AACC Practical Recommendations for Implementing and Interpreting SARS-CoV-2 EUA and LDT Serologic Testing in Clinical Laboratories

Y. Victoria Zhang1*, Joesph Wiencek2*, Qing H. Meng3, Elitza S. Theel4, Nikolina Babic5, Lusia Sepiashvili6, Nicole D. Pecora7, Patricia Slev8, Andrew Cameron9, Danijela Konforte10 on behalf of the AACC COVID 19 Serologic Testing Task Force.

*: Equal Contribution

1Y. Victoria Zhang
Department of Pathology and Lab Medicine
University of Rochester Medical Center
victoria_zhang@urmc.rochester.edu

2Joesph R. Wiencek
Department of Pathology, Microbiology, and Immunology
Vanderbilt University Medical Center
joe.wiencek@vumc.org

3Qing H. Meng
Department of Laboratory Medicine, Division of Pathology and Laboratory Medicine
The University of Texas/MD Anderson Cancer Center
QHMeng@mdanderson.org

4Elitza S. Theel
Department of Laboratory Medicine and Pathology
Mayo Clinic
theel.elitza@mayo.edu

5Nikolina Babic
Department of Pathology and Laboratory Medicine
Medical University of South Carolina
babic@musc.edu

6Lusia Sepiashvili
Departments of Biochemistry and Laboratory Medicine & Pathobiology
The Hospital for Sick Children/University of Toronto
lusia.sepiashvili@sickkids.ca

7Nicole D. Pecora
Department of Pathology and Laboratory Medicine
University of Rochester Medical Center
Nicole_Pecora@URMC.Rochester.edu

8Patricia R. Slev
Department of Pathology, University of Utah
ARUP Laboratories
patricia.slev@aruplab.com

9Andrew Cameron
Department of Clinical Microbiology
University of Rochester Medical Center
Andrew_Cameron@URMC.Rochester.edu

10Danijela Konforte
LifeLabs Medical Laboratories
Danijela.Konforte@lifelabs.com

© American Association for Clinical Chemistry 2021. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
Abstract
BACKGROUND: The clinical laboratory continues to play a critical role in managing the coronavirus pandemic. Numerous FDA emergency use authorization (EUA) and laboratory developed test (LDT) serologic assays have become available. The performance characteristics of these assays and their clinical utility continue to be defined in real-time during this pandemic. The American Association for Clinical Chemistry (AACC) convened a panel of experts from clinical chemistry, microbiology, and immunology laboratories, the in vitro diagnostics (IVD) industry, and regulatory agencies to provide practical recommendations for implementation and interpretation of these serologic tests in clinical laboratories.

CONTENT: The currently available EUA serologic tests and platforms, information on assay design, antibody classes including neutralizing antibodies, and the humoral immune responses to SARS-CoV-2 are discussed. Verification and validation of EUA and LDTs are described along with quality management approach. Four indications for serologic testing are outlined. Result interpretation, reporting comments, and the role of orthogonal testing are also recommended.

SUMMARY: This document aims to provide a comprehensive reference for laboratory professionals and healthcare workers to appropriately implement SARS-CoV-2 serologic assays in the clinical laboratory and interpret test results during this pandemic. Given the more frequent occurrence of outbreaks associated with either vector-borne or respiratory pathogens, this document will be a useful resource in planning for similar scenarios in the future.
1. Introduction

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome (SARS) coronavirus (CoV)-2 has resulted in millions of deaths worldwide and is continuing to spread at the time of this publication (3).

The Secretary of Health and Human Services (HHS) issued a public health emergency declaration for SARS-CoV-2 on January 31st, 2020 which allowed the FDA to grant emergency use authorization (EUA) of unapproved medical products or devices. While the FDA immediately required EUA for SARS-CoV-2 molecular tests, EUA was not required for serologic assays until May 4th, 2020. As of January 8, 2021, over 200 SARS-CoV-2 serologic tests are available, of which 64 have obtained EUA (4).

As a result of the limited FDA review process for EUA approval, numerous available tests, varied performance characteristics, and incomplete understanding of the humoral immune response in COVID-19, questions have arisen on how to best utilize and interpret these tests. Interim guidelines were published by several professional organizations (5–8), but no guidance to date provides comprehensive and practical recommendations for the selection, validation, implementation, and quality management of EUA or laboratory developed test (LDT) serologic tests. To provide assistance on these topics, a panel of clinical diagnostic laboratory and industry experts from AACC reviewed the current literature and developed this guidance and recommendation document.

This manuscript provides the most up-to-date understanding of host immune responses to SARS-CoV-2, the associated antibody kinetics, and the currently available EUA assays. Clinical utility and limitations are discussed to help laboratories select appropriate test(s) for their purposes and targeted population needs. The processes
and considerations to verify or validate either EUA or LDT serologic tests in a clinical setting are described. In addition, quality management, test interpretation, and orthogonal testing strategies are outlined.

