

Infectious Diseases Society of America Guidelines on the Diagnosis of COVID-19:

Serologic Testing

Kimberly E. Hanson¹, Angela M. Caliendo², Cesar A. Arias³, Janet A. Englund⁴, Mary K. Hayden⁵, Mark J. Lee⁶, Mark Loeb⁷, Robin Patel⁸, Osama Altayar⁹, Abdallah El Alayli¹⁰, Shahnaz Sultan¹⁵, Yngve Falck-Ytter¹¹, Valéry Lavergne¹², Rebecca L. Morgan¹³, M. Hassan Murad¹⁴, Adarsh Bhimraj¹⁶, Reem A. Mustafa¹⁷

¹Divisions of Infectious Diseases and Clinical Microbiology, University of Utah, Salt Lake City, Utah

²Department of Medicine, Warren Alpert Medical School of Brown University, Providence, Rhode Island

³Division of Infectious Diseases, Center for Antimicrobial Resistance and Microbial Genomics, University of Texas Health Science Center at Houston, McGovern Medical School and Center for Infectious Diseases, School of Public Health, Houston, TX

⁴Department of Pediatrics, University of Washington, Seattle Children's Research Institute, Seattle, Washington

⁵ Division of Infectious Diseases, Department of Medicine, Rush University Medical Center, Chicago, Illinois; Department of Pathology, Rush University Medical Center, Chicago, Illinois.

⁶Department of Pathology and Clinical Microbiology Laboratory, Duke University School of Medicine, Durham, North Carolina

⁷Division of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario

⁸Division of Clinical Microbiology, Mayo Clinic, Rochester, Minnesota

© The Author(s) 2020. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com.

⁹Department of Medicine, Washington University School of Medicine, St. Louis, Missouri

¹⁰Outcomes and Implementation Research Unit, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, Kansas

¹¹VA Northeast Ohio Healthcare System, Case Western Reserve University School of Medicine, Cleveland, Ohio

¹²Department of Pathology and Laboratory Medicine, Vancouver General Hospital, Vancouver, British Columbia, Canada

¹³Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario

¹⁴Division of Preventive Medicine, Mayo Clinic, Rochester, Minnesota

¹⁵Division of Gastroenterology, Hepatology, and Nutrition, University of Minnesota, Minneapolis VA Healthcare System, Minneapolis, Minnesota

¹⁶Department of Infectious Diseases, Cleveland Clinic, Cleveland, Ohio

¹⁷Division of Nephrology and Hypertension, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, Kansas

Corresponding Author: Kimberly Hanson

Email: Kim.hanson@hsc.utah.edu

Panel Members: Kimberly E. Hanson (lead), Angela M. Caliendo, Cesar A. Arias, Janet A. Englund (PIDS representative), Mary K. Hayden, Mark J. Lee, Mark Loeb (SHEA representative), Robin Patel (ASM representative), Adarsh Bhimraj

Methodologists: Reem A. Mustafa (lead), Osama Altayar, Abdallah El Alayli, Shahnaz Sultan, Yngve Falck-Ytter, Valery Lavergne, Rebecca L. Morgan, M. Hassan Murad

Abstract

For the most updated version of these guidelines, please go to

<https://www.idsociety.org/covid19guidelines>.

Background: The availability of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) serologic testing has rapidly increased. Current assays use a variety of technologies, measure different classes of immunoglobulin or immunoglobulin combinations and detect antibodies directed against different portions of the virus. The overall accuracy of these tests, however, has not been well-defined. The Infectious Diseases Society of America (IDSA) convened an expert panel to perform a systematic review of the coronavirus disease 2019 (COVID-19) serology literature and construct best practice guidance related to SARS-CoV-2 serologic testing. This guideline is the fourth in a series of rapid, frequently updated COVID-19 guidelines developed by IDSA.

Objective: IDSA's goal was to develop evidence-based recommendations that assist clinicians, clinical laboratories, patients and policymakers in decisions related to the optimal use of SARS-CoV-2 serologic tests in a variety of settings. We also highlight important unmet research needs pertaining to the use of anti-SARS-CoV-2 antibody tests for diagnosis, public health surveillance, vaccine development and the selection of convalescent plasma donors.

Methods: A multidisciplinary panel of infectious diseases clinicians, clinical microbiologists and experts in systematic literature review identified and prioritized clinical questions related to the use of SARS-CoV-2 serologic tests. Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology was used to assess the certainty of evidence and make testing recommendations.

Results: The panel agreed on eight diagnostic recommendations.

Conclusions: Information on the clinical performance and utility of SARS-CoV-2 serologic tests are rapidly emerging. Based on available evidence, detection of anti-SARS-CoV-2 antibodies may be useful for confirming the presence of current or past infection in selected situations. The panel identified three potential indications for serologic testing including: 1) evaluation of patients with a high clinical suspicion for COVID-19 when molecular diagnostic testing is negative and at least two weeks have passed since symptom onset; 2) assessment of multisystem inflammatory syndrome in children; and 3) for conducting serosurveillance studies. The certainty of available evidence supporting the use of serology for either diagnosis or epidemiology was, however, graded as very low to moderate.

Accepted Manuscript

IDSA Legal Disclaimer

It is important to realize that guidelines cannot account for individual variation among patients. They are not intended to supplant physician judgment with respect to particular patients or special clinical situations. IDSA considers adherence to these guidelines to be voluntary, with the ultimate determination regarding their application to be made by the physician in the light of a patient's individual circumstances. While IDSA makes every effort to present accurate and reliable information, the information provided in these guidelines is "as is" without any warranty of accuracy, reliability, or otherwise, either express or implied. Neither IDSA nor its officers, directors, members, employees, or agents will be liable for any loss, damage, or claim with respect to any liabilities, including direct, special, indirect, or consequential damages, incurred in connection with implementation of these guidelines or reliance on the information presented.

The guidelines represent the proprietary and copyrighted property of IDSA. Copyright 2020 Infectious Diseases Society of America. All rights reserved. No part of these guidelines may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of IDSA. Permission is granted to physicians and health care providers solely to copy and use the guidelines in their professional practices and clinical decision-making. No license or permission is granted to any person or entity, and prior written authorization by IDSA is required, to sell, distribute, or modify the guidelines, or to make derivative works of or incorporate the guidelines into any product, including but not limited to clinical decision support software or any other software product. Except for the permission granted above, any person or entity desiring to use the guidelines in any way must contact IDSA for approval in accordance with the terms and conditions of third-party use, in particular any use of the guidelines in any software product.

Executive Summary

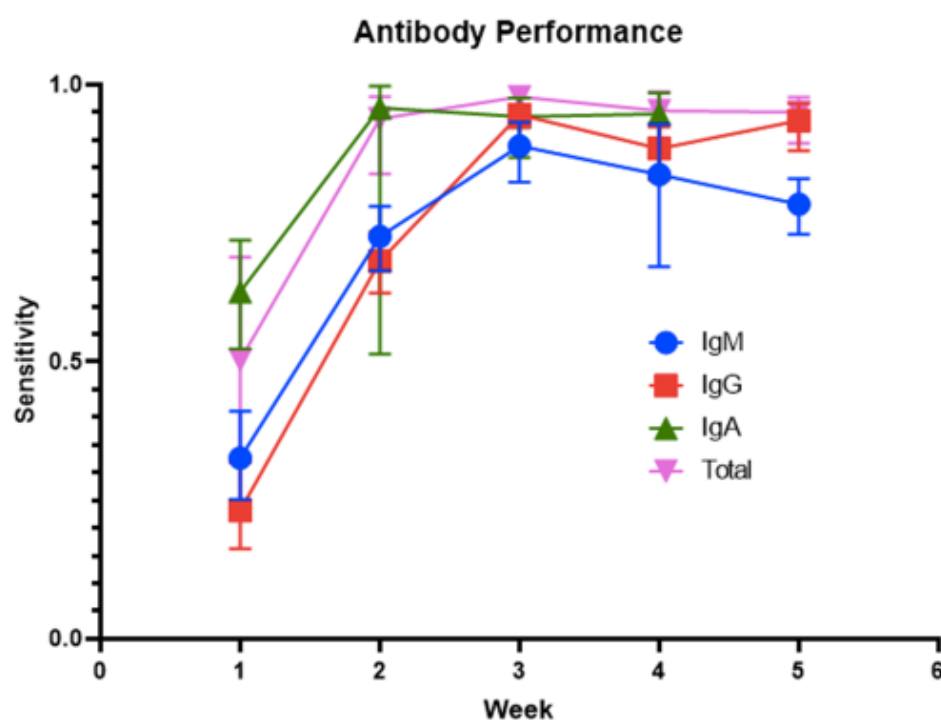
Serologic tests for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) are now widely available. Unlike nucleic acid amplification tests (NAAT) which detect viral RNA, antibody-based assays measure the host's humoral immune response to current or past infection. Anti-SARS-CoV-2 antibodies typically become detectable more than two weeks after the onset of symptoms ([Figure 1](#)). As a result, SARS-CoV-2 serology lacks sufficient sensitivity to confidently exclude the diagnosis of coronavirus disease 2019 (COVID-19) when antibodies are not detected in the acute phase of illness. NAAT remains the diagnostic modality of choice for acute infection. Antibody testing, however, may be useful as an adjunct to NAAT at later time points following infection. In general, IgM tests tend to have lower sensitivity to detect past infection than IgG or total antibody tests. Assays designed to detect and differentiate IgM and IgG in combination, where the detection of either IgM or IgG is used to define a positive test result, and IgA tests tend to have lower specificity to detect past infection compared to IgG only or total antibody tests. Test specificity is especially important for large serosurveillance studies when the prevalence of prior infection in the community is expected to be low. To be of value, anti-SARS-CoV-2 antibody tests are required to have high clinical sensitivity and specificity (i.e., $\geq 99.5\%$).

In addition to use in epidemiologic studies, the panel identified two clinical scenarios where antibody testing was felt to have potential utility for diagnosis. Serologic testing may be helpful in the evaluation of individual patients with a high clinical suspicion for COVID-19 when the results of molecular diagnostic testing are repeatedly negative or such testing was not performed. The sensitivity and specificity of IgG and total antibody is optimal three to four weeks after the onset of symptoms. At the current time, little data exists in the fifth week post-symptom onset to judge serologic test performance at later periods after infection. Detection of anti-SARS-CoV-2 antibodies is also useful for assessments of suspected multisystem inflammatory syndrome in children. For symptomatic patients, optimal serology result interpretation requires careful determination of the timing of testing relative to symptom onset

combined with assessments of disease severity. Based on the available evidence at this time, serologic tests should not be used to determine immunity or risk of re-infection. Thus, anti-SARS-CoV-2 antibody detection cannot inform decisions to discontinue physical distancing or lessen the use of personal protective equipment.

Summarized below are specific recommendations and comments related to the use of SARS-CoV-2 serologic testing in clinical practice and public health. A detailed description of background, methods, evidence summary and rationales that support each recommendation can be found online in the full text.

Figure 1. Antibody Sensitivity over Time



Footnote: This figure summarizes the pooled sensitivity with 95% confidence intervals of antibody classes per week post-symptom onset. The estimates were derived from the 24 studies and seven package inserts informing recommendations 1 through 5.

Recommendation 1: The IDSA panel suggests against using serologic testing to diagnose SARS-CoV-2 infection during the first two weeks (14 days) following symptom onset (conditional recommendation, very low certainty of evidence).

Recommendation 2: When SARS-CoV-2 infection requires laboratory confirmation for clinical or epidemiological purposes, the IDSA panel suggests testing for SARS-CoV-2 IgG or total antibody three to four weeks after symptom onset to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

- **Remark** – When serology is being considered as an adjunct to NAAT for diagnosis, testing three to four weeks post-symptom onset maximizes the sensitivity and specificity to detect past infection.
- **Remark** – Serosurveillance studies should use assays with high specificity (i.e., $\geq 99.5\%$), especially when the prevalence of SARS-CoV-2 in the community is expected to be low.

Recommendation 3: The IDSA panel makes no recommendation either for or against using IgM antibodies to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

Recommendation 4: The IDSA panel suggests against using IgA antibodies to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

Recommendation 5: The IDSA panel suggests against using IgM or IgG antibody combination tests to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

- **Remark** - IgM or IgG combination tests are those where detecting either antibody class is used to define a positive result.

