Accreditation and Quality Assurance
https://doi.org/10.1007/s00769-026-01721-w

RESEARCH

A playbook for harmonization of standards for emergent viral pathogens

Sara Suliman[1,2,3] · Sebastien Fuchs[4] · David Catoe[5] · Loren Hansen[5] · Deepa Eveleigh[6] · Michael A. Crone[7,8,9] · Paul S. Freemont[7,8,9] · Martin Kammel[10,11] · Heinz Zeichhardt[11,12] · Andrew T. Anfora[13] · Russell Garlick[13] · Neil Almond[14] · Mark Page[14] · Brian Beck[15] · Hui Wang[16] · Karuna Sharma[17] · Gary Pestano[18] · Joannes Vancann[19] · Jo Vandesompele[20] · Adam S. Corner[21] · Jeffrey Albrecht[22] · Deborah J. Boles[22] · Marcia Eisenberg[22] · Ayla B. Harris[22] · Brian J. Krueger[22] · Amanda L. Suchanek[22] · Thomas J. Urban[22] · Jonathan D. Williams[22] · Patrick Chain[23] · Alina Deshpande[23] · Attelia Hollander[23] · Sujata Chalise[24] · David R. Walt[24,25,26] · Jeffrey Germer[27] · Joseph D. Yao[27] · Bobbi S. Pritt[27] · Frederick S. Nolte[28] · Alexandra Bogožalec Košir[29] · Mojca Milavec[29] · Megan H. Cleveland[30] · Peter M. Vallone[30] · Eloise J. Busby[31,32] · Jim F. Huggett[31,32] · Denise M. O'Sullivan[31,32] · Elizabeth M. Marlowe[33] · Benjamin A. Pinsky[34] · Annaleise Kealiher[35] · Jeantine E. Lunshof[25,36] · Christopher E. Mason[37,38,39,40] · John Sninsky[41] · Cristina Tato[2] · Timothy Mercer[42] · Thomas J. White[43] · Marc Salit[5,35,44]

Received: 3 March 2026 / Accepted: 3 July 2026

Abstract
Effective clinical and public health decision-making during a pandemic depends on reliable and interoperable clinical diagnostic test results. To ensure trustworthy outcomes, we need widely available calibration standards harmonized to a shared scale. We present a "playbook" for an interlaboratory harmonization study that calibrates any available standards against a limited-availability standard issued by a global authority like the World Health Organization (WHO).

keywords Molecular Diagnostics · Playbook · Standards · Harmonization · SARS-CoV-2 · Pandemic Preparedness

