A playbook for harmonization of standards for emergent viral pathogens

- 3 Authors: Sara Suliman^{1,2,3,#,*}, Sebastien Fuchs^{4,#}, David Catoe⁵, Loren Hansen⁵, Deepa
- 4 Eveleigh⁶, Michael A. Crone^{7,8,9}, Paul S. Freemont^{7,8,9}, Martin Kammel^{10,11}, Heinz
- 5 Zeichhardt^{11,12}, Andrew T. Anfora¹³, Russell Garlick¹³, Neil Almond¹⁴, Mark Page¹⁴, Brian Beck¹⁵,
- 6 Hui Wang¹⁶, Karuna Sharma¹⁷, Gary Pestano¹⁸, Joannes Vancann¹⁹, Jo Vandesompele¹⁹, Adam
- 7 S Corner²⁰, Jeffrey Albrecht²¹, Deborah J. Boles²¹, Marcia Eisenberg²¹, Ayla B. Harris²¹, Brian J.
- 8 Krueger²¹, Amanda L. Suchanek²¹, Thomas J. Urban²¹, Jonathan D. Williams²¹, Patrick Chain²²,
- 9 Alina Deshpande²², Attelia Hollander²², Sujata Chalise²³, David R. Walt^{23,24,25}, Jeffrey Germer²⁶,
- 10 Joseph D. Yao²⁶, Bobbi S. Pritt²⁶, Frederick S. Nolte²⁷, Alexandra Bogožalec Košir²⁸, Mojca
- 11 Milavec²⁸, Megan H. Cleveland²⁹, Peter M. Vallone²⁹, Eloise J. Busby^{30,31}, Jim F. Huggett^{30,31},
- Denise M. O'Sullivan^{30,31}, Elizabeth M. Marlowe³², Benjamin A. Pinsky³³, Annaleise Kealiher³⁴,
- Jeantine E. Lunshof^{24,35}, Christopher E. Mason^{36,37,38,39}, John Sninsky⁴⁰, Cristina Tato⁴¹, Timothy
- 14 Mercer⁴², Thomas J. White⁴³, Marc Salit^{5,34,44,45,*}
- ¹Zuckerberg San Francisco General Hospital, Division of Experimental Medicine, University of
- 17 California San Francisco, San Francisco, CA, USA
- 18 ²Chan Zuckerberg Biohub, San Francisco, CA, USA
- 19 ³COVID Diagnostics Accelerator Lab, Mass General Brigham, Boston, MA, USA
- ⁴Western University of Health Sciences, College of Osteopathic Medicine of the Pacific,
- 21 Pomona, CA, USA

1

2

15

- ⁵Joint Initiative for Metrology in Biology, SLAC National Accelerator Laboratory, 2575 Sand Hill
- Road, Menlo Park, California 94025, United States.
- ⁶Bio-techne, Inc, Austin, TX, USA
- ⁷London Biofoundry, Imperial College Translation and Innovation Hub, London, UK

- ⁸Section of Structural and Synthetic Biology, Department of Infectious Disease, Imperial College
- 27 London, London, UK
- ⁹UK Dementia Research Institute Centre for Care Research and Technology, Imperial College
- 29 London, London, UK
- 30 ¹⁰INSTAND Society for Promoting Quality Assurance in Medical Laboratories, Düsseldorf,
- 31 Germany
- 32 ¹¹GBD Gesellschaft für Biotechnologische Diagnostik mbH, Berlin, Germany
- 33 ¹²IQVD GmbH Institut fuer Qualitaetssicherung in der Virusdiagnostik, Berlin, Germany
- 34 ¹³LGC Clinical Diagnostics, Product Management, Gaithersburg, MD, USA
- 35 ¹⁴Medicines and Healthcare Products Regulatory Agency, MHRA, Blanche Lane, Potters Bar,
- 36 United Kingdom
- 37 ¹⁵Microbiologics, St Cloud, MN, USA
- 38 ¹⁶Thermo Fisher Scientific, Waltham, MA, USA
- 39 ¹⁷ZeptoMetrix LLC; Buffalo NY, USA
- 40 ¹⁸Biodesix Inc. Louisville, CO, USA
- 41 ¹⁹Biogazelle, a CellCarta company, Technologiepark, Zwijnaarde, Belgium
- 42 ²⁰Digital Biology Group, Bio-Rad Laboratories Inc., Pleasanton, California
- 43 ²¹Labcorp, Burlington, NC, 27215, USA.
- 44 ²²Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA
- 45 ²³Mass General Brigham Center for COVID Innovation, Boston, MA, USA
- 46 ²⁴Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA
- 47 ²⁵Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston,
- 48 MA, USA
- 49 ²⁶Department of Laboratory Medicine and Pathology, Division of Clinical Microbiology, Mayo
- 50 Clinic, Rochester, MN, USA

