Journal of Veterinary Diagnostic Investigation
2021, Vol. 33(6) 1039–1051

DOI: 10.1177/10406387211029913
jvdi.sagepub.com

Interlaboratory comparison of SARS-CoV2 molecular detection assays in use by U.S. veterinary diagnostic laboratories

Kaiping Deng, Steffen Uhlig, Hon S. Ip, Mary Lea Killian, Laura B. Goodman, Sarah Nemser, Jodie Ulaszek, Shannon Pickens, Robert Newkirk, Matthew Kmet, Kirstin Frost, Karina Hettwer, Bertrand Colson, Kapil Nichani, Anja Schlierf, Andriy Tkachenko, Ravinder Reddy, Renate Reimschuessel[1]

Abstract. The continued search for intermediate hosts and potential reservoirs for SARS-CoV2 makes it clear that animal surveillance is critical in outbreak response and prevention. Real-time RT-PCR assays for SARS-CoV2 detection can easily be adapted to different host species. U.S. veterinary diagnostic laboratories have used the CDC assays or other national reference laboratory methods to test animal samples. However, these methods have only been evaluated using internal validation protocols. To help the laboratories evaluate their SARS-CoV2 test methods, an interlaboratory comparison (ILC) was performed in collaboration with multiple organizations. Forty-four sets of 19 blind-coded RNA samples in Tris-EDTA (TE) buffer or PrimeStore transport medium were shipped to 42 laboratories. Results were analyzed according to the principles of the International Organization for Standardization (ISO) 16140-2:2016 standard. Qualitative assessment of PrimeStore samples revealed that, in approximately two-thirds of the laboratories, the limit of detection with a probability of 0.95 (LOD95) for detecting the RNA was ≤20 copies per PCR reaction, close to the theoretical LOD of 3 copies per reaction. This level of sensitivity is not expected in clinical samples because of additional factors, such as sample collection, transport, and extraction of RNA from the clinical matrix. Quantitative assessment of Ct values indicated that reproducibility standard deviations for testing the RNA with assays reported as N1 were slightly lower than those for N2, and they were higher for the RNA in PrimeStore medium than those in TE buffer. Analyst experience and the use of either a singleplex or multiplex PCR also affected the quantitative ILC test results.

Keywords: interlaboratory comparison; real-time RT-PCR; SARS-CoV2.

Coronavirus infectious disease 2019 (COVID-19), caused by the novel virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV2), has spread rapidly throughout many countries, including the United States, since its discovery in December 2019.[3] On February 4, 2020, the Secretary of the U.S. Department of Health and Human Services (HHS) determined that there was a public health emergency that had a significant potential to affect national security or the health and security of U.S. citizens living abroad. To better understand the scope and spread of the COVID-19 pandemic, U.S. diagnostic laboratories must have accurate testing systems for human and animal specimens. To date, the most commonly used test strategy employs reverse-transcription real-time PCR (RT-rtPCR) to identify the viral RNA from patient specimens. Early in vitro analyses indicated that the RT-rtPCR tests developed in China and Germany were highly specific for SARS-CoV2.[4,28] Since then, additional test systems have been developed and evaluated for recovery and detection of the viral RNA from human samples by test originators. Many of these test systems have received Emergency Use Authorizations (EUAs) from the U.S. Food and Drug Administration (FDA), which has increased testing capacity in the United States.[24]

Based on the similarities of the novel virus with SARS-CoV–like coronaviruses, a zoonotic origin of the COVID-19 outbreak is likely.[1] The World Health Organization is

Division of Food Processing Science and Technology, U.S. Food and Drug Administration, Bedford Park, IL, USA (Deng, Newkirk, Kmet, Reddy); QuoData–Quality & Statistics, Dresden, Germany (Uhlig, Frost, Hettwer, Colson, Nichani, Schlierf); National Wildlife Health Center, U.S. Geological Survey, Madison, WI, USA (Ip); National Animal and Plant Health Inspection Service Laboratories, Veterinary Services, U.S. Department of Agriculture, Ames, IA, USA (Killian); College of Veterinary Medicine, Cornell University, Ithaca, NY, USA (Goodman); Center for Veterinary Medicine, U.S. Food and Drug Administration, Laurel, MD, USA (Nemser, Tkachenko); Institute for Food Safety and Health, Illinois Institute of Technology, Bedford Park, IL, USA (Ulaszek); Cheverly, MD, USA (Reimschuessel).

[1]Corresponding author: Renate Reimschuessel, 2403 Cheverly Ave, Cheverly, MD 20785, USA. rreimsch8888@gmail.com

conducting a large-scale survey of the origins of SARS-CoV2, while researchers throughout the world are also monitoring the risk of spillback from humans to new animal reservoirs.[13,27] Indeed, SARS-CoV2 transmission in mink farms and spillover events to humans has been reported.[14] Animals naturally infected with SARS-CoV2 have been reported in the United States and Europe, including cats, dogs, mink, tigers, and lions.[12,16,17,23,26] Many of these confirmed cases were diagnosed initially using RT-rtPCR methods.[24] RT-rtPCR was also used to test research animals challenged with the virus.[4-6,11,18,19] However, different instrument platforms, RNA isolation procedures based on magnetic bead–nucleic acid extraction or silica-based membrane binding, and a variety of gene targets (*N*, *E*, and *RdRp* genes, alone or combined) were used in these reports.

The tests currently being used by U.S. veterinary diagnostic laboratories have only been evaluated using internal validation protocols. To help laboratories evaluate SARS-CoV2 test methods used for animal samples, an interlaboratory comparison (ILC) was collaboratively conducted by: 1) the FDA–Center for Veterinary Medicine–Veterinary Laboratory Investigation and Response Network (Vet-LIRN), 2) the Moffett Proficiency Testing (PT) Laboratory located at the Institute for Food Safety and Health at the Illinois Institute of Technology (IIT-IFSH; Chicago, IL, USA) and the FDA Division of Food Processing Science and Technology, 3) the U.S. Geological Survey (USGS), 4) QuoData Quality and Statistics (Dresden, Germany), 5) U.S. Department of Agriculture (USDA)–National Animal and Plant Health Inspection Service (APHIS) Laboratories and National Animal Health Laboratory Network (NAHLN), 6) Cornell University (Ithaca, NY, USA), and 7) 42 participating U.S. veterinary diagnostic laboratories.

