Diagnostic accuracy of rapid antigen tests for COVID-19 compared to the viral genetic test in adults: a systematic review protocol

Megan DeArmond1,2 • Shelby Feliciano3 • Ellyn Hirabayashi3 • Karen M. Duus2,3 • Terry A. Else2,3 • Amy E.L. Stone2,3

1 Touro University Nevada Jay Sexter Library, Henderson, NV, USA; 2 Touro University Nevada: A JBI Affiliated Group, Touro University Nevada, Henderson, NV, USA; and 3 Touro University Nevada College of Osteopathic Medicine, Henderson, NV, USA

ABSTRACT

Objective: The objective of this diagnostic accuracy review is to evaluate the effectiveness of rapid antigen tests versus viral genetic PCR-based tests on COVID-19 diagnostic accuracy in adults 18 years and over.

Introduction: Due to the rapidly changing nature of the COVID-19 pandemic, it is imperative that clinicians have access to the most relevant and effective tools and information required to combat this disease. Testing strategies are being developed continuously and need to be evaluated to ensure their appropriate implementation into clinical practice.

Inclusion criteria: This systematic review will include publications that are in the English language (originally or translated) and any gray literature pertaining to the tests of interest. All races, ages over 18, and geographic locations will be considered.

Methods: MEDLINE (PubMed), Embase (Elsevier), Scopus (Elsevier), Qinsight (Quertle), and WHO COVID-19 database (World Health Organization) will be searched. Scopus, Qinsight, and WHO COVID-19 include gray literature. Studies in English published from November 2019 to the present will be considered. Animal studies and studies including pregnant women will be excluded. Retrieval of full-text studies, data extraction, and assessment of methodological quality will be performed independently by two reviewers. A custom data extraction table will be used. Findings will be graphically represented with two forest plots, one for sensitivity and the other for specificity. The strategy for meta-analysis includes producing a summary receiver operating characteristic curve and estimating the summary sensitivity/specificity for each threshold provided in the articles.

Systematic review registration number: PROSPERO CRD42020224250

Keywords: COVID-19; point of care; rapid antigen tests; respiratory infection; SARS-CoV-2


Introduction

SARS-CoV-2 is a β-coronavirus belonging to the Coronaviridae family. This family is a positive-sense, single stranded RNA virus family with several subfamilies: α, β, γ, and δ. The α and β subfamilies are associated with mammalian infections, whereas the γ and δ subfamilies are associated with pig and bird infections.1 The main genes products belonging to these viruses are the nucleocapsid protein (N), the spike protein (S), the receptor-binding domain of the spike protein (RBD), the small membrane associated protein (SM), and the membrane glycoprotein (M).1 These genes can be used to identify the virus using nucleic acid amplification techniques (NAATs) to diagnose COVID-19 in patients being tested. Alternatively, the proteins themselves can be used to detect the presence of the virus in a sample.

As of February 3, 2021, the World Health Organization states that there have been over 103 million confirmed cases of COVID-19 and over 2.2 million
Accurate COVID-19 point-of-care (POC) diagnostic tests are needed for health providers to advise infected individuals and their contacts about appropriate treatment and the need to quarantine themselves from others. Point-of-care diagnostic tests for COVID-19 provide results within minutes and can be performed in a variety of locations including temporary test centers, doctors’ offices, the emergency room, and schools.

Point-of-care testing includes traditional real-time RT-PCR (qRT-PCR) and rapid antigen testing. The number of confirmed cases drawn from POC and medical laboratory diagnostic tests also provide vital statistics that allow governments and health organizations to assess the effectiveness of their policies for controlling the spread of COVID-19 in their region.

The demand for POC diagnostic tests has been answered by the innovation of a variety of nucleic acid, protein, CRISPR-Cas, and antibody-based detection tests for SARS-CoV-2. Evaluation of the accuracy of POC diagnostic tests is needed in order to utilize these tests with confidence.

Reverse transcriptase polymerase chain reaction (RT-PCR)–based tests use a process by which small amounts of viral RNA are converted to cDNA and then amplified to qualitatively demonstrate the presence of the virus or to quantitatively determine the virus load in a given sample. For SARS-CoV-2, this test detects the viral genome. Reverse transcriptase polymerase chain reaction is considered the gold standard for COVID-19 diagnosis due to its high sensitivity and specificity.

Combined with the relatively low cost of PCR and the ease of use, RT-PCR is an attractive option for the diagnosis of COVID-19. The major drawbacks to using RT-PCR in clinical diagnoses are the risks of specimen contamination and degradation in the collection and transport of the sample to the facility where the tests will be performed. Additionally, the amount of time that is required from the collection of the sample to the return of the results is often too long for effective quarantine procedures. Reverse transcriptase polymerase chain reaction requires trained personnel and must be performed at an equipped, certified medical laboratory, further hindering its efficiency as a POC technique.