2. SARS-CoV-2 and the Humoral Immune Response

2.1. Antigenic Targets

SARS-CoV-2 encodes four structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N), among which the S and N proteins are most commonly used for SARS-CoV-2 serologic assays (9–11). The S protein is divided into S1 and S2 subunits, S1 contains the receptor binding domain (RBD), which binds the human angiotensin-converting enzyme 2 (ACE2) receptor, mediating host cell entry, and S2 facilitates fusion of the viral and host membranes (11). Distal regions of the S protein (S1, RBD) are the least conserved among members of Beta-CoV (e.g., SARS-CoV, MERS-CoV) and are likely to induce a SARS-CoV-2 specific antibody response. Overall, the SARS-CoV-2 S protein shares 76% homology with SARS-CoV-1 and only about 30% homology with seasonal Beta-CoVs (e.g., OC43 and HKU1) (12).

The N protein is the most abundantly expressed immuno-dominant protein among CoVs, functioning to stabilize viral RNA (12–16). It is highly conserved between SARS-CoV-2 and SARS-CoV with approximately 90% identity (14) but shares only 33% identity with seasonal Beta-CoVs (12).

2.2. Antibody Classes

Commercial SARS-CoV-2 serologic assays are available for detection of total antibodies, specific antibody subclasses (IgG, IgM, or IgA), or neutralizing antibodies (nAbs) using qualitative or semi-quantitative methods. There is no clear evidence to support the clinical utility of standalone IgM testing (15). IgA-based assays have been
reported to suffer from lower specificity as compared to IgG-based assays (12), and are currently not recommended for use by either the Centers for Disease Control and Prevention (CDC) or the Infectious Diseases Society of America (IDSA) (6,15). Detection of total antibodies may enhance sensitivity (16–19).

The antibody response to a virus can be split into two broad categories – binding and neutralizing. While binding antibodies inactivate the virus through mechanisms such as complement activation or opsonization, neutralizing antibodies (nAbs) inhibit by binding to regions of the virus that directly interact with host cell receptors, effectively blocking viral entry and inhibiting replication. Unlike the detection of binding antibodies, the detection of nABs requires functional assays. The “gold” standard is the plaque reduction neutralization test (PRNT), which is technically challenging to perform, requires live viral and cellular culture, has a prolonged turnaround time (days to weeks), and for SARS-CoV-2, requires biosafety level (BSL) 3 facilities.

To overcome these challenges, alternative methods have been developed, including pseudovirus-based live-cell neutralization assays or blockade-of-binding (BoB) immunoassays. Pseudovirus neutralization assays can be performed at BSL2 (20), though these assays are still complex, associated with significant analytical variability and challenging to support in most clinical laboratories. BoB immunoassays can be performed in a 96-well format and can be automated on different immunoassay processing platforms for high throughput analysis (16).

nAb assays have played an important role in the development and assessment of SARS-CoV-2 vaccines and in research studies probing the host immune response to infection (21). Given the challenges associated with assay maintenance, lack of
standardization, and the currently unknown correlation of nAb titers with protective immunity, their role in the clinical laboratory will likely be limited.

2.3. Antibody Kinetics

Understanding the kinetics of the antibody response to SARS-CoV-2 is a prerequisite for test selection and accurate result interpretation. The current understanding of the kinetics of the antibody responses against SARS-CoV-2 are depicted in Fig. 1. Of note, antibody kinetics in specific sub-populations, including immunosuppressed patients, cancer patients, and other sub-groups, may differ and continue to be studied.

Unlike viral RNA and antigens, detection of antibodies during the incubation phase is unlikely. Multiple published studies demonstrate that most individuals develop an IgM/IgA/IgG response within 7-14 days of symptom onset, with over 90% of individuals seropositive after three weeks (22,23). IgM/IgA peak and decline earlier than IgG, often within weeks of symptom onset (24–27). IgG antibodies correlate with disease severity, decline at varying rates, and may be detectable for months following infection (28–35). Notably, approximately 4-10% of the population with confirmed SARS-CoV-2 infection may have either an undetectable or delayed antibody response (36). Regarding antibody longevity, some studies indicate that up to 40% of confirmed individuals become IgG seronegative by the early convalescent phase (37) while others have demonstrated that antibodies decline, yet remain detectable for months post-infection (36,38,39). Given these inconsistencies, the precise kinetics of the SARS-CoV-2 antibody response requires further elucidation.
3. EUA Serologic Tests

3.1. Assay Designs

Several assay formats for detection of SARS-CoV-2 antibodies have received EUA. Lateral flow assays (LFAs) utilize immunochromatographic chemistry to detect antibodies, usually at the point-of-care. Manual or semi-automated 96-well enzyme-linked immunosorbent assays (ELISAs) are also available (40,41), as well as chemiluminescent immunoassays/chemiluminescent microparticle immunoassays (CIAs/CMIAs) for fully automated, high-throughput platforms. These methods are illustrated in Fig. 2.

