A roadmap to better COVID-19 testing from the Coronavirus Standards Working Group

Tim Mercer
The University of Queensland
Neil Almond
National Institute for Biological Standards and Control
Patrick Chain
Los Alamos National Laboratory https://orcid.org/0000-0003-3949-3634
Michael Crone
Section of Structural and Synthetic Biology, Department of Infectious Disease, Imperial College London
https://orcid.org/0000-0002-5675-0172
Alina Deshpande
Los Alamos National Laboratory
Deepa Eveleigh
Asuragen, Inc.
Paul Freemont
Imperial College London https://orcid.org/0000-0002-5658-8486
Sebastien Fuchs
Western Univeristy of Health Sciences, College of Osteopathic Medicine of the Pacific
Russell Garlick
LGC Clinical Diagnostics
Jim Huggett
National Measurement Laboratory, LGC
Martin Kammel
INSTAND https://orcid.org/0000-0002-2011-086X
Po-E Li
Los Alamos National Laboratory
Mojca Milavec
National Institute of Biology https://orcid.org/0000-0002-5794-2109
Elizabeth Marlowe
Quest Diagnostic Infectious Disease
Denise O'Sullivan
National Measurement Laboratory, LGC
Mark Page
National Institute for Biological Standards and Control

Gary Pestano
Biodesix, Inc.
Sara Suliman
Zuckerberg San Francisco General Hospital, Division of Experimental Medicine, University of California San Francisco https://orcid.org/0000-0002-5154-576X
Birgitte Simen
Ginkgo Bioworks, Inc.
John Sninsky
SPARK Stanford University
Lynne Sopchak
SPARK Stanford University
Cristina Tato
CZ Biohub https://orcid.org/0000-0002-0614-756X
Jo Vandesompele
Biogazelle
Peter Vallone
National Institute of Standards and Technology https://orcid.org/0000-0002-8019-6204
Thomas White
Human Rights Center, School of Law, University of California, Berkeley
Heinz Zeichhardt
INSTAND e.V., Duesseldorf, and IQVD GmbH Institut fuer Qualitaetssicherung in der Virusdiagnostik
Marc Salit (msalit@stanford.edu)
Stanford University
no mistake
none

Article

Keywords: COVID-19, SARS-CoV-2 testing, standards

Posted Date: November 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1066221/v1

Version of Record: A version of this preprint was published at Nature Biotechnology on November 2nd, 2022. See the published version at https://doi.org/10.1038/s41587-022-01538-1.

Abstract

Testing has been central to our response to the COVID-19 pandemic. However, the accuracy of testing relies on standards, including reference materials, proficiency testing schemes, and information and reporting guidelines. The use of standards is a simple, inexpensive, and effective method to ensure reliable test results that inform clinical and public health decisions. Here we describe the central role of standards during the COVID-19 pandemic, where they have enabled population-scale screening, genomic surveillance and measures of immune protection measures. Given these benefits, the Coronavirus Standards Working Group (CSWG) was formed to coordinate standards in SARS-CoV-2 testing. This network of scientists has developed best-practices, reference materials, and conducted proficiency studies to harmonize laboratory performance. We propose that this coordinated development of standards should be prioritized as a key early step in the public health response to future pandemics that is necessary for reliable, large-scale testing for infectious disease.

Introduction

The scale and diversity of testing performed in response to the COVID-19 pandemic is unprecedented. Testing has been needed to identify infectious individuals, diagnose patients, and measure the immune protection elicited by vaccines[1]. The pandemic has prompted innovations in molecular, antigen, and serology testing, and the global deployment of genomic surveillance. Given these key roles, SARS-CoV-2 testing has been established at unprecedented urgency and scale, and will likely remain established in future clinical diagnostics and public health testing.

Standards are required to ensure that testing is accurate and reliable. Standards include well-characterized reference materials that ensure a test is calibrated and fit-for-purpose, proficiency testing schemes that evaluate laboratory performance, and information standards for clear communication of test performance and results. Together, these standards underpin reliable and robust testing, however, despite their importance the development and implementation of SARS-CoV-2 standards has received little attention during the pandemic[2].

This importance of standards was recognized early during the pandemic by an *ad hoc* consortium of concerned scientists that formed the *Coronavirus Standards Working Group* (CSWG)[2]. This diverse group of scientists fulfilled differing roles during the pandemic, and represented a range of commercial, government, academic and other organizations. Together, this group has advocated for the importance of standards in SARS-CoV-2 testing, and led the development of reference materials, conducted proficiency studies, and critical consideration of testing methods.

This article describes the collective expertise, experiences and recommendations of the CSWG during the COVID-19 pandemic. The CSWG proposes that the development and dissemination of standards is a cost-effective strategy that can broadly improve SARS-CoV-2 testing worldwide, and is needed to address the evolving needs of testing and preparation for future public-health emergencies. The development and

implementation of standards should be prioritized within the pandemic response, and recognized as the necessary foundation for the development of a robust and reliable testing enterprise.

Main Text

The SARS CoV-2 testing process

Testing measures the presence of an analyte in a sample, such as the presence of the SARS-CoV-2 RNA genome within a respiratory sample. However, like any measurement, there is uncertainty in this process, and standards are needed to estimate and manage this uncertainty. Testing can also involve a quantitative measurement to determine the abundance of an analyte in a sample. However, viral load is rarely considered during SARS-CoV-2 testing, and quantitative measurements may be reduced to a simple presence or absence result when viral abundance exceeds a threshold[3].

The testing process can be usefully divided into three steps; the pre-analytical phase which includes preparing the sample specimen, the analytical phase which includes the actual testing for the analyte in the sample, and the post-analytical phase, which includes interpreting the data for an actionable result (see **Figure 1**). Testing performance, including the **sensitivity** (see **Glossary**) and **specificity** for detecting the analyte, is influenced by the performance of all these steps. While many manufacturer studies may focus on test performance during the analytical step, in practice, variables in the pre-analytical phase, such as sample source and amount, collection, transport and storage often markedly impact test performance and can be the least-controllable steps in the testing process[4–6].

