Evaluation of the Xiamen AmonMed Biotechnology rapid diagnostic test COVID-19 IgM/IgG test kit (Colloidal gold)

ANJA DÖRSCHUG¹, JULIAN SCHWANBECK¹, ANDREAS HAHN², ANKE HILLEBRECHT³, SABINE BLASCHKE³, UWE GROB¹, MARKUS M. HEIMESAAT⁴, HAGEN FRICKMANN²,⁵† and ANDREAS E. ZAUTNER¹†*

¹ Institute for Medical Microbiology, University Medical Center Göttingen, Göttingen, Germany
² Institute for Medical Microbiology, Virology and Hygiene, University Medicine Rostock, Rostock, Germany
³ Interdisciplinary Emergency Department, University Medical Center Göttingen, Göttingen, Germany
⁴ Department of Microbiology and Hygiene, Charité – University Medicine Berlin, Berlin, Germany
⁵ Department of Microbiology and Hospital Hygiene, Bundeswehr Hospital Hamburg, Hamburg, Germany

Received: August 28, 2020 ● Accepted: September 9, 2020
Published online: September 25, 2020

ABSTRACT

Introduction: To efficiently monitor the COVID-19 pandemic for surveillance purposes, reliable serological rapid diagnostic tests (RDTs) are desirable for settings where well-established high-throughput bench-top solutions are not available. Here, we have evaluated such an RDT. Methods: We have assessed the Xiamen AmonMed Biotechnology COVID-19 IgM/IgG test kit (Colloidal gold) and the EUROIMMUN benchtop assay with serum samples from patients with polymerase chain reaction (PCR)-confirmed COVID-19 disease. Samples from patients with Epstein-Barr-virus (EBV) infection and blood donors were used for specificity testing. Results: For the colloid gold rapid test and the EUROIMMUN assay, the study indicated overall sensitivity of 15.2% and 67.4%, respectively, while specificity of 99.0% and 97.9% with the blood donor sera, as well as 100% and 96.8% with the EBV-patients, were observed, respectively. An association of the time period between positive PCR results and serum acquisition with serological test positivity could be observed for the immunoglobulin G subclass of the EUROIMMUN assay only. Conclusions: In spite of acceptable specificity of the assessed RDT, the detected poor sensitivity leaves room for improvement. The test results remain difficult to interpret and therefore the RDT can currently not be recommended for routine diagnostic or surveillance use.

KEYWORDS
COVID-19, SARS-CoV-2, rapid diagnostic testing, point-of-care testing, surveillance, serology, epidemiology

BACKGROUND

COVID-19 (Corona Virus disease), associated with pandemic spread starting from Wuhan, China in December 2019 and caused by SARS-CoV-2 (Severe Acute Respiratory Syndrome-Corona Virus-2), was first observed in Europe at the beginning of 2020 [1]. In the meantime, surveillance and containment became a global concern.

As recently discussed for sexually transmitted infections (STIs), diagnostic approaches can help to limit the transmission of infectious agents [2–5]. Diagnostic results can further...
contribute to the estimation of risk behavior, which facilitates transmission of infectious agents, on the population level [6]. To support the diagnosis of COVID-19 and associated containment purposes, first real-time polymerase chain reaction (PCR) schemes targeting SARS-CoV-2 were rapidly provided [7].

To circumvent laborious and demanding in-house real-time PCR, cartridge-based [8–14] and other [15–18] fully-automated molecular tools for the detection of SARS-CoV-2 were introduced to allow more rapid or point-of-care testing (POCT). For mass testing purposes, pooling and sequencing were discussed; while CRISPR gene-editing tools combine good sensitivity and specificity with the easy-to-apply lateral flow technology [19].

Although molecular diagnostic approaches are required for the identification of infective individuals, copy numbers of viral pathogens decrease in the course of the infection [17, 20]. It is therefore unlikely that all individuals with SARS-CoV-2 infections will be detected within the replicative period. Accordingly, serological assessments are an alternative for surveillance purposes. In resource-limited settings, POCT approaches may provide readily available and easy-to-apply tools.

Previously described serological POCT tools have shown test characteristics with room for improvement. While reported specificity of immunochromatographic and enzyme linked immunosorbent assay (ELISA)-based was usually above 95%, heterogeneous sensitivity with widely varying values between less than 70% and more than 90% was described depending on the populations assessed [21–34]. In these studies, false-positive results were described to be associated with rheumatoid factor-immunoglobulin M (IgM), and urea dissociation was suggested to overcome this problem [21]. In a meta-analysis, sensitivity of gold immunochromatography assays (GICA) was slightly better than sensitivity of ELISA approaches [22]. Heat inactivation of sera was described as not relevantly interfering with the reliability of immunochromatographic test assays [28]. Time since positive PCR testing has been reported to be crucial for optimum serological POCT sensitivity, with best results at least 14 days after the positive PCR result [31]. At later post-infectious stages, serology can become negative again [35].

