In this issue: Current laboratory aspects of COVID-19

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Verification and comparison of qualitative serological assays for Anti-SARS-COV-2 IgM and IgG antibodies detection

Maša Štebih[1,2], Milan Skitek[1], Aleš Jerin[1,2]

[1] *Institute of Clinical Chemistry and Biochemistry, University Medical Centre Ljubljana, Ljubljana, Slovenia*
[2] *Faculty of Pharmacy, University of Ljubljana, Slovenia*

ARTICLE INFO

Corresponding authors:
Dr. Aleš Jerin
University Medical Centre Ljubljana
Institute of Clinical Chemistry and Biochemistry
Zaloška 2, 1000 Ljubljana
Slovenia
E-mail: ales.jerin@kclj.si

Key words:
SARS-CoV-2, COVID-19, serological assays, verification, antibodies

ABSTRACT

Background

Due to their wide application in the SARS-CoV-2 pandemic, we verified and compared three qualitative serological methods in order to select the most optimal that will best serve its purpose under laboratory conditions.

Methods

We assessed the diagnostic characteristics of two automated serological methods (Roche Elecsys® Anti-SARS-CoV-2 and Abbott SARS-CoV-2 IgG) and a POCT test (Colloidal Gold Method SARS-CoV-2 IgM/IgG Antibody Assay Kit). In the process of verification, analytical precision was also assessed for the automated assays.

Results

Diagnostic characteristics were determined by measuring antibodies against SARS-CoV-2 in 91 RT-PCR-negative and 60 RT-PCR-positive samples. The POCT test gave the highest number of false positive cases (8.61%). Roche Elecsys® Anti-SARS-CoV-2 gave only 2.65% false positivity and showed the highest diagnostic sensitivity of 98.33% (95% CI: 91.06–99.96), while Abbott SARS-CoV-2 IgG method showed 100.00% (95% CI: 96.03–100.00) diagnostic specificity and an almost perfect agreement with Roche Elecsys® Anti-SARS-CoV-2. When assessing the precision of the automated methods, we observed some variability in the positive control samples, but the values did not affect clinical interpretation.

Conclusion

Both automated methods demonstrate superior diagnostic characteristics compared to the Colloidal Gold Method, and this POCT test is not considered as an appropriate choice for routine testing. The two automated methods showed low variability without altering the results and their interpretation.

INTRODUCTION

In December 2019, a new type of β-coronavirus began to emerge, which has been named Severe Acute Respiratory Syndrome Coronavirus, better known as SARS-CoV-2. Due to its severe pathogenicity and ability to spread in March 2020 the World Health Organization (WHO) declared it a global epidemic (1). Symptoms of COVID-19 infection are often nonspecific and heterogeneous, and depend on sex, age, immune status, viral load, associated diseases or possible history of other coronavirus infections (2–5). Despite the variation of responses from patient to patient, dry cough, fever, dyspnoea, loss of smell and taste are the most common symptoms and, in severe cases the infection may lead to death. In the same way, the kinetics of the immune response is also highly variable depending on the same factors as seen in symptoms (5,6). The severity and the magnitude of the epidemic with the variability of SARS-CoV-2 infection, which can affect the response of each person differently makes rapid, reliable and effective diagnosis essential. Diagnostic methods in an epidemic situation are playing an important role, particularly in controlling the epidemic and limiting the spread of the SARS-CoV-2 virus (7).

Diagnostics in SARS-CoV-2 epidemics is divided into indirect, serological assays that are measuring the humoral immune response, and direct diagnostic methods, which are detecting the presence of the virus by detecting viral RNA with RT-PCR or by detecting viral antigen, where POCT methods are most commonly used (7–9). POCT methods are particularly important for rapid diagnosis, whereas RT-PCR methods are more time-consuming and complex. Diagnostic sensitivity is crucial for both methods, as we want to minimise false negative results. To achieve this goal, it is important to be aware of the variability in viral load that is highest at the onset of symptoms, and sampling at the appropriate time point is also crucial for POCT methods, as they have inherently lower diagnostic sensitivity compared with RT-PCR (7,8,10–12).

Most common assays for serology are based on chemiluminescence or enzyme-linked immunosorbent principle. Immunochromatographic methods have also been developed for the purpose of point-of-care testing (6,13). Serological methods, which most often involve qualitative or quantitative detection of IgG and IgM antibodies, usually are directed against the nucleocapsid or spike protein of SARS-CoV-2. The combination of antibodies has been shown to be a more sensitive technique as part of the diagnostic approach

to infection identification and epidemic control (11). Seroconversion of IgM antibodies starts soon after the appearance of the symptoms but declines rapidly, whereas IgG antibodies appear in detectable concentrations around day 5 after the appearance of the symptoms and can remain detectable for several months (6,13,14). In relation to antibody seroconversion, serological methods are therefore not suitable to diagnose active infection, yet they are in an important adjunct to molecular methods, especially when the clinical picture is not consistent with the results (6,9,15). The combination of the two diagnostic approaches strongly increases the sensitivity of detecting the presence of infection in the acute phase (11,13).