3.2. Characteristics of EUA Serologic Tests

As of January 8, 2021, 64 assays have received FDA EUA. Selected examples are listed in online Supplemental Table 1. The majority detect IgG, followed by IgM/IgG, total antibody, and IgM-only. All EUA assays use serum, some accept plasma, and less frequently, whole-blood or dried blood spots. Currently, serologic testing is not recommended for other sample types such as saliva and cerebrospinal fluid. The most frequent antigen targeted in these assays is the RBD, followed by S (including full S, S1, and S2), and N. Currently, only one assay uses all three antigens. Most current EUA assays are qualitative with a few being semi-quantitative.

Assessing the relative performance characteristics of each EUA assay is complicated, as the approach, the sample size, sample collection time, and disease prevalence in the population tested by each manufacturer vary widely. Clinical laboratory professionals should take these variables into consideration when evaluating assay performance.
4. Utility and Limitations of SARS-CoV-2 Serology

SARS-CoV-2 serologic testing is not recommended as the primary approach for diagnosis of SARS-CoV-2 infection. However, it can be used for: 1) supportive diagnosis of COVID-19, 2) manufacture of convalescent plasma, 3) epidemiologic and seroprevalence studies, and 4) vaccine response and efficacy studies (5,6,15) (Table 1).

4.1. Supporting Diagnosis of COVID-19

Serologic testing may be helpful to diagnose COVID-19 in symptomatic patients presenting later in disease (e.g., >9-14 days post symptom onset), who test negative by a molecular assay, with optimal assay sensitivity occurring at least 2-3 weeks post symptom onset (42,43). Total antibody or IgG testing may be more useful for evaluating patients presenting later in the disease course (18,43–47).

Serologic testing, alongside RT-PCR, has been recommended to support the diagnosis of multisystem inflammatory syndrome in children (MIS-C), including for hospitalized individuals <21 years presenting with fever, inflammation, and multi-system organ involvement following exclusion of other potential diagnoses (48–51). Serologic testing should precede intravenous immunoglobulin or blood product administration as these therapies may influence serologic results.

4.2. Convalescent Plasma Donor Identification and Manufacturing

Identification of potential convalescent plasma (CP) donors for COVID-19 CP therapy, which has received FDA EUA, is a recognized application of serologic testing. The FDA continues to refine donor eligibility criteria, identify serologic assays for the manufacture of COVID-19 CP units, and define acceptable antibody thresholds. Originally, the FDA recommended that the qualitative Ortho Clinical Diagnostics SARS-
CoV-2 IgG CIA be used in the manufacturing of CP, with signal/cutoff (S/CO) threshold values ≥12 considered “high titer” and preferred for infusion (52). Given that most currently available serologic assays are qualitative, there are limited mechanisms for distinguishing donors with high versus low titers (52). Recently, the FDA updated their COVID-19 CP EUA to include nine serologic assays for manufacture of CP, including two semi-quantitative assays (53).

**4.3. Epidemiologic and Seroprevalence Studies**

Determination of seroprevalence is important to characterize the epidemiology of COVID-19 in the community and support public health efforts (36,40,41,54). However, serologic assays have limitations that may lead to an underestimate of the true seroprevalence. First, most commercial assays were developed using symptomatic patients with moderate to severe disease. It is unknown whether the cutoffs based on these populations will detect antibodies in asymptomatic or mild disease cases. Second, a small proportion of the population may never develop detectable antibodies following infection. Third, the accuracy of this approach is dependent on the prevalence of the disease in the community, as the positive predictive value may be low in regions with little disease, even if using a highly specific assay (27,55).

**4.4. Vaccine Response and Efficacy**

Available and developing vaccines range from inactivated or live platforms to more novel DNA or RNA based preparations, such as the two recently authorized vaccines in the US: Moderna and Pfizer/BioNTech (56). Because the primary target of neutralizing antibodies is the S protein, the majority of the vaccines target the S protein (57–60). Vaccine trials have assessed vaccine efficacy by using endpoint outcome measures such as prevention of moderate or severe disease due to SARS-CoV-2.
infection in the placebo vs. vaccinated populations. Vaccine trials have also used several different approaches to assess vaccine response, including binding antibody ELISAs, and both PRNT and pseudovirus-based neutralization assays to determine that the majority of vaccinated individuals developed a robust antibody response, including neutralizing antibodies (13,14,55,56,58–61). To date, only one assay has received EUA for detection of nAbs. It is important to note that, although a detectable antibody response in a vaccinated individual (including immunosuppressed persons) indicates that an antibody response has developed in response to vaccination, there is no threshold on any assay that is indicative of vaccine efficacy. Therefore, at this point in time, even semi-quantitative or quantitative assays against S protein that can quantify the magnitude of the antibody response to vaccines should not be used to determine vaccine efficacy and protective immunity. This is true not only for binding antibody ELISAs but also for neutralization assays. Currently, there are no recommendations from any professional societies in the US for monitoring or assessing vaccine response in any population, including immunosuppressed individuals.