The objectives of our ILC were to determine if participants could reliably detect SARS-CoV2 RNA at various levels in buffer and virus transport medium, and to compare results and methods from participants. The ILC was designed to provide laboratories with a confidential and structured way to evaluate the method in use or planned to be used. Such a multi-laboratory study also allowed evaluation of the reliability of the results for the SARS-CoV2 RNA test.[7] The primary outcome of our ILC was the qualitative detected or non-detected result. However, the main measures of laboratory performance presented in our ILC were likelihood scores based on cycle threshold (Ct) values and calculated efficiency values. These scores were intended primarily as early indicators of possible future problems in the performance of the qualitative method. Information provided by participants on the test methods used were summarized to discuss possible correlations with the result variabilities. Each laboratory was expected to use the data to evaluate the performance of its own method and to compare its results with those of its peers. The need for this type of study was of major importance given that, during 2020, at least 22 of the laboratories participating in our study received provisional Clinical Laboratory Improvement Amendments (CLIA) certification and were testing human samples.

Materials and methods

SARS-CoV2 RNA preparation

Inactivated SARS-CoV2 RNA was isolated in a laboratory at USGS and shipped to the ILC sample-preparation laboratory on dry ice. Briefly, the SARS-CoV2 virus was propagated in Vero cells (CCL-81; ATCC),[8] and viral genomic RNA was extracted (Applied Biosystems MagMAX-96 viral RNA isolation kit, KingFisher Flex robotic extractor; Thermo Fisher). The extracted RNA was verified for non-infectivity by inoculation into Vero cells, and examined for cytopathic effects (CPEs) daily for 4 d. The cell culture was re-inoculated into a second flask (blind passage) and re-incubated for an additional 4 d. No infectious virus was found, as indicated by the absence of CPE at any stage.

The extracted RNA was quantified prior to shipment by comparing its reverse-transcription quantitative PCR (RT-qPCR) standard curve with that of quantitative synthetic SARS-CoV2 RNA: ORF, E, N (VR-3276SD; ATCC) of known target quantity. Ten-fold serial dilutions of the extracted RNA and the synthetic RNA at 1,000,000, 100,000, 10,000, 1,000, 100, and 10 copies were reverse-transcribed to cDNA and subsequently amplified with specific primers and probes (Integrated DNA Technologies, IDT) targeting 2 regions of the viral *N* gene, following the CDC 2019-nCoV EUA kit method.[2] The RT-rtPCR was run on the Applied Biosystems 7500 fast real-time PCR instrument (Thermo Fisher) with v.2.3 software. The concentration of the extracted RNA stock was determined as 1,000,000 copies/µL by comparing the Ct values of the extracted RNA and synthetic RNA at the same dilutions.

A reverse-transcription droplet-digital PCR (RT-ddPCR)-based back-titration of the samples of RNA in buffer (VM1–4 , Table 1) was performed on an extra kit prepared and stored the same as those that were shipped to the participating laboratories. This was performed by the Cornell University Genomics Facility using the QX200 instrument (Bio-Rad), which uses limiting dilutions of the target in up to 20,000 sub-nanoliter droplets to perform quantification without the use of a standard curve. The CDC N1 assay[2] was used for this analysis with the one-step RT-ddPCR advanced kit for probes (Bio-Rad). Each sample was measured in triplicate. The concentration ranges in copies/µL of the original samples were determined to be 133.2–139.2 (VM1); 50.8–62.8 (VM2); 9,070–9,500 (VM3); and 8,890–9,800 (VM4). The VM5 and VM6 negative controls and a no-template control were all undetected.

Sample homogeneity and stability studies

Acceptable homogeneity and stability were verified in 2 studies. For study 1, homogeneity and stability testing were

Table 1. Concentration and description of the interlaboratory comparison samples.

Sample ID	RNA concentration in sample	Description
VM1	100 copies/5 µL	RT-rtPCR only.
VM2		30-µL sample in Tris-EDTA buffer solution.
VM3	10,000 copies/5 µL	
VM4		
VM5	Negative control	
VM6		
VM7	100 copies/50 µL	Extraction followed by RT-rtPCR.
VM8		150-µL sample in PrimeStore medium.
VM9		
VM10		
VM19		
VM11	10,000 copies/50 µL	
VM12		
VM13		
VM16	100,000 copies/50 µL	
VM17		
VM18		
VM14	Negative control	
VM15		

performed by 2 analysts in 2 trials to ensure that shipped samples were appropriate for testing. During each trial, 5 sets (i.e., one set for each day of testing) of 10 samples (S1–S10) were prepared by adding SARS-CoV2 RNA into PrimeStore molecular transport medium (Longhorn Vaccines & Diagnostics) at levels of 0, 100, 10,000, or 100,000 copies/50 µL. The trial samples were prepared and stored at −80°C in the same manner as ILC shipment samples. Sample sets were tested on days 0, 3, 7, 10, and 15 after preparation. The RNA was isolated from the 50-µL samples (RNeasy mini kit; Qiagen), and 30 µL was eluted from the Qiagen purification column. The purified RNA was reverse-transcribed to cDNA and subsequently amplified as described above.

Study 2 was performed by analyzing 3 sets of randomly chosen ILC samples prepared for shipment to participants; the first set was analyzed in the ILC sample-preparation laboratory prior to the shipment day. Samples VM1–6 in Tris-EDTA (TE) buffer were used directly for the RT-PCR assays without an extraction step; VM7–19 were the PrimeStore samples from which RNA was extracted prior to RT-PCR. The ILC sample-preparation laboratory used the RNeasy mini kit (Qiagen) to purify RNA from samples VM7–19 (Table 1), and PCR was performed using AgPath-ID one-step RT-PCR (Thermo Fisher). The second set was analyzed in an independent laboratory prior to shipment. This laboratory used the MagMAX-96 viral RNA isolation kit for RNA purification and the TaqPath 1-step RT-qPCR kit for PCR (Thermo Fisher). The third set was analyzed in the ILC sample-preparation laboratory 2 d after shipment, using the same materials and RT-PCR method described for the first set.

ILC sample preparation and pre-shipment distribution

The inactivated SARS-CoV2 RNA (in MagMAX elution buffer) for our ILC was provided by USGS as described above and stored at −80°C in the ILC sample-preparation laboratory before use. The RNA was quantified by making 10-fold serial dilutions in TE buffer solution (pH 7.5). Diluted RNA samples (30 µL each) were aliquoted into 1.5-mL screw-top microfuge tubes for samples VM1–6, and 15 µL was added into 135-µL PrimeStore medium in 1.5-mL snap-top microfuge tubes for samples VM7–19 (Table 1). The PrimeStore samples were mixed by pipetting up and down 10 times, and snap-top microfuge tubes were sealed with Parafilm (Amcor) strips. All samples were stored at −80°C before shipping.

A pre-shipment temperature trial was conducted to ensure that sample packages remained frozen during transportation. Packaging configuration was tested by inserting a temperature monitoring device inside the sample container and packaging the samples per the International Air Transport Association Dangerous Goods Regulations (https://www.iata.org/en/publications/dgr/). After holding the container for 72 h at room temperature, the data from the temperature monitoring device were downloaded.