In summary, there are several limitations of presently available diagnostic POCT options [36–38] as well as gaps of knowledge regarding the appropriate interpretation of their results. For serological test assays, in particular, considerable variability of their results has been reported [39]. Patient age has been shown to play a role for measured antibody titers [40]. While binding antibodies seem to be more sensitive than neutralizing antibodies [40], antigenic cross-reactivity between SARS-CoV-2 and human coronaviruses 229E and OC43 has been described [41].

To provide another piece of the puzzle regarding the assessment of serological POCT assays for COVID-19, the rapid diagnostic test (RDT) COVID-19 IgM/IgG test kit (Colloidal gold, Xiamen AmonMed Biotechnology Co., Ltd., Haicang District, Xiamen, China) was assessed in direct comparison with the EUROIMMUN benchtop assay, which shows imperfect performance characteristics as well [31, 32, 42–53]. Tests were performed with serum samples from patients with PCR-confirmed COVID-19 disease. For specificity testing purposes, samples from patients with Epstein-Barr-virus (EBV) and blood donors were used.

**MATERIALS AND METHODS**

**Sample collections**

Three serum sample collections were included in the assessment. The first collection consisted of samples from 46 PCR-confirmed COVID-19 patients. PCR had been performed from nasopharyngeal swabs. The time between the positive PCR result and the acquisition of the serum samples was documented in 27 out of 46 (58.7%) cases, ranging from 0 to 60 days (median: 21 days, mean: 20 days, standard deviation (SD): 19.2 days). The second collection consisted of samples from 96 blood donors taken in 2015, before the onset of the COVID-19 pandemic. Finally, a third collection contained 31 EBV-positive serum samples, which were used to assess the effects of polyclonal B-cell stimulation.

Ethical clearance allowed a fully anonymized use of sample materials only. Accordingly, patient details cannot be provided.

**Serological assays**

Applied serological assays comprised the COVID-19 IgM/IgG test kit (Colloidal gold, Xiamen AmonMed Biotechnology Co., Ltd., Haicang District, Xiamen, China; referred to as “AmonMed assay” in the following), and the EUROIMMUN Covid-19 IgG/IgA assay (EUROIMMUN, Lübeck, Germany; referred to as “EUROIMMUN assay” in the following). All tests were performed as described by the manufacturer’s instructions. The AmonMed assay as the RDT of interest was performed in duplicate to confirm its positivity rates (see Fig. 1).

**Real-time PCR testing**

Respiratory sample materials from the patients with suspected or confirmed COVID-19 were subjected to real-time PCR for SARS-CoV-2. Screening was performed applying the Genesig Real-Time PCR Coronavirus (COVID-19) assay (Premierdesign Ltd., Chándlers Ford, UK) and first-time positive results were confirmed using automated Cepheid Xpert Xpress SARS-CoV-2 PCR (Cepheid, Sunnyvale, CA, USA). The assays were performed as described by the manufacturers.

**Statistical assessment**

Descriptive analysis was performed including all sample collections and test assays. Sensitivity of the serological assays was calculated with the sample collection from the suspected or PCR-confirmed COVID-19 patients. Thereby, patients with positive PCR results were considered as confirmed positive cases. The effect of the number of days between positive PCR results and the serum sample
acquisition for serology testing was calculated applying rank sum testing and binary logistic regression. Those calculations were performed applying the software Stata/IC 15.1 for macOS 64-bit Intel (College Station, Texas, USA).

Specificity of the serological assays was determined with the serum samples from the blood donors and the patients with EBV.

Individual missing data points were no exclusion criteria.

Ethics

Ethical clearance for this study was provided by the institutional ethics board of the University Medical Center Göttingen (Application number 21/05/20).

RESULTS

Sensitivity of the serological assays as calculated with data from sample collective 1 in total and by antibody sub-class

If a serological assay was rated positive in case of any positive signal in a COVID-19-PCR-positive patient, irrespective of the assessed antibody subclass, sensitivity of the EUROIMMUN assay was 67.4%, while sensitivity of the AmonMed assay was 15.2%. Assessed by antibody subclasses, EUROIMMUN IgG (immunoglobulin G) and AmonMed IgG sensitivity were 56.5 and 0%, respectively, EUROIMMUN IgA sensitivity was 54.3%, and AmonMed IgM sensitivity was 15.2% (Table 1). There were no differences between results of the first and second attempt of AmonMed assay application.

Focusing on the quantitative assessment of the EUROIMMUN optical densities (OD), the mean (±standard deviation (SD)) and median (interquartile range IQR) IgA ODs based on 25 positive results were 5.6 (±7.1) and 2.5 (1.6, 5.1), respectively, while for the IgGs based on 26 positive results, the values were 4.7 (±2.5) and 4.6 (2.6, 6.0), respectively.