As the S/RBD protein is primarily used for vaccines, the availability of antibody assays that detect N- versus S-specific antibodies may also be useful to distinguish between naturally infected versus vaccinated individuals but further studies are needed to understand the merits and limitations of this approach.

5. Performance Verification of EUA Assays

Verification studies for non-waived EUA assays are the same as those for FDA-approved/cleared assays. However, waived EUA tests should be verified in a similar manner as moderately complex, non-waived tests. Further resources for detailed
method verification protocols are available through the Clinical & Laboratory Standards Institute (CLSI) (online **Supplemental Table 2**).

5.1. Regulatory and Accreditation Requirements

Clinical laboratories in the United States are required by CLIA to verify assay performance of unmodified, FDA-approved/cleared and EUA assays, and must adhere to manufacturer instructions. Several accreditation organizations are available; labs should refer to the specific requirements. The College of American Pathology (CAP) is used as an example to discuss some specific requirements for EUA verification.

1. Ensure testing personnel are properly trained and qualified based on test complexity authorized by the FDA;
2. Perform testing as outlined in the EUA without modification;
   a. Any deviation from instructions for use will render the assay an LDT, which needs to be validated (Section 6).
3. Verify test method performance following the CAP’s All Common Checklist:
   a. COM.40300 - The laboratory must assess analytical accuracy, analytical precision, and reportable range (as appropriate).
   b. COM.40475 - Laboratory director must sign the laboratory’s written assay assessment.
   c. COM.40500 – Laboratory understands analytical interferences for each test and has a plan of action when present.
4. Update the laboratory’s activity menu.

Here, we consolidate and expand on prior recommendations to provide a systematic approach for EUA assay verification (62–64).
5.2. Sample Collection

Sample type, target population (e.g., symptomatic, asymptomatic, ambulatory, hospitalized, pediatric, pregnant patients), and number of positive samples may be difficult to discern early on in a public health emergency. Below are recommended strategies.

5.2.1. Positive Samples

1. Residual, unmodified patient samples (preferred) collected after testing positive on a comparative EUA assay. Comparator assays should be matched to sample matrix, antibody class(es), and antigenic targets for optimal evaluation. If a comparator EUA assay is unavailable, samples collected from RT-PCR-confirmed patients can be used, with knowledge of days post symptom onset or first RT-PCR-positive result.

2. Residual, positive samples with an increased S/CO may be mixed, at different ratios, with one or more confirmed negative patient samples to generate a range of S/CO positive samples.

3. Commercially verified materials (e.g., positive QC, patient or pooled patient samples) may be used in an emergent situation if residual, positive patient samples are not available. However, judicious selection of third-party materials must be performed to mitigate possible matrix effects.

4. Sample and antibody stabilities should be considered. Laboratories should use manufacturers’ package inserts as a guide and can also validate alternative stability timeframes. Limited studies are available in convalescent plasma settings (65–67).

5.2.2. Negative Samples

1. Residual, unmodified pre-pandemic patient samples collected and properly stored.
2. Residual, unmodified patient samples (matched to sample matrix, antibody class(es), and antigenic targets) collected after testing negative on a comparative EUA assay.

3. Commercially verified materials may be used if the other samples are limited or not available.

5.3. Accuracy

Accuracy is verified by assessing the result concordance with either another EUA assay or clinical correlate, reflecting assay clinical sensitivity and specificity. Below are recommendations for accuracy assessments.

5.3.1. Single Analyte or Total Analyte

A minimum of 10 negative and 10 positive samples per sample type should be used. For total antibody tests, it is optimal to use known positive patient samples from each antibody class.

5.3.2. Multiple Differentiated Analytes

Accuracy verification must be demonstrated with known positive samples for each antibody class that could be reported. Combinations of a minimum of 20 samples (e.g., IgM-/IgG+, IgM+/IgG-, IgM+/IgG+, IgM-/IgG-) should be used to assess class specific positive and negative agreement and to verify clinical specificity. For SARS-CoV-2, it may be challenging to identify IgM+/IgG- samples due to concurrent seroconversion.

5.4. Precision

The reproducibility and repeatability of an EUA assay around the positive cutoff must be verified. For qualitative assays, a positive and negative sample can be used, with the positive sample near the cutoff. Semi-quantitative assays should be evaluated
as quantitative assays, and samples should span low, mid, and high S/CO, with at least one sample near the cutoff. The intra-day and inter-day precision experiments should test both positive and negative samples over 10 replicates on the same day or over 10 runs on a minimum of five days and over multiple shifts, respectively. Precision for single use LFAs should be assessed for inter-day only over five days with multiple testing operators.