Reference materials for SARS CoV-2 testing

Standards include well-characterized reference materials with known properties that can be used to evaluate whether tests are fit-for-purpose (see **Figure 1**). **Primary reference standards** are issued by a recognized authority, and are assigned qualitative and quantitative properties without reference to other standards. For example, the WHO Expert Committee on Biological Standardization established primary International Standards for SARS-CoV-2 testing in December 2020, and these standards define the International Units[7,8]. These primary standards are then used to calibrate **secondary reference materials** which are more routinely used during test and laboratory validation.

Natural reference materials are derived from natural sources, such as clinical patient samples, that have been well-characterized using differing methods. Whilst natural materials match the complexity and challenges of a patient sample, they are often finite in quantity, and difficult to reliably manufacture at scale[9]. Natural reference materials should ideally match final use cases and target populations, given that viral and antibody titers can vary widely between individuals and across the course of infection, resulting in natural reference materials with varying properties[10,11]. SARS-CoV-2 tests that were originally

validated using hospitalized patient samples with high viral titers performed markedly worse when used to screen mildly or asymptomatic individuals for which there are few available reference materials[12].

During initial stages of the pandemic, laboratories faced a major challenge in sourcing reference materials to evaluate testing, with patient samples also needed for competing research and therapeutic needs. While some large laboratories could leverage established clinical collaborations to source patient reference materials, the majority of laboratories faced difficulties sourcing patient reference materials needed to verify tests[13]. The coordinated and equitable dissemination of reference patient materials would have accelerated the deployment of tests, and also provide an early opportunity to harmonize test performance amongst laboratories.

The rapid development of **synthetic reference materials** can provide an interim solution in the absence of reference patient materials. Synthetic reference materials, including synthetic SARS-CoV-2 DNA and RNA genomes and recombinant proteins, were rapidly developed following the publication of the SARS-CoV-2 reference genome, and enabled the analytical validation of tests in countries even before the first reported cases of SARS CoV-2 virus had emerged[14–16]. However, synthetic materials may be poor surrogates for assessing pre-analytical variables, and may need to be further spiked into buffer or negative respiratory specimens to contrive a specimen-like matrix that can recapitulate RNA extraction steps. While regulatory agencies, such as the Food and Drug Administration (FDA) and European Medicines Agency (EMA) permitted the use of synthetic reference material to validate tests at early stages of the pandemic, they now require natural reference materials for verification and approval[17,18].

Molecular testing for SARS-CoV-2 genes.

Molecular testing involves the detection of the SARS-CoV-2 RNA genome using methods such as reverse transcription quantitative (RT-qPCR) and digital PCR , loop-mediated isothermal amplification, next-generation sequencing, and other nucleic amplification methods[19]. Molecular tests use respiratory samples that are particularly susceptible to pre-analytical variables, including specimen type, collection media and transport conditions[20,21].

Numerous synthetic genomes and natural reference materials have been developed to verify molecular testing protocols. However, given the sensitivity of molecular tests, laboratories must take care to implement best practice guidance and negative controls, with false-positive test results due to contamination by synthetic SARS-CoV-2 reagents or previous tests leading to false-positive results[22–25]. At early stages of the pandemic the US Centre for Disease Control (CDC) distributed RT-qPCR tests that included faulty negative controls, ultimately requiring new test kits to be developed and issued that resulted in delays during critical early stages of the pandemic[21].

Whilst molecular tests have the potential to measure viral abundance, they must be first calibrated to reference materials to realize this quantitative potential. For example, although the RT-qPCR cycling

threshold (Ct, also known as Cq or Cp) score can be used to estimate viral abundance, this abundance can vary up ~1000-fold for a given Ct value between instruments and laboratories[3]. Accordingly, calibration to reference materials is needed for quantitative comparisons of RT-qPCR results[7], and viral load cannot be considered in clinical stratification and the assessment of analytical sensitivity due to the current lack of harmonized reporting of RT-qPCR results[26,27].

Reference materials must also be regularly updated to reflect the diversity of SARS-CoV-2 variants circulating within a population. For example, new genetic variants can interfere with RT-qPCR primer or probe binding and result in false-negative testing results[28,29]. Whilst multiplex testing for several gene targets can mitigate the impact of a single variant, probes and primers require ongoing verification to ensure continued validity of a molecular test, and the FDA routinely monitors the predicted impact of variants on the performance of EUA approved tests[30].

Molecular testing can be sufficiently sensitive to detect the presence of SARS-CoV-2 RNA in wastewater. This can enable surveillance for SARS-CoV-2 in sewage collected across a large catchment area, and provide leading indication of an outbreak[31]. However, the sensitive detection and interpretation of wastewater testing is challenging, and there is a pressing need to standardize the sampling, methods and analysis used to detect and quantify SARS-CoV-2 in sewage. These standards are needed to ensure comparability and consistency between different municipalities and across time, and support surveillance measures that inform a public health response.

Antigen testing for SARS-CoV-2 proteins

Antigen tests employ lateral-flow or enzyme-linked immunosorbent assays to directly detect the presence of viral proteins. These antigen tests are typically less sensitive than molecular tests, and detect SARS CoV-2 across a narrower window during the viral infectious course (although repeated serial testing may mitigate this lower sensitivity)[32,33]. However, antigen tests are inexpensive to manufacture, can be performed at point-of-care and quickly return results. Given this convenience, antigen tests can be sold direct to consumers, and are often used for rapid screening of individuals during travel, or for attendance to schools, workplaces or community events.

Many laboratories reported that the performance of antigen tests differed markedly from the manufacturers declarations, and independent validation with reference standards was needed to confirm test performance[34]. Antigen tests can be evaluated using inactivated viruses or recombinant expressed proteins, however, their performance is more typically measured by comparison to results from previously authorized RT-qPCR tests[18]. However, relying on evaluation by positive agreement to a comparator test can be problematic, as it can propagate inaccuracies, differences or limitations present in the benchmark RT-qPCR method[35].

Antigen tests have been promoted as a viable method to realize massive population-scale testing; however, this proposal remains controversial[36]. An antigen test widely used in a pilot program to evaluate whether population-scale testing could curb rates of infection in Liverpool, UK was criticized for poor sensitivity, and found to miss almost half of individuals who otherwise tested positive using RT-qPCR[37–39]. This field performance of this antigen test was markedly lower than the manufacturer's declaration, and demonstrated the need to independently verify test performance with appropriate reference materials to understand limitations.