Despite the limitations in detecting active infection, serological methods are of great importance for surveillance of the epidemiological situation, identification of patients who have been infected in the past and assessing the prevalence (8,9,12). Serological methods can also be a good approach to prognosis, since a correlation between the level of IgG antibodies against the nucleocapsid protein and the severity of the infection has been proven (12,16). In monitoring the immune response, serology may serve to identify those individuals who have developed a strong immune response and many of them can consequently be potential plasma donors for therapy of those, who have developed a more severe form of COVID-19 infection (8). The wide spectrum of use and importance of serological and other diagnostic methods, makes it essential to implement them as soon as possible especially during an epidemic. Despite the strong need for immediate implementation, a verification process is required before their use, mainly because of limitations, such as the impact of prevalence or disease stage, which may affect the sensitivity of the methods and consequently the quality of the results (8,10). Verification must cover the basic diagnostic and analytical properties of the method, as these are the characteristics that ensure the reliability of the results, and are crucial for the correct interpretation and comparison of the method with other methods, and also with other laboratories (17,18).

The aim of this paper is to verify three qualitative serological methods for the determination of specific antibodies against SARS-CoV-2 in order to determine which method gives the best results, best serves its purpose and consequently is the most optimal for early use in the laboratory. Repeatability, intermediate and intralaboratory precision (intra- and inter-daily) were assessed to automated methods according to the CLSI EP15-A3 protocol to determine whether variability affects the results and final clinical interpretation. Coefficients of variation (CVs) were also compared with the manufacturer's claims.

METHODS & MATERIALS

Study design

Serological analyses were performed on all three methods in the Hormone and Tumour Marker Laboratory and the Body Fluid Laboratory during the onset of the epidemic. RT-PCR analysis was performed at the Institute of Microbiology and Immunology. Serum samples were obtained from the staff at the University Medical Centre Ljubljana, the Clinic for Infectious Diseases and Febrile Conditions and the Clinical Institute of Clinical Chemistry and Biochemistry. A proportion of the samples also belonged to hospitalised patients infected with COVID-19. The samples were anonymised residues of routine diagnostic samples.

Defining diagnostical properties

As part of the verification of serological methods used, diagnostic specificity, diagnostic sensitivity and predictive values were determined. The results obtained by the serological methods were compared with the RT-PCR results considering

the cut-off values of the manufacturer. For ease of overview, a 2x2 contingency table was drawn to calculate the diagnostic characteristics for each method. The methods were compared according to the number of false results and Cohen's kappa coefficient (κ).

Assesing analytical precision of automated assays

The precision of the automated methods was assessed according to the CLSI EP15-A3 protocol. We performed a 5x5 experimental model and assessed the repeatability, intermediate precision and intra-laboratory precision (intra and inter-daily) (18). We used a laboratory-prepared negative and positive control samples. The negative control was a 'pool' of two samples that were negative for SARS-CoV-2 antibodies and negative based on the RT-PCR test. The positive control was prepared from a 'pool' of two other samples reactive to SARS-CoV-2 antibodies and positive on the RT-PCR test. The precision was calculated by using one-way ANOVA. By monitoring variability, we observed the possible impact on the results and data interpretation. CVs were also compared with manufacturer's precision results. In case of deviation, statistical comparisons were performed to demonstrate that there is no statistically significant difference between the values.

Samples

In order to determine diagnostic properties and assess precision, we collected a total of 151 serum samples from subjects for whom we had information that a previous RT-PCR test had been performed. Out of the 151 samples, we used samples from non-hospitalised random subjects who were RT-PCR positive (n = 41), hospitalised patients with COVID-19 who were also RT-PCR positive (n = 19) and random subjects who were RT-PCR negative (n = 91) to determine the diagnostic properties of the methods. Age, sex, other possible infections, immune status, symptoms, the time since possible infection and time since RT-PCR result were not considered when collecting the samples. Serum samples were appropriately aliquoted and prepared for individual analyses, which were performed consecutively on all three methods within one day, avoiding repeated freeze-thawing.

Serological methods

We used three qualitative serological methods to determine specific antibodies against SARS-CoV-2. The general characteristics of the methods are listed in Table 1. The Roche Elecsys® Anti-SARS-CoV-2 method performed on a Cobas e411 analyser detects total Ig (IgG and IgM) by electro-chemiluminescence (ECLIA), whereas the Abbott SARS-CoV-2 IgG method performed on an ARCHITECT i1000SR analyser detects only IgG antibodies by chemiluminescence paramagnetic immunochemical immunoassay (CMIA). Both automated assays detect antibodies directed against the nucleocapsid (N) protein of the virus. The last method manufactured by Maccura Biotechnology is the SARS-CoV-2 IgM/IgG Antibody Assay Kit by Colloidal Gold Method POCT, which detects separately IgG and IgM antibodies against the SARS-CoV-2 antigen using the principle of colloidal gold immunochromatography.

Statistical analysis

Descriptive statistics were run in Microsoft Office Excel 2016 (Microsoft Corporation, Washington, USA). We calculated the diagnostic parameters and Cohen's kappa coefficient (κ) in GraphPad Prism 9 (GraphPad Software Inc., San Diego, CA, USA). Diagnostic parameters were presented with 95% confidence interval (CI) determined by the Clopper-Pearson method.

