This indicated an individual immune response to SARS-CoV-2 and the influence of the assays used (Kohmer et al., 2020a). Antibodies that are exclusively detected by the N-protein-based Roche assay (Table 2) might not have neutralising capacity. In contrast, antibodies directed against the S-protein can interfere with the virus's ability to bind to the ACE receptor of the host cell and will most likely have neutralising capacity. Further analyses are needed to decipher the humoral immune response (Kreer et al., 2020).

The IgG-positive rate was consistently higher than the IgM-positive rate, and this phenomenon was also observed in another study (Zhang et al., 2020). The clinical value of IgM for early diagnosis of COVID-19 is currently unclear. SARS-CoV-2-specific IgM does not always appear before its IgG counterpart. Some studies have reported the detection of SARS-CoV-2-specific IgG even before IgM (Thevarajan et al., 2020; To et al., 2020; Zhao et al., 2020). However, even if IgM sensitivity is superior in a few cases, the low specificity of IgM antibodies is an obstacle to clear interpretations of test results. Specifically, a low prevalence of COVID-19 patients leads to significant numbers of false positives, not only for IgM assays but also for IgG assays with high specificity (Lisboa Bastos et al., 2020). Furthermore, the possibility of false negative results and the phenomenon of missing seroconversion will have to be discussed. The current study observed the repeated absence of SARS-CoV-2-specific IgG and IgM antibodies for one patient as long as 9 weeks after a positive qRT-PCR result (Figure 5). These negative results have been confirmed by independent measurements with the Roche IgG/IgM assay. A multiplexed assay format for simultaneous quantitation of IgG, IgM, and IgA immunoglobulins against four SARS-CoV-2 targets was recently described (Norman et al., 2020). It resulted in the monitoring of 12 antibody isotype-viral protein interactions, and this high-resolution profile of the immune response might be needed to

resolve rare cases of missing seroconversion. Another explanation for this phenomenon might be a false-positive test result. These cases are rare but have been reported in literature (fda.gov, 2020; Surkova et al., 2020).

This study had several limitations. First, clinical data for most patients were unavailable because all of the specimens from the MVZ Laboratory in Heidelberg were anonymised. Second, the humoral immune response might have been affected by the severity of COVID-19 infections or by various comorbidities (Long et al., 2020; To et al., 2020; Zhao et al., 2020). Furthermore, the serological diagnosis of COVID-19 is strongly influenced by the time of blood draw, as seroconversion is time-dependent. The date of positive qRT-PCR was taken as the reference timepoint in the current study. However, as virus elimination and positive qRT-PCR can be prolonged (Guo et al., 2020; Wolfel et al., 2020), the onset of symptoms might have been a more accurate reference point; this was only available for a small fraction of the samples. Finally, due to the use of leftover material, not all samples were passed through all platforms in this retrospective study. A further limitation was the collective 2B, where the Roche assay was used as a surrogate marker of an experienced SARS-CoV-2 infection. This risked positively biasing the study results in favour of the Roche assay, but had to be chosen due to the study design. Laboratories that are not directly integrated into a hospital infrastructure often do not receive the corresponding PCR results. The Roche test was chosen because the manufacturer could demonstrate a specificity collective of almost 5000 patients at that time. These limitations indicate the need for large, prospective and multicentre studies to elucidate the advantages and disadvantages of serological testing.

Ethics approval

The study project was approved by the local ethics committee (ethics committee II Medical Faculty Mannheim of the University of Heidelberg).

Funding

All study projects were carried out without funding.

Conflict of interest

None declare.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.ijid.2020.12.003>.

References

- Bryant JE, Azman AS, Ferrari MJ, Arnold BF, Boni MF, Boum Y, et al. Serology for SARS-CoV-2: apprehensions, opportunities, and the path forward. *Sci Immunol* 2020;5(47).
- Bundesärztekammer ZEBd. Die (Weiter-)Verwendung von menschlichen Körpermaterialien für Zwecke medizinischer Forschung. *Dtsch Arztebl* 2003;100(23): A1632.
- Cheng MP, Papenburg J, Desjardins M, Kanjilal S, Quach C, Libman M, et al. Diagnostic testing for severe acute respiratory syndrome-related coronavirus 2: a narrative review. *Ann Intern Med* 2020;172(11):726–34.
- Coucke W, Soumali MR. Demystifying EQA statistics and reports. *Biochem Med (Zagreb)* 2017;27(1):37–48.
- Farnsworth CW, Anderson NW. SARS-CoV-2 serology: much hype, little data. *Clin Chem* 2020;.
- fda.gov. False Positive Results with BD SARS-CoV-2 Reagents for the BD Max System-Letter to Clinical Laboratory Staff and Health Care Providers. 2020 Available from: <https://www.fda.gov/medical-devices/letters-health-care-providers/false-positive-results-bd-sars-cov-2-reagents-bd-max-system-letter-clinical-laboratory-staff-and>. [Accessed 20 October 2020].