Comparatively, rapid antigen testing is an attractive option for COVID-19 diagnosis due to its low cost, minimal training required for use, use of the same sample types as the RT-PCR test, and its ability to provide a result within a short timeframe. Rapid antigen tests are immunoassays that detect the presence of a specific viral protein, glycan, or nucleic acid, which implies current infection with SARS-CoV-2. Currently, there are seven rapid antigen diagnostic tests that have been granted emergency use authorization by the Food and Drug Administration (FDA) in the United States. However, there are other tests being developed and used that have not been granted FDA emergency-use authorization, and still more rapid antigen tests are used in other countries but are not readily available in the US. The accuracy of these tests compared to the gold standard RT-PCR appear to vary depending on the manufacturer; for example, the BD Veritor Plus System is reported to have an 84% positive percent agreement (PPA), while Quidel Sofia SARS Antigen FIA is reported to have a 96.7% PPA. In this systematic review, we will synthesize the current evidence regarding rapid antigen test accuracy for detection of SARS-CoV-2 and consider the overall performance of these techniques compared to the gold standard RT-PCR.

Our title was registered with the JBI Registration of Systematic Review Titles in June 2020, at which time no similar studies were available. A search of PROSPERO, DARE (Database of Abstracts of Reviews of Effects), PubMed, the Cochrane Database of Systematic Reviews, JBI Registration of Systematic Review Titles, and JBI Evidence Synthesis was conducted in November 2020. We identified one review in PROSPERO and two systematic reviews in the Cochrane Database of Systematic Reviews, each of which became available after our title registration. The PROSPERO review examined peer-reviewed publications for tests commercially available before August 15, 2020. Our systematic review will include additional sources for tests, including gray-literature available from the manufacturers, and will include search results from tests not yet commercially available. The Cochrane study’s literature searches ended in May 2020. In the time since then, significant amounts of research and numbers of tests have become available, warranting an additional review.

Dinnes et al. examined accuracy of tests to detect SARS-CoV-2 infection at the time of presentation in primary or secondary care settings, while our study asks if specific POC tests could replace RT-PCR tests.
in a primary care setting. Studies that contain results from secondary care settings that cannot be separated from primary care settings will be excluded. Further, the Dinnes et al. systematic review included secondary objectives, including participant characteristics (symptomology, disease severity, length of symptoms, virus load), to predict poor POC test performance in specific settings.\textsuperscript{13} Our study does not consider these influences on diagnostic accuracy.

Stegeman et al. focused on a triage setting for laboratory COVID-19 testing, not POC.\textsuperscript{14} As with Dinnes et al., the Stegeman et al. study included patient characteristics that may influence test results.\textsuperscript{13,14} Regionally, the other studies available now are focused on care in non-US settings, including the Philippines\textsuperscript{13} and the European Union/United Kingdom.\textsuperscript{12-14} Both Cochrane studies were published by the Cochrane COVID-19 Diagnostic Test Accuracy Group.\textsuperscript{13,14} We feel, with the rapidly changing environment around COVID-19, that our study will add to those published with the longer timeline, and our question is of a more general nature of POC diagnostic accuracy for a primary care setting anywhere in the world. Our study has important implications for health care providers caring for patients in both resource-rich and resource-poor regions.

We framed our review question using the Population Index test Reference test Diagnosis (PIRD) mnemonic, commonly used for diagnostic reviews. The objective of this systematic review is to synthesize the best available evidence related to diagnostic accuracy of the currently available POC rapid antigen tests (index test) relative to a certified medical laboratory viral genetic RT-PCR test (reference test) for the diagnosis of COVID-19/SARS-CoV-2 in adults (18+ years). The rationale for combining both test types in this systematic review is to provide a comprehensive comparison of RT-PCR with the POC rapid antigen tests. As the pandemic surrounding COVID-19 is rapidly evolving, the diagnostic accuracy, lowest cost, and quickest results are important considerations in the monitoring and management of disease spread for a primary care setting.

**Review question**

What is the diagnostic accuracy of the currently available and upcoming POC rapid antigen tests used in primary care settings relative to the viral genetic RT-PCR test as a reference for the diagnosis of COVID-19/SARS-CoV-2 in adults (18+ years)?

**Inclusion criteria**

**Participants**

The review will consider studies that include non-pregnant adults (18 years and older) with suspected SARS-CoV-2 infection, regardless of symptomology or disease severity. Any ethnicity or race in any geographic location will be considered. Studies that include pregnant women or children within the study population that can be separated from the overall study data will be included in this review. We will exclude studies that contain results from secondary care settings, such as inpatient care, that cannot be separated from results of primary care settings.