Recommendation 6: The IDSA panel suggests using IgG antibody to provide evidence of COVID-19 infection in symptomatic patients with a high clinical suspicion and repeatedly negative NAAT testing (weak recommendation, very low certainty of evidence).

- **Remark** – When serology is being considered as an adjunct to NAAT for diagnosis, testing three to four weeks post-symptom onset maximizes the sensitivity and specificity to detect past infection.

Recommendation 7: In pediatric patients with multisystem inflammatory syndrome, the IDSA panel suggests using both IgG antibody and NAAT to provide evidence of current or past COVID-19 infection (strong recommendation, very low certainty of evidence).

Recommendation 8: The IDSA panel makes no recommendation for or against using capillary *versus* venous blood for serologic testing to detect SARS-CoV-2 antibodies (knowledge gap).

Background

Since its emergence in December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused over 21 million known infections and nearly 770,000 deaths worldwide [1]. Definitive diagnosis of coronavirus disease 19 (COVID-19), the illness caused by SARS-CoV-2 infection, relies on the direct detection of virus-specific RNA or virus-specific glycoprotein antigens in respiratory tract specimens. Serologic tests that detect the host antibody response to SARS-CoV-2 may also help to confirm the presence of current or past infection using blood samples.

Coronavirus genomes encode four major structural proteins including spike (S), envelope (E), membrane (M) and nucleocapsid (N). Both the S and N proteins of SARS-CoV-2 have been shown to be immunogenic in humans and current serologic tests target antibodies directed against these antigens [2]. The S protein is the most exposed viral protein and is responsible for viral attachment and entry into the host cell via binding to the angiotensin-converting enzyme 2 (ACE-2) receptor [3]. The S protein is comprised of an N-terminal S1 subunit, involved in virus-receptor binding, and a C-terminal S2 subunit that is involved in fusion to the host cell membrane. The S1 subunit is further divided into the N terminal domain (NTD) and a receptor binding domain (RBD). There has been particular focus on the SARS-CoV-2 RBD for vaccine development and targeted antibody therapies because neutralizing antibodies against this region effectively block viral entry [4, 5]. The N protein is an RNA-binding protein that is abundantly expressed during infection and plays an important role in RNA transcription and replication [6].

There are two general types of antibodies, neutralizing antibodies (nAbs) and non-neutralizing (also known as binding antibodies) [7]. Neutralization is defined as the loss of infectivity that occurs when a nAb binds to a viral particle. Virus-specific or vaccine-induced nAbs can play a crucial role in controlling viral infection, but definitive data is lacking to know whether individuals with detectable anti-SARS-CoV-2 nAbs are protected against reinfection. In comparison, binding antibodies are characterized by their inability to prevent viral infection of

permissive cells. Regardless of their function, both types of virus-specific antibodies are potentially useful as diagnostic indicators of current or past infection.

Commercially available anti-SARS-CoV-2 antibody tests use different technologies to qualitatively measure single immunoglobulin classes (IgM, IgG or IgA) or total antibody, but do not differentiate nAbs from binding antibodies. IgM antibodies directed against microorganisms are typically produced first after infection and are used as a measure of recent infection. IgG antibodies generally develop later after IgM and remains elevated for months to years after infection. Although IgM antibodies can be detected within the first two weeks of symptoms in some patients, SARS-CoV-2 infection appears unusual in that IgM and IgG more commonly rise together, more than two weeks after the onset of symptoms [8]. Secretory IgA is important for mucosal immunity. IgA can also be detected systemically in certain types of infection including SARS-CoV-2, but comparatively little is known about the kinetics of IgA in blood. The components of “total antibody” presumably include IgM and IgG and theoretically other antigen-specific immunoglobulins as well.

Given that the majority of the population has previously been exposed to seasonal human coronaviruses (HCoV), and these viruses may share similar structure with SARS-CoV-2, an essential part of serologic test development and validation is to assure that the anti-SARS-CoV-2 antibodies detected by a given assay do not cross react with other coronaviruses (e.g., HCoV-229E, HCoV-NL63, HCoV-OC43 or HCoV-HKU1). Specificity studies typically involve analyzing archived sera obtained before the identification of COVID-19 as a clinical entity, as well as assessing for potential interfering substances such as auto-antibodies or heterophile antibodies.

The most common clinical diagnostic platforms utilized for SARS-CoV-2 include lateral flow (LF) devices, enzyme linked immunosorbent assays (ELISA) and chemiluminescent immunoassays (CIA). Lateral flow assays typically require a drop of blood (or serum or plasma) applied to a test strip, with results read in approximately 15-30 minutes. These devices are suitable for point-of-care testing and have potential to be deployed in the field as a part of large

serological surveys. ELISA comes in a variety of different formats. Typically, a bound antigen-antibody complex is detected using a type-specific secondary antibody linked to a substrate that generates a colorimetric or fluorescent signal. CIA methods are similar to ELISA but use chemical probes that emit light instead of enzymatic substrates. Both ELISA and CIA are clinical laboratory-based methods amenable to high throughput testing using serum, plasma or potentially dried blood spots. At this time, neutralization assays are mainly used in research settings or offered as laboratory developed tests by reference laboratories.

In the United States, the Food and Drug Administration (FDA) currently requires Emergency Use Authorization (EUA) to market a SARS-CoV-2 antibody test. This means that commercial manufacturers and clinical laboratories with laboratory-developed tests must submit performance data to the FDA for review. Early in the pandemic, however, official EUA review was voluntary. Test developers were only expected to internally validate their tests and notify the FDA of their intent to market. As a result, the market was flooded with poorly performing assays. In response the FDA subsequently issued a “removed” test list that includes tests where significant performance problems were identified, assays for which official EUA review was not appropriately submitted or assays voluntarily withdrawn by the developer.

Many different serologic tests for SARS-CoV-2 have become commercially available in a short amount of time. The incredible speed of development has significantly outpaced rigorous assessments of test performance. Therefore, IDSA convened an expert panel to systematically review the available serologic literature, compare pooled estimates of test accuracy and make evidence-based recommendations for informed use in clinical practice.

Methods

This guideline was developed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for evidence assessment. In addition, given the need for a rapid response to an urgent public health crisis, the methodological approach was modified according to the GIN/McMaster checklist for the development of rapid

recommendations [9]. To assess the positive and negative predictive value of serologic testing we considered prevalence 1% to represent communities with low levels of circulating SARS-CoV-2 infections, 10% [10-12] which is typically seen in “hot spots” and 40% which is typically seen in patients meeting the clinical definition for COVID-19 who were hospitalized or in the investigation of outbreaks in congregate settings or factories [13, 14].

Panel Composition

The panel was composed of clinicians and clinical microbiologists who are members of the IDSA, the American Society for Microbiology (ASM), the Society for Healthcare Epidemiology of America (SHEA), and the Pediatric Infectious Diseases Society (PIDS). They represented the disciplines of infectious diseases, pediatrics, clinical microbiology, hepatology, nephrology and gastroenterology. The Evidence Foundation provided technical support and guideline methodologists for the development of this guideline.

Disclosure and Management of Potential Conflicts of Interest

The conflict of interest (COI) review group included two representatives from IDSA who were responsible for reviewing, evaluating and approving all disclosures. All members of the expert panel complied with the COI process for reviewing and managing conflicts of interest, which required disclosure of any financial, intellectual, or other interest that might be construed as constituting an actual, potential, or apparent conflict, regardless of relevancy to the guideline topic. The assessment of disclosed relationships for possible COI was based on the relative weight of the financial relationship (i.e., monetary amount) and the relevance of the relationship (i.e., the degree to which an association might reasonably be interpreted by an independent observer as related to the topic or recommendation of consideration). The COI review group ensured that the majority of the panel and chair was without potential relevant (related to the topic) conflicts. The chair and all members of the technical team were determined to have no COIs relevant to the guidelines.

Question Generation

Clinical questions were developed into a PICO format (Population, Intervention, Comparison, Outcomes) [15] prior to the first panel meeting. Panel members prioritized questions with available evidence that met the minimum acceptable criteria (i.e., diagnostic test accuracy reported on at least a case-series design, case reports were excluded) (**Figure s1**). Panel members prioritized patient-important outcomes such as the role of serologic testing in diagnosing acute or recent infections, the role of point of care serologic tests and the role of serologic tests in the evaluation of pediatric patients with inflammatory multisystem syndromes.

Search Strategy

The National Institute of Health and Care Excellence (NICE) and the Centers for Disease Control and Prevention (CDC) highly sensitive search was reviewed by the methodologists in consultation with an experienced information specialist and was determined to have high sensitivity. Terms identified in the PICO questions and the term “COVID” were added to the search strategy. We searched Ovid Medline and Embase from 2019 through June 19, 2020. We also performed horizon scans periodically during the evidence assessment and recommendation process to locate additional grey literature and manuscript preprints from LitCovid, medRxiv/bioRxiv, and SSRN. Reference lists and literature suggested by the panelists were reviewed for inclusion as well. We also manually searched the manufacturers’ package inserts of the serologic tests that received EUA by the FDA.

Screening and Study Selection

Two independent reviewers screened the titles and abstracts of the references identified by the search strategy and then reviewed the full texts of the studies under consideration for inclusion. Disagreements were resolved by discussion to reach consensus in consultation with a

third reviewer when needed. Studies were reviewed if they reported data on the diagnostic test accuracy of anti-SARS-CoV-2 IgM, IgG, IgA, and/or total antibody tests compared to NAAT as the reference standard. We only included studies that evaluated one of the three most commonly used serology platforms (i.e., LF, ELISA and CIA). To be included, cohort, cross-sectional and case-control studies had to include at least 30 samples, describe the time of testing relative to symptom onset and provided information on the reference standard for comparison. A sample size of 30 specimens was set to mirror the minimal numbers of specimens required by the FDA for EUA, but assays were included in our review regardless of FDA authorization status. Assays that required both IgM and IgG to be detected to define a positive result were excluded. We also excluded studies that reported sensitivity only, specificity only and studies that did not provide enough information to extract the true positives, true negatives, false positives, and false negatives.

Data Collection and Analysis

The data extraction was completed by two independent reviewers in duplicate. Disagreements were resolved by reaching consensus and consulting with an expert clinician scientist member of the panel. We extracted baseline characteristics (authors, publication year, country, study design, inclusion criteria, age, gender), index test information (timing from onset of symptoms, sample type, target antigen, platform, immunoglobulin class, FDA EUA status, and European Economic Area CE marking status), reference test (name of test, sample type), and diagnostic test accuracy raw data (true and false positives and negatives).

We used the bivariate random effects model to pool the sensitivity and specificity using the logit transformation, when there were enough studies [16]. When the number of studies did not allow the use of the bivariate model, we pooled the sensitivity and specificity separately using the random effects generalized linear mixed models [17, 18]. To evaluate the between study heterogeneity, we examined the forest plots for each pooled estimate rather than relying on the I^2 statistic which does not take into account the variability resulting from different

positivity thresholds. We evaluated the graphs visually for factors that could explain the heterogeneity including the platform used, FDA EUA or European Conformity (CE) mark statuses, and target antigen. To evaluate the effects of our decision to exclude studies that included fewer than 30 samples, we performed sensitivity analyses by including all the studies that reported both sensitivity and specificity. We also performed sensitivity analyses by excluding studies that included less than 100 samples in the specificity group and studies that evaluated tests that were removed from FDA EUA list. The analyses were performed using the packages *mada 0.5.10* and *meta 4.11.0* in *R 3.6.3* [19-21].

Risk of Bias and Certainty of Evidence

We used the Quality Assessment of Diagnostic Studies (QUADAS)-2 revised tool to assess the risk of bias in the included studies [22]. We used the GRADE framework to assess the overall certainty by evaluating the body of evidence for each outcome on the following domains: risk of bias, imprecision, inconsistency, indirectness, and publication bias [23, 24]. We developed GRADE summary of findings tables using the GRADEpro Guideline Development Tool [25].

Evidence to Recommendations

The panel considered the core elements of GRADE evidence in the decision process, including certainty of evidence and balance between desirable and undesirable effects. Additional domains were acknowledged where applicable (e.g., feasibility, resource use, acceptability). For all recommendations, the expert panelists reached consensus. Voting rules were agreed on prior to the panel meeting for situations when consensus could not be reached.