Serology testing for a COVID-19 immune response

Serology testing measures the presence of antibodies in an individual's blood that are reactive to SARS-CoV-2 proteins. A range of serology tests have been designed using different methods (including lateral flow, enzyme-linked, and chemiluminescent immunoassays) that detect antibody isotypes (IgM, IgA, IgG and total) elicited by previous infection or vaccination[40,41]. Serology tests can measure the avidity, duration and composition of reactive antibody response to SARS-CoV-2 infection. However, serology assays must be standardized and calibrated to enable quantitative comparisons of antibody measurements between individuals, across time and in response to differing variants[42].

Reference materials for COVID-19 serology tests are largely derived from convalescent patient serum. The WHO International Standard, prepared and supplied by National Institute for Biological Standards and Control (NIBSC), comprises a pool of convalescent plasma from recovered COVID-19 patients, with plasma from healthy donors collected before the pandemic to serve as a negative control[8]. The WHO assigned an arbitrary unitage to the reference standards to establish international units for neutralizing antibodies (e.g. IU/mL) and binding assays (e.g. BAU/mL). Calibration to these international units can standardize quantitative serology measurements used in clinical trials, can define consensus antibody titer thresholds, and can establish comparable correlates of protection against COVID-19[43,44].

Standardization for serological testing also underpins reproducible research in epidemiology, and the development of vaccines and therapeutics. Large-scale studies have used serology testing to understand the transmission of SARS-CoV-2 though populations, as well as the impact of vaccination on this transmission[45–47]. However, without standardizing these serology tests, comparison among datasets and populations can be difficult, resulting in a lost opportunity for interoperable research. Indeed, standardization of key research methods, such as cell-based assays used to measure neutralizing antibodies, would is a next step for reproducible research results[48].

Proficiency testing schemes are used to harmonize testing across laboratories.

Proficiency studies (or external quality assessment studies) share samples amongst participating laboratories for testing, who then report their results back for evaluation and comparison. Proficiency testing can evaluate the performance of individual laboratories, and provide collective appraisal of diverse SARS CoV-2 testing across different laboratories.

Given the scale and diversity of COVID-19 testing, proficiency testing schemes are needed to harmonize results between laboratories with different capabilities, particularly given many laboratories were established or re-purposed for COVID-19 testing with little previous clinical experience. Systematic proficiency testing schemes also provide an opportunity to calibrate international standards and units amongst participating laboratories, and provide the technical basis for guidance documents for harmonizing results[49].

Proficiency testing schemes can also evaluate the ongoing performance of tests following their initial validation for regulatory approval. Numerous proficiency studies were launched at early stages of the pandemic for both genome detection and serology testing[50–52]. These studies proved particularly important given that many SARS CoV-2 tests were given accelerated regulatory approval (such as emergency use authorization, EUAs) with little demonstrated performance under real-world settings. These studies proved key in providing an independent validation of testing methods, with results shared widely among laboratories and organizations using the results from those labs.

Information standards are needed to describe materials, procedures, results and performance.

Information standards are rules or guidelines that define how test performances, processes and results should be described. Information standards ensure these descriptions use consistent, transparent and harmonized terminology that enables the necessary information to be clearly communicated and interpreted by labs, clinical, and public health authorities[53]. With information standards, a practitioner can reliably select the appropriate test according to their requirements.

Information standards provide a consistent basis to compare amongst SARS-CoV-2 tests, and promote communication, consistency and transparency in results shared between laboratories[54]. A key standard for describing and reporting SARS-CoV-2 RT-qPCR assays is the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guideline, which provide a checklist for the disclosure of all reagents, sequences and methods necessary for other laboratories to reproduce RT-qPCR methods and results[55,56].

Information standards can ensure that results and data-sets are harmonized, and accompanied by meta-data that ensure results can be easily searched, accessed and analyzed. Meta-data typically describes useful accompanying information about samples (such as patient clinical status, location, gender and age) to support downstream integrated data analysis. The CSWG encourages scientists, clinicians,

journals, and peer reviewers to request appropriate standardized annotation and reported test results in International Units where possible.

Information standards are needed to provide unambiguous descriptions of test performance that can be independently verfied[53]. Manufacturers of SARS-CoV-2 tests often used secondary reference materials without traceability to a primary reference that can be difficult to independently verify. Many metrics, such as sensitivity and specificity, are also not fixed test properties, and need to be considered in the context of the clinical samples or secondary reference material used[57]. As a result, comparisons of manufacturer's declarations of test performance often showed diverged markedly from independent real-world evaluations[58].

Genome surveillance of new SARS-CoV-2 variants

Novel SARS-CoV-2 variants impact the fitness of the virus, allowing the virus to spread more easily, cause more severe disease, or escape the body's immune response[59]. Genomic surveillance has proven useful in monitoring the emergence and circulation of variants of concern and informing public health response. Many countries now sequence a targeted fraction of positive samples to measure SARS-CoV-2 variants circulating within a population[60]. Sequencing is also used in genomic epidemiology where the presence of shared mutations can infer chains and clusters of transmission between individuals. This can assist contact tracing of infected individuals in an outbreak, and track the spread of SARS-CoV-2 strains across the world[61].

Genomic surveillance will likely become an established feature of global testing, with a requirement to monitor novel, seasonal or resistant strains. Accordingly, there is a pressing need to develop reference materials and bioinformatic standards to ensure quality and comparability of results between national and international surveillance laboratories. This includes the collection of reference materials for different variant strains in biorepositories, or the rapid synthesis of variant genomes under safe biosecurity restrictions[62].

Information standards are also needed to ensure that sequence data-sets, which can be large and complex, are standardized to enable subsequent querying and analysis. More than one million SARS-2-CoV genome sequences have been submitted to databases such as GISAID[63]. The assignment of consistent meta-data to these genome sequences facilitates data integration, accessibility and the re-use of data for insightful analysis in future studies[64]. The nomenclature used to describe different SARS-CoV-2 strains has also been standardized to consolidate naming schemas and avoid the stigma of naming variants according to origin[65].

The analysis of genomic data, from raw sequence data into actionable information, is often complex and lengthy[66,67]. The diverse range of sequencing technologies, bioinformatic tools and data formats used within the analytical workflow must be harmonized to ensure interoperability and best-practice across the

surveillance network[64]. Stable versioning, data freezes and workflow management tools can standardise bioinformatic protocols, data outputs and reference files. Reference genomic data-sets can also be used in bioinformatic proficiency schemes to test the ability of genomic surveillance networks to detect novel variants.