- 51 ²⁷Department of Pathology and Laboratory Medicine, Medical University of South Carolina,
- 52 Charleston, South Carolina, USA
- 53 ²⁸Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana,
- 54 Slovenia
- 55 ²⁹Applied Genetics Group, National Institute of Standards and Technology, Gaithersburg, MD,
- 56 USA
- 57 ³⁰National Measurement Laboratory (NML), LGC, Queens Road, Teddington, Middlesex, UK.
- 58 ³¹School of Biosciences & Medicine, Faculty of Health & Medical Science, University of Surrey,
- 59 Guildford, UK
- 60 ³²Quest Diagnostic Infectious Disease, San Juan Capistrano, CA, USA
- 61 ³³Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA
- 62 ³⁴MITRE Corporation, McLean, VA USA
- 63 ³⁵Department of Genetics, Harvard Medical School, Boston, MA, USA
- 64 ³⁶Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA
- 65 ³⁷The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational
- 66 Biomedicine, Weill Cornell Medicine, New York, NY, USA
- 67 ³⁸The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY,
- 68 USA
- 69 ³⁹The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY,
- 70 USA

76

- 71 ⁴⁰SPARK Stanford University, Palo Alto, CA, USA
- 72 ⁴¹Chan Zuckerberg Biohub, San Francisco, CA, USA
- 73 ⁴²University Of Queensland, Brisbane, Australia
- 74 ⁴³Human Rights Center, School of Law, University of California, Berkeley, CA, USA
- 75 ⁴⁴Department of Bioengineering, Stanford University, Stanford, CA USA

*Contributed Equally *Co-Corresponding Authors: Sara Suliman, UCSF, San Francisco, CA, USA, 94110, Email: sara.suliman@ucsf.edu Marc Salit, MITRE Corporation, McLean, VA, USA 22102 Email: msalit@mitre.org

Abstract (63 words)

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 and harmonized calibration standards on a shared scale. We present a 'playbook' for an inter-laboratory harmonization study that calibrates any available standards against a limited-availability standard issued by a global authority like the World Health Organization (WHO).

Main Text (1,250 words)

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. 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)⁴, 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 (Figure 1A and B and Supplementary Table 1B). Our intent was for widely available harmonized 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 Figure 1).

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 not intended to be a comparison of tests or laboratories, a survey of test performance, or an evaluation of commutability.

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 tests⁵. Comparable results enable quantitative, systematic diagnostic virology to understand pathogenesis and rapidly integrate data across space and time. Understanding pathogenesis⁶ 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.

What we did:

The study was coordinated by the Joint Initiative for Metrology in Biology, the convener of the CSWG³. 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 (Figure 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) 20/146⁴, the first WHO International Standard for SARS-CoV-2 RNA (WHO-IS) that defined the International Unit (IU, **Figure 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 Figures 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 Figure 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) and two of the 8 standards.

How it came out:

<u>Harmonization results:</u> The 8 standards were all successfully value-assigned to a common, shared reference: the IU (**Figure 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 lab for each of the 8 materials sortable by rank order or laboratory; deviations from median value for each assay; tabular results, and raw data.