Precision of automated methods was calculated using one-way ANOVA in Microsoft Excel 2016 version Analyse-it Software Method Validation edition (Ltd. The Tannery, 91 Kirkstall Rd., Leeds,

Table 1 **General characteristics of the three serological methods. Diagnostic properties of automated methods are presented with 95% confidence interval (CI)**

Characteristics	Roche Elecsys® Anti-SARS-CoV-2	Abbott SARS-CoV-2 IgG	Maccura Biotechnology (Colloidal Gold Method)
Method	ECLIA*	CLIA**	Immunochromatography
Target	Nucleocapsid protein	Nucleocapsid protein	Antigen
Detection	Total antibodies IgG/IgM	IgG antibodies	Separate IgM and IgG antibodies
Way of interpretation	Automated	Automated	Manually
Unit	Cut-off index COI (S/C)***	Index (S/C)***	Not Applicable
Result interpretation	Positive: COI ≥ 1.0 Negative: COI < 1.0	Positive: Index ≥ 1.4 Negative: Index < 1.4	Positive: Colour reaction on the control line and test line Negative: Colour reaction on the control line
Diagnostic properties (manufacturer)	Sensitivity 100% (95% CI: 88.1–100) Specificity: 99.81% (95% CI: 99.65–99.91)	Sensitivity 100% (95% CI: 95.89–100) Specificity: 99.60% (95% CI: 98.98–99.98)	True positive: 3 out of 5 for IgG and 2 out of 5 for IgM False positive: none for IgG and 2 out of 20 for IgM

*Keys: *electro-chemiluminescence immunoassay, **chemiluminescent magnetic microparticle immunoassay, ***Signal (Sample/Calibrator).*

UK) and IBM SPSS Statistics 28.0 for Windows (Armonk, New York: IBM Corp.). We presented the results of the 5x5 experimental model using the average value, standard deviation (SD) and coefficient of variation (CV). Precision values that differed from the manufacturer's results were evaluated by F-test to assess whether the difference was statistically significant. The limit of statistical significance was $\alpha<0.05$. Graphical representations were produced in Microsoft Excel 2016 and GraphPad Prism 9.

RESULTS

Diagnostical properties of the serological assays

A total of 151 serum samples were analysed by all three methods to determine diagnostic sensitivity, diagnostic specificity, positive predictive value (PPV) and negative predictive value (NPV). Out of 151 samples, the Colloidal Gold Method detected the presence of at least one antibody class in 61 samples and no antibodies were

detected in 90 samples. A control line was visible in all test plates and therefore it can be claimed that no invalid results were observed. The automated Abbott SARS-CoV-2 IgG method detected the presence of IgG antibodies in 55 samples, the remaining 96 were negative. On the Roche Elecsys® anti-SARS-CoV-2 method, which automates the detection of total IgG and IgM antibodies, 62 out of 151 samples were positive and 89 samples were negative for SARS-CoV-2 antibodies. For each method, a 2x2 contingency table was plotted based on the RT-PCR result that previously confirmed or rejected the suspicion of COVID-19 infection. This is how we defined false and true results and presented them in absolute value and as a proportion of all samples analysed. The POCT SARS-CoV-2 IgM/IgG Antibody Assay Kit by Colloidal Gold Method correctly detected the presence or absence of at least one of the SARS-CoV-2 antibody classes in 138 (91.39%) samples. Out of all false results, 7 (4.64%) were false positive and 6 (3.97%) were false negative. The Abbott SARS-CoV-2 IgG automated method produced slightly fewer false results, 5 (3.31%), all of which were false negative. The second automated method, the Roche Elecsys® anti-SARS-CoV-2 method, gave the highest number of true results. Out of all false results, 3 (1.99%) were false positive and 1 (0.66%) was a false negative, according to the previous RT-PCR results. For all three methods, diagnostic parameters were calculated from the results and given with 95% confidence intervals. The results with 95% confidence intervals (CI) for all three methods are presented in Table 2. The lowest diagnostic characteristics were estimated for the POCT SARS-CoV-2 IgM/IgG IgM/IgG Antibody Assay Kit by Colloidal Gold Method. The highest diagnostic specificity was exhibited by Abbott SARS-CoV-2 IgG method and the highest diagnostic sensitivity by the automated Roche Elecsys® anti-SARS-CoV-2 method.

Comparison of serological assays

The results were initially compared in terms of the number of true results (TN + TP) and false results (FN + FP), which is graphically shown in Figure 1. We found that the POCT Colloidal Gold

Table 2 Diagnostic characteristics for all three serological methods applied in our study

Method	Diagnostic specificity		Diagnostic sensitivity		*NPV		**PPV	
	Value (%)	95% CI (%)	Value (%)	95% CI (%)	Value (%)	95% CI (%)	Value (%)	95% CI (%)
Maccura Biotechnology (Colloidal Gold Method)	92.31	84.79–96.85	90.00	79.49–96.24	93.33	86.05–97.51	88.52	77.78–95.26
Abbott SARS-CoV-2 IgG	100.00	96.03–100.00	91.67	81.61–97.24	94.79	88.26–98.29	100.00	93.51–100.00
Roche Elecsys® anti-SARS-CoV-2	96.70	90.67–99.31	98.33	91.06–99.96	98.88	93.90–99.97	95.16	86.50–98.99

*Negative predictive value, **Positive predictive value.*

Method by Maccura Biotechnology had the highest number of false results and the Roche Elecsys® anti-SARS-CoV-2 automated method had the lowest, which is also reflected in the better diagnostic performance shown in Table 2.