As per GRADE methodology, recommendations were labeled as “strong” or “conditional.” The words “we recommend” indicate strong recommendations and “we suggest” indicate conditional recommendations. [Figure 2](#) provides the suggested interpretation of

strong and weak recommendations for patients, clinicians, and healthcare policymakers. For recommendations where the comparators are not formally stated, the comparison of interest was implicitly referred to as “not using the test.” Some recommendations acknowledge a current “knowledge gap” and aim at avoiding premature favorable recommendations for test use and to avoid encouraging the rapid diffusion of potentially non-useful tests.

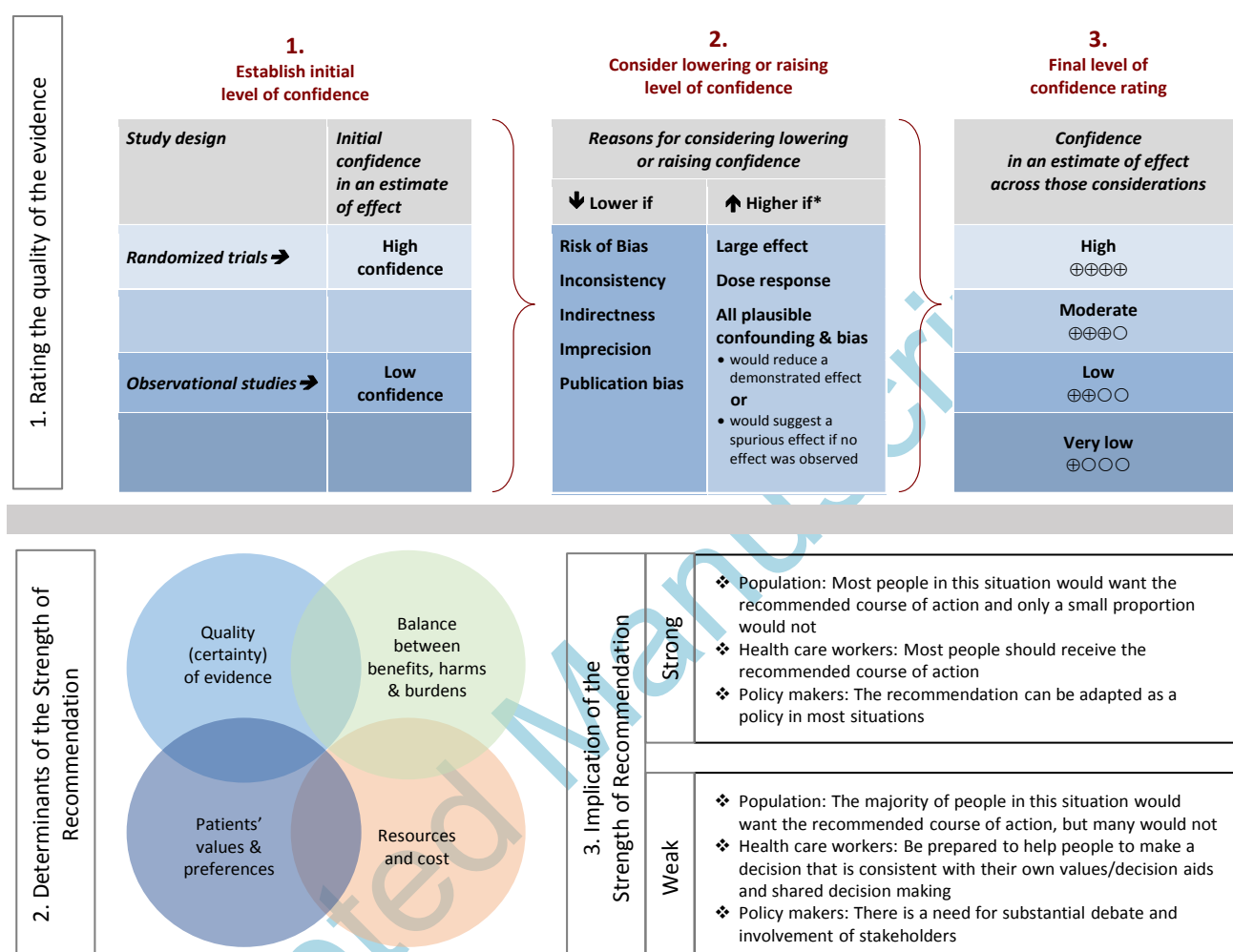
Revision Process

The draft guideline underwent a rapid review for approval by IDSA Board of Directors Executive Committee external to the guideline development panel. The guidelines was reviewed and endorsed by ASM, PIDS and SHEA. The IDSA Board of Directors Executive Committee reviewed and approved the guideline prior to dissemination.

Updating Process

Regular, frequent screening of the literature will take place to determine the need for revisions based on the likelihood that any new data will have an impact on the recommendations. If necessary, the entire expert panel will be reconvened to discuss potential changes.

Figure 2. Approach and implications to rating the quality of evidence and strength of recommendations using the GRADE methodology (unrestricted use of the figure granted by the U.S. GRADE Network)



Results

Systematic review and horizon scan of the literature identified 9468 references of which 47 informed the evidence base for these recommendations (**Figure s2 - PRISMA flow diagram**). Characteristics of the included studies can be found in (**Tables s1-s5**).

Recommendation 1: The IDSA panel suggests against using serologic testing to diagnose SARS-CoV-2 infection during the first two weeks (14 days) following symptom onset (conditional recommendation, very low certainty of evidence).

Summary of evidence

Using our search strategy, we identified 24 [26-49] studies and seven package inserts [50-56] that assessed the diagnostic test accuracy of serologic tests compared to SARS-CoV-2 reverse transcription polymerase chain reaction (RT-PCR) (**Table s2**). Studies used different controls including samples collected from healthy individuals, patients who had other respiratory or non-respiratory infections, patients with autoimmune diseases whose blood had been collected prior to the COVID-19 pandemic, symptomatic patients with negative NAAT for SARS-CoV-2, patients hospitalized for other reasons and asymptomatic individuals including pregnant women evaluated during the COVID-19 pandemic. All of the package insert information [50-56] and 21 studies [26-47] were case control studies. Two additional studies used a cohort design [48, 49]. Included studies reporting on three different testing methodologies, namely LF assays, ELISA, and CIA. Studies also assessed different antibody classes including IgM alone, IgG alone, IgA alone or total antibody, depending on the kit used. For the assays that detected and differentiated IgM and IgG in the same platform, results interpretation included assessments “IgM or IgG” approaches where the presence of one of them qualified as a positive test.

The total number of samples included ranged from 91 to 2,708 and 721 to 11,887, for the sensitivity and specificity analyses, respectively. The pooled sensitivity at week one after symptom onset ranged from 0.23 to 0.63 and at week two ranged from 0.68 to 0.96, while the pooled specificity ranged from 0.96 to 1 ([Tables 1a-1e](#)). The quality of evidence that informed sensitivity determinations ranged from very low to low, while specificity was low to moderate. Quality was rated down for serious risk of bias (case control study design), imprecision

(assuming the upper and lower limits of the confidence interval would lead to different decisions) and unexplained inconsistency.

Benefits and Harms

The panel placed a high value on reducing false negative results. During the first two weeks following infection the sensitivity for all serology tests, regardless of the platform and immunoglobulin detected, was inadequate to avoid a large number of false negative results. The concern with low sensitivity is that individuals who test negative would be classified as uninfected, when, in fact, they had been infected but have not fully developed an antibody response. Likewise, when assessing the seropositivity rate of a population, tests with poor sensitivity will provide an underestimation of the percentage of the population that is or was infected. Tests with low specificity can lead to false positive results at any time point of testing. A false positive result can lead to the incorrect conclusion that an individual has been infected, thus eliminating the search for the true etiology of their symptoms. When assessing the seroprevalence of the population, false positive results can lead to an overestimation of the percentage of the population that has been infected.

Waiting beyond two weeks after the onset of symptoms to test for an antibody response may delay the confirmation of infection in some cases. A portion of infected individuals (23% to 63%, depending on the test) will develop antibodies within the first week from onset of symptoms, such that early testing could be actionable if positive. However, depending on the specificity of the test, a significant number of those results could be false positives; this is particularly concerning when the pre-test probability of infection is low. While this may not be as serious of an issue for the surveillance studies, it may be important in the limited clinical situations where serologic tests are used for the diagnosis of COVID-19 (see recommendations [6](#) and [7](#)).

Other considerations

IgM antibody responses typically occur earlier after the onset of infection compared to IgG antibody responses in most microbial infections. In contrast, with SARS-CoV-2 infection, there does not appear to be a significant difference in the sensitivity of tests that detect IgM antibody compared to those that detect IgG antibodies in the first weeks following infection. At one week after onset of symptoms, the pooled sensitivity for IgM tests was 33% compared to 23% for IgG tests, and at two weeks post onset of symptoms, the pooled sensitivity of IgM tests was 73% compared to 68% for IgG tests ([Tables 1a-1b](#)). The specificity for IgM tests was slightly lower (98%) compared to IgG tests (99%), but the confidence intervals overlapped. Based on available evidence, there does not appear to be a substantially increased diagnostic accuracy when using IgM tests compared to IgG tests early in the course of illness. IgA tests, while more sensitive than IgG and IgM tests within the first two weeks of symptom onset, have a lower specificity (96%) and so their use would greatly increase false positive results. In a low prevalence population for example (i.e., 1%), 87% of the positive IgA results would be false positives ([Table 1c](#)); when the prevalence is 10% (for example in “hot spots”), 39% of the positive results would be false positives. Of note, the number of studies evaluating IgA are much fewer than for IgM and IgG.

There are LF assays that detect and distinguish between IgM and IgG. These platforms require either IgM or IgG to be detected for a positive result (referred to as IgM or IgG tests). This “either/or” interpretation increases test sensitivity, but also slightly reduces specificity from 98% to 97%. Overall, IgM or IgG detection does not improve the diagnostic accuracy relative to detection of either antibody class alone. In contrast, detection of total antibody within the first two weeks after infection does increase sensitivity while maintaining a high specificity. These differences may be due to the fact that IgM or IgG tests tend to be LF assays, which show more inconsistency than total antibody detected using ELISA or CIA (**Figure s3 - Forest Plots**). Total antibody platforms may have the highest diagnostic utility during the early

time period. However, the number of studies evaluating total antibody tests was substantially lower than the number of studies assessing IgM or IgG tests.

Conclusions and research needs for this recommendation

There is substantial variability in the performance of tests and different testing platforms (i.e., LF vs. CIA vs. ELISA). Overall, there is more inconsistency across LF platforms and with IgM compared to IgG tests. Regardless of the immunoglobulin detected or the testing platform used, anti-SARS-CoV-2 antibodies generally lack adequate sensitivity to rule out infection during the first two weeks of symptoms. Additional cohort studies are needed to truly understand the performance of serology tests including studies with larger numbers of well characterized patients, where the timing of symptom onset and the severity of illness are clearly defined. Ideally, cohort studies should be conducted assessing multiple tests and using the same well characterized specimens. The antibody response in special populations such as children, immunocompromised patients, and patients with autoimmune or rheumatologic disease also needs to be studied. Lastly, correlation of viral RNA shedding and culture positivity (or other surrogate for infectivity) should be studied relative to immunoglobulin titers over time.