<u>Proof-of-Concept results:</u> Calibrating the assays with harmonized standards allowed us to compare results from different laboratories and tests on the clinical samples (Figure 2A). Each laboratory used its own experiment design, representing real-life testing where labs 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 (Figure 2B). Critical analysis of these data and assessment of bias and variability is in the Supplement (Supplementary Figure 3A-C). Lessons learned from this experiment have been incorporated into the playbook recommendations. Results from these studies promise that test results can be aggregated across space-and-time for scientific knowledge development and public-health decision-making.

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 make 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 highlight 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 lab.

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 pathogens¹⁰. 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.

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.

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. SS is supported by grants from the Chan Zuckerberg Biohub, San Francisco and the National Institute of Allergy and Infectious Diseases (NIAID).

Figure Legends:

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

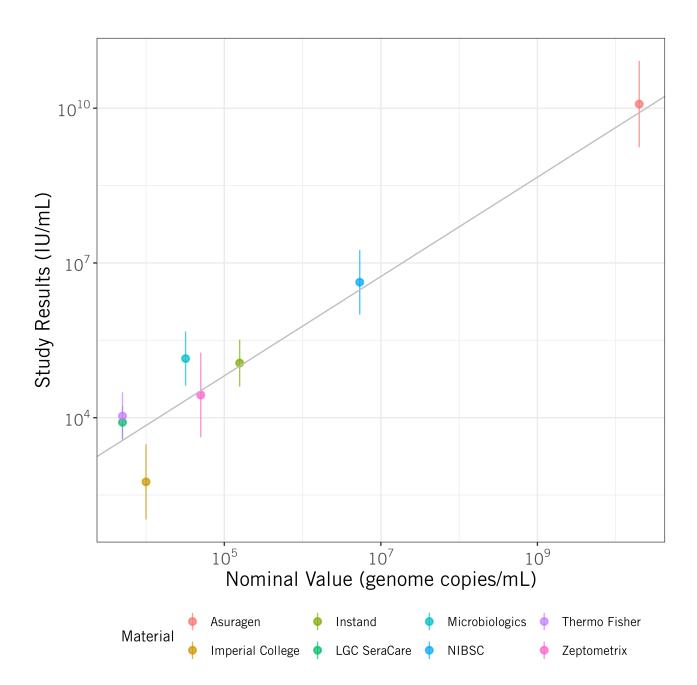
Figure 1: (A) Molecular Standards Harmonization Study Design: Schematic representation of the Coronavirus Standards Working Group (CSWG) molecular standards harmonization study. Four identical aliquots of each of the secondary reference materials were shipped to the 14 participating laboratories. The 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:assay are independent samples from the population. (C) Observed concentrations of each standard in log₁₀ IU mL⁻¹ versus the nominal value in log₁₀ genome copies mL⁻¹. The error bars are 95% Confidence Intervals. 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. Figure 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 labs (Los Alamos National Laboratory (LANL), Biogazelle, and Biodesix), along with 3 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. Labs 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 labs along with two calibrant materials: LGC Clinical Diagnostics, and AcroMetrix. Data is shown after calibration of the raw data from both qPCR and dPCR platforms (Supplementary Figure 3). Supplementary Figure 1: Calibration curves of each individual laboratory and assay using the 6-step dilution curves of the NIBSC 20/146 reference material. All data (RT-qPCR and RTdPCR) have generated standard curves. Curve statistics are in Supplementary table 3. Supplementary Figure 2: Expanded Box and whiskers represent median and interquartile ranges corresponding to the SARS-CoV-2 concentrations in IU, as depicted in Figure 1B in different types of standards: inactivated virus (A), recombinant virus (B), and bacteriophage (C). Supplementary Figure 3: Raw Cq values from LANL (A) and Biogazelle (B), and raw counts from Biodesix (C) of the 15 pools of samples along with the secondary reference standards. Red points correspond to samples with undetectable amplification by PCR.