ILC sample distribution

The final shipment samples were packaged (STP-309DI UN 3373 category B frozen insulated shipping system; Saf-T-Pak) according to the manufacturer's instructions and shipped via

FedEx priority overnight. A total of 44 sets of samples were shipped on dry ice to the 42 participating laboratories (2 laboratories requested duplicate sets of samples).

Sample analysis and data acquisition

Participants were instructed to use the SARS-CoV2 RNA extraction and detection method that was used routinely in their laboratory. They analyzed samples VM1–6 (pre-extracted RNA in TE buffer) without extraction steps and analyzed samples VM7–19 in PrimeStore medium after isolating the RNA from the samples. Sample handling and result reporting were discussed with the participants via 2 training sessions. To ensure confidentiality, each laboratory, as well as the laboratory responsible for stability testing, was assigned a laboratory identification number (LIN). Each analyst reported the results as "detected" (D), "not detected" (ND), or "inconclusive" (IN) for SARS-CoV2 viral RNA. The instructions also required the analysts to report Ct values, basic method information, and any modifications. Optionally, participants could specify Ct values for multiple targets. Multiple participants reported *N* gene results designated as "N1 and N2," which were grouped this way for analysis; however, these results were not necessarily generated using the same assays. Detailed methods from each participating laboratory were kept confidential to maintain anonymity.

Qualitative assessment, probability of detection curves

For VM7–19 samples in PrimeStore medium, qualitative assessment was based on the concept of *probability of detection* (POD).[20] For a given number of copies and a given laboratory, the POD is the *theoretical* probability that a gene target is detected. This theoretical probability cannot be observed directly; however, it is considered to govern the observed number of "detects." More specifically, dividing the number of "detects" by the number of replicate test results for a given number of copies of the gene in the sample yields *rate of detection* (ROD) values. For instance, if the laboratories were instructed to perform 5 independent tests at a given number of copies—as is the case for "level 1 (1,000 copies per PCR reaction)"—and if a given laboratory obtains 4 "detects," the corresponding ROD value is 4/5 = 80%. These ROD values are an *estimate* of the underlying POD.

Further details regarding the Poisson assumption and POD curves have been described previously.[10,21] For the qualitative analysis in our study, the cloglog model (i.e., the statistical model that corresponds to the Poisson assumption) was modified to take into consideration the laboratory-specific copy numbers in each PCR reaction because different laboratories used different volumes in the PCR assay.

For qualitative assessment of the results from VM1–6 samples in TE buffer, sensitivity (Se) and specificity (Sp) were calculated as follows:

$$\begin{aligned}\text{Se} &= \text{true-positives} / \begin{pmatrix}\text{true-positives} \\ + \text{ false-negatives}\end{pmatrix} \\ &= \text{true-positives} \\ &\quad / \text{test portions containing the gene.}\end{aligned}$$

$$\begin{aligned}\text{Sp} &= \text{true-negatives} / \begin{pmatrix}\text{true-negatives} \\ + \text{ false-positives}\end{pmatrix} \\ &= \text{true-negatives} \\ &\quad / \text{test portions not containing the gene.}\end{aligned}$$

We did not include any confounding organisms in Sp determination.

Quantitative assessment

The quantitative assessment was based on 3 parameters: submitted Ct values, adjusted Ct values, and calculated efficiency values. Mean values and standard deviations (SDs) for the 3 parameters were calculated according to the Q/Hampel method[9] and the approach described for the weighted Hampel mean.[22] The calculations were performed using the software PROLab Plus.[15]

Adjusting Ct values. Because each participating laboratory used a method that it had used or planned to use to test the ILC samples, and extraction sample size and RT-PCR volume varied across the participating laboratories, the *actual* copy numbers per well differed from laboratory to laboratory. Therefore, the Ct values reported represented different viral copy numbers. To ensure comparability of the Ct values and thus a meaningful assessment of laboratory performance, we adjusted the submitted Ct values. For a given laboratory, the *adjusted* Ct value corresponded to the value the laboratory would have obtained had the *nominal* number of copies per well coincided with the *actual* number. The adjustment factor was calculated on the basis of mean efficiency across laboratories.

Calculation of efficiency. Efficiency was calculated separately for each of the 2 sample types (TE, PrimeStore), and separately for the N1 and N2 markers as reported by the participants, on the basis of the *nominal* copy numbers, and on the basis of the submitted (non-adjusted) Ct values.

For a particular dataset, the efficiency was calculated as follows:

$$\text{efficiency} = \left(10^{-\frac{1}{b}} - 1\right) \times 100$$

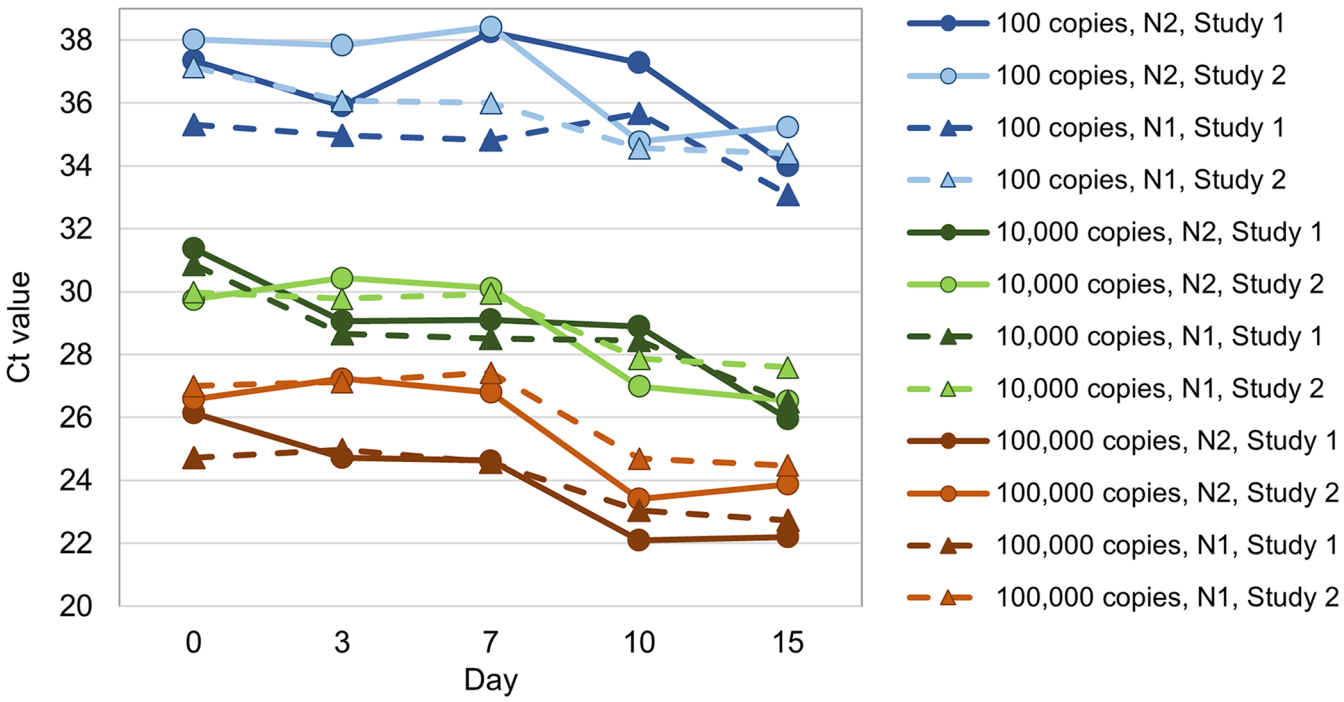

Figure 1. Test results (mean cycle threshold [Ct] values for markers N1 and N2) throughout the homogeneity and stability study.

where b denotes the slope parameter of a linear regression analysis of Ct versus $\log_{10}$ (nominal number of copies).