Table 1a. Antibody Performance, Weeks 1 and 2 – IgM

IgM	Week 1			Week 2		
Sensitivity	0.33 (95% CI: 0.25 to 0.41)			0.73 (95% CI: 0.66 to 0.78)		
Specificity	0.98 (95% CI: 0.97 to 0.99)					
Outcome	Effect per 1,000 patients tested					
	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c
True positives (patients with COVID-19)	3 (3 to 4)	33 (25 to 41)	132 (100 to 164)	7 (7 to 8)	73 (66 to 78)	292 (264 to 312)
False negatives (patients incorrectly classified as not having COVID-19)	7 (6 to 7)	67 (59 to 75)	268 (236 to 300)	3 (2 to 3)	27 (22 to 34)	108 (88 to 136)
Quality of the evidence	12 studies, 919 patients ⊕⊕○○ LOW ^{d, e}			16 studies, 2309 patients ⊕⊕○○ LOW ^{d, e}		
	pre-test probability of 1% ^a		pre-test probability of 10% ^b		pre-test probability of 40% ^c	
True negatives (patients without COVID-19)	970 (960 to 980)		882 (873 to 891)		588 (582 to 594)	
False positives (patients incorrectly classified as having COVID-19)	20 (10 to 30)		18 (9 to 27)		12 (6 to 18)	
Quality of Evidence	21 studies, 7165 patients ⊕⊕⊕○ MODERATE ^d					

a. Typically seen in general population in areas that are not hotspots

b. Typically seen in general population in high risk populations

c. Typically seen in general population in exposed and nursing homes

d. The case-control design leads to a serious risk of bias

e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C1: 0.06-0.62, C2: 0.33-1.00

Table 1b. Antibody Performance, Weeks 1 and 2 – IgG

IgG	Week 1			Week 2		
Sensitivity	0.23 (95% CI: 0.16 to 0.32)			0.68 (95% CI: 0.62 to 0.73)		
Specificity	0.99 (95% CI: 0.99 to 0.99)					
Outcome	Effect per 1,000 patients tested					
	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c
True positives (patients with COVID-19)	2 (2 to 3)	23 (16 to 32)	92 (64 to 128)	7 (6 to 7)	68 (62 to 73)	272 (248 to 292)
False negatives (patients incorrectly classified as not having COVID-19)	8 (7 to 8)	77 (68 to 84)	308 (272 to 336)	3 (3 to 4)	32 (27 to 38)	128 (108 to 152)
Quality of the evidence	13 studies, 1343 patients ⊕○○○ VERY LOW ^{d,e}			16 studies, 2708 patients ⊕⊕○○ LOW ^{d,e}		
	pre-test probability of 1% ^a		pre-test probability of 10% ^b		pre-test probability of 40% ^c	
True negatives (patients without COVID-19)	980 (980 to 980)		891 (891 to 891)		594 (594 to 594)	
False positives (patients incorrectly classified as having COVID-19)	10 (10 to 10)		9 (9 to 9)		6 (6 to 6)	
Quality of Evidence	25 studies, 11887 patients ⊕⊕⊕○ MODERATE ^d					

a. Typically seen in general population in areas that are not hotspots

b. Typically seen in general population in high risk populations

c. Typically seen in general population in exposed and nursing homes

d. The case-control design leads to a serious risk of bias

e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C1: 0.00,0.69, C2: 0.27-0.91

Table 1c. Antibody Performance, Weeks 1 and 2 – IgA

IgA	Week 1			Week 2		
Sensitivity	0.63 (95% CI: 0.52 to 0.72)			0.96 (95% CI: 0.51 to 1.00)		
Specificity	0.96 (95% CI: 0.91 to 0.99)					
Outcome	Effect per 1,000 patients tested					
	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c
True positives (patients with COVID-19)	6 (5 to 7)	63 (52 to 72)	252 (208 to 288)	10 (5 to 10)	96 (51 to 100)	384 (204 to 400)
False negatives (patients incorrectly classified as not having COVID-19)	4 (3 to 5)	37 (28 to 48)	148 (112 to 192)	0 (0 to 5)	4 (0 to 49)	16 (0 to 196)
Quality of the evidence	2 studies, 91 patients ⊕⊕○○ LOW ^{d,e}			2 studies, 102 patients ⊕○○○ VERY LOW ^{d,e}		
	pre-test probability of 1% ^a		pre-test probability of 10% ^b		pre-test probability of 40% ^c	
True negatives (patients without COVID-19)	950 (901 to 980)		864 (819 to 891)		576 (546 to 594)	
False positives (patients incorrectly classified as having COVID-19)	40 (10 to 89)		36 (9 to 81)		24 (6 to 54)	
Quality of Evidence	4 studies, 760 patients ⊕⊕○○ LOW ^{d,e}					

- Typically seen in general population in areas that are not hotspots
- Typically seen in general population in high risk populations
- Typically seen in general population in exposed and nursing homes
- The case-control design leads to a serious risk of bias
- Considering the Upper vs Lower limits of the sensitivity's confidence interval would lead to different clinical decisions.

Table 1d. Antibody Performance, Weeks 1 and 2 – Total Antibodies

Total Antibodies	Week 1			Week 2		
Sensitivity	0.50 (95% CI: 0.32 to 0.69)			0.94 (95% CI: 0.84 to 0.98)		
Specificity	1.00 (95% CI: 0.99 to 1.00)					
Outcome	Effect per 1,000 patients tested					
	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c
True positives (patients with COVID-19)	5 (3 to 7)	50 (32 to 69)	200 (128 to 276)	9 (8 to 10)	94 (84 to 98)	376 (336 to 392)
False negatives (patients incorrectly classified as not having COVID-19)	5 (3 to 7)	50 (31 to 68)	200 (124 to 272)	1 (0 to 2)	6 (2 to 16)	24 (8 to 64)
Quality of the evidence	7 studies, 418 patients ⊕○○○ VERY LOW ^{d,e,f}			6 studies, 359 patients ⊕⊕○○ LOW ^{d,e}		
	pre-test probability of 1% ^a		pre-test probability of 10% ^b		pre-test probability of 40% ^c	
True negatives (patients without COVID-19)	990 (980 to 990)		900 (891 to 900)		600 (594 to 600)	
False positives (patients incorrectly classified as having COVID-19)	0 (0 to 10)		0 (0 to 9)		0 (0 to 6)	
Quality of Evidence	8 studies, 4521 patients ⊕⊕⊕○ MODERATE ^d					

a. Typically seen in general population in areas that are not hotspots

b. Typically seen in general population in high risk populations

c. Typically seen in general population in exposed and nursing homes

d. The case-control design leads to a serious risk of bias

e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C1: 0.03-0.75, C2: 0.74-1.00

f. Considering the Upper vs Lower limits of the sensitivity's confidence interval would lead to different clinical decisions.

Table 1e. Antibody Performance, Weeks 1 and 2 – IgM or IgG

IgM or IgG	Week 1			Week 2		
Sensitivity	0.51 (95% CI: 0.42 to 0.59)			0.81 (95% CI: 0.77 to 0.84)		
Specificity	0.97 (95% CI: 0.95 to 0.98)					
Outcome	Effect per 1,000 patients tested					
	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c	pre-test probability of 1% ^a	pre-test probability of 10% ^b	pre-test probability of 40% ^c
True positives (patients with COVID-19)	5 (4 to 6)	51 (42 to 59)	204 (168 to 236)	8 (8 to 8)	81 (77 to 84)	324 (308 to 336)
False negatives (patients incorrectly classified as not having COVID-19)	5 (4 to 6)	49 (41 to 58)	196 (164 to 232)	2 (2 to 2)	19 (16 to 23)	76 (64 to 92)
Quality of the evidence	7 studies, 830 patients ⊕⊕○○ LOW ^{d,e}			7 studies, 1996 patients ⊕⊕○○ LOW ^{d,e}		
	pre-test probability of 1% ^a		pre-test probability of 10% ^b		pre-test probability of 40% ^c	
True negatives (patients without COVID-19)	960 (941 to 970)		873 (855 to 882)		582 (570 to 588)	
False positives (patients incorrectly classified as having COVID-19)	30 (20 to 50)		27 (18 to 45)		18 (12 to 30)	
Quality of Evidence	11 studies, 5660 patients ⊕⊕○○ LOW ^{d,e}					

a. Typically seen in general population in areas that are not hotspots

b. Typically seen in general population in high risk populations

c. Typically seen in general population in exposed and nursing homes

d. The case-control design leads to a serious risk of bias

e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C1: 0.09-0.79, C2: 0.46-0.94

Recommendation 2: When SARS-CoV-2 infection requires laboratory confirmation for clinical or epidemiological purposes, the IDSA panel suggests testing for SARS-CoV-2 IgG or total antibody three to four weeks after symptom onset to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

- **Remark** – When serology is being considered as an adjunct to NAAT for diagnosis, testing three to four weeks post-symptom onset maximizes the sensitivity and specificity to detect past infection.
- **Remark** – Serosurveillance studies should use assays with high specificity (i.e., $\geq 99.5\%$), especially when the prevalence of SARS-CoV-2 in the community is expected to be low.

Recommendation 3: The IDSA panel makes no recommendation either for or against using IgM antibodies to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

Recommendation 4: The IDSA panel suggests against using IgA antibodies to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

Recommendation 5: The IDSA panel suggests against using IgM or IgG antibody combination tests to detect evidence of past SARS-CoV-2 infection (conditional recommendation, very low certainty of evidence).

- **Remark** - IgM or IgG combination tests are those where detecting either antibody class is used to define a positive result.

Summary of the evidence

The same studies [26-49] and package inserts [50-56] that were used to inform recommendation 1 were examined to assess test performance at weeks three, four, and five following the onset of COVID-19 signs or symptoms (**Table s2**). The total number of samples included in the sensitivity analyses ranged from 163 to 2,298 samples and the number included in the specificity analyses ranged from 721 to 11,887 samples. The pooled sensitivity at week three after symptom onset ranged from 0.89 to 0.98, at week four from 0.84 to 0.95, and at week five from 0.78 to 0.95, while the pooled specificity ranged from 0.96 to 1 ([Tables 2a-2e](#)). The quality of evidence that informed sensitivity analyses ranged from very low to moderate; the quality of evidence that informed specificity was also low to moderate. There was a serious risk of bias (case control study design), imprecision (assuming the upper and lower limits of the confidence interval would lead to different decisions) and unexplained inconsistency.

Benefits and Harms

The panel placed a high value on reducing false positive test results and determining the optimal timing of testing to confidently assess previous infection. Detection of IgG or total antibody at three to four weeks after the onset of symptoms provides the highest sensitivity, and thus the lowest rate of false negative results, compared to other immunoglobulin classes or earlier time points. IgG or total antibody tests also provide high specificity, and thus, reduce the rate of false positive results relative to other antibody types.

The specificity of IgM tests in the three to four weeks post-symptom onset timeframe was equivalent to IgG tests. However, in contrast to other viral infections where IgM tests show high sensitivity shortly after onset of symptoms compared to IgG, IgM sensitivity against SARS-CoV-2 is relatively low early on and there is no significant increase over time as is seen with IgG or total antibody ([Table 2a](#), [Table 2b](#), [Table 2d](#)). The use of the IgM tests alone would result in increased

false negative rates compared to IgG or total antibody tests. Given these tradeoffs, the panel suggested neither for nor against the use of IgM tests to assess previous infection.

IgA tests or LF devices with both IgM and IgG targets, where detection of either antibody is considered a positive result, all suffer from lower specificity compared to tests involving IgG or total antibody targets, leading to increased false positive rates. These tests would falsely increase seroprevalence and potentially mislead public health officials, policymakers, and the general public. In addition, false positive results might detract from pursuing alternative diagnoses in symptomatic patients. Therefore, the panel suggested against the use of any of these tests at this time.

Other considerations

The predictive value of diagnostic test results depends on the performance characteristics of the test (i.e., sensitivity and specificity) and on the prevalence of the disease in the population tested. In general, the higher the prevalence, the higher the false negative rate, and the lower the prevalence, the higher the false positive rate. To illustrate the importance of this concept in interpreting antibody test results, the panel calculated the number of false positive and false negative results in populations with 1% and 10% SARS-CoV-2 seroprevalence, using pooled sensitivity and specificity from the data review. The 1% prevalence was chosen to represent a population with few prior cases of COVID-19, and the 10% prevalence was chosen to represent a COVID-19 “hot spot” population. As can be seen in [Figure 3](#), even tests with seemingly high sensitivity and specificity can yield a large proportion of erroneous results at the extremes of prevalence. In this example, in a population with a 1% prevalence of SARS-CoV-2, a test with 96% sensitivity and 99% specificity would generate an equal number of true positive and false positive results.

Limited data are available on duration of the IgG response, but in the studies we reviewed it appeared to be sustained for at least three to five weeks. The panel recommends testing for IgG or

total antibody at three to four weeks because there are limited data on IgG or total antibody responses five weeks or longer after onset of symptoms. The IgM response begins to decrease by week four after symptom onset, such that IgM appears to be less sensitive over time than IgG for detection of recent past infection with SARS-CoV-2.

Overall, the panel identified substantial variability in sensitivity and specificity among tests (**Figure s3 - Forest Plots**). While inconsistent performance characteristics were observed across the three test methodologies included in the data summary, more inconsistency was seen for LF methods than for CIA or ELISA assays. Similar to the acute timeframe, IgM assays showed more variable performance than did IgG assays at later time posts after infection. Interestingly, the specific antigen targeted in the antibody test (e.g., S protein or N protein) did not appear to impact test performance, although some studies did not report the antigen that was targeted.