What this is

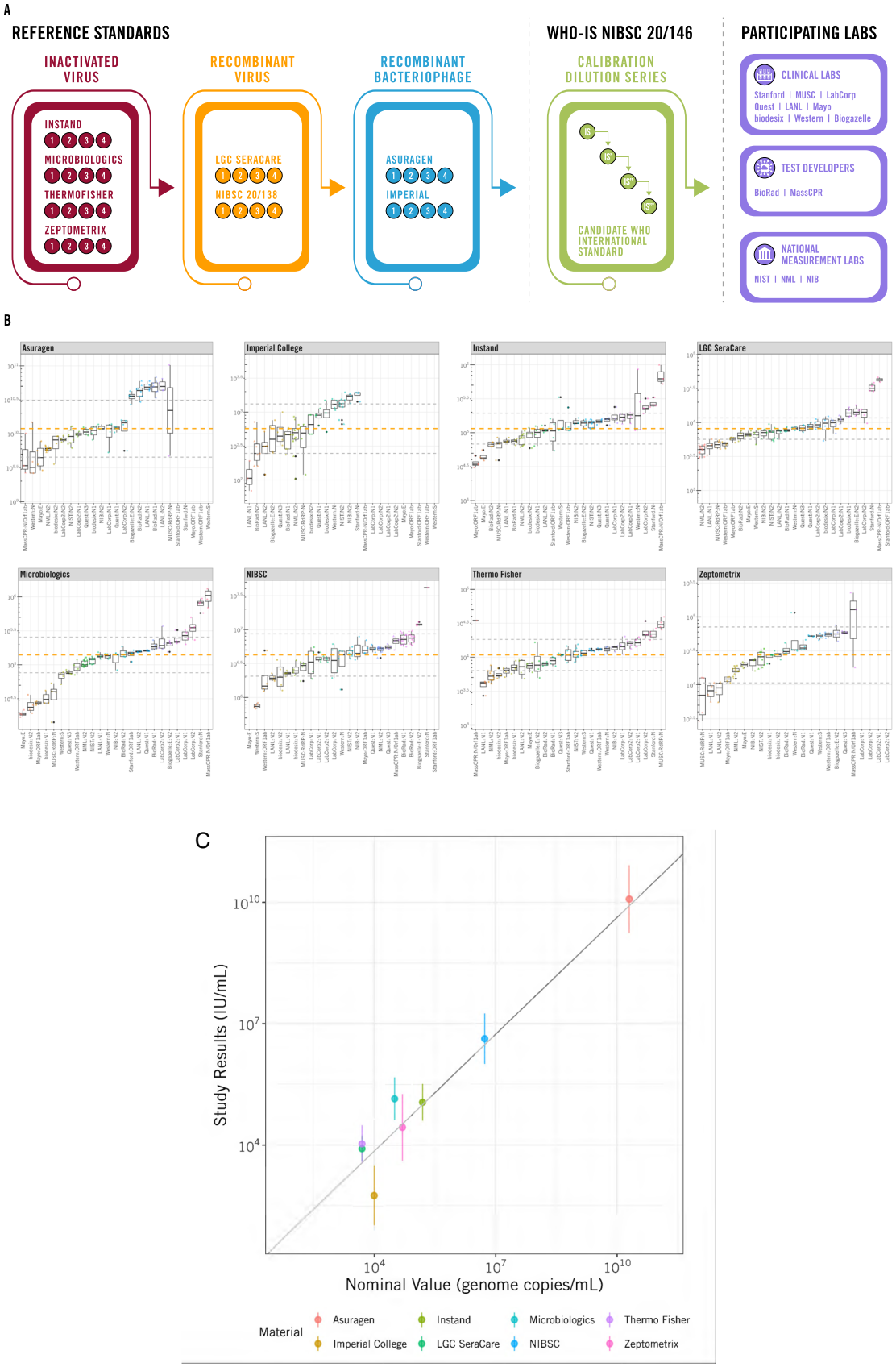
As part of the global response to the COVID-19 pandemic [1, 2], various organizations rapidly made available a wide variety of standards and controls to develop, calibrate, and validate molecular diagnostics for SARS-CoV-2. Our *Coronavirus Standards Working Group* (CSWG) [3] designed and conducted a "harmonization study" to establish the comparability of a set of these standards and controls (hereafter collectively referred to as "standards"). Harmonized standards can be used as a basis for meaningful quantitative comparability of test performance, and can underpin quantitative virology, therapeutic actions, and vaccinology. Historically, molecular diagnostic results are frequently reported using assay-dependent measurements, including cycle threshold (Ct), crossing point (Cp), or quantification cycle (Cq) values. Although these measurements are useful within individual assays, they are not comparable across laboratories and depend on assay designs, reagents, instrumentation, and other analytical conditions [4]. Calibration against unified standards, such as the WHO International Standards, enables reporting on a common quantitative scale, such as genome copies/mL or international units (IU)/mL, which allows for meaningful comparisons across heterogeneous assays and over time [4]. Harmonization balances the tension between the widespread need for standards and the limited availability of standards from global authorities. Our study is presented here in the form of a "playbook" as a recipe to conduct a robust harmonization study when it is needed again.

We calibrated 8 widely available molecular standards to the rapidly developed World Health Organization (WHO) International Standard (IS) [5], putting all the quantitative values of these standards on this common scale (Supplementary Table 1A). We recruited 14 laboratories internationally to participate in a designed experiment to "value assign" the standards on the scale (Fig. 1A and B and Supplementary Table 1B). Our intent was for widely available harmonized

Sara Suliman and Sebastien Fuchs have contributed equally to this work.

Extended author information available on the last page of the article

Published online: 13 July 2026


A
REFERENCE STANDARDS
INACTIVATED VIRUS
INSTAND
MICROBIOLOGICS
THERMOFISHER
ZEPTOMETRIX
RECOMBINANT VIRUS
LGC SERACARE
NIBSC 20/138
RECOMBINANT BACTERIOPHAGE
ASURAGEN
IMPERIAL
WHO-IS NIBSC 20/146
CALIBRATION DILUTION SERIES
CANDIDATE WHO INTERNATIONAL STANDARD
PARTICIPATING LABS
CLINICAL LABS
Stanford | MUSC | LabCorp
Quest | LANL | Mayo
biodesix | Western | Biogazelle
TEST DEVELOPERS
BioRad | MassCPR
NATIONAL MEASUREMENT LABS
NIST | NML | NIB
B
Asuragen
Imperial College
Instand
LGC SeraCare
Microbiologics
NIBSC
Thermo Fisher
Zeptometrix
C
Study Results (IU/mL)
Nominal Value (genome copies/mL)
Material
Asuragen
Imperial College
Instand
LGC SeraCare
Microbiologics
NIBSC
Thermo Fisher
Zeptometrix