Coordinated efforts amongst organizations for SARS CoV-2 testing and standards.

The pandemic response has demanded close cooperation between commercial, academic, non-governmental, and governmental organizations. The scale, urgency and uncertainty of the pandemic has required organizations to assume new roles, share information and pool resources to build testing capacity, and this coordination has been considered key to the success of pandemic responses[13].

Government and regulatory organizations were able to leverage reference laboratories to develop standards, evaluate tests, and disseminate information and best practices. The WHO and its collaborating reference laboratories at the Paul Ehrlich Institute (Germany), the FDA's Centre for Biologics Evaluation and Research (CBER) and the UK's National Institute for Biological Standards and Control (NIBSC), quickly initiated development of **international standards**[7,8]. However, while these international standards constitute a valuable global resource, their dissemination and adoption needs ongoing support from regional organizations to calibrate secondary standards to these international standards, and promote widespread harmonization of testing[68].

At early stages of the pandemic, many regulatory bodies allowed expedited validation of COVID-19 tests due to concerns about access and scaling of testing, and prior to widespread availability of reference standards[18]. While this accelerated the deployment of tests, it has also resulted in uneven testing performance. The FDA developed a reference panel that was available to test developers intending to submit tests for EUA submission to assess limit of detection[69]. Additionally, laboratories or organizations, such as FIND[70], were needed to independently evaluate test performance with interim reference materials.

In addition to developing reference materials, global organizations also provided expertise and guidance to countries that may lack their own established regulatory or standards organizations. Global dissemination of reference materials is needed to support the implementation of testing in low- and middle-income countries with greater dependence on point-of-care tests performed under heterogenous conditions. One innovation proposed dissemination of materials for decentralized production of "open source" secondary standards that could be provided by plasmid repositories and distributed under open-source terms to empower regional centers to manufacture their own secondary standards to validate local testing workflows[71]. Global imbalances in vaccination and testing have contributed to a global disparity in the impact of COVID-19; standards have a key role in mitigating these imbalances, and ensuring that testing is performed effectively and efficiently worldwide.

Conclusions

The standardization of SARS-CoV-2 testing remains an ongoing priority, and is part of the normalisation of epidemic prevention and control. Standards are needed for evolving testing methods and the increasing use of serological testing, genomic and wastewater surveillance. Worldwide testing must also be calibrated to international primary standards to harmonize performance and results across countries, enabling the global research enterprise to meaningfully share results. All of the standards must be maintained and updated in response to the emergence of novel SARS-CoV-2 variants. Whilst SARS-CoV-2 testing will likely remain a feature of public health, it is likely to become reduced in scope to seasonal and targeted testing of outbreaks, vulnerable populations or for international travel.

An independent review of the processes by which the FDA authorizes tests in the pandemic has recommended establishing a framework to validate test performance in preparation for and during a public health emergency[13]. This framework includes strengthening the communication between regulatory, government organizations and test developers, developing an independent capability to evaluate test performance quickly, and the better development and deployment of reference standards, such as clinical samples that were needed for test validation. During the pandemic, the CSWG contributed to many to these capabilities, and we recommend that this expertise and experience be institutionalized as part of pandemic preparedness (see **Box 1**). An institutionalized capacity to respond to standards needs will improve our pandemic response and establish the foundation for a reliable testing infrastructure for emerging diseases.

Many of these recommendations for standards are generalizable, and could similarly benefit the testing of pathogens, such as influenza, that are currently monitored for seasonal variants. These proposals can also be extended to developing standards for monitoring and responding to viral outbreaks in agriculture and livestock populations, which can further act as reservoirs for SARs-CoV-2 and other viruses that undergo zoonotic transfer[72].

The public are the ultimate beneficiaries of better standards. Standards ensure that patients will receive consistent and reliable results that inform their treatment, regardless of how and where they are tested. Standards permit informed evaluation of consistent measures of test performance, which may be otherwise considered secondary to cost and convenience. During the pandemic, numerous governments awarded contracts to the manufacturers of tests that were subsequently shown to perform poorly when independently validated[73]. Reliable testing and surveillance are also needed to inform policy, potentially avoiding the costs of alternative non-pharmaceutical interventions, and ultimately reducing healthcare costs. Previous research has demonstrated that standards are a cost-effective solution to improve testing and health outcomes that ultimately benefit the broader economy[74].

The pandemic has focused media, government and community attention on the importance of testing. However, while extensive resources have been invested in new testing methods[75], relatively fewer resources have been invested in the development of standards, despite their proven effectiveness. There

is an opportunity to ensure that the new widespread appreciation of testing for public health is accompanied by a matched appreciation of standards. Accordingly, we propose similar consideration and investment be afforded to standards, commensurate with their strategic, far-reaching and impactful benefits. Standards are a simple, proven method to assure test performance and a robust, reliable, and effective testing enterprise at the massive scale and diversity we have witnessed during the COVID-19 pandemic.

Boxes

Box 1. *Emerging Infectious Disease Standards Working Group*

We propose to establish an enduring working group, termed the *Emerging Infectious Disease Standards Working Group* (EIDSWG) to advocate for testing standards in emerging infectious diseases and future public health emergencies. The proposed Terms of Reference are:

Scope and objective.

- The EIDSWG serves as an independent, international, standing network of members that coordinate efforts and advocate for standards in testing undertaken in preparation or response to an infectious disease public health emergency.

Membership.

- The EIDSWG comprises general members that are self-selected and unrestricted. Membership includes representatives from differing technical backgrounds and expertise, including research and clinical scientists, government, public-health organizations, regulatory authorities, non-profit organizations and companies developing tests or reference standards. The EIDWG strives to ensure an open, inclusive diverse membership needed to respond to the global challenges of a public health emergency.
- The EIDSWG also comprises core members selected on technical expertise that represent an established network of laboratories readied to develop, implement and disseminate standards in preparation for and immediately in response to the emergence of a public health emergency. In addition, the EIDWSG may establish sub-groups dedicated to specific objectives as required.
- The EIDSW will work closely and partner with regional and international third-party partners involved in the development, implementation or validation of tests or standards.