Calculation of likelihood scores. Likelihood scores were computed by Bayesian statistics as a statistical estimate of the success rate in future proficiency testing. This is true as long as the analytical methods and procedures are not modified. Likelihood scores do not provide an indication of success in this ILC. *Likelihood scores for sensitivity* were computed on the basis of the adjusted Ct values. These scores were calculated separately for each marker and for each sample type, resulting in 4 different scores per laboratory. *Likelihood scores for efficiency* were computed on the basis of the efficiency values. For each laboratory, one score was provided and pooled across assay markers and sample types.

Evaluation of factorial effects. Factorial effects were evaluated on the basis of Ct values, using additional information provided by the participating laboratories, such as the instrument used and whether the test was performed routinely. This calculation was performed using the least absolute shrinkage and selection operator method (regression analysis with penalization).

Results

Homogeneity and stability

For the homogeneity and stability study prior to ILC sample preparation (study 1), qualitative data indicated that all inoculated samples were detected and that blank samples were not detected. The Ct values obtained from the homogeneity and stability study (study 1) were subjected to quantitative analysis (Fig. 1). The homogeneity SD represents the variability of the samples adjusted for analytical variability and variation between days. In all cases, it fell within a range of 0.14–0.61 Ct (Table 2). Hence, the samples could be considered homogeneous for all 3 concentration levels, and for both N1 and N2 assay markers.

Regardless of the testing day (day 0–15), a decrease by 2–4 Ct values was observed for all series. Decreases corresponded to an increase in the nominal concentration level by a factor of 2–16. The corresponding trends were statistically significant ($p < 0.05$; critical value for 3 df = ±3.18). It is unlikely that these trends corresponded to actual differences in sample concentrations. We concluded that the storage of samples for 15 d did not cause an upward trend of the measured Ct values as a result of decreases in sample concentration. The samples were deemed sufficiently homogeneous and stable, and the inoculation process was suitable to produce the targeted ILC samples.

For marker N1, post-shipment sample Ct values were always higher than the respective pre-shipment sample value (Suppl. Table 1; Suppl. Fig 1). The average difference between post-shipment and pre-shipment values was 0.4 Ct. For marker N2, only slight random fluctuations were observed. The average difference was <0.1 Ct. This was consistent with the observation in study 1, in which the trend from day 0 to day 15 did not continue for either N1 or

Table 2. Standard deviation analysis of homogeneity and stability, study 1.

Marker	Copies	s_{sample}	s_e			s_{day}	
			(A1,M1)	(A1,M2)	(A2,M2)	(A1,M1)	(A2,M2)
N1	100	0.34	1.25	0.45	0.73	0.44	1.01
N2		0.61	0.69	0.37	0.49	0.65	1.29
N1	10,000	0.49	1.21	0.10	0.34	0.94	1.02
N2		0.41	0.51	0.21	0.18	1.10	1.33
N1	100,000	0.32	0.41	0.17	0.31	0.91	1.17
N2		0.14	1.27	0.36	0.14	1.28	1.33

s_e = SD of the variation of the duplicate determinations (within samples, within day, specific for each analyst [A1/A2] and machine [M1/M2]); S_{day} = SD between days (specific for each analyst [A1/A2] and machine [M1/M2]); s_{sample} = SD between samples (on the same day).

Table 3. Rate of detection (ROD) depending on the level for N1 and N2 markers as well as the "overall detection."

Level	No. of PCR replicates	ROD across laboratories		
		N1 35 laboratories (%)	N2 35 laboratories (%)	Overall detection 42 laboratories (%)
Blank	2	0	0	0
Low	5	96.0	97.7	96.2
Medium	3	98.1	98.1	98.4
High	3	100	100	100

Low, medium, and high levels were 100, 10,000, and 100,000 copies per 50 µL of PrimeStore, respectively.

N2. This confirmed our assumption that the observed downward trends between days 0 and 15 were not related to the copy numbers in the samples and that the samples should be considered stable.

ILC result submission

The data of the pre-shipment temperature trial showed that the packaging configuration kept the primary sample container frozen for 72 h. All participants confirmed the successful delivery of the frozen packages without any delay.

We shipped 44 sets of samples to 42 laboratories; 2 laboratories requested duplicate sets of samples, but only 1 of those laboratories provided 2 sets of results. One other laboratory withdrew from the ILC. Therefore, we received 42 datasets from 41 laboratories.

It should be noted that the submitted qualitative results for "overall detection" (Supplemental Table 2) were based on different markers and criteria for Ct values selected by the individual laboratories. Results for targets other than N1 and N2, such as the *E* gene, may have also been used to interpret whether a sample was "detected" or "non-detected." Although all of the laboratories reported results for the "overall detection," only 35 of 41 laboratories submitted Ct values attributed to N1 and N2 markers (Suppl. Table 3). The Ct values for the detection of other gene markers are not shown.

Qualitative assessment of TE samples

For the VM1–6 samples in TE buffer solution, the qualitative test results were all "not detected" for the blank samples and all "detected" for the samples containing SARS-CoV2 inactivated RNA (Suppl. Table 2). Therefore, the qualitative assessment was limited to the calculation of the Se and Sp. Se is the ability of a method to detect the target organism, whereas Sp is the method's ability to discriminate between target and non-target organisms. Both Se and Sp were 100% for the TE samples.

Qualitative assessment of PrimeStore samples

For the VM7–19 samples in PrimeStore medium, the qualitative assessment was performed separately for detection of N1, detection of N2, and overall detection. For each of the 3 evaluations, the basis of the assessment was the number of positive results ("detects") per laboratory and the number of copies. Given differences in volumes used for nucleic acid extraction and volumes used for each PCR reaction across the participating laboratories, *actual* copy numbers per well differed from laboratory to laboratory (Table 3).

