Title: Comparative Evaluation of Three Serologic Assays for the Identification of SARS-CoV-2 Antibodies

Authors: Keenan O. Hogan, MD, Dave Klippel, BS, Fred V. Plapp, MD PhD, Rachael M. Liesman, PhD

Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas, USA

Corresponding Author: Rachael Liesman
Department of Pathology and Laboratory Medicine
3901 Rainbow Boulevard, Kansas City, KS 66160
Email: rliesman@kumc.edu
Telephone: (913) 588-1101
Fax: (913) 588-8780

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Interest: None.

Brief Title: Evaluation of Serologic Assays for SARS-CoV-2

Key words: SARS-CoV-2; Serology; Immunoassays; Diagnostic Accuracy

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background and aims

Serologic assays for the detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibodies are being developed and approved rapidly with limited external validation. Accurate diagnostics are an essential component to pandemic management and public health.

Materials and methods

Residual serum samples (N=113) from patients who were evaluated for SARS-CoV-2 infection status by polymerase chain reaction (PCR) were retrospectively tested in parallel across three automated SARS-CoV-2 serologic assays: Liaison SARS-CoV-2 S1/S2 IgG, Elecsys anti-SARS-CoV-2 total antibody, and Access SARS-CoV-2 IgG.

Results

Testing of 51 PCR-positive and 62 PCR-negative patients demonstrated qualitative inter-test agreement of 96% overall, 100% in PCR-negative patients, 88% in early positive samples (0-13 days post positive PCR), and 100% in convalescent samples (14+ days post positive PCR). Calculated kappa values for paired inter-test agreement ranged 0.93-0.96. Compared to PCR, overall percent positive agreement ranged from 82-86% (100% for convalescent samples) and percent negative agreement was 100% for each assay.

Conclusion

This study demonstrates high diagnostic accuracy and inter-test agreement for three automated SARS-CoV-2 serologic assays. External validation of serologic assays is critical to ensure diagnostic accuracy and appropriate utilization of critical resources.
1. Introduction

Documented cases of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) exceeded 12 million by July 12, infecting individuals from 213 countries/regions and killing more than 560,000 [1]. Management of the coronavirus disease 2019 (COVID-19) pandemic relies on accurate diagnostics, routinely provided by viral RNA detection through reverse transcriptase polymerase chain reaction (RT-PCR) and antibody detection through immunoassays. Although seroconversion has been detected as early as 4 days post symptom onset, PCR remains the current standard for identification of acute infection [2-3].

Infectious disease and microbiology organizations recommend targeted use of serologic assays for epidemiologic surveillance, identification of potential convalescent plasma donors, and evaluation of experimental vaccine efficacy [4-5]. As of the writing of this manuscript, the Food and Drug Administration (FDA) has issued emergency use authorization for 25 different serology assays [6]. These assays vary significantly in methodology, antibody targets, and manufacturer validation with few published peer-reviewed studies available for independent performance evaluation.

In this study, we assessed the performance characteristics of three automated SARS-CoV-2 serologic immunoassays using divided serum samples from patients with positive and negative SARS-CoV-2 PCR results: Liaison SARS-CoV-2 S1/S2 IgG, Elecsys anti-SARS-CoV-2 total antibody, and Access SARS-CoV-2 IgG. We hypothesized that parallel testing would demonstrate good qualitative agreement and high negative and positive percent agreement (>98%).

2. Material and Methods

2.1 Specimen Collection

Between April 15 and June 1, 2020, residual serum samples ordered for routine medical management of inpatients at the University of Kansas Hospital were retrospectively collected. Samples were collected for two groups: (1) serum samples from patients who tested positive for SARS-CoV-2 by an RT-PCR assay, and (2) serum samples from randomly selected patients who had tested negative for SARS-CoV-2 by an RT-PCR assay within 48 hours prior to collection. SARS-CoV-2 diagnostic testing was performed using a commercially available FDA EUA RT-PCR assay. Briefly, nasopharyngeal swabs collected in either UTM or PBS were tested with the
Abbott RealTime SARS-CoV-2 assay (Abbott Diagnostics Inc, Scarborough, ME), performed on the Abbott m2000 instrument, or the Simplex COVID-19 Direct assay (DiaSorin Molecular LLC, Cypress CA), following manufacturer’s instructions.