Roles.

- Establish standing relationships between key stakeholders to coordinate the development and implementation of standards in preparation for and immediately in response to the emergence of a public health emergency.

- Develop and communicate policy and recommendations relating to standards and their use in test validation.
- Provide representatives to champion use of standards within the local, national, and international testing organizations.
- Support the development and use of standards (including reference materials, bioinformatic and information standards) by members as well as third-party organizations. This includes conducting harmonization studies to propagate international standards and units.
- Provide training and education on the development, role and adoption of standards. This includes the dissemination of standards best practices for the use of standards in test and laboratory validation.
- Prepare for and coordinate the equitable dissemination of standards in response to the emergence of a public health emergency. This includes dissemination of clinical specimens that are needed for optimal test validation.
- Support the development and use of interim reference materials in the absence of an international standard, and support the harmonization of interim references to international standard when available.
- Identify appropriate partners and funding to support the roles of the EIDSWG.

Meetings.

- The EIDSWG will regularly convene its membership with annual meetings, seminars, conferences and communications amongst members.

Box 2. GLOSSARY

Diagnostic performance – evaluates the ability of a test to discriminate between two binary outcomes. In the case of testing for infectious disease, diagnostic performance is typically the ability of a test to distinguish between the presence or absence of an analyte (such as viral genome, protein, or reactive antibodies) within a sample. For quantitative tests, the binary discrimination between presence or absence of a biomarker is based on the abundance of the biomarker exceeding a decision threshold (for example, a positive RT-qPCR result corresponds to a Ct level exceeding a given threshold).

Sensitivity – the probability that the test result is positive when the analyte is present (also referred to as the true positive rate).

Specificity – the probability that the test result Is negative when the analyte is absent (also referred to as the true negative rate).

Analytical performance – evaluates the ability of a test to measure the analyte of interest. This evaluation can be quantitative, and considers the difference between the result and the target reference

value. Analytical performance includes considerations such as the accuracy, precision, limit of detection of the test, as well as the reproducibility of test performance.

Reference materials – physical materials with well-characterized properties that can be used to validate or calibrate a testing process.

Information standards – Guidelines and rules that define how information (about process, performance, or result) is defined and described.

Primary standards - Reference materials whose metrological qualities are assigned without reference to other standards.

Secondary standards – reference materials whose metrological qualities are assigned by comparison to a primary standard. Secondary standards are more routinely used for within laboratory.

Natural materials – Reference samples that are derived from a natural source, such as a patient respiratory or blood samples, whose meteorological properties have been established through extensive characterization.

Synthetic materials – reference materials that have been manufactured to mimic patient sample properties using an artificial process. For example, the production of synthetic of RNA or DNA genomes, of the preparation of contrived sample matrices.

Proficiency (or external quality assurance) study – study conducted by external agency to evaluate the performance of one or more laboratories.

Accreditation - Formal evaluation and recognition that a testing laboratory is sufficiently competent to carry out specific tests.

TYPES OF REFERENCE MATERIALS

International reference standard - reference materials that are issued by an authorized body (such as WHO, NIST, JRC, NMI Australia, or other national and international bodies), and whose whose metrological qualities (possibly including the definitions of a quantitative unit) are assigned without reference to other standards and thereby provides the highest level of traceability.

Certified Reference Material – reference material accompanied by a certificate that provides the value of the specified property, its associated uncertainty, and a statement of metrological traceability, typically to a higher order standard (such as an SI unit or international reference material). Certified reference materials are produced and disseminated in accordance with ISO17034 and ISO17025 guidance.

Reference Material – a fit-for-purpose material used to verify, validate, or calibrate a test procedure. Reference materials are often produced and disseminated by an accredited reference material producer

(according to ISO17034). Certificates of Analysis are often provided according to ISO guidelines as best practice.

Declarations

DISCLAIMER.

Points of view in this document are those of the authors and do not necessarily represent the official position or policies of listed affiliated organizations. Specifically, the manuscript does not o not necessarily represent the official position or policies of National Institute of Standards and Technology or the U.S. Department of Commerce.

Certain commercial equipment, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by authors or affiliated organizations, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

References

1. Mercer, T. R. & Salit, M. Testing at scale during the COVID-19 pandemic. *Nature Reviews Genetics* vol. 22 415–426 (2021).

2. Page, M., Almond, N., Rose, N. J. & Schneider, C. K. Diagnostics and the coronavirus: don't let the standards slip. *Nature Biotechnology* vol. 38 673–674 (2020).

3. Evans, D. *et al.* The dangers of using Cq to quantify nucleic acid in biological samples; a lesson from COVID-19. *Clin. Chem.* (2021) doi:10.1093/CLINCHEM/HVAB219.

4. Lippi, G., Simundic, A. M. & Plebani, M. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). *Clinical Chemistry and Laboratory Medicine* vol. 58 1070–1076 (2020).

5. Engelmann, I. *et al.* Preanalytical Issues and Cycle Threshold Values in SARS-CoV-2 Real-Time RT-PCR Testing: Should Test Results Include These? *ACS Omega* **6**, 6528–6536 (2021).

6. Göhring, K., Dietz, K., Hartleif, S., Jahn, G. & Hamprecht, K. Influence of different extraction methods and PCR techniques on the sensitivity of HCMV-DNA detection in dried blood spot (DBS) filter cards. *J. Clin. Virol.* **48**, 278–281 (2010).

7. WHO/BS.2020.2402 Collaborative Study for the Establishment of a WHO International Standard for SARS-CoV-2 RNA. https://www.who.int/publications/m/item/WHO-BS-2020.2402.

8. Kristiansen, P. A. *et al.* WHO International Standard for anti-SARS-CoV-2 immunoglobulin. *The Lancet* vol. 397 1347–1348 (2021).

9. Bossuyt, P. M. Testing COVID-19 tests faces methodological challenges. *Journal of Clinical Epidemiology* vol. 126 172–176 (2020).

10. Woloshin, S., Patel, N. & Kesselheim, A. S. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *N. Engl. J. Med.* **383**, e38 (2020).

11. Iacobucci, G. Covid-19: Tests must be more rigorously regulated to protect public, say statisticians. *BMJ* **373**, (2021).