5.5. Reportable Range

For semi-quantitative or quantitative serologic assays with EUA, the reportable range must be verified. Verification should be done by using non-diluted, known standards of anti-SARS-CoV-2 antibodies, such as the recently available standard from the World Health Organization (68), or if unavailable, an alternate calibrator lot or patient samples that span the analytical measuring range. Future standardization of quantitative assays to a single international standard will be essential for accurate assessment of antibody levels once a protective immunity threshold is established.

6. Validation of Laboratory Developed Tests (LDTs)

An in vitro diagnostic test that is designed and used in a clinical setting by a single laboratory is considered an LDT. Clinical laboratories authorized to perform high-complexity testing under CLIA must perform thorough LDT validation studies before patient testing. This section will discuss minimum validation requirements of LDTs that go beyond those necessary for EUA assay verification with respect to sensitivity and specificity, the establishment of assay result cutoffs, class specificity, and carryover.

6.1. Regulatory and Accreditation Requirements
Typically, following HHS declaration of a public health emergency, any clinical test used to diagnose that condition, regardless of type (i.e., molecular or serologic or antigen), requires EUA. On August 19th, 2020, the EUA requirement for COVID-19 laboratory assays was removed to ease regulatory burdens placed on high-complexity CLIA laboratories capable of developing LDTs (69). The CAP and other accreditation organizations provide specific requirements for clinical laboratories that must be used in the implementation of LDTs (70).

6.2. Sample Collection

Generally, LDTs require additional samples to establish assay performance as compared to EUA assays that require verification only. The FDA recommends at least 30 positive and 75 antibody negative (or pre-COVID-19) samples in their guidance for EUA applications (4). In situations where an assay using 75 negative specimens does not demonstrate greater than 95% specificity, or if 75 specimens are not available, the FDA recommends specific cross-reactivity studies with samples known to be positive for a variety of potentially cross-reactive antibodies or those directed against other respiratory pathogens. It is also our recommendation to collect at least 30 positive (with known days from symptom onset) and 75 negative samples (ideally 100-200).

6.3. Analytical Sensitivity and Specificity

6.3.1 Sensitivity

Assay sensitivity can be evaluated with well characterized RT-PCR-positive samples, ideally with chart data that indicate the days from a patient’s symptom onset. In this way, assay sensitivity as a function of time can be assessed. It is also valuable to compare performance with another EUA assay, if available.
6.3.2 Specificity, Cross-Reactivity, and Interfering Substances

Samples from patients with known acute respiratory infections should be included for any LDT assay assessing the serologic response to SARS-CoV-2. Ideally, these would include samples from patients infected with one of the circulating seasonal human CoVs (NL63, OC43, HKU1, 229E), although data indicating that these are not a source of significant cross-reactivity on SARS-CoV-2 serologic tests have begun to accumulate. A disclaimer to the effect that cross-reactivity cannot be ruled out should be included if such samples were not evaluated in the validation (72,73). Additionally, samples from those diagnosed with other infectious and autoimmune conditions known to give false positive results in immunoassays (e.g., syphilis, Lyme disease, cytomegalovirus, rheumatoid arthritis, etc.) should be included.

Investigation of other interfering substances is another core component to determine an assay’s analytical specificity (e.g., hemoglobin, lipids, bilirubin). For laboratories validating an LDT, it is necessary to investigate potential interferences based on assay design and devise interference validation studies.

6.4. Establishing Assay Cutoffs

Cutoffs for a qualitative/semi-quantitative LDT can be established with limit of blank studies using known negative samples tested repeatedly over several runs (e.g., 20 known negative samples tested by multiple operators on five separate runs). The mean optical density (OD) (or equivalent readout) and standard deviations from the mean should be calculated, with the assay threshold determined as the mean readout plus 3 to 5 times the standard deviations (SD). Further refinement of cutoffs can be performed using Receiver-Operating Characteristic (ROC) analysis to optimize
sensitivity and specificity. Alternatively, if risk assessment dictates an overriding concern, then cutoffs can be set accordingly (e.g., for 100% specificity).

Assays that report quantitative results, as well as those that indicate neutralization levels, are less commonly used in clinical laboratories, and require additional layers of validation. Once a cutoff is established, it is also recommended to verify the cutoff as required by the respective accreditation agencies.

6.5. Antibody Class Specificity

If a claim about antibody class specificity is made for an LDT, it must be validated. Methods for this include the use of a detection or capture antibody with a known class specificity or class-specific antibody depletion of the sample. As an alternative, the FDA recommends treating samples with dithiothreitol (DTT) (74) which effectively removes IgM-class antibodies (4).