**Index test**

The index tests that will be investigated in this review are any currently available SARS-CoV-2 POC rapid antigen tests. Rapid antigen tests are qualitative or semi-quantitative diagnostics that provide a result within a short timeframe, typically within the hour following sample collection.\textsuperscript{8} Tests can use the following bodily fluids for sampling: saliva, mucus, blood, urine, and feces. Currently, authorized antigen tests primarily use nasopharyngeal, nasal, or oropharyngeal swab specimens, however, bronchoalveolar lavage and endotracheal aspirate are also used in more severe cases of COVID-19.\textsuperscript{3} Rapid antigen tests do not require specialized personnel to perform them, and therefore have been used not only in traditional laboratories but also at primary care/urgent care settings, patients’ bedsides, temporary test centers, schools, long-term care facilities, and pharmacies.\textsuperscript{15} Rapid antigen tests include a variety of techniques such as ELISA, chromogenic-based or fluorescence-based detection, and lateral flow-based detection, but as a common denominator all detect viral antigens from presently infected fluids and cells.\textsuperscript{8} In addition, CRISPR-Cas based testing, such as SHERLOCK and STOP, will be considered in this review as this technique provides rapid detection of viral nucleic acid without the need for traditional PCR equipment and PCR-trained personnel.\textsuperscript{3} Tests that detect immunoglobulin against SARS-CoV-2 will be excluded from this review, as antibodies develop upon resolution of SARS-CoV-2 infection or from vaccination and, therefore, are not used in the POC setting for diagnosing acute infection.\textsuperscript{3}
**Reference test**
The reference test will be commercially distributed RT-PCR–based tests that detect the RNA genome of SARS-CoV-2 and have been validated by an independent third party. Any RT-PCR test will be considered; for example, on February 4, 2020, the FDA issued an emergency use authorization for the 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. This panel is meant to provide qualitative results regarding the presence of the SARS-CoV-2 nucleic acid in a specimen collected from the upper or lower respiratory tract of the person being tested. These tests must be performed in certified laboratories where personnel have been trained to perform qRT-PCR assays. The primers and probes were chosen from regions of the nucleocapsid (N) gene.

**Diagnosis of interest**
The diagnoses of interest are COVID-19 disease and SARS-CoV-2 infection.

**Types of studies**
This review will consider any English-language cross-sectional study that examines the diagnostic accuracy (sensitivity and specificity, positive predictive value, negative predictive value) of COVID-19/SARS-CoV-2 infection where the participants have had both index and reference tests performed.

Studies published in English or translated into English will be included. Studies published from November 2019 to the present date will be included as SARS-CoV-2 emerged in late 2019 and is the cause of a continuing pandemic.

**Methods**
The proposed systematic review will be conducted in accordance with JBI methodology for systematic reviews of diagnostic test accuracy.

**Search strategy**
The search strategy will aim to locate both published and unpublished studies. An initial limited search of PubMed, PROSPERO, *JBI Evidence Synthesis*, Cochrane Database of Systematic Reviews, DARE, and the Cochrane Study Register was undertaken to identify articles on the topic, review other search strategies, and search for published articles on the same topic. The text words contained in the titles and abstracts of relevant articles, and the index terms used to describe the articles were used to develop a full search strategy for PubMed (see Appendix I). In addition, the strategy for COVID-19 was adapted from the Canadian Agency for Drugs and Technologies (CADTH) COVID-19 search string for PubMed. The search strategy was peer reviewed by a second librarian following the Peer Review of Electronic Search Strategy (PRESS) Guideline Statement. The search strategy, including all identified keywords and index terms, will be adapted for each included information source. The reference list of all studies selected for critical appraisal will be screened for additional studies.

MEDLINE (PubMed), Embase (Elsevier), Scopus (Elsevier), Qinsight (Quertle), and WHO COVID-19 database (World Health Organization) will be searched. Scopus, Qinsight, and WHO COVID-19 include gray literature.

**Study selection**
Following the search, all identified citations will be collated and uploaded into EndNote v.X9.3.3 (Clarivate Analytics, PA, USA) and duplicates removed. Titles and abstracts will then be screened by two independent reviewers for assessment against the inclusion criteria for the review. Potentially relevant studies will be retrieved in full and their citation details imported into a spreadsheet using Google Sheets (Alphabet Inc., CA, USA). The full text of selected citations will be assessed in detail against the inclusion criteria by at least two independent reviewers. Reasons for exclusion of full text studies that do not meet the inclusion criteria will be recorded and reported in the systematic review. Any disagreements that arise between the reviewers at each stage of the study selection process will be resolved through discussion or with an additional reviewer. The results of the search will be reported in full in the final systematic review and presented in a Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram.