◂**Fig. 1** **A** Molecular standard harmonization study design: schematic representation of the Coronavirus Standards Working Group (CSWG) molecular standard harmonization study. Four identical aliquots of each of the secondary reference materials were shipped to the 14 participating laboratories. The NIBSC code 20/146 WHO-IS was also shipped in 4 aliquots to benchmark the 8 secondary reference standards. **B** Box and whiskers represent median and interquartile ranges corresponding to the SARS-CoV-2 concentrations in IU, calibrated using the NIBSC 20/146 reference material. Each measurement represents a unique laboratory and assay, with the orange line spanning the median concentration across the different platforms and assays. Measurements are ordered from the lowest to highest extrapolated values. Data are summarized assuming each laboratory and assay is independent samples from the population. Orange lines represent population medians, and the dashed gray lines are 95% confidence intervals estimated from a median absolute deviation. **C** Observed concentrations of each standard in $\log_{10}$ IU mL^{-1} versus the nominal value in $\log_{10}$ genome copies mL^{-1}. The points are population medians and 95% confidence intervals (same as summary lines in Fig. 1B). The line represents the IS values for the international standard serial dilution. The line is the multiplicative factor of the IS value to copies, showing that this proportionality is consistent across the population of the 8 standards

standards to be part of the infrastructure to confidently deploy testing at scale. We prototyped a "proof-of-concept study" to test whether calibrating with the harmonized standards would yield comparable results when measuring clinical samples (Supplementary Fig. 1).

The study was not intended to be a comparison of tests or laboratories, a survey of test performance, or an evaluation of commutability. This interlaboratory study had the sole aim to establish the quantitative comparability of these widely available SARS-CoV-2 standards for molecular diagnostics. The study was designed to provide a comprehensive combined estimate of the sources of variability arising from variation among laboratories, assays, and standards.

Why it matters

The pandemic response (including development of diagnostics) was based on best practices developed in the moment. We offer here a study design pre-positioned for a better response to emerging pathogens with pandemic potential. Harmonizing standards with this design will assure comparability of results from diagnostic tests [4]. Comparable results enable quantitative, systematic diagnostic virology to understand pathogenesis and rapidly integrate data across space and time. Understanding pathogenesis [6] and transmission [2, 7–9] is critical for clinical and public health decision-making, and well-calibrated controls are essential for adjudicating accuracy of diagnostics, especially tests using new or nascent technology. We present and support this approach here for a viral pathogen but recognize that the principles would apply to non-viral pathogens. While we did not deploy quantitative virology at scale in the SARS-CoV-2 pandemic, the playbook presented here is aspirational, because quantitation has proved to be useful in other pandemics (e.g., HIV/AIDS and HCV). This harmonization framework supports the broader transition from qualitative interpretation of Ct/Cq values toward quantitative viral load reporting using traceable units. Such calibrated reporting has been recommended to improve interpretation of molecular diagnostics and facilitate comparison across laboratories, clinical studies, and public health surveillance.

What we did

The study was coordinated by the Joint Initiative for Metrology in Biology, the convener of the CSWG [3]. The coordinating laboratory worked with the CSWG to design the study, develop a standard operating procedure (SOP), and create a reporting template for the study. The study design ensured that the results meaningfully represented the variety of standard materials and laboratories and their protocols (Fig. 1A). The SOP ensured that we obtained results from a common experimental design that permitted assessment of variability (Supplement 1), and the template ensured consistent analysis of results (Supplementary Table 2 and Methods). We developed an open public web-based analysis package to transparently present and share results (Shiny App: https://msalit.shinyapps.io/RNAstudy/).