Evaluation of PCR amplification rate

With few exceptions, for the N1 and N2 markers and for the overall detection, observed results are compatible with the assumption that the PCR method is able to amplify all of the RNA copies present in the samples. This assumption applies to our ILC and would not be expected in a clinical setting. In the case of N1, statistically significant exceptions were noted for 3 of the 35 laboratories. The same was seen for N2 (not the same laboratories). In the case of "overall detection," the

Table 4. Overview of laboratory results in which false-negative results were obtained.

LIN	No. of copies used for PCR	Detection of N1		Detection of N2		Overall detection	
		Detected/total	*p*	Detected/total	*p*	Detected/total	*p*
1	2.5	0/5	<0.001				
	250	3/3	1				
	2,500	3/3	1				
2	2.5	4/5	0.348	3/5	0.057	4/5	0.348
	250	3/3	1	3/3	1	3/3	1
	2,500	3/3	1	3/3	1	3/3	1
4	13.333	No Ct values submitted				0/5	<0.001
	1,333.3					3/3	1
	13,333					3/3	1
24	3.571	5/5	1	5/5	1	5/5	1
	357.1	1/3	<0.001	1/3	<0.001	1/3	<0.001
	3,571	3/3	1	3/3	1	3/3	1
39	5	No Ct values submitted				4/5	0.033
	500					3/3	1
	5,000					3/3	1
40	10			4/5	<0.001		
	1000			3/3	1		
	10,000			3/3	1		
42	6.667	4/5	0.006	4/5	0.006	4/5	0.006
	666.7	3/3	1	3/3	1	3/3	1
	6,667	3/3	1	3/3	1	3/3	1

LIN = laboratory identification number.
Note that $p < 0.05$ indicates that not all copies were detected.

exceptions concern 4 of 42 laboratories. Seven laboratories submitted results containing false-negatives (Table 4).

The exceptions mentioned in the previous paragraph are statistically significant if the *p* value is <0.05. The calculation of *p* values takes into account the actual volume used for the PCR procedure. Three important observations can be made from the evaluation (Table 4):

1) For 1 of the 7 laboratories (LIN 2), there was not enough evidence to reject the assumption of 100% amplification rate. In other words, the result is consistent with 100% amplification rate (percentage of gene markers in a sample that are amplified).
2) For 5 of the 7 laboratories, the observed results could be explained by an amplification rate <100%.
3) For LIN 24, the observed results could not be attributed to any single amplification rate, given that all 5 replicates at the low level were correctly detected, but 2 of 3 replicates at the medium level (with 100 times more copies than the low level) were false-negatives.

Apart from the 6 laboratories listed in Table 4 (excluding LIN 2), all laboratories identified all positive samples. This is not to suggest that the PCR method had a perfect level of performance. Instead, the very high level of performance should be understood in relation to the fact that, for the great majority of laboratories, the copy numbers in the PCR reaction were so high that—assuming high Se—negative results were very unlikely.

POD and laboratory-specific upper limit for the LOD95

Given the very limited data basis and the considerable differences between volumes used for nucleic acid extraction and volumes for each PCR reaction, it was not possible to determine the exact POD curve of each laboratory and thus the exact laboratory-specific level of detection with a probability of 0.95 (LOD95). However, it was possible to calculate an upper limit for the LOD95 for each laboratory. In the ideal case, all copies in a well are amplified, and the LOD95 is 3 copies per PCR reaction (Fig. 2).

For the N1 marker, the upper limit for the LOD95 was 17–88 copies per PCR reaction, apart from one laboratory, and for the N2 marker, 9–88 copies per PCR reaction (Table 5). For the "overall detection," the LOD95 was 9–88 copies per PCR reaction. It should also be noted that, for one laboratory, the upper limit for the LOD95 values could not be calculated because its results for the low and medium levels were highly inconsistent. It is apparent that differences in the elution volume can affect Se, and laboratories may consider adjusting volumes to maximize detection.

For almost two-thirds of all laboratories, the upper limit for the LOD95 is no higher than 21 copies per PCR reaction

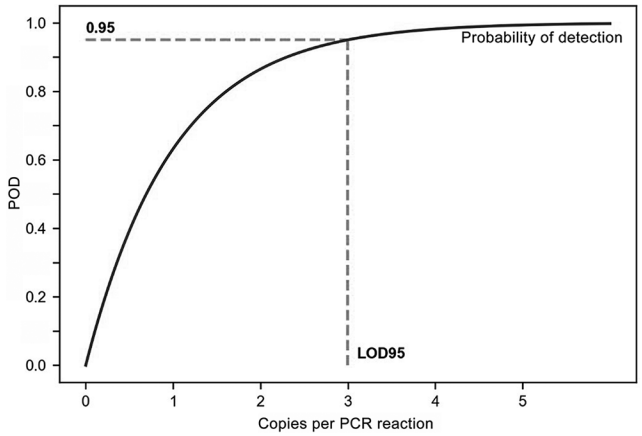

Figure 2. Calculation of LOD95 (vertical gray line) on the basis of a hypothetical probability of detection (POD) curve. LOD95 is the number of copies at which a POD of 95% is achieved.

(Table 5). The upper limit for the LOD95 does not characterize the performance of a laboratory. Rather, it is the upper limit of the interval in which the actual LOD95 of this laboratory lies. For the majority of laboratories, the upper limit for the LOD95 is already very close to the theoretical best value of 3 copies per PCR reaction.

The upper limit for the LOD95 value can also lie well above 21 copies per PCR reaction, even in the absence of false-negative results (Table 5). These higher upper limits for the LOD95 values are caused by higher copy numbers per PCR reaction given different extraction, elution, and PCR reaction volumes. It should also be noted that, given a higher copy number in the PCR reaction, detection of lower concentrations may be possible despite the higher limit for LOD95.

Quantitative assessment

The least absolute shrinkage and selection operator method (regression analysis with penalization) was applied to identify influence factors for the Ct values. The following factors were identified: instrument, routine versus non-routine analysis, singleplex versus multiplex, and extraction kit (Table 6). Measurements using singleplex recorded a higher Ct value for all samples of the N1 target region, for both PrimeStore (on average higher by 0.693) and TE (on average higher by 0.368). There may be no obvious technical explanation for such an effect. However, from a statistical point of view, a reason may be that the number of laboratories using multiplex assays is much lower (~5) than the number using singleplex assays (~30). Therefore, it is possible that other confounding factors may lead to such results.

Ct values for the samples prepared using the viral extraction kit for the N1 marker for all samples were, on average, 0.12 lower than the results from using the pathogen extraction kit. The reason for this is unknown. Furthermore, the Ct values for samples measured in PrimeStore medium for the 3 sample levels were on average 0.2 lower for analysts who perform the test routinely compared to the ones who do not perform the test routinely. An explanation for this effect may be that an analyst who performed the assay routinely had more experience handling small volumes in the low to single-digit microliter range typical for PCR methods.