**Assessment of methodological quality**
Selected studies will be critically appraised by at least two independent reviewers for methodological quality in the review using the standardized critical appraisal instrument from the QUADAS2. QUADAS2 provides a series of yes/no questions to appraise studies. We will require the following
questions to be answered “Yes” for a study to be included in the systematic review:

#2 – Was a case control design avoided?
#3 – Did the study avoid inappropriate exclusions?
#6 – Is the reference standard likely to correctly classify the target condition?
#8 – Was there an appropriate interval between index test and reference standard?

Any disagreements that arise will be resolved through discussion or with an additional reviewer. Following critical appraisal, studies that do not meet a certain quality threshold will be excluded. The decision to exclude will be based on the consensus of the two independent reviewers and, if needed, the additional reviewer. Studies will be excluded if the data cannot be used to calculate specificity and sensitivity, lacks statistical power, or is missing key relevant information based on our data extraction tool.

**Data extraction**

Data will be extracted from papers included in the review by at least two independent reviewers using a customized data extraction tool (see Appendix II). The data extracted will include specific details about the tests, populations, study methods, and outcomes of significance to the review question and specific objectives. Any disagreements that arise between the reviewers will be resolved through discussion or with an additional reviewer. Authors of papers will be contacted to request missing or additional data where required.

**Data synthesis**

Where possible, the results of papers will be pooled in statistical meta-analysis using the JBI System for the Unified Management, Assessment and Review of Information (JBI SUMARI; JBI, Adelaide, Australia). If meta-analysis is performed, pooled sensitivity and specificity values will be presented with 95% confidence intervals. These will be displayed on paired forest plots if the same diagnostic threshold values are used across studies, or summary receiver operating characteristic curves (SROC) if they vary. Heterogeneity will be assessed visually based on how closely the studies on the paired forest plot align, or how closely they map to the curve of the SROC. The strategy for meta-analysis will be based on the guidance by Campbell et al. If heterogeneity is suspected, it will be investigated through subgroup analysis. Potential subgroups that will be investigated include index tests that detect proteins, index tests that detect nucleic acids, index test methodological techniques, and country of origin. Where statistical pooling is not possible, the findings will be presented in narrative form, including tables and figures, to aid in data presentation where appropriate.

**Assessing certainty in the evidence**

A Summary of Findings will be created using GRADEPro GDT software (McMaster University, ON, Canada). The GRADE approach for grading the quality of evidence for diagnostic test accuracy will be followed.

The following outcomes will be included in the Summary of Findings: the review question; the index test names and types; the reference tests used; the population; the estimates of true negatives, true positives, false negatives, and false positives; the absolute difference between the index and reference tests for these values per 1000 patients; the sample size; the number of studies contained within the sample set; the GRADE (Grading of Recommendations Assessment, Development and Evaluation) quality of evidence for each finding; and any comments associated with the finding.

**Acknowledgments**

Elizabeth Hinton, librarian and assistant professor, and Alexa Stoneman, library director, for reviewing and assisting with the search strategy. Cheryl Vanier, chief research officer at Touro University Nevada, for her support of the research team and statistical expertise. Catie Chung, RN, and Kelly Mecham, DNP, APRN, for their support of this work and their critical review of the text.

**Funding**

SF and EH were supported by Touro University Nevada’s Federal Work Study Program.

**References**


Appendix I: Search strategy

MEDLINE (PubMed)
Search conducted on February 15, 2021

Filters: English language; publication date October 31, 2019 to present

<table>
<thead>
<tr>
<th>Search Number</th>
<th>Query</th>
<th>Results retrieved</th>
</tr>
</thead>
</table>
### Continued

<table>
<thead>
<tr>
<th>Search Number</th>
<th>Query</th>
<th>Results retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>(“Predictive Value of Tests”[Mesh] OR “Predictive value of tests”[all fields]) OR (“Sensitivity and Specificity”[Mesh] OR “sensitivity and specificity”[all fields])</td>
<td>638,528</td>
</tr>
<tr>
<td>#4</td>
<td>(#1 AND #2) AND (#3)</td>
<td>99</td>
</tr>
</tbody>
</table>
Appendix II: Data extraction instrument

<table>
<thead>
<tr>
<th>First author</th>
<th>Article title</th>
<th>DOI/PMID/other identifier</th>
<th>Country</th>
<th>Setting/context</th>
<th>Year/timeframe for data collection</th>
<th>Participant characteristics</th>
<th>Index test description and sample</th>
<th>Reference test description and sample</th>
<th>Description of main results (include adverse events from tests)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Predictive value</th>
<th>Exclusion reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>