Effects of the time between positive PCR results and serum sample acquisition on the observed sensitivity

For the 27 COVID-19-positive patients, for which the time between positive PCR result and serological testing had been recorded, the influence of time on serological test positivity was calculated using rank sum testing. As Cohen’s kappa for the two attempts of AmonMed assay testing was one, meaning that their results were perfectly associated and identical, they were also considered as identical for the calculation. As shown in table two, significance of time-dependence could be shown for EUROIMMUN IgG-testing, but not for any other immunoglobulin sub-class or subclass-independent overall test-positivity (Table 2).

Table 1. Sensitivity of the assessed test serological assays

<table>
<thead>
<tr>
<th>Positive controls</th>
<th>n (%)</th>
<th>EUROIMMUN assay</th>
<th>AmonMed assay first attempt</th>
<th>AmonMed assay second attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>Sensitivity (0.95 CI)</td>
<td>n</td>
</tr>
<tr>
<td>Overall positives</td>
<td>46 (100)</td>
<td>31</td>
<td>0.674 (0.523, 0.796)</td>
<td>7</td>
</tr>
<tr>
<td>IgA</td>
<td>46 (100)</td>
<td>25</td>
<td>0.543 (0.396, 0.684)</td>
<td>n.a.</td>
</tr>
<tr>
<td>IgG</td>
<td>46 (100)</td>
<td>26</td>
<td>0.565 (0.417, 0.703)</td>
<td>n.a.</td>
</tr>
<tr>
<td>IgM</td>
<td>46 (100)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

0.95 CI = 95% confidence interval. n = number, n.a. = not applicable.
Similar results were achieved by binary logistic regression testing. Performing a binary logistic regression with the test results as the dependent variables and the days after a first positive PCR result as the independent variable, an increased likelihood for an overall positive test result of the Euroimmun assay is driven by EUROIMMUN IgG subclass testing (Table 3).

Specificity of the serological assays as calculated with data from the sample collectives 2 and 3 in total and by antibody sub-class

As assessed with the blood donor samples, total specificity was 97.9% for the EUROIMMUN assay and 99.0% for the AmonMed assay. Thereby, non-specificity was driven by the IgA subclass analysis of the EUROIMMUN assay with 2 false-positive results (OD values 1.1; 3.4) and IgM subclass analysis of the AmonMed assays, while all IgG tests remained negative for both assays (Table 4).

With the sera from the 31 EBV-positive samples, similarly good specificity could be observed with only a single false positive EUROIMMUN IgA (OD value 8.0). All other test results were correctly negative, resulting in a total specificity of 96.8% for the EUROIMMUN assay and 100% for AmonMed assay (Table 5).

Table 2. Rank sum testing assessment of the effects of time on serological test positivity

<table>
<thead>
<tr>
<th>Days after first record</th>
<th>EUROIMMUN assay</th>
<th>AmonMed assay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Overall Positive</td>
<td>19</td>
<td>25.79 (19.72)</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>7.13 (9.36)</td>
</tr>
<tr>
<td>Negative</td>
<td>8</td>
<td>7.13 (9.36)</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>7.13 (9.36)</td>
</tr>
<tr>
<td>IgA</td>
<td>Positive</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>11</td>
</tr>
<tr>
<td>IgG</td>
<td>Positive</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>12</td>
</tr>
<tr>
<td>IgM</td>
<td>Positive</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>22</td>
</tr>
</tbody>
</table>

*Significance in rank sum testing.

n.a. = not applicable. n.e. = not estimable. n = number. SD = standard deviation. IQR = interquartile range.

Table 3. Binary logistic regression assessment of the effects of time on serological test positivity

<table>
<thead>
<tr>
<th>Binary logistic regression</th>
<th>EUROIMMUN assay</th>
<th>AmonMed assay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (SE)</td>
<td>P-value</td>
</tr>
<tr>
<td>overall</td>
<td>27</td>
<td>0.074 (0.036)</td>
</tr>
<tr>
<td>IgA</td>
<td>27</td>
<td>0.035 (0.023)</td>
</tr>
<tr>
<td>IgG</td>
<td>27</td>
<td>0.137 (0.047)</td>
</tr>
<tr>
<td>IgM</td>
<td>27</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

SE = standard error. n.a. = not applicable. n.e. = not estimable.
### DISCUSSION

The study was performed in order to assess the performance characteristics of the serological RDT COVID-19 IgM/IgG test kit (Colloidal gold/AmonMed assay) for the surveillance of recent SARS-CoV-2 infections compared with a standard benchtop approach by EUROIMMUN. In a group of 46 COVID-19-PCR positive patients, however, sensitivity of the AmonMed assay was as poor as 15.2% and only IgM-driven, in spite of time periods up to 60 days between positive PCR-testing and serum sample acquisition, with a median of three weeks. Due to a complete lack of positive IgG results in AmonMed testing, Cohen’s kappa [54] for a comparison of the AmonMed assay with the EUROIMMUN assay could not be calculated. No time dependency of positive results in AmonMed assay testing could be shown, which was only detectable for EUROIMMUN IgG. The previously described EUROIMMUN IgA-sensitivity problems could be confirmed by this study as well [42–53]. The AmonMed assay's specificity, however, was excellent as assessed both with blood donor sera and sera taken from EBV patients.