The methods were also compared with each other in terms of the level of agreement, which was determined by Cohen's kappa coefficient (κ). When comparing the POCT SARS-CoV-2 IgM/IgG Antibody Assay Kit by Colloidal Gold Method and the Abbott SARS-CoV-2 IgG method, we only observed an agreement between the presence or absence of IgG antibodies, as the automated method does not identify IgM antibodies. The level of agreement with the given Cohen's kappa coefficient (κ) and 95% CI are shown in Table 3. We found that the automated Roche Elecsys® anti-SARS-CoV-2 and Abbott SARS-CoV-2 IgG methods differed only in 7 results (7 results were negative by Abbott SARS-CoV-2 IgG but were positive with Roche Elecsys® anti-SARS-CoV-2), that is why these methods had the highest level of agreement. The weakest agreement was observed between POCT Colloidal Gold Method and the automated Roche Elecsys® anti-SARS-CoV-2 method where the methods differed in 17 results (9 results were positive with Roche Elecsys® anti-SARS-CoV-2 but were negative with Colloidal Gold Method, while 8 results were negative with the automated method but were positive by Colloidal Gold Method).

Figure 1 **Graphical presentation of the number of false and true results among all three serological methods where FN means false negative, FP false positive, TN true negative and TP means true positive data**

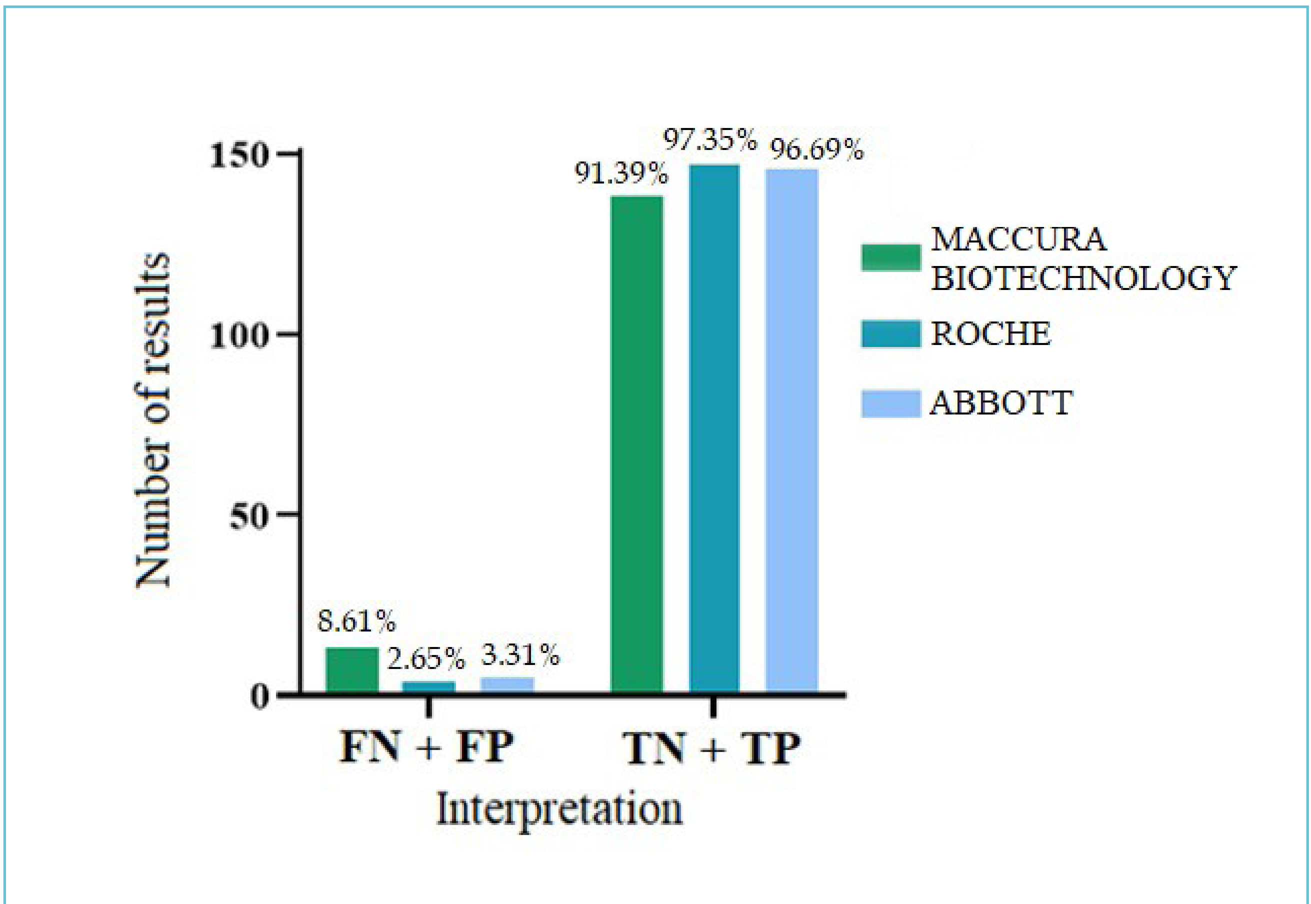


Table 3 **Statistical agreement between serological methods demonstrated by Cohen's kappa coefficient (κ). The value of κ showed whether there was any agreement between two methods (19): none (κ = 0-0.20), minimal (κ = 0.21-0.39), weak (κ = 0.40-0.59), moderate (κ = 0.60-0.79), strong (κ = 0.80-0.90) or almost perfect (κ > 0.90)**