In all cases, the results for samples at particular concentration levels confirm the outcome seen when all samples are considered together. It is important to note that, for the interpretation drawn from the values presented in Table 6, the values are an estimate.

Statistical parameters

Although the repeatability SD characterizes how a parameter varies inside a laboratory under near-identical testing conditions on average across all laboratories, the reproducibility SD characterizes the way a given parameter varies between laboratories (Tables 7–9). In general, both parameters are method parameters. In our study, these parameters characterized the variation not only for one specific method but for several PCR methods. However, because the distribution of the efficiency and Ct values is almost normal, it is statistically acceptable to characterize the ILC result by these overall parameters (i.e., by repeatability SD [$s_{r;ILC}$] and reproducibility SD [$s_{R;ILC}$]).

For the efficiency values, reproducibility SD values were greater for the TE samples than for the PrimeStore samples. This is because the narrow range of copy numbers for the TE samples resulted in considerably greater random variability.

The repeatability SD fell in the range 0.18–0.69, which is very similar to the range of the sample SD calculated in the homogeneity study 1 (0.14–0.61; Table 2).

Reproducibility was greater for the PrimeStore medium samples. This was to be expected because an extra extraction step was required. The SDs assigned to the extraction step were 0.7–1.0 for the N1 marker and 1.1–1.5 for the N2 marker. It is evident that reproducibility was better for N1 than for N2.

For both types of samples, the Ct values allow a semi-quantitative determination of copy numbers.

Likelihood scores

The likelihood scores for efficiency and sensitivity are presented in Supplemental Table 4. The *likelihood score for efficiency* represents the probability that the amplification efficiency in a laboratory is significantly different from the other participants. The *likelihood score for sensitivity* represents the probability that the adjusted Ct values of a laboratory are significantly higher than those of the other participants. Such a significantly higher Ct value indicates poorer reproducibility and/or poorer sensitivity than the other participants. These likelihood scores are intended primarily as an early indicator of possible future problems with the performance of the qualitative method.

Table 5. Overview of qualitative results, the actual copy number in the PCR reaction, and the upper limit for the LOD95.

	No. of copies in PCR reaction & no. of detects of N1/N2/Overall						Upper limit for LOD95 (copies in PCR reaction)		
LIN	Level 1 (5 replicates)		Level 2 (3 replicates)		Level 3 (3 replicates)		N1	N2	Overall detection
1	2.5	0/5/5	250	3/3/3	2,500	3/3/3	1,630	9	9
2	2.5	4/3/4	250	3/3/3	2,500	3/3/3	18	36	18
3	11.7	-/-/5	1,167	-/-/3	11,667	-/-/3	-	-	44
4	13.3	-/-/0	1,333	-/-/3	13,333	-/-/3	-	-	8,700
5	20	5/5/5	2,000	3/3/3	20,000	3/3/3	75	75	75
6	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
7	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
8	20	-/-/5	2,000	-/-/3	20,000	-/-/3	-	-	75
9	17	5/5/5	1,667	3/3/3	16,667	3/3/3	63	63	63
10	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
11	10	5/5/5	1,000	3/3/3	10,000	3/3/3	38	38	38
12	8.3	5/5/5	833	3/3/3	8,333	3/3/3	31	31	31
13	16	-/-/5	1,600	-/-/3	16,000	-/-/3	-	-	60
14	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
15	8.9	5/5/5	889	3/3/3	8,889	3/3/3	33	33	33
16	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
17	10	5/5/5	1,000	3/3/3	10,000	3/3/3	38	38	38
18	4.4	5/5/5	444	3/3/3	4,444	3/3/3	17	17	17
19	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
20	23	5/5/5	2,333	3/3/3	23,333	3/3/3	88	88	88
21	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
22	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
23	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
24	3.6	5/5/5	357	1/1/1	3,571	3/3/3	-*	-*	-*
25	8.9	5/5/5	889	3/3/3	8,889	3/3/3	33	33	33
26	5	5/5/5	500	3/3/3	5,000	3/3/3	19	19	19
27	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
28	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
29	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
30	5	5/5/5	500	3/3/3	5,000	3/3/3	19	19	19
31	10	5/5/5	1,000	3/3/3	10,000	3/3/3	38	38	38
32	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
33	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
34	20	-/-/5	2,000	-/-/3	20,000	-/-/3	-	-	75
35	9.3	-/-/5	933	-/-/3	9,333	-/-/3	-	-	35
36	5	5/5/5	500	3/3/3	5,000	3/3/3	19	19	19
37	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
38	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
39	5	-/-/4	500	-/-/3	5,000	-/-/3	-	-	36
40	10	5/4/5	1,000	3/3/3	10,000	3/3/3	38	71	38
41	5.6	5/5/5	556	3/3/3	5,556	3/3/3	21	21	21
42	6.7	4/5/5	667	3/3/3	6,667	3/3/3	48	48	48
43	17	5/5/5	1,667	3/3/3	16,667	3/3/3	63	63	63

LOD95 = limit of detection with a probability of 0.95. The notation 0/5/5 means 0 "detects" for N1, 5 "detects" for N2, and 5 "detects" for overall. Dash (-) indicates cases where the test for the corresponding marker was not performed (a Ct value not reported). Cells in which at least 1 false-negative is observed are shaded (e.g., in the cell corresponding to Level 1 and LIN 40, only 4 "detects" for N2 correspond to 1 false-negative).

* The upper limit for LOD95 cannot be calculated given statistically inconsistent results at the low and medium levels.

The very uniform efficiency in almost all laboratories simplified the check for deviating results. Such deviations can indicate methodologic problems, even if Ct values determined are inconspicuous when viewed from the outside. Laboratory LIN 21 was particularly noticeable in this context. The efficiency for their TE samples differed significantly from the

Table 6. Effects of influence factors for Ct values for **A.** Tris-EDTA samples and **B.** PrimeStore samples.