We selected 8 widely available standards (including materials derived from inactivated virus, recombinant virus, and recombinant bacteriophage) that could be processed in the laboratory workflow as clinical specimens. We recruited 14 laboratories from different sectors (clinical, academic, public, commercial, and government) to each calibrate the standards to Medicines and Healthcare Products Regulatory Agency (MHRA), formerly known as National Institute for Biological Standards and Control (NIBSC) 20/146 [4], the first WHO International Standard for SARS-CoV-2 RNA (WHO-IS) that defined the international units (IU, Fig. 1C, Supplementary Table 1A). Our study design used 4 replicate serial dilutions of the IS across 6 orders of magnitude, with 4 replicates of each dilution of the WHO-IS and each standard (Supplementary Figs. 2A–K). After calibration of these 8 standards, they were harmonized to a common scale by their value assignment in IU.

We operated a companion study to evaluate the hypothesis that the PCR tests calibrated with harmonized standards yield comparable results when measuring clinical samples (Supplementary Fig. 1). We used pooled biobanked samples (15 pools including negative controls and clinical positives at 3 viral load levels), evaluated by a subset of laboratories (3 laboratories using a total of 5 different assays). Three calibrated standards representing the three material classes were distributed to participating laboratories. However, due

to lot changes in the Asuragen standards, we only used two of the 8 standards: Thermo Fisher (inactivated virus), and LGC (recombinant virus), now shown in Fig. 2B.

How it came out

Harmonization results

The 8 standards were all successfully value-assigned to a common, shared reference: the IU (Fig. 1C and *Results v. Nominal* in analysis dashboard of the Shiny app). The analysis dashboard also contains figures showing the measurement results for each standard by assay within laboratory for each of the 8 materials sortable by rank order or laboratory; deviations from median value for each assay; tabular results, and raw data.

Proof-of-Concept results

Calibrating the assays with harmonized standards allowed us to compare results from different laboratories and tests on the clinical samples (Fig. 2A). Each laboratory used its own experiment design, representing real-life testing where laboratories use different tests. This proof-of-concept experiment demonstrated the use of calibration to bring clinical sample results from disparate assay modalities (reverse transcription quantitative real-time PCR (RT-qPCR) and reverse transcription digital PCR (RT-dPCR)) to the same scale, plotted on the same axis (Fig. 2B). Critical analysis of these data and assessment of bias and variability is in the Supplement (Supplementary Fig. 3A–C). Lessons learned from this experiment have been incorporated into the playbook recommendations. Results from these studies suggest that harmonized calibration has the potential to enable aggregation of quantitative test results across laboratories and over time, although broader validation across additional laboratories, assays, and clinical specimens will be needed.

What we learned

It is practical and possible to harmonize standards and controls early in an outbreak or pandemic by strategically bringing together a diverse set of specialists. A thoughtful study design supported by clear SOPs and consistent data reporting makes harmonization possible. Our CSWG experience cultivated a community of practitioners working collectively to respond to a global emergency, fostering new fruitful collaborations.

Our harmonization study and proof-of-concept study were conducted at different times, a year apart. We recommend that these studies be done simultaneously, with the same assays and lots of standards used in the harmonization. This improved design (see "playbook" in Supplement 3) can be periodically repeated to maintain comparability as the pathogens, assays and tests, sampling, and standards evolve.

The heterogeneity of laboratories, assays, and materials highlights the influence of these variables on test performance. The most significant artifacts we observed arose from standards developed from different materials. Full-genome inactivated virus standards and recombinant whole-virus samples behaved differently than the much smaller genomes with limited PCR target coverage of the bacteriophage materials. These different materials also had very different genome concentrations, requiring different dilution and handling in the laboratory.

The study design permits robust estimates of measurement uncertainty. Such estimates promise the ability to propagate uncertainties in calibrated results. This design ensures broad applicability to the heterogeneous tests and materials deployed internationally.

How to do it next time

It is inevitable that new pathogens (viral, bacterial, fungal or parasites) will emerge. The WHO provides guidance on preparing secondary standards to test for emerging pathogens [10]. Our experience with the harmonization study conducted by the CSWG for SARS-CoV-2 is captured and presented (Supplement 3) in a generalized Playbook for Harmonization of Standards to Calibrate Pathogen Molecular Diagnostics. In future implementations of the playbook, we recommend conducting the clinical sample evaluation concurrently with the harmonization study, using the same breadth of participating laboratories, assays, and instruments. In the present study, however, access to the clinical samples became available only after completion of the harmonization study. Consequently, the proof-of-concept evaluation was performed in a smaller subset of participating laboratories and assays than the primary harmonization study.

Summary

Both the harmonization and proof-of-concept studies demonstrate that it is possible to contribute to the standardization of measurement results for diagnostic laboratories through international and interdisciplinary cooperation. Our data provide a framework to break the glass when needed and rapidly harmonize molecular standards.