Maccura Biotechnology vs Abbott		**Abbott SARS-CoV-2 IgG**		
		Positive	Negative	Total
Maccura Biotechnology (Colloidal Gold Method)	Positive	**43**	**0**	43
	Negative	**12**	**96**	108
	Total	55	96	151
κ = 0.82 (0.72–0.92)				

Maccura Biotechnology vs Roche		**Roche Elecsys® anti-SARS-CoV-2**		
		Positive	Negative	Total
Maccura Biotechnology (Colloidal Gold Method)	Positive	**52**	**8**	60
	Negative	**9**	**82**	91
	Total	61	90	151
κ = 0.76 (0.66–0.87)				

Roche vs Abbott		**Abbott SARS-CoV-2 IgG**		
		Positive	Negative	Total
Roche Elecsys® anti-SARS-CoV-2	Positive	**55**	**7**	62
	Negative	**0**	**89**	89
	Total	55	96	151
κ = 0.90 (0.83–0.97)				

Analytical precision of automated assays

The results for repeatability, intermediate precision and intra-laboratory precision (intra- and inter-daily) for the automated Roche Elecsys® anti-SARS-CoV-2 and Abbott SARS-CoV-2 IgG methods are presented in Table 4, together with the manufacturer's values given. The within-run and between-run signal variation is shown in Figures 2A and B for Abbott SARS-CoV-2 and Figures 2C and D for Roche Elecsys® anti-SARS-CoV-2. Based on the values listed in Table 4 and graphical representation, we can estimate that on both methods the variability of the between-run signal is slightly higher in positive control samples. Repeatability is slightly poorer for negative control samples on both methods. Despite the smaller variability, we can conclude that the precision of both automated methods was satisfactory, and the variability was too small to affect the results given by the method based on the values of coefficients of variation. Variability of the Roche Elecsys® anti-SARS-CoV-2 method did not exceed the manufacturer's values, so it can be concluded that both automated methods meet the manufacturer's criteria in terms of precision.

DISCUSSION

Despite some limitations of the serological methods, especially in identifying infection at an early stage, they are a very important complement to molecular methods and an important tool for epidemic surveillance, determining seroprevalence in the general population, understanding the

Table 4 Precision results for automated methods. The manufacturer's values are coloured in blue

		Abbott SARS-CoV-2 IgG				Roche Elecsys® anti-SARS-CoV-2			
		Negative control		Positive control		Negative control		Positive control	
Average		**0.07 Index**	0.04 Index	**3.51 Index**	3.53 Index	**0.08 COI**	0.059 COI	**66.03 COI**	2.97 COI
Repeatability	SD	**0.004**		**0.08**		**0.02**		**0.798**	
	CV (%)	**5.9**	5.9	**2.3***	1.1	**2.6**	2.6	**1.2**	1.3
Intermediate precision	SD	**0.002**		**0.109**		**0.000**		**0.971**	
	CV (%)	**2.8**		**3.1**		**0.0**	5.0	**1.5**	2.2
Intra-laboratory precision	SD	**0.004**		**0.136**		**0.002**		**1.257**	
	CV (%)	**6.5***	5.9	**3.9***	1.2	**2.6**		**1.9**	

** Imprecision value is higher than that declared by manufacturer. After further statistical analysis we concluded that there is no statistical difference between the values.*

immune response of individuals to infection, understanding the virus and the development, and monitoring the response to vaccines (8,9,12).

Several serological methods have been developed in recent years, which, like other diagnostic methods during an epidemic, need to provide rapid and, above all, high-quality and reliable results. In order to meet these requirements, irrespective of the urgency for a particular method, the laboratory should ensure that an appropriate verification step is performed before implementing the method, in which the user is informed about the properties and limitations of the method and an assessment is made as to whether the method serves its purpose under laboratory conditions (8,17,18).

In order to implement the most appropriate method in the laboratory, three qualitative serological methods were verified and compared - the automated Roche Elecsys® Anti-SARS-CoV-2 and Abbott SARS-CoV-2 IgG methods and the POCT SARS-CoV-2 IgM/IgG Antibody Assay Kit by Colloidal Gold Method. The most obvious difference between the methods was the class of antibodies detected, with the Roche Elecsys® Anti-SARS-CoV-2 measuring total IgG and IgM, the POCT Colloidal Gold Method analysing IgG and IgM separately, while the automated Abbott SARS-CoV-2 IgG method detected only IgG antibodies. The property of the POCT Colloidal Gold Method, which therefore detects the two types of antibodies separately, may be an advantage over the automated methods in terms of predicting the stage of disease. It is known that the separate identification of IgM and IgG antibodies together with molecular methods can predict whether an infection is acute or in a late-phase or convalescent, considering the kinetics of the immune response, the patient's status and the method's ability (11).