		A. 30-µL sample in Tris-EDTA buffer solution							
		N1	N2	N1			N2		
		All samples	All samples	100 copies/5 µL		10,000 copies/5 µL	100 copies/5 µL		10,000 copies/5 µL
Instrument used	ABI 7500 (vs. other)	NE	–0.003	NE		–0.402	NE		NE
Analyst routinely	Yes	NE	NE	NE		–0.290	NE		NE
Singleplex or multiplex	Singleplex	0.368	0.244	NE		1.056	NE		NE
Extraction kit	Pathogen	NE	NE	NE		–0.230	NE		NE
	Viral	NE	NE	NE		NE	NE		NE
		B. 150-µL sample in PrimeStore medium							
		N1	N2	N1			N2		
		All samples	All samples	100 copies/50 µL	10,000 copies/50 µL	100,000 copies/50 µL	100 copies/50 µL	10,000 copies/50 µL	100,000 copies/50 µL
Instrument used	ABI 7500 (vs. other)	–0.097	–0.191	NE	–1.354	–1.173	NE	–0.448	–0.258
Analyst routinely	Yes	–0.204	NE	–0.362	–0.665	–0.317	NE	NE	NE
Singleplex or multiplex	Singleplex	0.693	0.551	0.320	1.972	1.469	NE	0.606	0.046
Extraction kit	Pathogen	NE	NE	NE	NE	NE	NE	NE	NE
	Viral	–0.118	–0.016	–0.069	–0.500	–0.283	NE	–0.009	NE

NE = no effect of the factor in question could be observed. The values in the rows corresponding to the 4 influence factors listed in the first column characterize the magnitude of the associated effect. For example, for the *N1* gene in 150-µL sample in PrimeStore medium, the cycle threshold (Ct) values were on average 0.204 lower for analysts who routinely perform the test in comparison to the ones who do not routinely perform the test.

Table 7. Statistical parameters for the efficiency in both sample types and for both markers.

	30-µL Tris-EDTA samples		150-µL PrimeStore samples	
Efficiency	N1	N2	N1	N2
No. of laboratories included in the calculation	35	35	35	35
Hampel mean (%)	91.3	86.3	98.1	92.0
Reproducibility SD $s_{R;ILC}$ (% abs.)	12.5	11.0	4.3	8.0

Table 8. Statistical parameters for adjusted and non-adjusted cycle threshold (Ct) values in 30-µL Tris-EDTA samples.

	100 copies/5 µL		10,000 copies/5 µL	
Statistical parameter	N1	N2	N1	N2
Non-adjusted Ct values				
No. of laboratories that submitted results	35	35	35	35
Weighted Hampel mean	29.46	30.01	22.67	22.99
Repeatability SD $s_{r;ILC}$	0.69	0.54	0.37	0.27
Reproducibility SD $s_{R;ILC}$	1.44	1.83	1.38	1.55
Adjusted Ct values				
No. of laboratories that submitted results	35	35	35	35
Weighted Hampel mean	29.54	30.05	22.78	23.06
Repeatability SD $s_{r;ILC}$	0.69	0.54	0.37	0.27
Reproducibility SD $s_{R;ILC}$	1.45	1.79	1.35	1.55

Table 9. Statistical parameters for adjusted and non-adjusted cycle threshold (Ct) values in 150-µL PrimeStore samples.

	100 copies/50 µL		10,000 copies/50 µL		100,000 copies/50 µL	
Statistical parameter	N1	N2	N1	N2	N1	N2
Non-adjusted Ct values						
No. of laboratories that submitted results	34	35	35	35	35	35
Weighted Hampel mean	34.38	34.72	27.40	27.67	23.97	24.08
Repeatability SD $s_{r;ILC}$	0.47	0.46	0.18	0.25	0.18	0.22
Reproducibility SD $s_{R;ILC}$	1.68	2.25	1.64	2.21	1.62	2.15
Adjusted Ct values						
No. of laboratories that submitted results	34	35	35	35	35	35
Weighted Hampel mean	32.70	32.97	25.92	26.44	22.44	22.39
Repeatability SD $s_{r;ILC}$	0.47	0.46	0.18	0.25	0.18	0.22
Reproducibility SD $s_{R;ILC}$	1.59	2.22	1.69	2.15	1.62	2.11

efficiencies of the other laboratories (Supplemental Figs. 2–13). The average difference in the Ct value between the 2 concentration levels (100 copies and 10,000 copies/50 µL TE) was ~4.4 (29–24.6; Supplemental Table 3). This value was significantly lower than the corresponding values of all other laboratories and implied an efficiency that was significantly >100% (i.e., significantly higher than can be explained theoretically). Whether the cause was from an error in the PCR or in the sample preparation is not clear. In any case, it is noticeable that the discrepancy that occurred manifested itself not only in 2 individual samples but also in the respective replicates.

For the adjusted Ct values themselves, the likelihood score was determined separately for each marker and for each sample type. In most cases, only small differences between markers and between sample types were observed, which indicated that performance was less dependent on marker and sample type and more on the PCR step itself.

Discussion

The results of our ILC demonstrate excellent results obtained by veterinary diagnostic laboratories testing for SARS-CoV2 in animals. The results of all inoculated samples in TE buffer

being "detected," and all blank samples being "not detected," indicate that Se and Sp were 100%. However, it should be noted that the blank samples did not contain any other non–SARS-CoV2 viral RNA; we did not include confounding organisms in this ILC. Thus, our study design did not allow calculation of the Sp of the method as the ability to discriminate between target and non-target organisms.

Qualitative assessment of PrimeStore samples by POD analysis revealed that, in approximately two-thirds of the laboratories, the LOD95 was 3–20 copies per PCR reaction. It is apparent that differences in the elution volume can affect Se, and laboratories may consider adjusting volumes to maximize detection, which is also of importance considering the analysis of pooled samples. Within our ILC, the differences in the volumes used resulted in differences in copy numbers per PCR reaction by a factor of up to 9.3.

Quantitative assessment of the submitted data indicates that each laboratory seemed to have adapted its own nucleic acid extraction method and PCR assay in such a way that, despite large differences in the volumes and methods used across laboratories, the resulting Ct values were similar. Therefore, the Ct values can be considered as semi-quantitative results of the number of copies in the samples. A more accurate determination of the virus concentration could be achieved by calibration using a PCR control sample.

The reproducibility SD ($s_{R;ILC}$) of both non-adjusted and adjusted Ct values was ~2 Ct values. The SD was slightly lower for N1 and slightly higher for N2. The reproducibility SD ($s_{R;ILC}$) for PrimeStore medium samples was generally higher than that for TE buffer solution samples. This can be explained by the fact that, for PrimeStore samples, nucleic acid extraction was performed by the laboratories, which is an additional source of variability. In contrast, the TE buffer solution samples contained pre-extracted RNA.

General differences in the sensitivities for *N* markers have been reported previously.[25] The outcomes of our study provide further evidence about the effect of extraction kits, laboratory analyst, and instrument affecting the variation of Ct values for different markers. Average PCR amplification efficiency values fell within the range of 86–98%. It is particularly noticeable that the variability of the laboratory-specific efficiency values of the PrimeStore samples for N1 was very low. Thus, almost no laboratory had an efficiency value <90%. In comparison, the variability of the efficiency values for N2 was clearly greater.

Most laboratories achieved good test Se, a necessary and reassuring result. However, insights gained from our study point to the importance of reporting the LOD and the volumes used along with the Ct values. A virtue of LOD values accompanying the test results is that it can potentially help to explain conflicting (negative) test results, especially when LOD values vary greatly from laboratory to laboratory.