Residual serum samples were centrifuged, aliquoted, and frozen at -30 °C for 1 to 46 days. Samples were sequentially thawed and maintained at 2-8 °C for ≤14 days prior to testing.

Patient medical records were reviewed to determine sex, age, and the date of first positive SARS-CoV-2 RT-PCR test. Samples were grouped by number of days after the first positive molecular result into 3 pre-specified categories: 0-6 days, 7-13 days, 14+ days. All available serum samples from PCR-positive patients and randomly selected PCR-negative patients that were older than 18 years with adequate residual volume for parallel testing were included. Samples for a single PCR-positive patient from sequential days were combined when required for adequate testing volume, as long as the combined samples fell within the same pre-specified category. Clinical Laboratory Scientists were not specifically blinded to the clinical status or PCR results of the patients. This study was approved by the University of Kansas Medical Center institutional review board prior to testing.

2.2 Serologic Assays

Table 1 provides a breakdown of key attributes for each serologic assay. The Elecsys anti-SARS-CoV-2 antibody test (Roche Diagnostics, Rotkreuz, Switzerland) is a double antigen sandwich electro-chemiluminescence immunoassay that detects total antibody against the nucleocapsid (N) antigen and is performed on the Cobas e801 immunoassay analyzer. Testing was performed in accordance with manufacturer’s instructions. The patient sample is incubated with biotinylated SARS-CoV-2 recombinant N antigen labeled with a ruthenium complex. Streptavidin-coated magnetic microparticles are added and the antigen-antibody-antigen complex binds to the solid phase by interaction of biotin with streptavidin. Microparticles are magnetically captured onto the surface of an electrode where application of a voltage induces chemiluminescence which is measured by a photomultiplier. Results are expressed as a cut-off index (COI). A result of ≥1.0 is considered non-reactive while a result of <1.0 is considered reactive.

The Liaison SARS-CoV-2 S1/S2 assay (DiaSorin S.p.A., Saluggia, Italy) detects IgG antibody against the S1 and S2 subunits of the spike protein using a chemiluminescent indirect immunoassay (CLIA) and is performed on the DiaSorin Liaison XL immunoassay analyzer. Testing was performed in accordance with manufacturer’s instructions. Patient serum is
combined with magnetic particles coated with biotinylated recombinant S1/S2 antigen, which
binds to SARS-CoV-2 antibody, if present. Following incubation, isoluminol labeled mouse anti-
human IgG are added, which will bind to the SARS-CoV-2 antibody. The amount of isoluminol
antibody conjugate is determined as a flash chemiluminescence reaction and reported in arbitrary
units per milliliter (AU/mL). A result of <15 is considered negative while a result of ≥15.0 is
considered positive.

The Access SARS-CoV-2 IgG assay (Beckman Coulter, Inc., Minnesota, USA) is a
chemiluminescent indirect immunoassay that detects IgG antibody to SARS-CoV-2 and is
performed on the Beckman Coulter DXI 800 immunoassay analyzer. Testing was performed in
accordance with manufacturer’s instructions. A sample is incubated with paramagnetic particles
coated with recombinant SARS-CoV-2 protein specific for the receptor binding domain (RBD)
of the S1 Protein. Monoclonal anti-human IgG alkaline phosphatase conjugate is added to bind
IgG antibodies captured on the particles. Light generated by chemiluminescent substrate is
measured with a luminometer. The light signal is compared to the cut-off value and is expressed
as a signal to cutoff ratio (S/CO). A result of <0.8 is interpreted as non-reactive while a result of
≥1.0 is considered reactive. Results between 0.8 and 1.0 (inclusive) are considered equivocal.