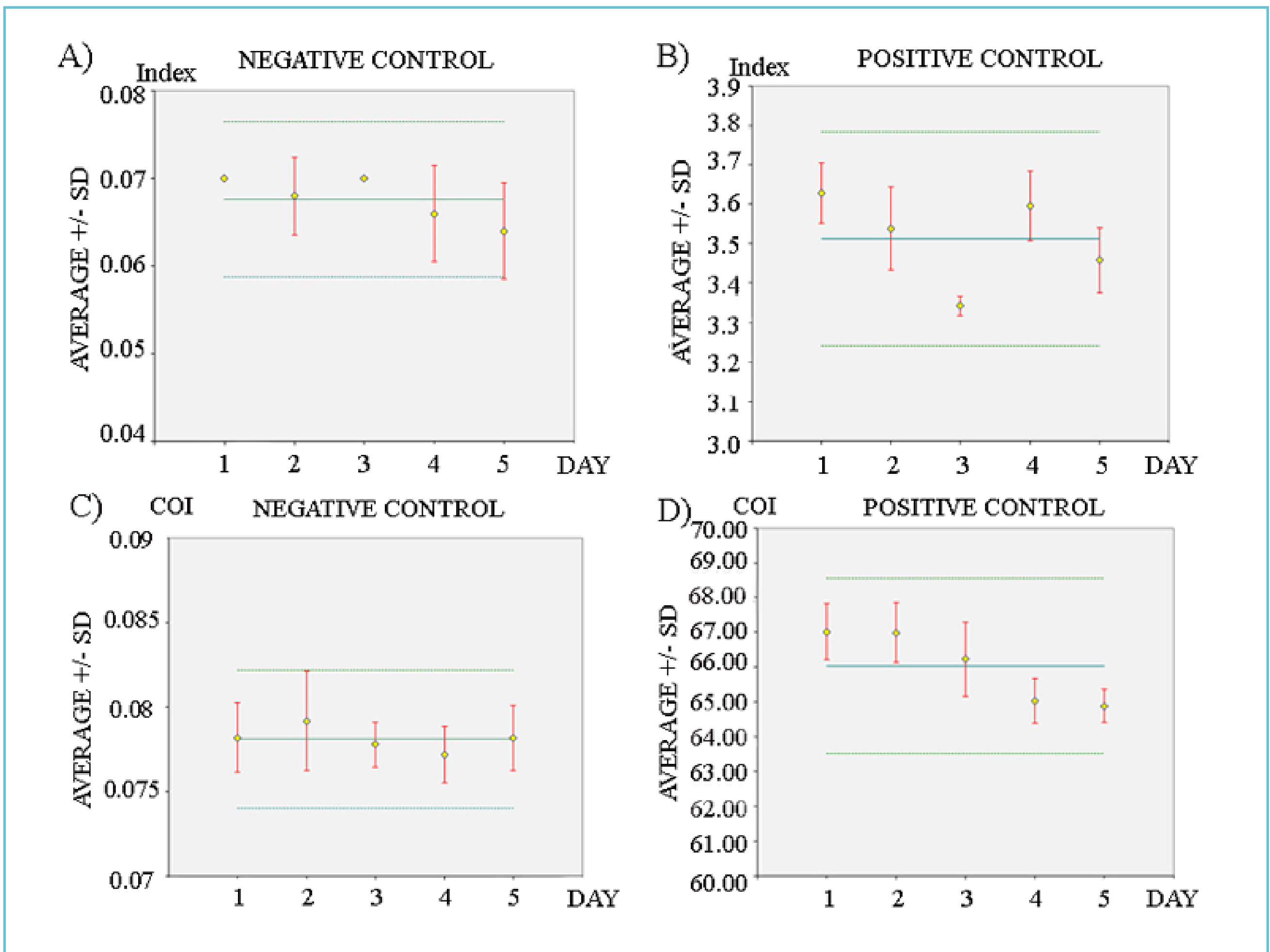
For this reason, the Abbott SARS-CoV-2 IgG method cannot be used in addition to molecular methods to detect early-phase disease, as this requires information on IgM antibodies as well. Nevertheless, the result obtained with the latter method is useful for demonstrating the presence of a history of COVID-19 infection (6,10). The limitations of IgG antibody detection in the early stages of infection were confirmed by Chew et al. who showed that the method had the highest clinical sensitivity after 14 days from the onset of symptoms (20). The known general properties of the selected qualitative serological methods already suggest that they are optimal in their use and performance. In order to implement the optimal method in the routine laboratory, we performed a verification study to determine the diagnostic characteristics of all three methods and to assess the precision of the two automated methods, in addition to the known properties.

The results of diagnostic sensitivity, diagnostic specificity and predictive values were obtained by measuring a total of 151 serum samples for which we had information on the result of a previously performed RT-PCR test (Table 2). We found that the POCT Colloidal Gold Method gave the highest number of false results (Figure 1), which was expected, as the sample preparation can influence the accuracy of the result and the visual reading makes the interpretation non-objective. We found that the most common cause for false-positive results was a reaction in the IgM antibody detection test line. The cause for false detection of IgM antibodies in POCT methods was investigated by Wang and his co-workers, who found that the presence of rheumatoid factor significantly increased the chance of false-positive results for IgM antibodies (21). In our case, this finding cannot be rejected or confirmed, as rheumatoid factor was not measured in these samples. In the case of false-negative results given by the POCT Colloidal Gold Method, we assumed that the reason was the low concentration, which was not detected

by the method because of its limited sensitivity compared to automated methods.

Since the diagnostic sensitivity is crucial, especially when using serological methods in the early phase of infection, it can be concluded that the automated Roche Elecsys® Anti-SARS-CoV-2 method is the most useful method with the lowest false-negative results according to the estimated sensitivity (11). The method is most likely to give the best results due to the identification of both classes of antibodies, which reduces the impact of the time elapsed since the onset of symptoms or a positive RT-PCR result. It can be assumed that this result could be further improved if this time were known and limited to a maximum of 14 days when seroconversion is usually definitely detectable (11,14).