6.6. Carryover

Clinical laboratories should perform an assessment to verify that a positive result was not due to positivity from a nearby high-titer positive sample (e.g., for probe-based instruments). This is commonly performed by alternating testing of a negative sample before and after a positive sample with a high index or S/CO value. If carryover cannot be eliminated from the assay, it is recommended to assess the impact on accuracy of a positive and negative result. Carryover should not exceed 20% of the lower limit of quantitation according to the FDA and additional details are available through CLSI EP10-A3-AMD (online Supplemental Table 2) and CAP.
7. Other EUA Assay Implementation Considerations
7.1. Quality Management

QC must be identified, verified, and implemented for routine SARS-CoV-2 testing based on test complexity and manufacturer’s instructions. A minimum of two levels of quality control (positive and negative) should be included with each run of the specified assay. For qualitative and semi-quantitative assays, a negative QC and a positive QC near the cutoff must be run at least daily. For SARS-CoV-2 serologic assays, controls may be provided as part of the assay kit or may need to be sourced separately. For the latter, laboratories can purchase separate controls provided by the assay manufacturer or third-party vendors or use pooled patient samples. Use of assay calibrator material to create assay controls is discouraged, but if needed, the calibrator material must be from a different kit lot. QC material should also match the analyte detected by the specific assay and patient matrix.

Typically, 20 QC data points on separate days are used to determine the target control mean and SD to establish the range. For vendor material with assigned QC ranges, the laboratory should verify the product. QC performance should be monitored in real-time to identify shifts and trends.

Laboratories should participate in proficiency testing (PT) either using vendor products or an alternative assessment program. Finally, because EUA assays were not extensively evaluated, laboratories may implement a more rigorous quality management system until assay reliability is established. This may include analysis of additional QC material, performing additional lot-to-lot comparisons, and identifying a partner laboratory for more frequent sample exchanges than the bi-annual PT requirement.
7.2. Pre-analytical Considerations

Pre-analytical variables should be noted for SARS-CoV-2 antibody tests and thoroughly reviewed to determine possible limitations. These include sampling time, sample stability, as well as potential endogenous and exogenous interferences (e.g., hemoglobin). Time of sample collection is important when selecting positive samples for verification studies. In order to verify test performance at the optimal reported sensitivity for most current EUA assays, samples collected $\geq 14$ days post symptom onset / PCR positivity should be used for verification. The limitation of test performance in patients tested $<14$ days prior to symptom onset/ PCR positivity should be clearly stated. For LDTs, clinical sensitivity relative to days from symptom onset needs to be determined during the validation (Section 6).

If an assay is performed with several sample types, including dried blood spots, laboratories should define and specify collection device, transportation, and pre-analytical requirements prior to patient testing.

8. Interpretation of Serologic Test Results

The majority of SARS-CoV-2 serologic assays are qualitative in design and generally, positive results indicate recent or prior SARS-CoV-2 infection. Negative results indicate that SARS-CoV-2 antibodies were not present or are below defined detection limits.

Negative results cannot rule out active or prior infection. Results should be interpreted in the context of antibody class(es) detected (Section 4.1) and antigenic target(s), time of sample collection (Section 2.3), disease severity, and assay analytical
performance characteristics (Section 6.3). Understanding the clinical sensitivity, clinical specificity, and disease prevalence are also key considerations for interpretation of serologic test results.

8.1. Impact of Clinical Sensitivity, Specificity and Disease Prevalence

To minimize potential false positives and to be of clinical value, the CDC and IDSA have suggested using tests with clinical sensitivity and specificity of 99.5% or greater.

Positive (or negative) predictive values (PPV/NPV) depend on disease prevalence in the target population and on assay clinical sensitivity and specificity. They indicate the percent probability that a positive (or negative) test result will correctly identify individuals with (or without) antibodies in a given population. An assay with 95% sensitivity and 90% to 99% specificity was used to illustrate this relationship (Fig. 3A).

The PPV increases as specificity increases. Using these sensitivity and specificity values, the PPV increases very rapidly with increased disease prevalence until it plateaus at ≥20% prevalence. The NPV, however, changes minimally with different levels of assay specificity and drops markedly when disease prevalence increases.

A test that has 95% sensitivity and 95% specificity within a population of 20% or 5% antibody prevalence (2000 or 500 individuals have antibodies assuming a population of 10,000, respectively) is used as an example to show the impact of disease prevalence on PPV/NPV (Fig. 3B). In a population with 20% prevalence, the test would correctly detect 1,900 of the 2000 positive individuals resulting in 400 false positive results. With 5% prevalence, the test would correctly identify 475 positive individuals
resulting in 475 false positive results. The PPVs are 82% and 50%, respectively, for these two populations.

8.2. Results Reporting

Clear and concise comments are needed to aid in interpretation of serologic test results. It is advisable for clinical laboratories to include a statement that diagnosis of COVID-19 should be performed using molecular tests. Per FDA EUA requirements, reports should include the assay name, and clinicians and patients should have access to the respective assay ‘Fact Sheets.’