2.3 Statistics

Each sample represented a unique patient. For the purposes of analysis, equivocal results were
treated as negative and combined samples were represented by the day farthest from the patient’s
positive PCR test. Percent positive agreement (PPA) and percent negative agreement (PNA)
were calculated with 95% confidence intervals (CI) using the Clopper-Pearson exact method for
sensitivity and specificity and stratified by length of time from positive PCR test. Paired inter-
test agreement was calculated with 95% CI and standard error (SE) using Fleiss kappa. Level of
agreement was pre-specified as almost perfect (0.81-0.99), substantial (0.61-0.80), moderate
(0.41-0.60), fair (0.21-0.40), and slight (0.00-0.20). Statistical analysis was performed using
GraphPad Prism 8 statistics software (San Diego, USA).

3. Results

3.1 Patient Demographics and Specimen Properties
Fifty-one specimens from unique PCR-positive patients and 62 specimens from unique PCR-negative patients were available for antibody testing. Table 2 describes patient demographics stratified by PCR status. Patients were 60% female with a median age of 59 years (range: 19-92). Date of symptom onset was not routinely documented so patients were grouped temporally by the period between the first positive PCR result and serology specimen collection. Each temporal group contained 17 specimens. The number of days post positive PCR ranged from 1-45 days overall (median: 9) with a median of 5, 9, and 18 for 0-6 days, 7-13 days, and 14+ days, respectively. Samples were combined for 14 patients 0-6 days, 3 patients 7-13 days, and 2 patients 14+ days.

### 3.2 Diagnostic Accuracy

Of 113 patients, there was consensus between the platforms in 96% (41 positive and 68 negative). There was 100% consensus in the convalescent positive samples (14+ days post positive PCR) and negative samples but only 88% consensus in early positive samples (0-13 days post positive PCR). The only equivocal result was returned by the Beckman assay on a single non-consensus sample. Table 3 details the results of each assay compared to PCR with diagnostic accuracy and estimates of precision. No false positive results were detected and the PNA was calculated at 100% for each platform. Convalescent positive samples also demonstrated zero false negative results for a calculated 100% PPA. The PPA ranged from 59-94% for the early positive samples.

Two patient specimens were reported as negative by at least 1 platform in the 7-13 days post positive PCR group and both specimens were collected at 7 days. One tested negative by all three platforms and review of the medical record revealed the patient was receiving immunosuppressive therapy following renal transplantation. The second patient tested positive on the Roche and Beckman platforms and negative on the DiaSorin platform. This patient was moderately symptomatic with an elevated lactate and required supplemental oxygen but not intubation during their disease course. This result may have represented early seroconversion as the patient was 8 days post symptom onset with low-positive reactivity for the Roche and Beckman assays (14.1 COI and 11.9 S/CO, respectively) and high-negative reactivity for the DiaSorin assay (11.3 AU/mL).
Eight patient samples collected 0-6 days post positive PCR demonstrated at least one negative serology result. The samples ranged from 1-6 days post positive PCR (median: 4). Out of the 8 samples, 5 were consensus negative, 1 was reactive by Roche alone, 1 was reactive by Beckman alone, and 1 was reactive by DiaSorin and Beckman.

3.3 Assay Comparison

Direct comparison of the three assays is detailed in Table 4. Calculated kappa values for inter-test agreement between Roche/DiaSorin, Roche/Beckman, and DiaSorin/Beckman were 0.94, 0.93, and 0.96, respectively. The confidence interval for each measure was within the bounds of near perfect agreement and included perfect agreement (1.00).

4. Discussion

4.1 Comparative Analysis

This study describes high PNA and temporally dependent PPA for three automated SARS-CoV-2 serologic assays in an independent evaluation using an inpatient cohort verified by PCR. The results of this study confirm the conclusions of the systematic review by La Marca, et al. in which they suggest the choice of serology assay should be made by individual labs based on experience with the manufacturer, expected workflow and volume, and thoroughness of assay guidance and validation [7]. They report on 49 serology studies primarily conducted in China with a range of 78-100% and 0-100% for specificity and sensitivity, respectively.