References:

272

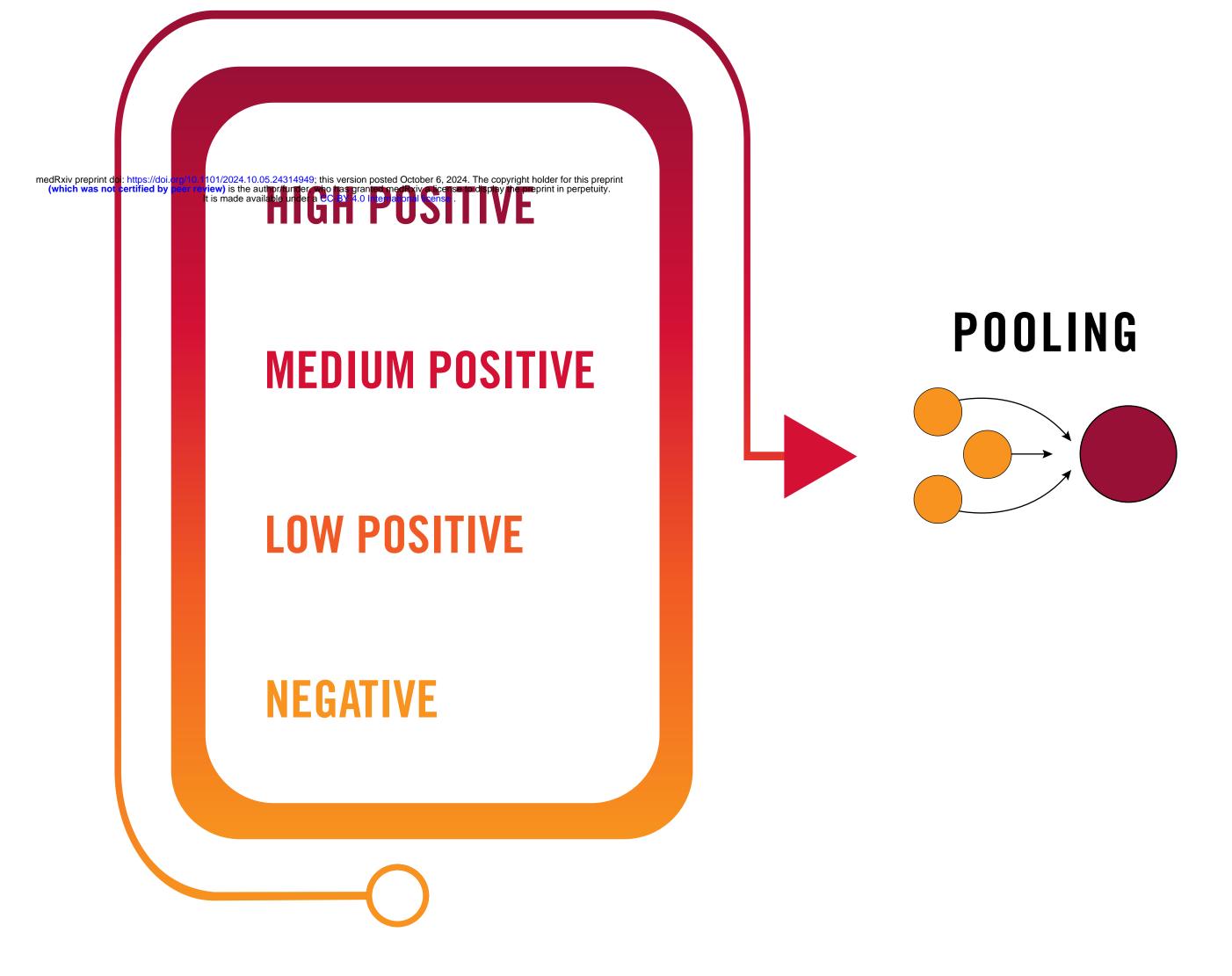
- 273 1 Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* **20**, 533-534 (2020). https://doi.org/10.1016/S1473-3099(20)30120-1
- 276 2 Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860-868 (2020). https://doi.org/10.1126/science.abb5793
- Mercer, T. *et al.* The Coronavirus Standards Working Group's roadmap for improved population testing. *Nat Biotechnol* **40**, 1563-1568 (2022). https://doi.org/10.1038/s41587-022-01538-1
- Nguyen, L. T. *et al.* 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 (2022). https://doi.org/10.1002/jcla.24355
- Evans, D. et al. The Dangers of Using Cq to Quantify Nucleic Acid in Biological Samples: A Lesson From COVID-19. Clin Chem 68, 153-162 (2021). https://doi.org/10.1093/clinchem/hvab219
- Dadras, O. *et al.* The relationship between COVID-19 viral load and disease severity: A systematic review. *Immun Inflamm Dis* **10**, e580 (2022). https://doi.org/10.1002/iid3.580
- 291 7 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-292 Infected Pneumonia. *N Engl J Med* **382**, 1199-1207 (2020). 293 https://doi.org/10.1056/NEJMoa2001316
- 294 8 Lipsitch, M., Krammer, F., Regev-Yochay, G., Lustig, Y. & Balicer, R. D. SARS-CoV-2 295 breakthrough infections in vaccinated individuals: measurement, causes and impact. *Nat* 296 *Rev Immunol* 22, 57-65 (2022). https://doi.org/10.1038/s41577-021-00662-4
- 297 9 Salvagno, G. L. *et al.* SARS-CoV-2 Omicron infection is associated with high nasopharyngeal viral load. *J Infect* (2022). https://doi.org/10.1016/j.jinf.2022.02.025
- WHO. 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. *Annex 6* **Health Product Policy and Standards (HPS)** (2017).