12. Takahashi, S., Greenhouse, B. & Rodríguez-Barraquer, I. Are Seroprevalence Estimates for Severe Acute Respiratory Syndrome Coronavirus 2 Biased? *J. Infect. Dis.* **222**, 1772–1775 (2020).

13. Emergency Use Authorization of COVID-19 Tests: Independent Assessment of the FDA's Response | FDA. https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/emergency-use-authorization-covid-19-tests-independent-assessment-fdas-response.

14. Madala, B. S., Reis, A. L. M., Deveson, I. W., Rawlinson, W. & Mercer, T. R. Chimeric synthetic reference standards enable cross-validation of positive and negative controls in SARS-CoV-2 molecular tests. *Sci. Rep.* **11**, 2636 (2021).

15. Corman, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **25**, 2000045 (2020).

16. Molecular assays to diagnose COVID-19: Summary table of available protocols. https://www.who.int/publications/m/item/molecular-assays-to-diagnose-covid-19-summary-table-of-available-protocols.

17. Ravi, N., Cortade, D. L., Ng, E. & Wang, S. X. Diagnostics for SARS-CoV-2 detection: A comprehensive review of the FDA-EUA COVID-19 testing landscape. *Biosens. Bioelectron.* **165**, (2020).

18. Mitchell, S. L. *et al.* Understanding, verifying, and implementing emergency use authorization molecular diagnostics for the detection of sars-cov-2 RNA. *Journal of Clinical Microbiology* vol. 58 (2020).

19. Esbin, M. N. *et al.* Overcoming the bottleneck to widespread testing: A rapid review of nucleic acid testing approaches for COVID-19 detection. *RNA* vol. 26 771–783 (2020).

20. Butler-Laporte, G. *et al.* Comparison of Saliva and Nasopharyngeal Swab Nucleic Acid Amplification Testing for Detection of SARS-CoV-2: A Systematic Review and Meta-analysis. *JAMA Intern. Med.* **181**, 353–360 (2021).

21. Lee, R. A., Herigon, J. C., Benedetti, A., Pollock, N. R. & Denkinger, C. M. Performance of saliva, oropharyngeal swabs, and nasal swabs for SARS-CoV-2 molecular detection: A systematic review and meta-analysis. *J. Clin. Microbiol.* **59**, (2021).

22. Mögling, R. *et al.* Delayed laboratory response to COVID-19 caused by molecular diagnostic contamination. *Emerg. Infect. Dis.* **26**, 1944–1946 (2020).

23. Wang, C. Y. T., Buckley, C., Bletchly, C., Harris, P. & Whiley, D. Contamination of SARS-CoV-2 RT-PCR probes at the oligonucleotide manufacturer. *Pathology* vol. 52 814 (2020).

24. Davidi, D. *et al.* Amplicon residues in research laboratories masquerade as COVID-19 in surveillance tests. *Cell Reports Methods* **1**, 100005 (2021).

25. Huggett, J. F. *et al.* Cautionary Note on Contamination of Reagents Used for Molecular Detection of SARS-CoV-2. *Clin. Chem.* **66**, 1369–1372 (2020).

26. Fajnzylber, J. *et al.* SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat. Commun. 2020 111* **11**, 1–9 (2020).

27. Goyal, A., Reeves, D. B., Fabian Cardozo-Ojeda, E., Schiffer, J. T. & Mayer, B. T. Viral load and contact heterogeneity predict sars-cov-2 transmission and super-spreading events. *Elife* **10**, 1–63 (2021).

28. SARS-CoV-2 Viral Mutations: Impact on COVID-19 Tests | FDA. https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-viral-mutations-impact-covid-19-tests.

29. Vogels, C. B. F. *et al.* Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. *PLoS Biol.* **19**, e3001236 (2021).

30. SARS-CoV-2 Viral Mutations: Impact on COVID-19 Tests | FDA. https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-viral-mutations-impact-covid-19-tests.

31. Peccia, J. *et al.* Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* **38**, 1164–1167 (2020).

32. Mina, M. J., Parker, R. & Larremore, D. B. Rethinking Covid-19 Test Sensitivity — A Strategy for Containment. *N. Engl. J. Med.* **383**, e120 (2020).

33. NR, P. *et al.* Performance and Implementation Evaluation of the Abbott BinaxNOW Rapid Antigen Test in a High-Throughput Drive-Through Community Testing Site in Massachusetts. *J. Clin. Microbiol.* **59**, (2021).

34. Fitzpatrick, M. C., Pandey, A., Wells, C. R., Sah, P. & Galvani, A. P. Buyer beware: inflated claims of sensitivity for rapid COVID-19 tests. *The Lancet* vol. 397 24–25 (2021).

35. Bossuyt, P. M., Irwig, L., Craig, J. & Glasziou, P. Comparative accuracy: Assessing new tests against existing diagnostic pathways. *Br. Med. J.* **332**, 1089–1092 (2006).

36. Mina, M. J., Parker, R. & Larremore, D. B. Rethinking Covid-19 Test Sensitivity — A Strategy for Containment. *N. Engl. J. Med.* (2020) doi:10.1056/nejmp2025631.

37. Crozier, A., Rajan, S., Buchan, I. & McKee, M. Put to the test: Use of rapid testing technologies for Covid-19. *BMJ* **372**, (2021).

38. Wise, J. Covid-19: Lateral flow tests miss over half of cases, Liverpool pilot data show. *BMJ* **371**, m4848 (2020).

39. Mahase, E. Covid-19: Innova lateral flow test is not fit for “test and release” strategy, say experts. *BMJ* (2020) doi:10.1136/bmj.m4469.

40. Lisboa Bastos, M. *et al.* Diagnostic accuracy of serological tests for covid-19: Systematic review and meta-analysis. *BMJ* **370**, 2516 (2020).

41. Whitman, J. D. *et al.* Evaluation of SARS-CoV-2 serology assays reveals a range of test performance. *Nat. Biotechnol.* **38**, 1174–1183 (2020).

42. Lisboa Bastos, M. *et al.* Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* **370**, 2516 (2020).

43. Earle, K. A. *et al.* Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* **39**, 4423–4428 (2021).

44. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med. 2021 277* **27**, 1205–1211 (2021).