Two peer-reviewed articles and three studies released as pre-prints specifically reported on the accuracy of the DiaSorin SARS-CoV-2 serologic assay in comparison to PCR, with a range of 96.8-100% for specificity and 77.8-95.7% for sensitivity in convalescent patients [8-12]. With 100% PNA and 100% PPA in the convalescent group, this study describes a more favorable diagnostic accuracy. One of the studies provided the median positive quantitative results for comparison, which ranged from 8-96.3 AU/mL, with 96.3 AU/mL representing their most acutely ill subjects [11]. In comparison, the median positive result in this study was 190 AU/mL, which may suggest our high PPA is a consequence of high patient acuity. Correlation between disease severity and antibody titer and detection has been reported [13-14]. However, direct comparison of the quantitative results from these qualitative assays is limited in significance,
especially given the heterogeneity in study design and reporting as identified in a recent Cochrane review [15].

Median negative results were comparable across the studies (below the limit of detection to 5.7 AU/mL), but two of the studies included pre-pandemic cohort samples of greater than 1000 patients. The difference in sampling likely explains the discordance in reported specificity. Additionally, two of these studies and an additional study released as a pre-print report on the Roche assay with a range of 98-99.7% and 89.2-100% for specificity and sensitivity, respectively, and similar comparisons to this study may be drawn [10,12,16].

To our knowledge, no study has reported independent evaluation of the Beckman assay diagnostic accuracy. With kappa values of 0.93 and 0.96 in comparison to the Roche and DiaSorin assay, respectively, 100% PNA, and 100% PPA in convalescent specimens, we demonstrate no significant difference in the performance between these three platforms.

Including several weak positive results, seroconversion was detected across all PCR-positive patients by day 8 as shown in Figure 1. Although the time of symptom onset for our patient population was not routinely recorded, our data likely coincides with the reported median seroconversion for total antibody and IgG antibody alone of 11 and 14 days post symptom onset, respectively [17]. Although the Roche assay detects total antibody and DiaSorin and Beckman assays detect IgG antibody, there was no identifiable difference in the detection of early seroconversion.

Clinical context will be essential in identifying pre-test probabilities based on local seroprevalence, patient symptomatology, and stage of disease course [18]. Although our population did not include any false positive results, testing of large pre-pandemic cohort samples have reported variable specificity. The use and interpretation of serologic assays is also complicated by the continued uncertainty of antibody durability and the relationship between positive serologic testing and the presence of neutralizing antibodies [14]. One study comparing the DiaSorin assay to a viral microneutralization test reported a sensitivity and specificity of 43.8% and 94.9%, respectively, for 70 PCR-positive patients and 81 pre-pandemic patients [19]. Overall, evidence is limited and laboratories offering serology testing for SARS-CoV-2 must make these limitations clear. For patients with a complicated medical history or atypical presentation, antibody testing may provide important diagnostic information. It has been reported that overall sensitivity for serology is greater than nucleic acid testing for SARS-CoV-2 starting
day 8 after symptom onset and combined serology and RT-PCR testing provides significantly improved sensitivity compared to RT-PCR alone early in the disease course [17,20].

4.2 Limitations

This study has several limitations. The testing of a relatively small number of retrospective convenience samples is at risk for confirmation basis and underpowered calculations. Specifically, our cohort of inpatients likely includes a greater average patient acuity, which would tend to inflate sensitivity. This may also increase the number of false positive results, which was not identified. In the absence of thorough chart review for the PCR-negative cohort, the presence of confounding factors contributing to an artificially high true negative rate cannot be excluded but is considered unlikely. By combining specimens from sequential days and using the farthest date, we likely reduced the reportable sensitivity. Longitudinal samples from individual patients were not tested to track seroconversion. The samples were not tested to ascertain the presence of potential cross-reactive viruses.