Conclusions and research needs for this recommendation

Sensitivity and specificity differences between available assays require that clinicians understand the performance characteristics of the particular test that is used. A general sense of prevalence (or pre-test probability) in the population being tested is also helpful to ensure accurate interpretation of test results. While NAAT remains the recommended approach for diagnosis of COVID-19, detection of IgG or total antibody directed against SARS-CoV-2 at three to four weeks after symptom onset may be useful for determining past infection in selected clinical situations (see recommendations [6](#) and [7](#)). IgG alone is typically used for determining seroprevalence, but information on the duration of a detectable antibody response beyond four weeks is limited.

It is important to realize that the presence of SARS-CoV-2 IgG or total antibody cannot be inferred to represent protective immunity against SARS-CoV-2 reinfection. Future research should include investigation as to whether SARS-CoV-2 reinfection occurs, and whether the presence of antibodies – including neutralizing antibodies – or other measurable immunological responses (e.g.,

T-cell responses) are markers of protection. Additional knowledge gaps include the duration of detectable IgG and total antibody; relationship of seropositivity to shedding of infectious virus in the convalescent phase; effect of seropositivity on COVID-19 vaccine response and factors that affect antibody responses (e.g., age, comorbid medical conditions, immunocompromised state).

Larger studies of well-characterized and varied populations will be needed to begin to answer these questions. Availability of a reference standard for seropositivity (e.g., presence of neutralizing antibody) would also be valuable.

Accepted Manuscript

IgM	Week 3			Week 4			Week 5		
Sensitivity	0.89 (95% CI: 0.82 to 0.93)			0.84 (95% CI: 0.67 to 0.93)			0.78 (95% CI: 0.73 to 0.83)		
Specificity	0.98 (95% CI: 0.97 to 0.99)								
Outcome	Effect per 1,000 patients tested								
	Pre-test Probability								
	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c
True positives (patients with COVID-19)	9 (8 to 9)	89 (82 to 93)	356 (328 to 372)	8 (7 to 9)	84 (67 to 93)	336 (268 to 372)	8 (7 to 8)	78 (73 to 83)	312 (292 to 332)
False negatives (patients incorrectly classified as not having COVID-19)	1 (1 to 2)	11 (7 to 18)	44 (28 to 72)	2 (1 to 3)	16 (7 to 33)	64 (28 to 132)	2 (2 to 3)	22 (17 to 27)	88 (68 to 108)
Quality of the evidence	14 studies, 1730 patients ⊕○○○ VERY LOW ^{d,e,f}			6 studies, 619 patients ⊕○○○ VERY LOW ^{d,e,f}			2 studies, 260 patients ⊕⊕○○ LOW ^{d,e}		
	pre-test probability of 1% ^a			pre-test probability of 10% ^b			pre-test probability of 40% ^c		
True negatives (patients without COVID-19)	970 (960 to 980)			882 (873 to 891)			588 (582 to 594)		
False positives (patients incorrectly classified as having COVID-19)	20 (10 to 30)			18 (9 to 27)			12 (6 to 18)		
Quality of Evidence	21 studies, 7165 patients ⊕⊕⊕○ MODERATE ^d								
a. Typically seen in general population in areas that are not hotspots b. Typically seen in general population in high risk populations c. Typically seen in general population in exposed and nursing homes d. The case-control design leads to a serious risk of bias e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C3: 0.55-1.00, C4: 0.36-1.00, C5: 0.76-0.92. f. Considering the Upper vs Lower limits of the sensitivity's confidence interval would lead to different clinical decisions.									

	Week 4
	0.88 (95% CI: 0.83 to 0.92)
	0.99 (95% CI: 0.99 to 0.99)

10 (9 to 10)	95 (92 to 96)	380 (370 to 390)
0 (0 to 1)	5 (4 to 8)	20 (15 to 25)
16 studies, 2298 patients ⊕⊕⊕○ MODERATE ^d		

Table 2c. Antibody Performance, Weeks 3 to 5 – IgA

IgA	Week 3			Week 4			Week 5		
Sensitivity	0.94 (95% CI: 0.87 to 0.98)			0.95 (95% CI: 0.83 to 0.99)			NR		
Specificity	0.96 (95% CI: 0.91 to 0.99)								
Outcome	Effect per 1,000 patients tested								
	Pre-test Probability								
	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c
True positives (patients with COVID-19)	9 (9 to 10)	94 (87 to 98)	376 (348 to 392)	10 (8 to 10)	95 (83 to 99)	380 (332 to 396)	NR	NR	NR
False negatives (patients incorrectly classified as not having COVID-19)	1 (0 to 1)	6 (2 to 13)	24 (8 to 52)	0 (0 to 2)	5 (1 to 17)	20 (4 to 68)	NR	NR	NR
Quality of the evidence	4 studies, 163 patients ⊕⊕○○ LOW ^{d,e}			3 studies, 191 patients ⊕⊕○○ LOW ^{d,e}			NR		
	pre-test probability of 1% ^a			pre-test probability of 10% ^b			pre-test probability of 40% ^c		
True negatives (patients without COVID-19)	950 (901 to 980)			864 (819 to 891)			576 (546 to 594)		
False positives (patients incorrectly classified as having COVID-19)	40 (10 to 89)			36 (9 to 81)			24 (6 to 54)		
Quality of Evidence	4 studies, 760 patients ⊕⊕○○ LOW ^{d,e}								

a. Typically seen in general population in areas that are not hotspots

b. Typically seen in general population in high risk populations

c. Typically seen in general population in exposed and nursing homes

d. The case-control design leads to a serious risk of bias

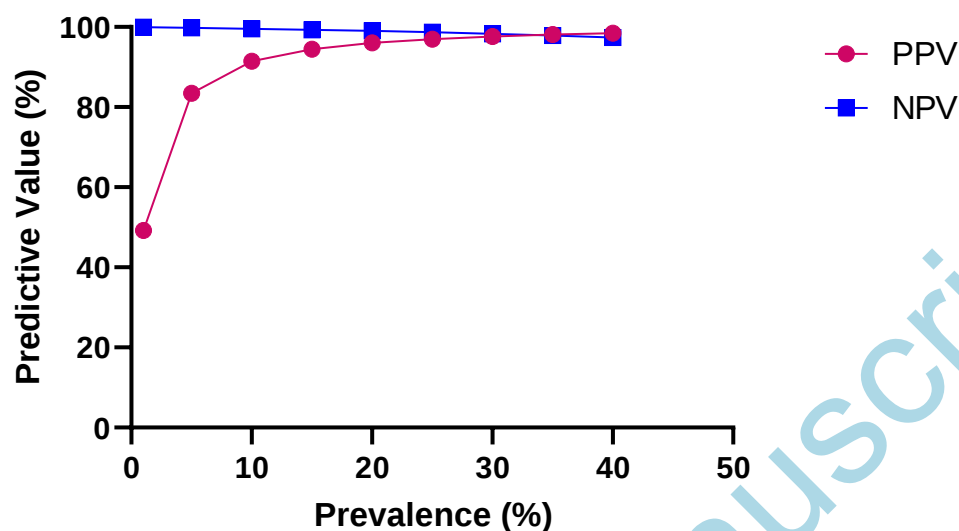
e. Considering the Upper vs Lower limits of the sensitivity's confidence interval would lead to different clinical decisions.

Total	Week 3			Week 4			Week 5		
Sensitivity	0.98 (95% CI: 0.89 to 1.00)			0.95 (95% CI: 0.84 to 0.99)			0.95 (95% CI: 0.90 to 0.98)		
Specificity	1.00 (95% CI: 0.99 to 1.00)								
Outcome	Effect per 1,000 patients tested								
	Pre-test Probability								
	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c	1% ^a	10% ^b	40% ^c
True positives (patients with COVID-19)	10 (9 to 10)	98 (89 to 100)	392 (356 to 400)	5 (4 to 5)	95 (84 to 99)	380 (336 to 396)	5 (5 to 5)	95 (90 to 98)	380 (360 to 392)
False negatives (patients incorrectly classified as not having COVID-19)	0 (0 to 1)	2 (0 to 11)	8 (0 to 44)	0 (0 to 1)	5 (1 to 16)	20 (4 to 64)	0 (0 to 0)	5 (2 to 10)	20 (8 to 40)
Quality of the evidence	6 studies, 472 patients ⊕⊕⊕○ MODERATE ^d			2 studies, 289 patients ⊕⊕○○ LOW ^{d,e}			1 study, 121 patients ⊕⊕○○ LOW ^{d,e}		
	pre-test probability of 1% ^a			pre-test probability of 10% ^b			pre-test probability of 40% ^c		
True negatives (patients without COVID-19)	990 (980 to 990)			900 (891 to 900)			600 (594 to 600)		
False positives (patients incorrectly classified as having COVID-19)	0 (0 to 10)			0 (0 to 9)			0 (0 to 6)		
Quality of Evidence	8 studies, 4521 patients ⊕⊕⊕○ MODERATE ^d								
a. Typically seen in general population in areas that are not hotspots b. Typically seen in general population in high risk populations c. Typically seen in general population in exposed and nursing homes d. The case-control design leads to a serious risk of bias e. Unexplained inconsistency observed with considerably variable sensitivity. Sensitivity Ranges: C3: 0.78-1.00, C4: 0.88-0.94, C5: 0.95									

	Week 4
	0.88 (95% CI: 0.78 to 0.94)
	0.97 (95% CI: 0.95 to 0.98)

9 (9 to 10)	94 (92 to 96)	376 (36
1 (0 to 1)	4 (6 to 8)	24 (16
6 studies, 1038 patients $\oplus\oplus\circ\circ$ LOW ^{d, e}		

Figure 3. Impact of Prevalence on Test Predictive Values



Footnote: This graph models the predictive value of a theoretic serologic test with 96% sensitivity and 99% specificity across a spectrum of prevalence. When the true prevalence of SARS-CoV-2 infection in a population is 1%, the positive predictive value of the test is only 49%. In other words, there is a 49% chance (close to a flip of the coin) that individuals with a positive screening test truly have the disease. As the prevalence increases, so does the predictive value. In contrast, when the prevalence of infection is 1%, the negative predictive value high (i.e., there is close to 100% probably the disease is truly absent when the result of the test is negative).

Abbreviations: positive predictive value (PPV); negative predictive value (NPV)

Recommendation 6: The IDSA panel suggests using IgG antibody to provide evidence of COVID-19 infection in symptomatic patients with a high clinical suspicion and repeatedly negative NAAT testing (weak recommendation, very low certainty of evidence).

- **Remark** – When serology is being considered as an adjunct to NAAT for diagnosis, testing three to four weeks post-symptom onset maximizes the sensitivity and specificity to detect past infection.

Summary of the evidence

In addition to the evidence that informed recommendations 1 through 5, we identified seven studies that compared the diagnostic test accuracy of serologic tests among patients with high clinical suspicion for COVID-19 and either negative NAAT throughout the course of their illness [57-61], or negative NAAT on repeat testing after initial positive tests [62, 63]. The seven case studies showed an increase in diagnostic yield when antibodies were added to NAAT after 14 days of symptom onset. In addition, IgG identified a higher number of positive cases compared to IgM (**Table s3**). It is again important to note that individual test sensitivity varied across studies. Variable test performance may have been due to sub-optimal timing of serology, poorly performing assays or misclassification of the reference standard (i.e., labeling non-infected patients as true positive cases based on serology alone). The overall quality of evidence was low due to risk of bias (case control study design) and imprecision as a result of small sample size.

Benefits and Harms

The potential benefit of true positive serologic results in patients with clinical signs and symptoms of COVID-19, but repeatedly negative NAAT, is confirmation of diagnosis for epidemiological and prognostic purposes. Additionally, positive results may help reduce unnecessary ancillary testing or empiric therapies directed towards alternative diagnoses. False negative serologic results, especially those obtained early in the course of symptomatic illness, could lead to inadequate isolation or quarantine as well as an under estimation of community

prevalence. For example, assuming a prevalence of 40% in 1,000 symptomatic patients with negative NAAT, based on the studies and package inserts that informed the diagnostic accuracy of IgG for recommendations 1 and 2 (above) [26-56], there will be 128 to 308 false negatives within the first two weeks of symptom onset compared to 20-48 false negatives during weeks three to five after symptoms onset.