Regarding the practical diagnostic application of the AmonMed assay, the following aspects should be considered. In case of high pandemic activity, the challenge is the rapid identification of subpopulations, within which at least one individual has been infected with COVID-19. In particular, rapid diagnoses are desired in order not to unnecessarily restrict the freedom of the examined individuals on the one hand and on the other hand to protect the society against the risk of infection. Indeed, the POCT tests examined are particularly suitable for this, because they have a very high specificity, which is the basic requirement for screening tests in order to serve the goal as outlined above, i.e. the protection of both the individual and the society. At the same time, a low sensitivity can be accepted as long as the test results are considered only to refer to the examined population as a whole, while the individual’s diagnosis is set applying a sufficiently sensitive test. Sensitivity on population level, which is the probability of detecting at least one infected individual within the tested population, is defined by the algorithms of multiple-testing and thus much higher than sensitivity for an individual sample [5]. The same applies to specificity on population level, which is the probability of not detecting a single positive result in the tested population, if there is no infected individual there. Considering this, the assessed POCT tests are therefore very suitable for rapid decision making on whether or not there have been infected individuals in a population of interest. However, the test is unsuitable for individual diagnoses. The resulting decision algorithm could be: A population can be considered as infected, if at least one POCT test is positive. In contrast, a population can be considered as free of infections beyond the diagnostic window period if all tests are negative. For the latter issue, the observed perfect specificity is crucial.

The study has a number of limitations. First of all, patients infected with SARS-CoV-2 were identified by PCR-based tests, which may also produce false positive results. However, the possibility of a false positive result was minimized by sequential confirmation of the results with a second PCR-based test. A total of three different targets were detected in the virus genome. Secondly, the EUROIMMUN assay, which was used for comparison purposes, cannot be considered as a gold standard, as its sensitivity and specificity are different from 100% themselves [31, 32, 42–53]. Thirdly, the EBV-positive samples, which were included in the specificity assessment, were collected during the time period when the COVID-19 pandemic had already reached Germany. Considering the very low COVID-19 prevalence...
of less than 1% in Germany when the study was conducted, as well as the very low number of assessed EBV-positive patients, co-infection with SARS-CoV-2 among the EBV-patients is virtually excluded. This is also the case for the single individual with a positive IgA in the EUROIMMUN assay. Finally, ethical clearance did not allow any assessment and presentation of patient data, which necessarily resulted in disagreement with the STARD reporting standard for test comparisons [55] and made comparisons with patient characteristics, which has recently been shown to be promising [40], impossible.

In spite of those limitations, the assessment confirmed considerable sensitivity weakness of the assessed RDT in spite of good specificity.

CONCLUSIONS

Due to low sensitivity, the assessed RDT AmonMed assay cannot be recommended for diagnostic or surveillance purposes. If such applications are, however, nevertheless considered, the observed very good specificity at least suggests reliability of positive results in an epidemic setting with acceptable pre-test probability, while negative results should always be confirmed by another assay.

Funding sources: The APC was funded by the Open Access Support Program of the Deutsche Forschungsgemeinschaft and the publication fund of the Georg-August-Universität Göttingen.

Authors’ contribution: Conceptualization, A.E.Z.; methodology, A.E.Z.; software, A.E.Z.; validation, A.E.Z.; formal analysis, A.H.; investigation, A.D., A.H., J.S., A.E.Z.; resources, A.H., S.B., U.G., A.E.Z.; data curation, A.E.Z.; writing—original draft preparation, A.D., H.F., A.E.Z.; writing—review and editing, A.D., J.S., H.F., A.E.Z.; visualization, A.H., A.E.Z.; supervision, A.E.Z.; project administration, U.G., S.B., A.E.Z.; funding acquisition, A.E.Z. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

ACKNOWLEDGMENTS

Patient data collection was supported by the study nurse Irina Pavlova, Interdisciplinary Emergency Department, University Medical Center Göttingen.

SUPPLEMENTARY MATERIAL

Supplementary data to this article can be found online at https://doi.org/10.1556/1886.2020.00029.

REFERENCES


Open Access. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (https://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.