One of the strengths of this study is the direct comparison of the three assays by parallel testing, including an assay without currently published external validation. Each sample was collected from a single patient, which avoids bias introduced by multiple measurements. Although our specificity population was smaller, the patient samples are representative of the current local circulating viruses among individuals with healthcare contacts, which lends external validity. The results also include early seroconversion specimens in addition to convalescent specimens. The performance characteristics of serologic assays for patients in this window period is critical for clinical management decisions, particularly for acutely ill patients. We reported no false positive results that may have otherwise been attributed to remote exposure with the use of a post-pandemic specificity cohort.

4.3 Conclusion

We report high diagnostic accuracy and inter-test agreement for three SARS-CoV-2 serologic assays. The choice of assay and its clinical utility will depend on the specific needs of individual laboratories and the local public health conditions. External validation of serologic assays is critical to ensure high diagnostic accuracy and appropriate utilization of critical resources.
Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of Interest: None.
References


<table>
<thead>
<tr>
<th>TABLE 1. Attributes for the evaluated serologic assays</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Roche</strong></td>
</tr>
<tr>
<td>Assay method</td>
</tr>
<tr>
<td>Antigen</td>
</tr>
<tr>
<td>Specimen type</td>
</tr>
<tr>
<td>Total volume</td>
</tr>
<tr>
<td>Result units</td>
</tr>
<tr>
<td>Positive cutoff</td>
</tr>
<tr>
<td>Equivocal cutoff</td>
</tr>
<tr>
<td>Throughput</td>
</tr>
</tbody>
</table>

Abbr.: Ab, Antibody; CLIA, Chemiluminescent Immunoassay; RBD, Receptor Binding Domain
TABLE 2. Patient demographics by PCR status

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>PCR (+)</th>
<th>PCR (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0-6 days</td>
<td>7-13 days</td>
</tr>
<tr>
<td>Number of patients, n</td>
<td>113</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Median age in years (IQR)</td>
<td>59 (43-72)</td>
<td>71 (52-77)</td>
<td>64 (42-74)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>69 (61)</td>
<td>12 (71)</td>
<td>9 (53)</td>
</tr>
</tbody>
</table>
### TABLE 3. Performance characteristics of each platform by PCR status

<table>
<thead>
<tr>
<th></th>
<th>PCR (+), n</th>
<th>PCR (-), n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-6 days</td>
<td>7-13 days</td>
</tr>
<tr>
<td><strong>Roche</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactive</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>Non-reactive</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>PPA, % (95% CI)</td>
<td>59 (33-82)</td>
<td>94 (71-99)</td>
</tr>
<tr>
<td>PNA, % (95% CI)</td>
<td>100 (94-100)</td>
<td>100 (94-100)</td>
</tr>
<tr>
<td><strong>DiaSorin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Negative</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>PPA, % (95% CI)</td>
<td>59 (33-82)</td>
<td>88 (64-99)</td>
</tr>
<tr>
<td>PNA, % (95% CI)</td>
<td>100 (94-100)</td>
<td>100 (94-100)</td>
</tr>
<tr>
<td><strong>Beckman</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reactive</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>Non-reactive</td>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
</tr>
<tr>
<td>PPA, % (95% CI)</td>
<td>65 (38-86)</td>
<td>94 (71-99)</td>
</tr>
<tr>
<td>PNA, % (95% CI)</td>
<td>100 (94-100)</td>
<td>100 (94-100)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Abbr.: PPA, percent positive agreement; PNA, percent negative agreement

<sup>a</sup> Includes 1 equivocal result
TABLE 4. Serologic assay comparison

<table>
<thead>
<tr>
<th></th>
<th>DiaSorin (n)</th>
<th>Beckman (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(+)</td>
<td>(-)</td>
</tr>
<tr>
<td>Roche (n)</td>
<td>(+)</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$K^a$</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.88-1.0, 0.032)</td>
</tr>
<tr>
<td>Beckman (n)</td>
<td>(+)</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$K^a$</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.91-1.0, 0.026)</td>
</tr>
</tbody>
</table>

$K^a$ Fleiss kappa (95% CI, SE)
FIGURE 1. Quantitative results for each assay by days post positive PCR