Other Considerations

Studies in NAAT negative symptomatic patients mainly reported on the use of IgM and IgG antibodies. Given the overall evidence supporting the use of total antibody in the later time periods after symptom onset, it is likely total antibody could also be useful for the evaluation of symptomatic patients presenting late (i.e., three to four weeks after symptom onset) and have repeatedly negative NAAT. Although the timing of testing relative to symptom onset was reported in these case series, exact timing may be difficult to ascertain in routine clinical practice. In addition, most of the studies included in our review involved hospitalized patients with radiographic abnormalities suggestive of viral pneumonia. Severity of illness may correlate with the likelihood of seropositivity [64]. In addition, it is unclear whether RNA detection would have occurred in patients with pneumonia if repeat NAAT testing had been performed using a lower respiratory tract specimen (i.e., sputum or bronchoalveolar lavage fluid) as follow-up of an initial negative upper respiratory swab.

Conclusions and research needs for this recommendation

In conclusion, assessing anti-SARS-CoV-2 IgG antibodies provides evidence of COVID-19 in patients with high clinical suspicion and negative NAAT results. The timing of antibody testing is important, with optimal sensitivity observed at least two weeks from the time of symptom onset. Ideally, serologic studies of NAAT negative patients should evaluate quantitative antibody

responses at defined time points, such that seroconversion and/or neutralizing antibodies can be measured as evidence of true infection. Furthermore, outpatients with mild illness, children and immunocompromised patients should also be included in future studies.

Recommendation 7: In pediatric patients with multisystem inflammatory syndrome, the IDSA panel suggests using both IgG antibody and NAAT to provide evidence of current or past COVID-19 infection (strong recommendation, very low certainty of evidence).

Summary of the evidence

We identified 9 case series [65-73] that evaluated the use of SARS-CoV-2 serologic tests in patients presenting with signs of multisystem inflammatory syndrome in children (MIS-C) (**Table s4**). The case definitions of this inflammatory syndrome, typically manifesting with fever and shock similar to a Kawasaki-like syndrome, varied across studies and most did not require laboratory evidence of SARS-CoV-2 infection to define the illness. Instead, included studies showed a higher detection rate of MIS-C when IgG antibodies were added to NAAT compared to NAAT alone for the diagnosis. NAAT was positive in approximately 30-50% of cases in these studies, while IgG antibodies were positive in a majority (>80%) of cases. The major limitations of the evidence include the fact that not all case series specified the timing of testing (NAAT or serology) with respect to symptom onset. In addition, the type of serological test utilized and epidemiological links to COVID-19 cases were also not uniformly specified. The overall quality of evidence was very low due to the limitations of the case series study design and serious risk of bias.

Benefits and Harms

Children infected with SARS-CoV-2 may have no history of symptomatic disease or known exposure to COVID-19. Therefore, viral and serological testing may be especially important in this population. In fact, positive anti-SARS-CoV-2 antibody testing is now part of the MIS-C case definition. However, false positive results may detract from identifying the true cause of illness. Given the overlap in clinical features and difference in treatment of multisystem inflammatory syndrome diseases, it is important to differentiate MIS-C from Kawasaki Disease and other inflammatory processes such as bacterial infections and rheumatic fever.

Other Considerations

In clinical practice, young patients with presumed MIS-C are often critically ill and therefore NAAT and IgG serology should be obtained simultaneously. Although we found no studies reporting on the use of total antibody in the setting of MIS-C, it seems likely that these assays could also be useful for defining the syndrome. Evaluation of serology prior to the administration of IVIG or blood products is also important because these therapeutic modalities have the potential to alter the serologic response.

Conclusions and research needs for this recommendation

MIS-C is an emerging syndrome. Although causality due to COVID-19 is strongly suspected, it is not proven. Serologic testing should be performed to help establish a diagnosis when patients have signs or symptoms consistent with the late complications of COVID-19. However, relatively little is known about the performance of SARS-CoV-2 antibody testing in children. Therefore, prospective pediatric studies are needed to define diagnostic sensitivity and specificity relative to a reference standard as well as measure the timing and durability of antibody responses in children.

Studies to better define the pathophysiology and risks for developing MIS-C following SARS-CoV-2 infection are also needed.

Recommendation 8: The IDSA panel makes no recommendation for or against using capillary *versus* venous blood for serologic testing to detect SARS-CoV-2 antibodies (knowledge gap).

Summary of evidence

We identified three studies [48, 74, 75] and one package insert [76] that compared the diagnostic test accuracy of serologic tests using fingerstick (capillary) blood to venous blood samples obtained via venipuncture (**Table s5**). Identified publications included both a cohort study [48] and case control studies [74-76]. Only one study [76] compared fingerstick to venous blood directly, showing 100% positive and negative percent agreement between the two sampling methods. However, this study was excluded from the analyses because it did not provide diagnostic test accuracy results by week after symptom onset. The remaining studies [48, 74, 75] reported diagnostic test accuracy using fingerstick sampling relative to SARS-CoV-2 NAAT.

Due to the lack of evidence comparing the two sampling methods directly, we relied on indirect evidence by using conventional subgroup analysis methods to compare the pooled sensitivities and specificities from LF studies that used fingerstick sampling to the pooled sensitivities and specificities from LF studies that used venous blood [28, 29, 36, 38, 43, 44, 46, 51, 59, 60, 77-84]. As the timing of sampling seems crucial to evaluate the antibody responses, comparisons for this recommendation were made using data obtained at least two to three weeks after symptom initiation.

The included studies provided diagnostic test accuracy results for LF assays using IgM and/or IgG on fingerstick blood collected during week two and three post symptom onset. The pooled sensitivities of IgG collected during week three and the presence of either IgM or IgG collected during weeks two and three using fingerstick blood were comparable to those with venous blood, except for wider confidence intervals with fingerstick blood. IgM collected during week three showed a lower pooled sensitivity when collected using fingerstick compared to venous blood. Of note, the main test modality available for capillary blood testing was LF assays, which may have variations in performance by manufacturer (**Figure s3 - Forest Plots**). In sum, there was very low certainty of evidence overall due to a serious risk of bias (i.e., evidence from case-control studies), indirectness (i.e., indirect comparisons as outlined above) and imprecision (due to the small number of studies and/or samples).

Benefits and Harms

Capillary blood is usually obtained from a dermal puncture of the fingertip capillary beds and represents a mixture of venous and arterial blood. In general, capillary blood is an ideal specimen for “point-of-care testing” and is often the specimen of choice for those with difficult venous access such as very young infants, elderly patients or severely burned patients. In addition, capillary blood can be used for dried blood spot testing, which can be mailed to the laboratory or preserved for future testing. Since this approach avoids large needles, it may be preferred by individuals who fear needles. Fingerstick collection also facilitates self-collection, which can save on PPE and healthcare worker-facing time. In contrast, a potential harm of fingerstick blood is that the test accuracy (sensitivity and specificity) may be lower compared to venous sampling. The harms of venous blood draw are that it is considered by some to be an unpleasant experience, potentially leading to local side effects such as bleeding or bruising, requires specialized personnel, and use of PPE and healthcare worker-facing time.

Other considerations

There is precedent for using capillary blood sampling for the detection of antibodies against other respiratory viruses, such as influenza [85]. Accurate finger stick testing for SARS-CoV-2 antibodies would be ideal when evaluating post-inflammatory complications of COVID-19 such as MIS-C. In addition, point-of-care testing would be useful for deployment in large surveillance projects and contact tracing efforts. Capillary blood testing could also be utilized in clinical trials for monitoring serological response to anti-SARS-CoV-2 vaccines.

Conclusions and research needs for this recommendation

In summary, there was inadequate information to compare the performance of capillary *versus* venous blood for the detection of antibodies against SARS-CoV-2. Due to the lack of high-quality evidence, the panel reached the conclusion not to recommend for or against use of capillary blood to measure SARS-CoV-2 antibodies. If equivalent performance of capillary blood and venous blood is ultimately confirmed, then there is potential to deploy a simple point-of-care strategy for large-scale assessments of previous SARS-CoV-2 infection. This may also be the serologic test method of choice when larger volumes of blood cannot be obtained (e.g., very young infants).

Further studies are needed to demonstrate equivalence of SARS-CoV-2 serological testing using capillary versus venous blood as well as to compare direct sample analysis to dried blood spots. It is likely that the potential impact of sample hemolysis, tissue clotting factors, and specimen volume will be extremely method-dependent. Therefore, understanding each individual assay's test performance will be essential. Accuracy studies should be designed such that venous and capillary blood are taken from the same patient, at the same time, with a robust sample size and excellent clinical characterization of study subjects. Since diagnostic test accuracy may also be affected by

clinical parameters (i.e., presence or absence of symptoms, signs and/or radiologic findings, duration of symptoms and disease severity), monitoring over time will be needed to assess waning antibody responses. Understanding the value of capillary blood testing in the context of assessing immunity after vaccination will also be warranted. Finally, beyond fingerstick blood collection, other non-venipuncture-based blood collection devices (especially self-collection) are being developed. As these become available, validation of individual serologic tests with blood collected from individual collection devices will need to be performed.

Narrative Summaries of Serodiagnostics Undergoing Evaluation

In addition to the clinical questions addressed above, the panel identified several diagnostic approaches currently undergoing evaluation for which additional data are needed to formulate recommendations. Narrative summaries for these approaches are provided below.

Neutralizing antibody and cellular immune responses

Neutralizing antibody titers vary among recovered patients [86], indicating that other parts of the immune system such as T-cells and cytokines are likely to contribute to viral clearance. However, in some COVID-19 recovered individuals, detectable antibodies do correlate with the number of virus-specific T-cells [87]. Standardized assays designed to accurately quantitate nAbs or measure T-cell mediated immunity directed against SARS-CoV-2 do not currently exist but will be needed for comprehensive studies of immunity going forward. Whether nAb titers correlate with protection and/or infectiousness still needs to be determined. Assessing protection also has important implications for vaccine trials, selecting convalescent plasma donors as well as determining which components of the plasma confer anti-viral activity.

Classical nAb assays rely on culturing live virus. For SARS-CoV-2, this procedure requires appropriate biocontainment facilities certified for work with BSL-3 pathogens. Pseudovirus assays

have been developed to circumvent the safety concerns associated with SARS-CoV-2 culture and, previously, other highly pathogenic viruses [88]. These methods typically use retroviruses capable of integrating the envelope glycoproteins of another virus to form a “pseudovirus.” Compared with natural viruses, pseudoviruses can only infect cells in a single round, produce high titer infections and are not easily inactivated by serum complement. Whether nAb titers determined by different pseudovirus assays produce similar results is not yet known and the FDA has yet to authorize use of neutralization tests for SARS-CoV-2. Of note, one small COVID-19 study observed excellent correlation between nAb titers measured by a pseudovirus compared to a live virus assay [86].

Detecting antibodies directed against different SARS-CoV-2 antigens and multiple-test algorithms

In our systematic review, there was no difference in the performance of assays designed to detect the N protein or various portions of the S protein. However, patient sera that reacts with N protein-based tests may be different than the sera that reacts with S protein-based tests [89]. Combining the two antigens in a single test, or sequential testing with two tests targeting different proteins when the first assay is negative, may increase the number of true positive results. Similarly, the CDC has recommended using an orthogonal testing algorithm (i.e., employing two different tests when the first yields a positive result) to increase the positive predictive value of testing when the likelihood of previous exposure to SARS-CoV-2 in the community is low [90]. The logistics, cost and impact of reflexive testing strategies need to be determined prospectively.

Discussion

The immune response to SARS-CoV-2 infection has not yet been fully elucidated. Early evidence suggests that, unlike other infectious diseases, anti-SARS-CoV-2 IgM antibodies become detectable later after symptom onset and rise nearly simultaneously with IgG after two weeks of infection ([Figure 1](#)). Therefore, detection of IgM without IgG is uncommon. In line with other

systematic reviews, our antibody comparisons stratified by time post-symptom onset indicate that serologic testing has limited utility for ruling out SARS-CoV-2 infection in the acute phases of illness [91, 92]. Pooled sensitivity estimates for IgM in week two post-symptom onset was 73% ([Table 1a](#)). Thus, relying on IgM for the diagnosis of COVID-19 at two weeks would miss 108 true cases out of 1000 patients when the clinical suspicion for infection is high. The sensitivity of other immunoglobulin classes is equally as poor early after infection.