Figure 2 Graphical representation of in series and between series variability of each control sample where yellow dots represents average value in one day, dotted red line represents total average ± 2SD and continuous line represents total average of a signal.
A) Abbott SARS-CoV-2 IgG variability of negative control sample
B) Abbott SARS-CoV-2 IgG variability of positive control sample
C) Roche Elecsys® anti-SARS-CoV-2 variability of negative control sample
D) Roche Elecsys® anti-SARS-CoV-2 variability of positive control sample



Despite the highest diagnostic sensitivity of the Roche Elecsys® Anti-SARS-CoV-2 method, it is important to be aware of that the method is still not suitable for the detection of acute infection. This was confirmed by Brochot and his coworkers, who investigated the issue of diagnostic sensitivity in their study, where, in particular for the detection of IgG class antibodies, false-negative results were detected at an early stage and also in asymptomatic patients. Because of these limitations, the study suggested that negative results of serological methods should be interpreted together with the patient's status and the method's capabilities (22).

Compared to diagnostic sensitivity, the Abbott SARS-CoV-2 IgG method had the highest diagnostic specificity, suggesting that there was no cross-reactivity with other respiratory viruses, which is the most common cause of false-positive results. Slightly lower diagnostic specificity was observed with the POCT Colloidal Gold Method and Roche Elecsys® Anti-SARS-CoV-2 methods, which could be explained by interferences that may cause false-positive data. The results could also be explained by the actual presence of antibodies in the presence of an otherwise negative RT-PCR result as is in the case of copresence of IgM antibodies in POCT Colloidal Gold Method and Roche Elecsys® Anti-SARS-CoV-2 methods. A situation can occur in the case of a false-negative RT-PCR result due to a low viral load at the time of collection (8). In this case, if the RT-PCR test was repeated and the serological results with true positive IgM were confirmed, an acute phase of infection could be inferred, as IgG antibody seroconversion has not yet occurred (13,14).

As part of the method verification and comparison, the level of agreement was assessed using Cohen's kappa coefficient (κ) and almost perfect agreement was found between the automated Roche Elecsys® anti-SARS-CoV-2 and Abbott SARS-CoV-2 methods (Table 3). This was the result we expected, based on the diagnostic properties found. In contrast to our assessment of agreement between the automated methods, Parai and colleagues found much poorer agreement between the Roche Elecsys® anti-SARS-CoV-2 and Abbott SARS-CoV-2 IgG methods (κ = 0.694; 0.641–0.746) in their study where they compared three chemiluminescent methods (23). According to our criteria, such level of agreement is considered to be moderate (19). Despite the lower level of agreement in some other studies, our results and those of other studies on the diagnostic performance of serological methods confirm that automated methods, in particular the Roche Elecsys® anti-SARS-CoV-2 method, have very good diagnostic characteristics.

In addition to the diagnostic properties, the precision of the two automated methods was assessed in the verification process according to the CLSI EP15-A3 protocol using a 5x5 experimental model. We assessed the repeatability, intermediate precision and intra-laboratory precision (intra and inter-daily) and estimated the possible impact on the results. In addition, we also compared the precision results with the manufacturer's data (Table 4). Despite the slightly higher variability in positive controls observed with both automated methods (Figure 2), we did not detect any major deviations that would affect the interpretation of the result. Based on the precision results, we can conclude that both methods also meet the manufacturer's criteria. Conflicting results were obtained by Padoan et al., who observed the highest variability in negative controls and concluded that the Abbott SARS-CoV-2 IgG method did not meet the manufacturer's criteria (24). The difference between the results compared with ours could be explained by the use of the 5×4 experimental model used in Padoan's study, as this model may give poorer results and may not capture all variability factors (18). Due to the difference in signals, we could not compare the precision of

the two methods. The problem of comparing qualitative methods due to signal differences was highlighted by Lee in her study. She also studied the importance of the signal and found a correlation between the CMIA-based method Index value and the severity of infection (25). Despite the satisfactory results, our assessment underestimated the variability between series, as we did not change reagents during the experimental work. The reagent replacement with different lot numbers or repeated calibrations could have been affected by random error.

Despite encouraging results, the paper has some limitations. First, the biggest limitation of our study is the relatively small number of samples, with which we verified all three serological methods. This reduced the statistical power of the results we obtained. Second, in the absence of information on the prevalence of SARS-CoV-2 virus in the studied population, we did not compare the predictive values with the manufacturer's data and with the data from other studies using the same serological methods. Third, when assessing analytical precision, we also compared the coefficients of variation with the manufacturer's results, which can often underestimate or overestimate the variability of the signal as shown by Martinello and colleagues (26). Also, even more accurate results on variability in the laboratory would be obtained, if the experimental model was extended over several days or several repetitions, as this would capture more of the potential causes of variation, such as changing reagents and performing calibrations (18).

Despite some limitations of our work, we can conclude that the automated methods have better diagnostic properties than POCT methods and we can also state that their precision is satisfactory, as the variability does not affect the results and CVs meet the manufacturer's criteria. Due to the better diagnostic sensitivity and performance, it can be concluded that the Roche Elecsys® Anti-SARS-CoV-2 automated method is better than the other methods for the identification of infected and recovered persons, as it gives fewer false results. We can conclude that, in terms of diagnostic properties and precision, the automated methods produce high-quality results that can be trusted and interpreted correctly. Despite the satisfactory results, further investigations could improve our work by including more patients in order to increase statistical power and via obtaining more data on the tested subjects to evaluate diagnostic methods more accurately. At the same time expanding the experimental model and comparing the results with other laboratories are also necessary (27).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Ethical approval

We obeyed the ethical rules while obtaining and testing the samples.