- Guo L, Ren L, Yang S, Xiao M, Chang, Yang F, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clin Infect Dis* 2020; health JHBSop. Serology-Based Tests for COVID-19. 2020 Available from: <https://www.centerforhealthsecurity.org/resources/COVID-19/serology/Serology-based-tests-for-COVID-19.html>. [Accessed 1 September 2020].
- Kohmer N, Westhaus S, Rühl C, Ciesek S, Rabenau HF. Brief clinical evaluation of six high-throughput SARS-CoV-2 IgG antibody assays. *J Clin Virol* 2020a;129:104480.
- Kohmer N, Westhaus S, Rühl C, Ciesek S, Rabenau HF. Clinical performance of different SARS-CoV-2 IgG antibody tests. *J Med Virol* 2020b;.
- Krammer F, Simon V. Serology assays to manage COVID-19. *Science* 2020;368(6495):1060–1.
- Kreer C, Zehner M, Weber T, Ercanoglu MS, Giesemann L, Rohde C, et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* 2020;.
- Kruttgen A, Cornelissen CG, Dreher M, Horneff M, Imohl M, Kleines M. Comparison of four new commercial serologic assays for determination of SARS-CoV-2 IgG. *J Clin Virol* 2020;128:104394.
- Lassaunière R, Frische A, Harboe Z, Nielsen A, Fomsgaard A, Krogfelt K, et al. Evaluation of nine commercial SARS-CoV-2 immunoassays. *medRxiv* 2020; preprint.
- Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui L-P, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* 2020;370:m2516.
- Liu L, Liu W, Zheng Y, Jiang X, Kou G, Ding J, et al. A preliminary study on serological assay for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in 238 admitted hospital patients. *Microbes Infect* 2020;22(4–5):206–11.
- Loeffelholz MJ, Tang YW. Laboratory diagnosis of emerging human coronavirus infections—the state of the art. *Emerg Microbes Infect* 2020;9(1):747–56.
- Long QX, Liu BZ, Deng HJ, Wu GC, Deng K, Chen YK, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* 2020;26(6):845–8.
- Lou B, Li TD, Zheng SF, Su YY, Li ZY, Liu W, et al. Serology characteristics of SARS-CoV-2 infection since exposure and post symptom onset. *Eur Respir J* 2020;.
- Lu H, Stratton CW, Tang YW. An evolving approach to the laboratory assessment of COVID-19. *J Med Virol* 2020;.
- Montesinos I, Gruson D, Kabamba B, Dahma H, Van den Wijngaert S, Reza S, et al. Evaluation of two automated and three rapid lateral flow immunoassays for the detection of anti-SARS-CoV-2 antibodies. *J Clin Virol* 2020;128:104413.
- Nicol T, Lefevre C, Serri O, Pivert A, Joubaud F, Dubee V, et al. Assessment of SARS-CoV-2 serological tests for the diagnosis of COVID-19 through the evaluation of three immunoassays: two automated immunoassays (Euroimmun and Abbott) and one rapid lateral flow immunoassay (NG Biotech). *J Clin Virol* 2020;129:104511.
- Norman M, Gilboa T, Ogata AF, Maley AM, Cohen L, Cai Y, et al. Ultra-sensitive high-resolution profiling of anti-SARS-CoV-2 antibodies for detecting early seroconversion in COVID-19 patients. *medRxiv* 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.
- Patel R, Babady E, Theel ES, Storch GA, Pinsky BA, St George K, et al. Report from the American society for microbiology COVID-19 international summit, 23 march 2020: value of diagnostic testing for SARS-CoV-2/COVID-19. *mBio* 2020;(2):11.
- Plebani M, Padoan A, Negrini D, Carpinteri B, Sciacovelli L. Diagnostic performances and thresholds: the key to harmonization in serological SARS-CoV-2 assays?. *Clin Chim Acta* 2020;509:1–7.
- Surkova E, Nikolayevskyy V, Drobniewski F. False-positive COVID-19 results: hidden problems and costs. *Lancet Resp Med* 2020;.
- Tang YW, Schmitz JE, Persing DH, Stratton CW. Laboratory diagnosis of COVID-19: current issues and challenges. *J Clin Microbiol* 2020;58(6).
- Theel ES, Slev P, Wheeler S, Couturier MR, Wong SJ, Kadkhoda K. The role of antibody testing for SARS-CoV-2: is there one?. *J Clin Microbiol* 2020;.
- Thevarajan I, Nguyen THO, Koutsakos M, Druce J, Caly L, van de Sandt CE, et al. Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nat Med* 2020;26(4):453–5.
- 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.
- Weidner L, Gansdorfer S, Unterweger S, Weseslindtner L, Drexler C, Farcet M, et al. Quantification of SARS-CoV-2 antibodies with eight commercially available immunoassays. *J Clin Virol* 2020;129:104540.
- Winter AK, Hegde ST. The important role of serology for COVID-19 control. *Lancet Infect Dis* 2020;20(7):758–9.
- Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020;581(7809):465–9.
- Yan Y, Chang L, Wang L. Laboratory testing of SARS-CoV, MERS-CoV, and SARS-CoV-2 (2019-nCoV): current status, challenges, and countermeasures. *Rev Med Virol* 2020;30(3):e2106.
- 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).
- Zhang W, Du RH, Li B, Zheng XS, Yang XL, Hu B, et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. *Emerg Microbes Infect* 2020;9(1):386–9.
- Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clin Infect Dis* 2020;.