45. Alter, G. & Seder, R. The Power of Antibody-Based Surveillance. *N. Engl. J. Med.* (2020) doi:10.1056/nejme2028079.

46. Gudbjartsson, D. F. *et al.* Spread of SARS-CoV-2 in the Icelandic Population. *N. Engl. J. Med.* (2020) doi:10.1056/nejmoa2006100.

47. Bonanni, P. *et al.* The Role of Serology Testing to Strengthen Vaccination Initiatives and Policies for COVID-19 in Europe. *COVID 2021, Vol. 1, Pages 20-38* **1**, 20–38 (2021).

48. Muruato, A. E. *et al.* A high-throughput neutralizing antibody assay for COVID-19 diagnosis and vaccine evaluation. *Nat. Commun.* **11**, (2020).

49. Milavec, M. *et al.* The performance of human cytomegalovirus digital PCR reference measurement procedure in seven external quality assessment schemes over four years. *Methods* (2021) doi:10.1016/J.YMETH.2021.03.016.

50. Matheeussen, V. *et al.* International external quality assessment for SARSCoV-2 molecular detection and survey on clinical laboratory preparedness during the COVID-19 pandemic. *Eurosurveillance* **25**, (2020).

51. Edson, D. C., Casey, D. L., Harmer, S. E. & Downes, F. P. Identification of SARS-CoV-2 in a Proficiency Testing Program. *Am. J. Clin. Pathol.* **154**, 475–478 (2020).

52. Tacker, D. H. *et al.* Interlaboratory agreement of anti–severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) serologic assays in the expedited College of American Pathologists proficiency testing program. *Arch. Pathol. Lab. Med.* **145**, 536–542 (2021).

53. Cohen, J. F. *et al.* STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* **6**, e012799 (2016).

54. Doust, J. A. *et al.* Guidance for the design and reporting of studies evaluating the clinical performance of tests for present or past SARS-CoV-2 infection. *BMJ* **372**, (2021).

55. SA, B. *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).

56. ISO - ISO 20395:2019 - Biotechnology — Requirements for evaluating the performance of quantification methods for nucleic acid target sequences — qPCR and dPCR. https://www.iso.org/standard/67893.html.

57. Irwig, L., Bossuyt, P., Glasziou, P., Gatsonis, C. & Lijmer, J. Evidence base of clinical diagnosis: Designing studies to ensure that estimates of test accuracy are transferable. *British Medical Journal* vol. 324 669–671 (2002).

58. MacKay, M. J. *et al.* The COVID-19 XPRIZE and the need for scalable, fast, and widespread testing. *Nature Biotechnology* (2020) doi:10.1038/s41587-020-0655-4.

59. Krause, P. R. *et al.* SARS-CoV-2 Variants and Vaccines. *N. Engl. J. Med.* (2021) doi:10.1056/NEJMsr2105280.

60. Rockett, R. J. *et al.* Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nat. Med.* (2020) doi:10.1038/s41591-020-1000-7.

61. Gardy, J. L. & Loman, N. J. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics* (2018) doi:10.1038/nrg.2017.88.

62. West, R. & Gronvall, G. K. California shows the way for biosecurity in commercial gene synthesis. *Nature Biotechnology* vol. 38 1021 (2020).

63. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Challenges* **1**, 33–46 (2017).

64. Griffiths, E. J. *et al.* The PHA4GE SARS-CoV-2 Contextual Data Specification for Open Genomic Epidemiology. (2020) doi:10.20944/PREPRINTS202008.0220.V1.

65. Konings, F. *et al.* SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nature Microbiology* vol. 6 821–823 (2021).

66. Black, A., MacCannell, D. R., Sibley, T. R. & Bedford, T. Ten recommendations for supporting open pathogen genomic analysis in public health. *Nat. Med. 2020 266* **26**, 832–841 (2020).

67. Gwinn, M., MacCannell, D. & Armstrong, G. L. Next-Generation Sequencing of Infectious Pathogens. *JAMA* **321**, 893–894 (2019).

68. ENGLISH ONLY 4 5 MANUAL FOR THE PREPARATION OF SECONDARY REFERENCE MATERIALS FOR IN VITRO DIAGNOSTIC ASSAYS DESIGNED FOR. (2016).

69. SARS-CoV-2 Reference Panel Comparative Data | FDA. https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-reference-panel-comparative-data.

70. FIND evaluations of SARS-CoV-2 assays - FIND. https://www.finddx.org/covid-19/sarscov2-eval/.

71. Kahl, L. *et al.* Opening options for material transfer. *Nat. Biotechnol. 2018 3610* **36**, 923–927 (2018).

72. Munnink, B. B. O. *et al.* Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science (80-.).* **371**, 172–177 (2021).

73. Covid-19: government must urgently rethink lateral flow test roll out - The BMJ. https://blogs.bmj.com/bmj/2021/01/12/covid-19-government-must-urgently-rethink-lateral-flow-test-roll-out/.

74. Leech, D. P. & Belmont, P. A. The Economic Impacts of NIST'S Cholesterol Standards Program. (2000).

75. Tromberg, B. J. *et al.* Rapid Scaling Up of Covid-19 Diagnostic Testing in the United States — The NIH RADx Initiative. *N. Engl. J. Med.* (2020) doi:10.1056/nejmsr2022263.