Although these tests have limited value for diagnosing acute infection, serology complements molecular testing for individuals presenting later in the course of illness. The optimal performance of serologic testing occurs approximately three to four weeks post-symptom onset and is achieved using IgG or total antibody assays. Overall, the sensitivity of IgG or total antibody at four weeks was 88% and 95%, respectively. Using IgG or total antibody reduces, but does not eliminate, false negative results at four weeks [20 to 48 false negatives out of 1000 individuals with high clinical suspicion for COVID-19 ([Table 2b](#), [Table 2d](#))]. Data on time points beyond four weeks is limited.

Interpretation of test specificity also requires caution because little information is available from individuals who were tested due to clinical suspicion of SARS-CoV-2. Moreover, using NAAT as the reference standard may spuriously increase “false positive” serology results, when it is the molecular diagnostic test that is falsely negative. Identifying a better reference standard for future serology studies will help to define the performance of antibody tests. Relying on seroconversion with a documented four-fold rise in IgG titers is suggestive of recent infection. However, most commercial antibody assays do not provide quantitative results and false positive IgG detection is still possible. Correlation of the circulating antibodies detectable with available tests with nAb titers can also provide confidence in positive immunoglobulin results. Whether all individuals, including special populations such as neonates, the elderly or immunocompromised, generate a detectable nAbs response to SARS-CoV-2 also needs to be established.

It is important to emphasize that serologic test performance is highly variable and sometimes the same manufacturer's assay performed quite differently across studies. In subgroup analysis, we found no substantive differences in test characteristics (i.e., sensitivity and specificity) based on viral antigen or whether the test has an EUA from the FDA (data not shown). Alternatively, the clinical performance of LF assays was more variable than ELISA or CIA tests. In addition to poorly performing LF tests, factors that may have contributed to heterogeneity across studies include differences in the patient population tested, the exact timing of testing relative to symptom onset and the use of different NAAT assays as the reference standard for comparison. Patients with mild symptoms, or those who are asymptomatic [64], may display weaker immune responses to SARS-CoV-2 compared to those with severe illness. Standardized reporting of disease severity was not included in most studies. Going forward, serologic assay studies should include a description of the severity of illness using well-defined grading criteria. The exact timing of testing was also not specified in all studies. For example, patients with symptoms for "less than 21 days" likely bridged various time points but were lumped in this guideline in the three-week time period. Uncertainty around the timing of testing could have affected the time-stratified analyses.

Perhaps the greatest interest in serologic testing has been for tracking SARS-CoV-2 exposure in the community. Accurate estimates of seroprevalence depend on the true prevalence of past infection in a given population along with the sensitivity and specificity of the test used to detect antibodies. When the true prevalence of infection is low, small decreases from a test sensitivity of 100% have a minimal impact on the negative predictive value of a test, while reductions in specificity artificially inflate measures of seroprevalence. As the true prevalence increases, similar deviations in sensitivity are more impactful, while reductions in specificity are less noticeable. Across the U.S., including regions significantly impacted by COVID-19, the prevalence of past infection is still expected to be relatively low (i.e., <7%) [93, 94]. At the current time, choosing a test with high specificity is especially important to reduce false positive results.

Seroprevalence in long term care facilities, congregate settings or factories which have experienced an outbreak are likely to have much higher seroprevalence, but these are situations that do not mirror the broader pandemic in most communities.

The magnitude and duration of humoral immune responses to SARS-CoV-2 infection has not been defined. Early evidence suggests that IgG and nAb antibodies begin to decline in the first three months following onset of symptoms in patients with mild COVID-19 similar to seasonal coronaviruses [64, 95]. These observations could have important implications for point prevalence studies. Future studies are needed to define the antibody dynamics and to determine whether detection of antibodies (and if so at what titers) confer protection against reinfection.

Conclusions

Clinicians and public health officials need to understand the performance of serology assays used in their settings in order to accurately interpret anti-SARS-CoV-2 antibody test results. Whenever possible, serologic assays with established high sensitivity and specificity (i.e., $\geq 99.5\%$) should be employed. IgG and total antibody tests appear to have better sensitivity and specificity than other immunoglobulin classes and perform best when used between three to four weeks after symptom onset. The clinical indications for antibody testing to support a diagnosis of COVID-19 are limited at this time. Testing can be considered for the evaluation of patients with a high clinical suspicion for SARS-CoV-2 infection despite repeatedly negative NAAT (especially those presenting late after symptom onset) and should be included in the assessment of MIS-C. Serologic tests also have potential utility for tracking the course of the SARS-CoV-2 pandemic in the community. The effectiveness and durability of anti-SARS-CoV-2 antibody responses, however, have not yet been defined. Thus, serologic testing cannot be used to determine immune status.

NOTES

Financial Support

This project was funded in part by a cooperative agreement with the Centers for Disease Control and Prevention (grant number 6 NU50CK000477-04-01). The Centers for Disease Control and Prevention is an agency within the Department of Health and Human Services (HHS). The contents of this guideline do not necessarily represent the policy of CDC or HHS and should not be considered an endorsement by the Federal Government.

COI Summary

The following list displays what has been reported to the IDSA. To provide thorough transparency, the IDSA requires full disclosure of all relationships, regardless of relevancy to the guideline topic. Evaluation of such relationships as potential conflicts of interest is determined by a review process which includes assessment by the Board of Directors liaison to the Standards and Practice Guideline Committee and, if necessary, the Conflicts of Interest (COI) and Ethics Committee. The assessment of disclosed relationships for possible COI is based on the relative weight of the financial relationship (i.e., monetary amount) and the relevance of the relationship (i.e., the degree to which an association might reasonably be interpreted by an independent observer as related to the topic or recommendation of consideration). The reader of these guidelines should be mindful of this when the list of disclosures is reviewed. **K.H.** serves as an advisor for BioFire and Quideland and receives research funding from the National Institutes of Health (NIH). **A.C.** serves as an advisor for Roche Diagnostics, Danaher, Quidel, First Light, Day Zero, Visby, and Chroma Code; receives research funding from ArcBio and Hologic; and has served as an advisor for Luminex. **C.A.** receives royalties from UpToDate and receives research funding from Merck, MeMed Diagnostics, Entasis Pharmaceuticals and the National Institute of Allergy and Infectious Diseases (NIAID)/NIH. **J.E.** serves as a consultant for Sanofi Pasteur; an

advisor/consultant for Meissa Vaccines; and receives research funding from the Centers for Disease Control and Prevention (CDC), Brotman Baty Research Institute, Merck, Novavax, GlaxoSmithKline, and AstraZeneca. **M.L.** serves as an advisor for Sanofi, Seqirus, Medicago and Roche; has served as an advisor for Pfizer, Sunovion, and MD Brief; and receives research funding from the Canadian Institutes of Health Research and the Medical Research Council (United Kingdom). **R.P.** receives grants from Shionogi, CD Diagnostics, Merck, Hutchison Biofilm Medical Solutions, Accelerate Diagnostics, ContraFect, and TenNor; serves as a consultant for Curetis, Specific Technologies, Next Gen Diagnostics, Pathoquest, Selux Diagnostics, 1928 Diagnostics, PhAst, and Qvella; holds patent for *B. pertussis/parapertussis* PCR, device/method for sonification, and an anti-biofilm substance; receives research funding from the NIH, the National Science Foundation and the U.S. Department of Defense; and receives monies/reimbursement from the American Society for Microbiology (ASM), the Infectious Diseases Society of America (IDSA), the National Board of Medical Examiners, UpToDate, and the Infectious Disease Board Review Course; **Y.F.Y.** receives honoraria for evidence reviews and teaching from the Evidence Foundation, honoraria for evidence reviews for the American Gastroenterological Association, and serves as a Director for the Evidence Foundation and for the U.S. GRADE Network; and **M.H.M** receives research funding from the Agency for Healthcare Research and Quality (AHRQ), the Endocrine Society, the Society for Vascular Surgery, and The American Society of Hematology and is a Board member for the Evidence Foundation. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. World Health Organization. Coronavirus disease (COVID-19) situation reports. Accessed 27 July 2020.
2. Premkumar L, Segovia-Chumbez B, Jadi R, et al. The receptor-binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Science Immunology* **2020**; 5(48): eabc8413.
3. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veasley D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**; 181(2): 281-92.e6.
4. Poh CM, Carissimo G, Wang B, et al. Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. *Nature Communications* **2020**; 11(1): 2806.
5. Jiang S, Hillyer C, Du L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol* **2020**; 41(5): 355-9.
6. Dutta NK, Mazumdar K, Gordy JT. The Nucleocapsid Protein of SARS-CoV-2: a Target for Vaccine Development. *Journal of Virology* **2020**; 94(13): e00647-20.
7. Hangartner L, Zinkernagel RM, Hangartner H. Antiviral antibody responses: the two extremes of a wide spectrum. *Nature Reviews Immunology* **2006**; 6(3): 231-43.
8. Lee CY-P, Lin RTP, Renia L, Ng LFP. Serological Approaches for COVID-19: Epidemiologic Perspective on Surveillance and Control. *Front Immunol* **2020**; 11: 879-.
9. Morgan RL, Florez I, Falavigna M, et al. Development of rapid guidelines: 3. GIN-McMaster Guideline Development Checklist extension for rapid recommendations. *Health Res Policy Syst* **2018**; 16(1): 63.
10. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. *N Engl J Med* **2020**; 382(24): 2302-15.
11. Huang G, Zeng W, Wang W, et al. Triaging patients in the outbreak of the 2019 novel coronavirus. *medRxiv* **2020**: 2020.03.13.20035212.
12. Tolia VM, Chan TC, Castillo EM. Preliminary Results of Initial Testing for Coronavirus (COVID-19) in the Emergency Department. *The western journal of emergency medicine* **2020**; 21(3): 503-6.
13. Ai J-W, Zhang H-C, Xu T, et al. Optimizing diagnostic strategy for novel coronavirus pneumonia, a multi-center study in Eastern China. **2020**: 2020.02.13.20022673.
14. Liu R, Han H, Liu F, et al. Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clin Chim Acta* **2020**; 505: 172-5.
15. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* **2011**; 64(4): 395-400.

16. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* **2005**; 58(10): 982-90.
17. Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. *Res Synth Methods* **2019**; 10(3): 476-83.
18. Stijnen T, Hamza TH, Özdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine* **2010**; 29(29): 3046-67.
19. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
20. Philipp Doebler (2020). mada: Meta-Analysis of Diagnostic Accuracy. R package version 0.5.10. <https://CRAN.R-project.org/package=mada>.
21. Balduzzi S, Rücker G, Schwarzer G (2019), How to perform a meta-analysis with R: a practical tutorial, Evidence-Based Mental Health.
22. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* **2011**; 155(8): 529-36.
23. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* **2020**; 122: 129-41.
24. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* **2020**; 122: 142-52.
25. GRADEpro GDT. GRADEpro Guideline Development Tool [Software]. McMaster University, 2015 (developed by Evidence Prime, Inc.). Available at: grade.pro.org.
26. 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**.
27. Egger M, Bundschuh C, Wiesinger K, et al. Comparison of the Elecsys(R) Anti-SARS-CoV-2 immunoassay with the EDI enzyme linked immunosorbent assays for the detection of SARS-CoV-2 antibodies in human plasma. *Clin Chim Acta* **2020**; 509: 18-21.
28. Herroelen PH, Martens GA, De Smet D, Swaerts K, Decavele A-S. Kinetics of the humoral immune response to SARS-CoV-2: comparative analytical performance of seven commercial serology tests. *medRxiv* **2020**: 2020.06.09.20124719.
29. Imai K, Tabata S, Ikeda M, et al. Clinical evaluation of an immunochromatographic IgM/IgG antibody assay and chest computed tomography for the diagnosis of COVID-19. *Journal of Clinical Virology* **2020**; 128: 104393.