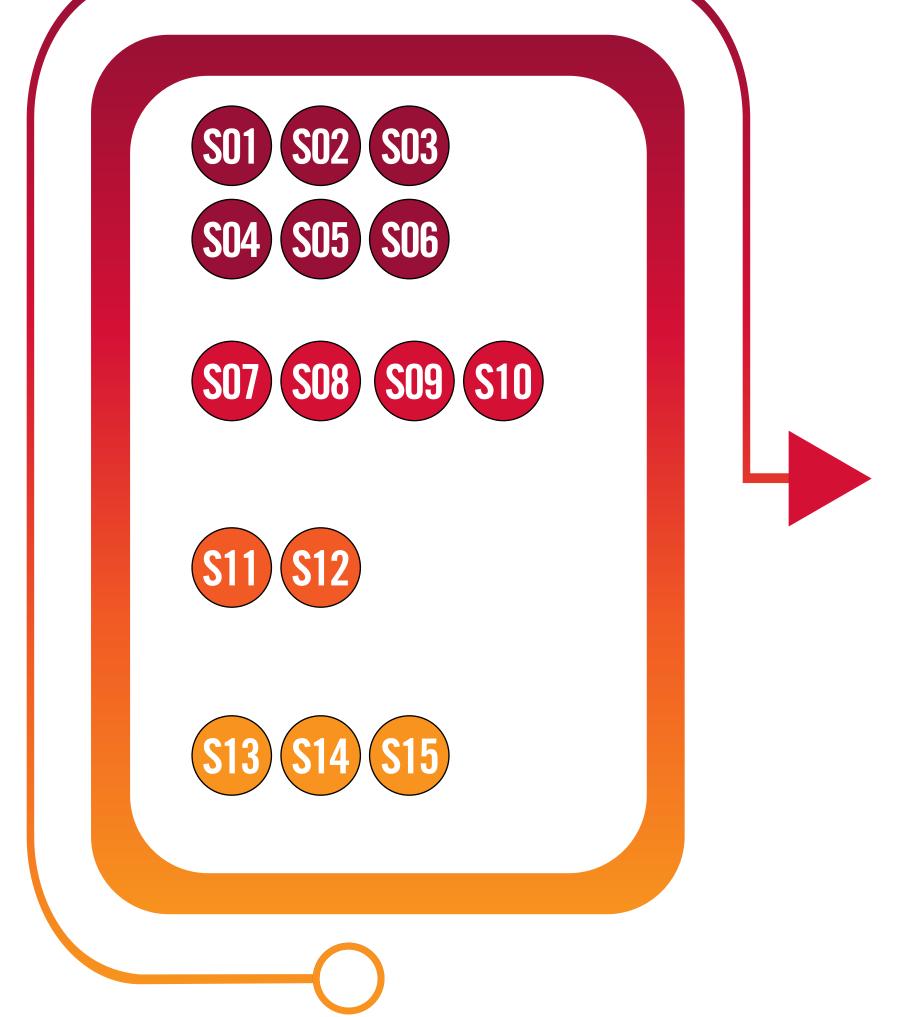


CLINICAL SAMPLES

CLINICAL SAMPLES

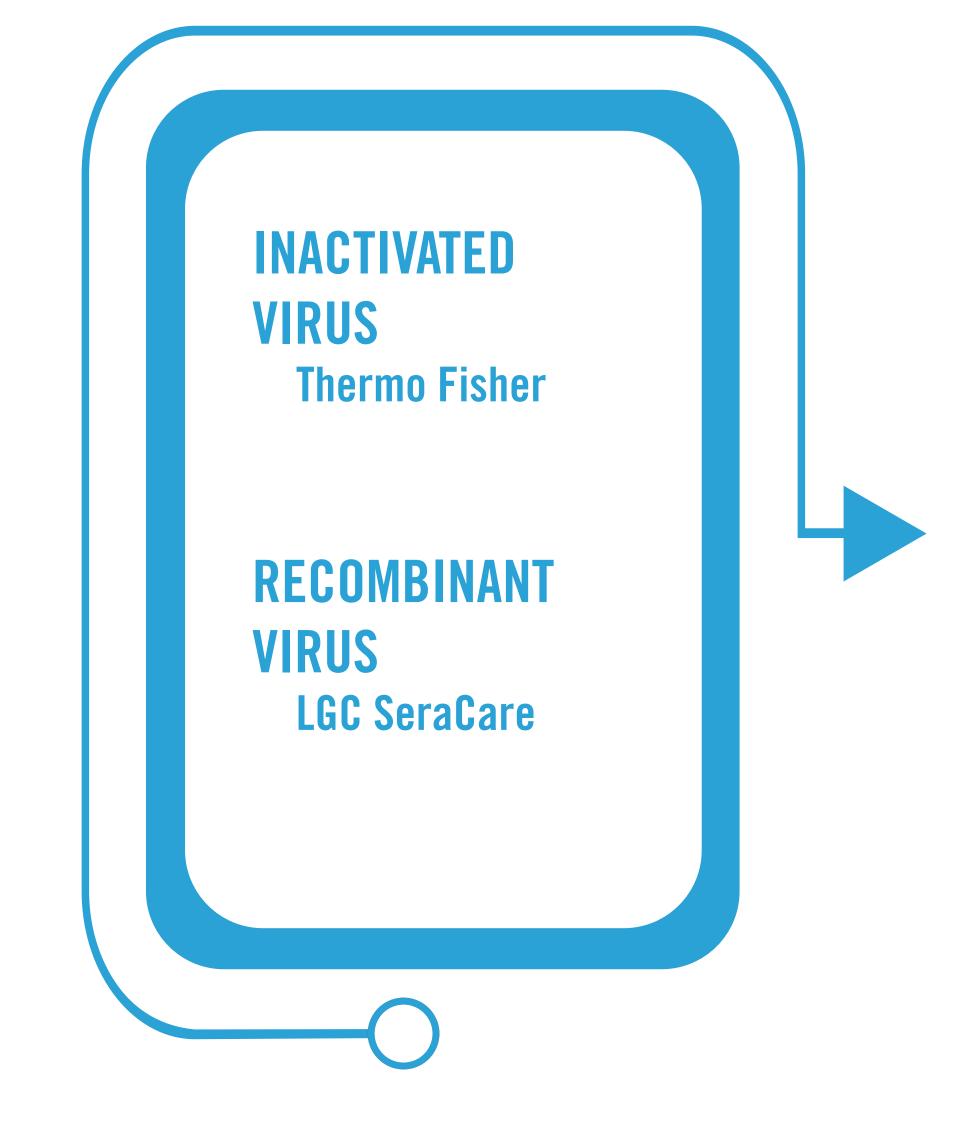


POOLED CLINICAL SAMPLES*



HARMONIZED CALIBRATION

CALIBRANTS



CLINICAL TESTS

PARTICIPATING LABS