Overall, our ILC facilitated the laboratories' validation efforts by providing standardized test materials and statistical comparison of results with those of peer laboratories. The data show that, for RNA in TE buffer, 100% of samples were detected. The data also show that, where extraction was required, the PCR methods used provided almost perfect results. In a future ILC, more technically challenging samples (samples with a low viral load) or in different matrices could be chosen to simulate more complex human and animal specimens and facilitate the determination of LOD95 of the RT-rtPCR method. Demonstrating that their assays are sensitive and provide accurate results is important to the veterinary laboratory community to ensure reliable diagnosis of SARS-CoV2–infected animals. Establishing test reliability became even more urgent in the summer of 2020 because veterinary diagnostic laboratories began testing human patient samples in addition to animal samples. Veterinary laboratories obtained provisional CLIA certification to support the public health efforts with respect to the COVID-19 pandemic and have since then tested a substantial number of human samples, exemplifying the bonds within the One Health continuum.

By conducting our ILC, the collaborating groups across government agencies, universities, and private industry made a profound contribution to protect human and animal health by demonstrating the validity and performance of current SARS-CoV2 RNA tests during the challenging times of the COVID-19 pandemic. Although animal-adapted methods were used here, many observations made are applicable in the context of clinical diagnosis using human specimens. Prompted by our ILC study, future studies can be designed to get an unequivocal picture of key factors affecting test results, ultimately contributing to the advancement of pathogen detection.

Acknowledgments

We acknowledge the following individuals for excellent technical assistance and administrative support: Angelica Jones, Olgica Ceric, Ellen Hart, and Dave Rotstein from FDA, and Mia Torchetti and Christina Loiacono from USDA. We also thank Kirsten Simon from QuoData for supporting and facilitating this work. The Genomics Facility of the Biotechnology Resource Center at the Cornell University Institute of Biotechnology assisted with copy number quantification, and we thank Peter Schweitzer for facilitating this. We also acknowledge the diligence and hard work of the laboratory scientists who rapidly developed SARS-CoV-2 assays for animals and participated in this collaborative study.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The ILC was organized and conducted by the FDA Vet-LIRN network in collaboration with USDA and USGS. Laboratories were not charged to participate in this exercise.

Disclaimer

The use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

ORCID iDs

Laura B. Goodman https://orcid.org/0000-0002-8327-3092
Anja Schlierf https://orcid.org/0000-0002-9955-557X
Andriy Tkachenko https://orcid.org/0000-0002-0048-9243
Renate Reimschuessel https://orcid.org/0000-0002-3979-8617

Supplemental material

Supplemental material for this article is available online.

References

1. Andersen KG, et al. The proximal origin of SARS-CoV-2. Nat Med 2020;26:450–452.
2. Centers for Disease Control and Prevention (CDC). CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel: for Emergency Use Only: Instructions for Use. CDC, 2020. Document CDC-006-00006, revision 3. https://www.fda.gov/media/134922/download
3. Centers for Disease Control and Prevention (CDC). Get the facts about coronavirus. [cited 2021 Jan 21]. https://www.cdc.gov/coronavirus/2019-nCoV/index.html
4. Corman VM, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020;25:2000045.
5. Deng W, et al. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. Science. 2020;369 (6505):818-823.
6. Gaudreault NN, et al. SARS-CoV-2 infection, disease and transmission in domestic cats. Emerg Microbes Infect 2020;9:2322–2332.
7. Görzer I, et al. First results of a national external quality assessment scheme for the detection of SARS-CoV-2 genome sequences. J Clin Virol 2020;129:104537.
8. Harcourt J, et al. Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, United States. Emerg Infect Dis 2020;26:1266–1273.
9. International Organization for Standardization (ISO). ISO 13528:2015. Statistical methods for use in proficiency testing by interlaboratory comparisons. ISO, 2016. https://www.iso.org/standard/56125.html
10. International Organization for Standardization (ISO). ISO 16140-2:2016. Microbiology of the food chain—method validation—part 2: protocol for the validation of alternative (proprietary) methods against a reference method. ISO, 2016. https://www.iso.org/standard/54870.html
11. Meekins DA, et al. Susceptibility of swine cells and domestic pigs to SARS-CoV-2. Emerg Microbes Infect 2020;9:2278–2288.
12. Newman A, et al. First reported cases of SARS-CoV-2 infection in companion animals—New York, March-April 2020. MMWR Morb Mortal Wkly Rep 2020;69:710–713.
13. Olival KJ, et al. Possibility for reverse zoonotic transmission of SARS-CoV-2 to free-ranging wildlife: a case study of bats. PLoS Pathog 2020;16:e1008758.
14. Oude Munnink BB, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. Science 2021;371:172–177.
15. QuoData. ProLab Plus software version 2020.9.15.0. [cited 2021 Jan 26]. https://quodata.de/software/prolab-interlaboratory-studies-quodata/ring.html?lang=en#0
16. Ruiz-Arrondo I, et al. Detection of SARS-CoV-2 in pets living with COVID-19 owners diagnosed during the COVID-19 lockdown in Spain: a case of an asymptomatic cat with SARS-CoV-2 in Europe. Transbound Emerg Dis 2021;68:973–976.
17. Sailleau C, et al. First detection and genome sequencing of SARS-CoV-2 in an infected cat in France. Transbound Emerg Dis 2020;67:2324–2328.
18. Schlottau K, et al. SARS-CoV-2 in fruit bats, ferrets, pigs, and chickens: an experimental transmission study. Lancet Microbe 2020;1:e218–e225.
19. Shi J, et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus. Science 2020;368:1016–1020.
20. Uhlig S, et al. Can the usual validation standard series for quantitative methods, ISO 5725, be also applied for qualitative methods? Accred Qual Assur 2011;16:533.
21. Uhlig S, et al. Validation of qualitative PCR methods on the basis of mathematical-statistical modelling of the probability of detection. Accred Qual Assur 2015;20:75–83.
22. Uhlig S, et al. Taking laboratory uncertainties into account in the Hampel estimator. Accred Qual Assur 2019;24:25–32.
23. U.S. Department of Agriculture. Cases of SARS-CoV-2 in animals in the United States. [cited 2021 Jan 21]. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/sa_one_health/sars-cov-2-animals-us
24. U.S. Food and Drug Administration (FDA). Emergency use authorization. [cited 2021 Jan 21] https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization
25. Vogels CBF, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. Nat Microbiol 2020;5:1299–1305.
26. Wang L, et al. Complete genome sequence of SARS-CoV-2 in a tiger from a U.S. zoological collection. Microbiol Resour Announc 2020;9:e00468-20.
27. World Health Organisation. WHO-convened Global Study of the Origins of SARS-CoV-2. 2020. https://www.who.int/publications/m/item/who-convened-global-study-of-the-origins-of-sars-cov-2
28. Zitek T. The appropriate use of testing for COVID-19. West J Emerg Med 2020;21:470–472.