Fig. 2 **A** Study design of a proof-of-concept application of the harmonized secondary reference materials to real clinical samples. 15 pools of biobanked samples (from 3 patients each) were aliquoted and shipped to three participating laboratories (Los Alamos National Laboratory (LANL), Biogazelle, and Biodesix), along with 2 calibrated standards representing the 3 classes: Inactivated virus: AcroMetrix/ThermoFisher Scientific, recombinant virus: AccuplexTM/LGC Clinical Diagnostics, and recombinant bacteriophage: Armored RNA Quant SARS-CoV-2/Asuragen. Laboratories performed PCR using their described testing platforms and sent raw Cq or count data from their RT-qPCR or RT-dPCR platforms, respectively. **B** Harmonized measurements of viral load in 12 pooled clinical samples s01–s12: Calibrated measurements in clinical samples reported in IU mL^{-1}. Pooled samples with different viral loads were measured by different laboratories along with two calibrant materials: LGC Clinical Diagnostics, and AcroMetrix. Data are shown after calibration of the raw data from both qPCR and dPCR platforms (Supplementary Fig. 3)

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00769-026-01721-w.

Acknowledgements We acknowledge Lynne Sopchak for critically reading the manuscript and providing comments. We also acknowledge Andrew Hyatt for support with graphic design. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Commerce. Certain commercial software, 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 the National Institute of Standards and Technology, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Author contributions All authors contributed to data generation and manuscript writing and editing.

Funding The funding was provided by CZ Biohub.

Data availability We developed an open public web-based analysis package to transparently present and share results (Shiny App: https://msalit.shinyapps.io/RNAstudy/). All code is available on this github repository: https://github.com/msalit/harmonization_dashboard/tree/main

Declarations

Competing Interests The authors declare no competing interests.

References

1. Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20:533–534. https://doi.org/10.1016/S1473-3099(20)30120-1
2. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M (2020) Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368:860–868. https://doi.org/10.1126/science.abb5793
3. Mercer T et al (2022) The Coronavirus Standards Working Group's roadmap for improved population testing. Nat Biotechnol 40:1563–1568. https://doi.org/10.1038/s41587-022-01538-1
4. Evans D et al (2021) The dangers of using Cq to quantify nucleic acid in biological samples: a lesson from COVID-19. Clin Chem 68:153–162. https://doi.org/10.1093/clinchem/hvab219
5. Nguyen LT et al (2022) Establishment of an in-house real-time RT-PCR assay for the detection of severe acute respiratory syndrome coronavirus 2 using the first World Health Organization international standard in a resource-limited country. J Clin Lab Anal 36:e24355. https://doi.org/10.1002/jcla.24355
6. Dadras O et al (2022) The relationship between COVID-19 viral load and disease severity: a systematic review. Immun Inflamm Dis 10:e580. https://doi.org/10.1002/iid3.580
7. Li Q et al (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 382:1199–1207. https://doi.org/10.1056/NEJMoa2001316
8. Lipsitch M, Krammer F, Regev-Yochay G, Lustig Y, Balicer RD (2022) SARS-CoV-2 breakthrough infections in vaccinated individuals: measurement, causes and impact. Nat Rev Immunol 22:57–65. https://doi.org/10.1038/s41577-021-00662-4
9. Salvagno GL et al (2022) SARS-CoV-2 Omicron infection is associated with high nasopharyngeal viral load. J Infect. https://doi.org/10.1016/j.jinf.2022.02.025
10. WHO (2017) WHO manual for the preparation of secondary reference materials for in vitro diagnostic assays designed for infectious disease nucleic acid or antigen detection: calibration to WHO International Standards. In: Annex 6 Health Product Policy and Standards (HPS)