REFERENCES

1. World Health Organization: Listings of WHO's response to COVID-19. World Health Organization 29 Jan 2021. https://www.who.int/news/item/29-06-2020-covidtimeline.

2. Hasöksüz M, Kiliç S, Saraç F: Coronaviruses and SARS-COV-2. Turk J Med Sci 2020; 50(SI-1): 549–556.

3. Bhardwaj A, Sapra L, Saini C, et al.: COVID-19: Immunology, Immunopathogenesis and Potential Therapies. Int Rev Immunol 2022; 41(2): 171–206.

4. Mittal A, Manjunath K, Ranjan RK, et al.: COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2. PLoS Pathog 2020; 16(8): e1008762.

5. Cevik M, Kuppalli K, Kindrachuk J, Peiris M: Virology, transmission, and pathogenesis of SARS-CoV-2. BMJ (Clinical research ed.) 2020; 371:m3862.

6. Ong DSY, Fragkou PC, Schweitzer VA, et al.: How to interpret and use COVID-19 serology and immunology tests. Clin Microbiol Infect: the official publication of the European Society of Clinical Microbiology and Infectious Diseases 2021; 27(7): 981–986.

7. European Commission: EU Common list of COVID-19 rapid antigen tests and a list of mutually recognised COVID-19 laboratory based antigenic assays. European Commission 10 Jun 2022. https://ec.europa.eu/health/system/files/2022-05/covid-19_rat_common-list_en.pdf.

8. Omersel J, Šmid A, Zore T, et al.: Laboratory diagnosis of Covid-19: from virus detection to immune response measurement. Farm Vest 2020; 71(2): 112–120.

9. European Commission: Recommendations for a common EU testing approach for COVID-19. European Commission 17 Sep 2020. https://ec.europa.eu/health/system/files/2020-09/common_testingapproach_covid-19_en_0.pdf.

10. Ghaffari A, Meurant R, Ardakani A. COVID-19 Serological Tests: How Well Do They Actually Perform? Diagnostics 2020; 10(7):453.

11. European Commission: Current performance of COVID-19 test methods and devices and proposed performance criteria: working document of Commission services. European Commission 16 Apr 2020. https://ec.europa.eu/docsroom/documents/40805.

12. Lai CC, Wang CY, Ko WC, Hsueh PR: In vitro diagnostics of coronavirus disease 2019: Technologies and application. J Microbiol Immunol Infect 2021; 54(2): 164–174.

13. Jacofsky D, Jacofsky EM, Jacofsky M: Understanding Antibody Testing for COVID-19. J Arthroplasty 2020; 35(7): 574–581.

14. La Marca A, Capuzzo M, Paglia T, et al: Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide to molecular and serological in-vitro diagnostic assays. Reprod Biomed Online. september 2020;41(3):483–99.

15. Pan American Health Organization: Interpretation of laboratory results for COVID-19 diagnosis. Pan American Health Organization, 6 May 2020. https://iris.paho.org/handle/10665.2/52138.

16. Sun B, Feng Y, Mo X, et al.: Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. Emerg Microbes Infect 2020; 9:1: 940–948.

17. Chesher D: Evaluating assay precision. Clin Biochem Rev 2008; 29 Suppl 1(Suppl 1): 23–26.

18. Clinical and Laboratory Standards Institute (CLSI). User Verification of Precision and Estimation of Bias; Draft Guideline—Third Edition. CLSI document 2013; EP15-A3.

19. McHugh ML: Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012; 22(3): 276–82.

20. Chew KL, Tan SS, Saw S, et al.: Clinical evaluation of serological IgG antibody response on the Abbott Architect for established SARS-CoV-2 infection. Clin Microbiol Infect 2020; 26(9):1256.e9–1256.e11.

21. Wang Q, Du Q, Guo B, et al.: A Method To Prevent SARS-CoV-2 IgM False Positives in Gold Immunochromatography and Enzyme-Linked Immunosorbent Assays. J Clin Microbiol 2020; 58(6):e00375-20.

22. Brochot E, Demey B, Handala L, et al.: Comparison of different serological assays for SARS-CoV-2 in real life. J Clin Virol 2020; 130: 104569.

23. Parai D, Dash GC, Choudhary HR, et al.: Diagnostic accuracy comparison of three fully automated chemiluminescent immunoassay platforms for the detection of SARS-CoV-2 antibodies. J Virol Methods 2021; 292:114121.

24. Padoan A, Bonfante F, Pagliari M, et al.: Analytical and clinical performances of five immunoassays for the detection of SARS-CoV-2 antibodies in comparison with neutralization activity. EbioMedicine 2020; 62:103101.

25. Lee N, Jeong S, Park M-J, Song W: Comparison of three serological chemiluminescence immunoassays for SARS-CoV-2, and clinical significance of antibody index with disease severity. PLoS ONE 2021; 16(6): e0253889.

26. Martinello F, Skitek M, Jerin A: Verifying and assigning own target values and ranges for internal quality control. Accreditation and Quality Assurance 2019; 24(4): 305–312.

27. Skitek M, Martinello F, Jerin A. How to really understand and improve the system of internal quality control and external quality assessment in the accreditation process of the medical laboratory?. eJIFCC 2022; 33: 23-27.