The qualitative nature of most EUA tests prohibits reporting of actual S/CO values due to the following: 1) the assay-specific values are not standardized and do not indicate an actual concentration of SARS-CoV-2 antibodies; 2) currently no S/CO cutoffs are available to correlate with protective immunity and 3) different dynamic ranges are available for current platforms. Some examples of result comments are provided in online Supplemental Table 3.

10. Orthogonal Testing

If a desired PPV cannot be achieved using a single assay, the CDC recommends use of an orthogonal testing algorithm (OTA), a two-step testing strategy where all initially positive results are tested with a second independent serologic test (5). Studies on the effectiveness of this approach are still scarce (34,75).

10.1. OTA Test Selection

Both tests should ideally have high sensitivities (>90%, ideally > 95%). The test with higher specificity should be selected as the first-line test to minimize the number of
discordant results while still retaining optimal PPV. OTA may include tests that use different methods/same antigenic target, same methods/same antigenic target but different domains or methods that detect antibodies against different antigenic targets (34). OTAs incorporating IgM or IgA serologic tests are not recommended as there is a higher likelihood of discordant results.

The relationship of PPV/NPV and discordant rate in different OTA designs is illustrated in Table 2 with a 2% disease prevalence. If evaluated by a single test (test 1) with specificity of 98%, about 50% false positive results would be expected. Adding a sequential test (test 2) with specificity of 95% will, however, result in PPV of >90% (Fig. 4). OTA1 and OTA2 represents different designs from test 1 and test 2. Both OTA1 and OTA2 have the same combined PPV of 95%. However, OTA1, which uses the highest specificity test first, results in a lower discordant rate without affecting combined PPV/NPV.

10.2. OTA Result Reporting and Interpretation

If the initial result is negative, the second test is not needed and a negative report is issued; if both tests are positive, a positive report is issued. The interpretive challenge arises when the first result is positive and second is negative, granting discordant or indeterminate results (online Supplemental Table 4). In this case, OTA results should be interpreted in the context of disease prevalence, sensitivity and specificity of each test, assay methodology, and antigenic targets. If different antigenic targets are used, a discordant result may be attributed to 1) initial false-positive, 2) early recovery and/or differences in antibody kinetics, 3) skewed immune response towards one antigen, or 4)
waning immunity. To rule out the contribution of differences in antibody kinetics, retesting in 2-4 weeks may be attempted.

11. Conclusions and Perspectives

Though not recommended as first-line testing for diagnosis of SARS-CoV-2, serologic testing can play important functions in the management of SARS-CoV-2 infection and the pandemic. It is AACC’s position that clinical laboratories should only use EUA assays or LDTs that have been developed and properly verified or validated in the clinical laboratory, respectively. Laboratorians should recognize the utility and limitations of serologic tests, and carefully select and implement EUA or LDT assays and interpret test results. The authors have provided expert opinions and practical recommendations based on the current research on these topics for the clinical laboratory community.

Many studies are underway to gain deeper and broader understanding of SARS-CoV-2 and the tests used to detect and manage the infection will continue to improve. Clinical laboratory professionals, in collaboration with their clinical colleagues, will continue to play an indispensable role in reviewing the evolving scientific literature and adjusting testing strategies to best serve patient and public health needs during this pandemic.
Acknowledgments

The authors would like to thank other members of the Task Force including Drs. David D. Koch, Stacy Melanson, Hubert Vesper, Ted Schutzbank, Caitlin Ondracek, Gyorgy Abel, and Khosrow Adeli and thank Drs. Sara Brenner, Brittany Schuck, Courtney H Lias, and Toby A Lowe from the Center for Devices and Radiological Health at the FDA for discussions about the FDA EUA regulations, Drs. Roger L. Bertholf and David Grenache for their guidance and support from the AACC Academy and AACC executive leadership, and Dr. Sol Green from BD Diagnostics for comments from the IVD industry perspective. Special thanks to Dr. Caitlin Ondracek for her administrative support for this project, organizing the references, and proofreading the manuscript.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: J.R. Wiencek, AACC.
Consultant or Advisory Role: E.S. Theel, Roche Diagnostics, Accelerate Diagnostics.
Stock Ownership: None declared.
Honoraria: N.D. Pecora, Elsevier; P.R. Slev, MD Anderson.
Research Funding: N.D. Pecora, Luminex.
Expert Testimony: None declared.
Patents: None declared.
References