30. Infantino M, Grossi V, Lari B, et al. Diagnostic accuracy of an automated chemiluminescent immunoassay for anti-SARS-CoV-2 IgM and IgG antibodies: an Italian experience. *J Med Virol* **2020**.
31. Lin D, Liu L, Zhang M, et al. Evaluations of serological test in the diagnosis of 2019 novel coronavirus (SARS-CoV-2) infections during the COVID-19 outbreak. *medRxiv* **2020**: 2020.03.27.20045153.
32. Liu W, Liu L, Kou G, et al. Evaluation of Nucleocapsid and Spike Protein-Based Enzyme-Linked Immunosorbent Assays for Detecting Antibodies against SARS-CoV-2. *Journal of clinical microbiology* **2020**; 58(6): e00461-20.
33. Liu Z-L, Liu Y, Wan L-G, et al. Antibody profiles in mild and severe cases of COVID-19. *Clin Chem* **2020**.
34. Lou B, Li TD, Zheng SF, et al. Serology characteristics of SARS-CoV-2 infection since exposure and post symptom onset. *Eur Respir J* **2020**.
35. Ma H, Zeng W, He H, et al. Serum IgA, IgM, and IgG responses in COVID-19. *Cell Mol Immunol* **2020**.
36. Minter C, Casanovas-Massana A, Li T, et al. Multi-site Validation of a SARS-CoV-2 IgG/IgM Rapid Antibody Detection Kit. *medRxiv* **2020**: 2020.05.25.20112227.
37. Ou J, Tan M, He H, et al. Study on the expression levels of antibodies against SARS-CoV-2 at different period of disease and its related factors in 192 cases of COVID-19 patients. *medRxiv* **2020**: 2020.05.22.20102525.
38. Pickering S, Betancor G, Pedro Galao R, et al. Comparative assessment of multiple COVID-19 serological technologies supports continued evaluation of point-of-care lateral flow assays in hospital and community healthcare settings. *medRxiv* **2020**: 2020.06.02.20120345.
39. Sarina Yang H, Racine-Brzostek SE, Lee WT, et al. SARS-CoV-2 antibody characterization in emergency department, hospitalized and convalescent patients by two semi-quantitative immunoassays. *Clin Chim Acta* **2020**.
40. Shen L, Wang C, Zhao J, et al. Delayed specific IgM antibody responses observed among COVID-19 patients with severe progression. *Emerging Microbes & Infections* **2020**; 9(1): 1096-101.
41. Suhandynata RT, Hoffman MA, Kelner MJ, McLawhon RW, Reed SL, Fitzgerald RL. Longitudinal Monitoring of SARS-CoV-2 IgM and IgG Seropositivity to Detect COVID-19. *J Appl Lab Med* **2020**.
42. Theel ES, Harring J, Hilgart H, Granger D. Performance Characteristics of Four High-Throughput Immunoassays for Detection of IgG Antibodies against SARS-CoV-2. *Journal of clinical microbiology* **2020**.
43. Van Elslande J, Houben E, Depypere M, et al. Diagnostic performance of seven rapid IgG/IgM antibody tests and the Euroimmun IgA/IgG ELISA in COVID-19 patients. *Clinical*

microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **2020**: S1198-743X(20)30300-1.

44. Whitman JD, Hiatt J, Mowery CT, et al. Test performance evaluation of SARS-CoV-2 serological assays. medRxiv **2020**: 2020.04.25.20074856.
45. Xiang F, Wang X, He X, et al. Antibody Detection and Dynamic Characteristics in Patients with COVID-19. Clin Infect Dis **2020**.
46. Dellièrè S, Salmona M, Minier M, et al. Evaluation of COVID-19 IgG/IgM Rapid Test from Orient Gene Biotech. Journal of clinical microbiology **2020**: JCM.01233-20.
47. Meyer B, Torriani G, Yerly S, et al. Validation of a commercially available SARS-CoV-2 serological Immunoassay. medRxiv **2020**: 2020.05.02.20080879.
48. Prazuck T, Colin M, Giache S, et al. Evaluation of performance of two SARS-CoV-2 Rapid whole-blood finger-stick IgM-IgG Combined Antibody Tests. medRxiv **2020**: 2020.05.27.20112888.
49. Shen B, Zheng Y, Zhang X, et al. Clinical evaluation of a rapid colloidal gold immunochromatography assay for SARS-Cov-2 IgM/IgG. Am J Transl Res **2020**; 12(4): 1348-54.
50. Abbott Architect SARS-CoV-2 IgG (Package insert). Available at: <https://www.fda.gov/media/137383/download>. Accessed 07/10/2020.
51. Autobio Anti-SARS-CoV-2 Rapid Test (package insert). Available at: <https://www.fda.gov/media/137367/download>. Accessed 09/07/2020.
52. Anti-SARS-CoV-2 ELISA (IgG) Instruction for use (package insert). Available at: <https://www.fda.gov/media/137609/download>. Accessed 07/10/2020.
53. Shen C, Wang Z, Zhao F, et al. Treatment of 5 Critically Ill Patients With COVID-19 With Convalescent Plasma. JAMA **2020**.
54. VITROS Immunodiagnostic Products Anti-SARS-CoV-2 IgG Instructions for use (package insert). Available at: <https://www.fda.gov/media/137363/download>. Accessed 07/10/2020.
55. Siemens Healthcare Diagnostics Dimension EXL SARS-CoV-2 Total antibody assay (CV2T) (package insert). Available at: <https://www.fda.gov/media/138757/download>. Accessed 07/10/2020.
56. Siemens Healthcare Diagnostics Dimension Vista SARS-CoV-2 Total antibody assay (COV2T) (package insert). Available at: <https://www.fda.gov/media/138753/download>. Accessed 07/10/2020.
57. Guo L, Ren L, Yang S, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2020**; 21.

58. Liu L, Liu W, Zheng Y, et al. A preliminary study on serological assay for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in 238 admitted hospital patients. *Microbes and Infection* **2020**; 22(4): 206-11.
59. Pan Y, Li X, Yang G, et al. Serological immunochromatographic approach in diagnosis with SARS-CoV-2 infected COVID-19 patients. *Journal of Infection* **2020**; 81(1): e28-e32.
60. Perez-Garcia F, Perez-Tanoira R, Romanyk J, Arroyo T, Gomez-Herruz P, Cuadros-Gonzalez J. Alltest rapid lateral flow immunoassays is reliable in diagnosing SARS-CoV-2 infection from 14 days after symptom onset: A prospective single-center study. *J Clin Virol* **2020**.
61. Xiang J, Yan M, Li H, et al. Evaluation of Enzyme-Linked Immunoassay and Colloidal Gold-Immunochromatographic Assay Kit for Detection of Novel Coronavirus (SARS-Cov-2) Causing an Outbreak of Pneumonia (COVID-19). *medRxiv*, **2020**.
62. Zhang G, Nie S, Zhang Z, Zhang Z. Longitudinal Change of SARS-Cov2 Antibodies in Patients with COVID-19. *J Infect Dis* **2020**.
63. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2020**; 28.
64. Long Q-X, Tang X-J, Shi Q-L, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nature Medicine* **2020**.
65. Belhadjer Z, Méot M, Bajolle F, et al. Acute heart failure in multisystem inflammatory syndrome in children (MIS-C) in the context of global SARS-CoV-2 pandemic. *Circulation* 0(0).
66. Dufort EM, Koumans EH, Chow EJ, et al. Multisystem Inflammatory Syndrome in Children in New York State. *New England Journal of Medicine* **2020**.
67. Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem Inflammatory Syndrome in U.S. Children and Adolescents. *New England Journal of Medicine* **2020**.
68. Grimaud M, Starck J, Levy M, et al. Acute myocarditis and multisystem inflammatory emerging disease following SARS-CoV-2 infection in critically ill children. *Ann Intensive Care* **2020**; 10(1): 69.
69. Perez-Toledo M, Faustini SE, Jossi SE, et al. Serology confirms SARS-CoV-2 infection in PCR-negative children presenting with Paediatric Inflammatory Multi-System Syndrome. *medRxiv* **2020**: 2020.06.05.20123117.
70. Riphagen S, Gomez X, Gonzalez-Martinez C, Wilkinson N, Theocharis P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet (London, England)* **2020**; 395(10237): 1607-8.
71. Toubiana J, Poirault C, Corsia A, et al. Kawasaki-like multisystem inflammatory syndrome in children during the covid-19 pandemic in Paris, France: prospective observational study. *BMJ* **2020**; 369: m2094.

72. Verdoni L, Mazza A, Gervasoni A, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *The Lancet* **2020**; 395(10239): 1771-8.
73. Whittaker E, Bamford A, Kenny J, et al. Clinical Characteristics of 58 Children With a Pediatric Inflammatory Multisystem Syndrome Temporally Associated With SARS-CoV-2. *JAMA* **2020**.
74. Black MA, Shen G, Feng X, et al. Analytical performance of lateral flow immunoassay for SARS-CoV-2 exposure screening on venous and capillary blood samples. *medRxiv* **2020**: 2020.05.13.20098426.
75. Pellanda LC, Wendland EM, McBride AJ, et al. Sensitivity and specificity of a rapid test for assessment of exposure to SARS-CoV-2 in a community-based setting in Brazil. *medRxiv* **2020**: 2020.05.06.20093476.
76. ACCELERATED EMERGENCY USE AUTHORIZATION (EUA) SUMMARY VIBRANT COVID-19 AB ASSAY ((VIBRANT AMERICA CLINICAL LABS). Available at: <https://www.fda.gov/media/138629/download>. Accessed 07/09/2020.
77. Adams ER, Ainsworth M, Anand R, et al. Antibody testing for COVID-19: A report from the National COVID Scientific Advisory Panel. *medRxiv* **2020**: 2020.04.15.20066407.
78. Candel FJ, Viñuela-Prieto JM, González Del Castillo J, et al. Utility of lateral flow tests in SARS-CoV-2 infection monitorization. *Rev Esp Quimioter* **2020**.
79. Fafi-Kremer S, Bruel T, Madec Y, et al. Serologic responses to SARS-CoV-2 infection among hospital staff with mild disease in eastern France. *medRxiv* **2020**: 2020.05.19.20101832.
80. Gao HX, Li YN, Xu ZG, et al. Detection of serum immunoglobulin M and immunoglobulin G antibodies in 2019 novel coronavirus infected patients from different stages. *Chin Med J (Engl)* **2020**; 133(12): 1479-80.
81. Lassaunière R, Frische A, Harboe ZB, et al. Evaluation of nine commercial SARS-CoV-2 immunoassays. *medRxiv* **2020**: 2020.04.09.20056325.
82. Shen B. Clinical evaluation of a rapid colloidal gold immunochromatography assay for SARS-Cov-2 IgM/IgG. *American Journal of Translational Research* **2020**; 12(4): 1348-54.
83. Tuailon E, Bolloré K, Pisoni A, et al. Detection of SARS-CoV-2 antibodies using commercial assays and seroconversion patterns in hospitalized patients. *Journal of Infection* **2020**.
84. Yongchen Z, Shen H, Wang X, et al. Different longitudinal patterns of nucleic acid and serology testing results based on disease severity of COVID-19 patients. *Emerg Microbes Infect* **2020**; 9(1): 833-6.
85. de Bruin E, Loeber JG, Meijer A, et al. Evolution of an influenza pandemic in 13 countries from 5 continents monitored by protein microarray from neonatal screening bloodspots. *Journal of Clinical Virology* **2014**; 61(1): 74-80.

86. Ni L, Ye F, Cheng M-L, et al. Detection of SARS-CoV-2-Specific Humoral and Cellular Immunity in COVID-19 Convalescent Individuals. *Immunity* **2020**; 52(6): 971-7.e3.
87. Thevarajan I, Nguyen THO, Koutsakos M, et al. Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nature Medicine* **2020**; 26(4): 453-5.
88. Li Q, Liu Q, Huang W, Li X, Wang Y. Current status on the development of pseudoviruses for enveloped viruses. *Reviews in medical virology* **2018**; 28(1).
89. Schnurra C, Reiners N, Biemann R, Kaiser T, Trawinski H, Jassey C. Comparison of the diagnostic sensitivity of SARS-CoV-2 nucleoprotein and glycoprotein-based antibody tests. *Journal of Clinical Virology* **2020**; 129: 104544.
90. Centers for Disease Control and Prevention. Interim Guidelines for COVID-19 Antibody Testing in Clinical and Public Health Settings. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html>. Accessed July 27, 2020.
91. Deeks JJ, Dinnes J, Takwoingi Y, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst Rev* **2020**; 6: Cd013652.
92. Lisboa Bastos M, Tavaziva G, Abidi SK, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *Bmj* **2020**; 370: m2516.
93. Havers FP, Reed C, Lim T, et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. *JAMA Intern Med* **2020**.
94. Menachemi N, Yiannoutsos CT, Dixon BE, et al. Population Point Prevalence of SARS-CoV-2 Infection Based on a Statewide Random Sample - Indiana, April 25-29, 2020. *MMWR Morb Mortal Wkly Rep* **2020**; 69(29): 960-4.
95. Ibarondo FJ, Fulcher JA, Goodman-Meza D, et al. Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N Engl J Med* **2020**.