Authors and Affiliations

Sara Suliman[1,2,3] · Sebastien Fuchs[4] · David Catoe[5] · Loren Hansen[5] · Deepa Eveleigh[6] · Michael A. Crone[7,8,9] · Paul S. Freemont[7,8,9] · Martin Kammel[10,11] · Heinz Zeichhardt[11,12] · Andrew T. Anfora[13] · Russell Garlick[13] · Neil Almond[14] · Mark Page[14] · Brian Beck[15] · Hui Wang[16] · Karuna Sharma[17] · Gary Pestano[18] · Joannes Vancann[19] · Jo Vandesompele[20] · Adam S. Corner[21] · Jeffrey Albrecht[22] · Deborah J. Boles[22] · Marcia Eisenberg[22] · Ayla B. Harris[22] · Brian J. Krueger[22] · Amanda L. Suchanek[22] · Thomas J. Urban[22] · Jonathan D. Williams[22] · Patrick Chain[23] · Alina Deshpande[23] · Attelia Hollander[23] · Sujata Chalise[24] · David R. Walt[24,25,26] · Jeffrey Germer[27] · Joseph D. Yao[27] · Bobbi S. Pritt[27] · Frederick S. Nolte[28] · Alexandra Bogožalec Košir[29] · Mojca Milavec[29] · Megan H. Cleveland[30] · Peter M. Vallone[30] · Eloise J. Busby[31,32] · Jim F. Huggett[31,32] · Denise M. O'Sullivan[31,32] · Elizabeth M. Marlowe[33] · Benjamin A. Pinsky[34] · Annaleise Kealiher[35] · Jeantine E. Lunshof[25,36] · Christopher E. Mason[37,38,39,40] · John Sninsky[41] · Cristina Tato[2] · Timothy Mercer[42] · Thomas J. White[43] · Marc Salit[5,35,44]

✉ Sara Suliman
sara.suliman@ucsf.edu

✉ Marc Salit
msalit@mitre.org

[1] Zuckerberg San Francisco General Hospital, Division of Experimental Medicine, University of California San Francisco, San Francisco, CA 94110, USA

[2] Chan Zuckerberg Biohub, San Francisco, CA, USA

[3] COVID Diagnostics Accelerator Lab, Mass General Brigham, Boston, MA, USA

[4] College of Osteopathic Medicine of the Pacific, Western University of Health Sciences, Pomona, CA, USA

[5] Joint Initiative for Metrology in Biology, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

[6] Bio-Techne, Inc, Austin, TX, USA

[7] London Biofoundry, Imperial College Translation and Innovation Hub, London, UK

[8] Section of Structural and Synthetic Biology, Department of Infectious Disease, Imperial College London, London, UK

[9] UK Dementia Research Institute Centre for Care Research and Technology, Imperial College London, London, UK

[10] INSTAND Society for Promoting Quality Assurance in Medical Laboratories, Düsseldorf, Germany

[11] GBD Gesellschaft Für Biotechnologische Diagnostik mbH, Berlin, Germany

[12] IQVD GmbH Institut Fuer Qualitaetssicherung in Der Virusdiagnostik, Berlin, Germany

[13] National Measurement Laboratory GB, Gaithersburg, MD, USA

[14] Medicines and Healthcare Products Regulatory Agency, MHRA, Blanche Lane, Potters Bar, UK

[15] Microbiologics, St Cloud, MN, USA

[16] Thermo Fisher Scientific, Waltham, MA, USA

[17] ZeptoMetrix LLC, Buffalo, NY, USA

[18] Biodesix Inc., Louisville, CO, USA

[19] OHMX.Bio, an Anacura Company, Evergem, Belgium

[20] Biogazelle, a CellCarta Company, Technologiepark, Zwijnaarde, Belgium

[21] Digital Biology Group, Bio-Rad Laboratories Inc., Pleasanton, CA, USA

[22] Labcorp, Burlington, NC 27215, USA

[23] Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA

[24] Mass General Brigham Center for COVID Innovation, Boston, MA, USA

[25] Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA

[26] Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[27] Department of Laboratory Medicine and Pathology, Division of Clinical Microbiology, Mayo Clinic, Rochester, MN, USA

[28] Department of Pathology and Laboratory Medicine, Medical University of South Carolina, Charleston, SC, USA

[29] Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia

[30] Applied Genetics Group, National Institute of Standards and Technology, Gaithersburg, MD, USA

[31] National Measurement Laboratory (NML), LGC, Queens Road, Teddington, Middlesex, UK

[32] School of Biosciences & Medicine, Faculty of Health & Medical Science, University of Surrey, Guildford, UK

[33] Quest Diagnostic Infectious Disease, San Juan Capistrano, CA, USA

[34] Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA

[35] MITRE Corporation, McLean, VA 22102, USA

[36] Department of Genetics, Harvard Medical School, Boston, MA, USA

[37] Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

[38] The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA

[39] The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA

[40] The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY, USA

[41] SPARK Stanford University, Palo Alto, CA, USA

[42] University of Queensland, Brisbane, Australia

[43] Human Rights Center, School of Law, University of California, Berkeley, CA, USA

[44] Department of Bioengineering, Stanford University, Stanford, CA, USA