Table 1. Recommended Use of Serologic Testing and Limitations

<table>
<thead>
<tr>
<th>Recommended use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serologic testing may be offered as an approach to support diagnosis of</td>
</tr>
<tr>
<td>COVID-19 illness in symptomatic patients and late phase negative molecular</td>
</tr>
<tr>
<td>testing or for patients presenting with late complications such as multisystem</td>
</tr>
<tr>
<td>inflammatory syndrome in children (MIC-C).</td>
</tr>
<tr>
<td>Serologic testing can help identify people who may have been infected with or</td>
</tr>
<tr>
<td>have recovered from the SARS-CoV-2 infection.</td>
</tr>
<tr>
<td>Serologic testing can be used to screen potential convalescent plasma donors</td>
</tr>
<tr>
<td>and in the manufacture of convalescent plasma.</td>
</tr>
<tr>
<td>Serologic testing can be used for epidemiology and seroprevalence studies.</td>
</tr>
<tr>
<td>Serologic testing can be used for vaccine response and efficacy studies.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>False positive results may occur.</td>
</tr>
<tr>
<td>Negative results do not preclude acute SARS-CoV-2 infection or viral shedding.</td>
</tr>
<tr>
<td>Serologic tests may not differentiate between natural infection and vaccine response.</td>
</tr>
</tbody>
</table>
The durability and kinetics of the humoral immune response continue to be elucidated.

Serologic results should not be used for
- Determining individual protective immunity
- Return to work decisions
- Cohorting individuals in congregate settings
- Assessment of convalescent plasma recipients
- Use of Personal Protective Equipment
- Placement of high-risk job functions
<table>
<thead>
<tr>
<th></th>
<th>Test 1 Sensitivity</th>
<th>TEST 1 Specificity</th>
<th>TEST 2 Sensitivity</th>
<th>TEST 2 Specificity</th>
<th>% Initial Positive (TEST 1)</th>
<th>% Discordant</th>
<th>TEST 1 PPV</th>
<th>T1 + T2 PPV</th>
<th>T1 + T2 NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTA 1</td>
<td>95%</td>
<td>98%</td>
<td>99%</td>
<td>95%</td>
<td>3.9%</td>
<td>1.9%</td>
<td>49.2%</td>
<td>95.0%</td>
<td>99.9%</td>
</tr>
<tr>
<td>OTA 2</td>
<td>99%</td>
<td>95%</td>
<td>95%</td>
<td>98%</td>
<td>6.9%</td>
<td>4.9%</td>
<td>28.8%</td>
<td>95.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

T1: Test 1 with sensitivity of 95% and specificity of 98%
T1: Test 2 with sensitivity of 99% and specificity of 95%

PPV/NPV, Positive/Negative Predictive Value; OTA, Orthogonal Testing Algorithm

An online calculator from the FDA (https://www.fda.gov/media/137612/download) is a helpful tool to assess the combined PPV/NPV.
Figure Legends

**Figure 1.** Generalized kinetic and dynamic model of antibody responses to SARS-CoV-2 with the expected positivity of qualitative and semi-quantitative assays.
Figure 2. General strategies for the serologic detection of SARS-CoV-2 antibodies. Common assay formats include lateral flow assays (LFAs), fluorescent microsphere immunoassays (FMIA), chemiluminescent immunoassays (CIA), enzyme-linked fluorescent assays (ELFA), and enzyme-linked immunoassays (ELISA). **LFAs:** Specific antibodies (e.g., IgG) to SARS-CoV-2 present in the added sample flow through the membrane, bind antigen-gold conjugated nanoparticles, and are captured by antibodies immobilized in the ‘test line’. This leads to the generation of a colored band that the user visually interprets. **ELISA:** Specific antibodies (e.g., IgG) to SARS-CoV-2 are captured by antigens immobilized on coated wells or other solid phase surfaces. Detection is the result of enzymatic production of a chromogen, chemiluminescence, or fluorescence, which may be measured using spectrophotometry, luminometry, or fluorometry, respectively. **FMIA and CIA/CMIA:** Specific antibodies (e.g., IgG) to SARS-CoV-2 are captured by antigens immobilized on paramagnetic microparticles or similar technology. Detection strategies often involve fluorescent or chemiluminescent antibody conjugates.
Figure 3. The relationship between assay sensitivity, specificity, disease prevalence, and positive and negative predictive values (PPV/NPV). (A) PPV, the proportion of true positives, is strongly influenced by the specificity of an assay when a disease is low prevalence. NPV, the proportion of true negatives, declines as disease prevalence increases. (B) A visual comparison of PPV for the same assay (sensitivity, 95%; specificity, 95%) performed in two test populations of 10,000 people with high (20%, PPV = 82.6%) and low (5%, PPV = 50.0%) disease prevalence (upper and lower panels).
Figure 4. Effect of an Orthogonal Testing Algorithm (OTA) on Positive Predictive Value (PPV). Shown is the performance of two tests with different characteristics (Test 1: sensitivity, 95%; specificity, 98%. Test 2: sensitivity, 99%; specificity, 95%) and their combined performance. Regardless of the order in which the tests are performed, sequential testing can increase PPV in testing populations with low disease prevalence.