Figures

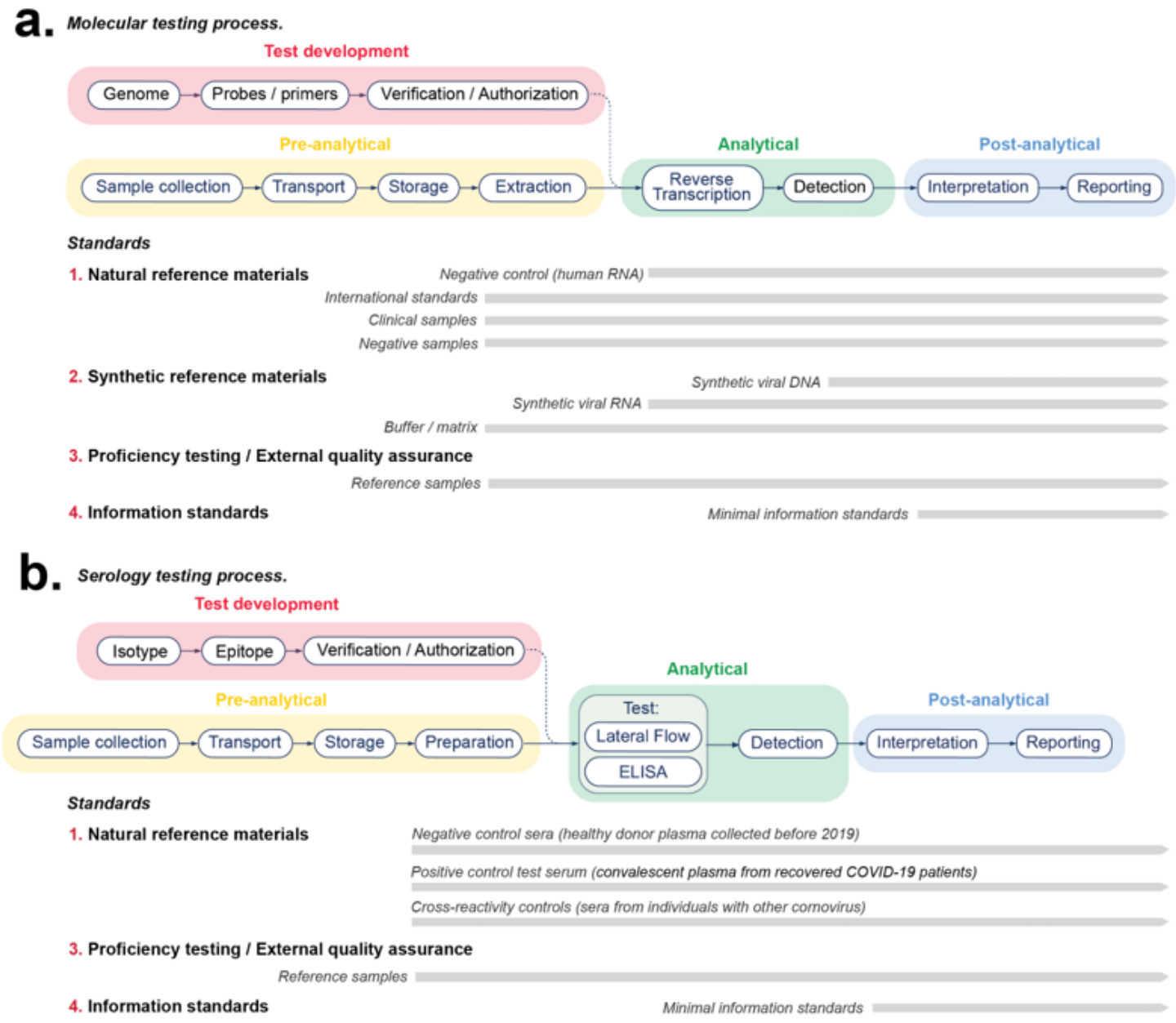

Figure 1

Standards needed for the SARS-CoV-2 testing process. Schematic diagram illustrating the steps in the molecular (a) and serological (b) testing process which can be classified into pre-analytical (yellow), analytical (green) and post-analytical (blue) stages. Additional test development (yellow) is performed prior be manufacturers. Lower panel describes the range of standards available for validating test development, processes and results.

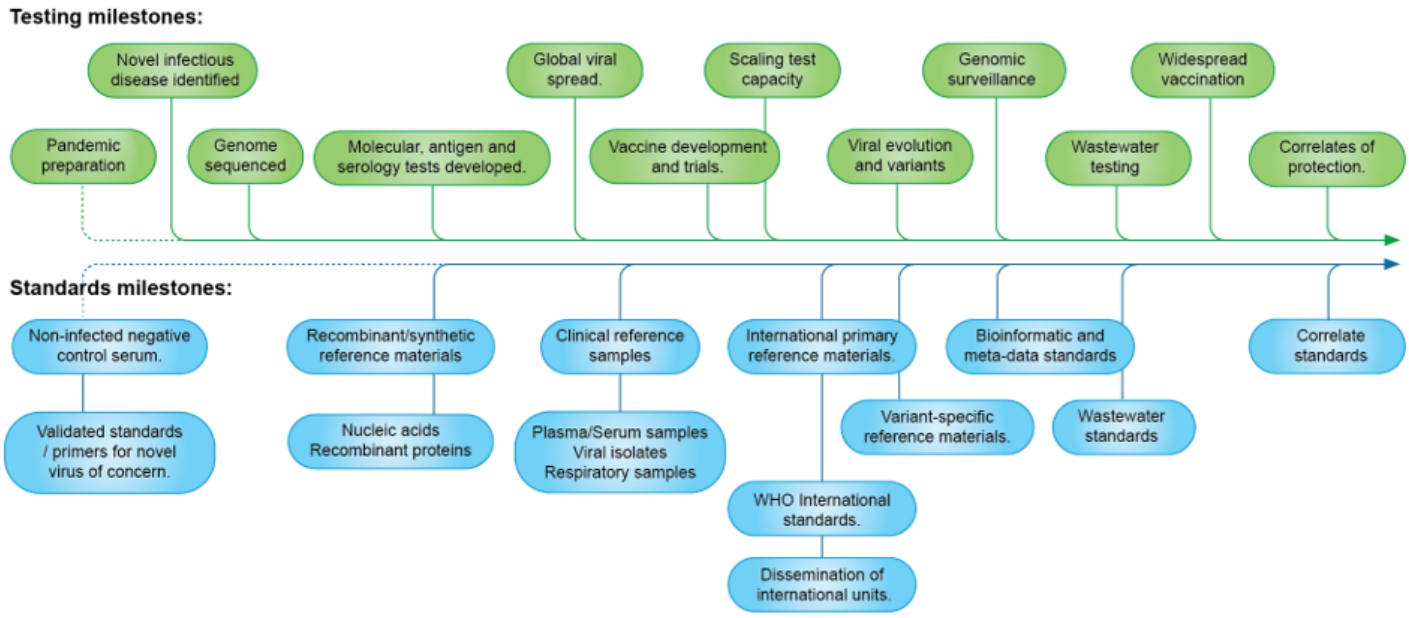

Figure 2

Schematic diagram illustrates key milestones in the development of testing and standards during the COVID-19 pandemic. Understanding these milestones can assist in preparing for current and future public health emergencies.