



Second round of an interlaboratory comparison of SARS-CoV2 molecular detection assays used by 45 veterinary diagnostic laboratories in the United States

Kaiping Deng, Steffen Uhlig, Laura B. Goodman,^{id}
 Hon S. Ip, Mary Lea Killian, Sarah M. Nemser, Jodie Ulaszek,
 Shannon Kiener, Matthew Kmet, Kirstin Frost, Karina Hettwer,
 Bertrand Colson, Kapil Nichani, Anja Schlierf,^{id} Andriy Tkachenko,
 Mothomang Mlalazi-Oyinloye, Andrew Scott, Ravinder Reddy,
 Gregory H. Tyson¹^{id}

Abstract. The COVID-19 pandemic presents a continued public health challenge. Veterinary diagnostic laboratories in the United States use RT-rtPCR for animal testing, and many laboratories are certified for testing human samples; hence, ensuring that laboratories have sensitive and specific SARS-CoV2 testing methods is a critical component of the pandemic response. In 2020, the FDA Veterinary Laboratory Investigation and Response Network (Vet-LIRN) led an interlaboratory comparison (ILC1) to help laboratories evaluate their existing RT-rtPCR methods for detecting SARS-CoV2. All participating laboratories were able to detect the viral RNA spiked in buffer and PrimeStore molecular transport medium (MTM). With ILC2, Vet-LIRN extended ILC1 by evaluating analytical sensitivity and specificity of the methods used by participating laboratories to detect 3 SARS-CoV2 variants (B.1; B.1.1.7 [Alpha]; B.1.351 [Beta]) at various copy levels. We analyzed 57 sets of results from 45 laboratories qualitatively and quantitatively according to the principles of ISO 16140-2:2016. More than 95% of analysts detected the SARS-CoV2 RNA in MTM at ≥ 500 copies for all 3 variants. In addition, for nucleocapsid markers N1 and N2, 81% and 92% of the analysts detected ≤ 20 copies in the assays, respectively. The analytical specificity of the evaluated methods was $>99\%$. Participating laboratories were able to assess their current method performance, identify possible limitations, and recognize method strengths as part of a continuous learning environment to support the critical need for the reliable diagnosis of COVID-19 in potentially infected animals and humans.

Keywords: interlaboratory comparison; COVID-19; real-time RT-PCR; SARS-CoV2.

Emergence of severe acute respiratory syndrome-related coronavirus (SARS-CoV) in 2003¹³ and the Middle East respiratory syndrome-related coronavirus (MERS-CoV) in 2012⁴ showed the zoonotic potential of animal coronaviruses. SARS-CoV2 transmission has been documented among animals, from humans to animals, and from animals to humans in the COVID-19 pandemic.^{2,9,13,14,20,21} SARS-CoV2 has been detected in wild and domestic animals around the world.^{8,15–18,23–25,28,30} A 2021 U.S. animal surveillance study indicated high prevalence of SARS-CoV2 among domestic and wild free-roaming animals tested on mink farms.¹¹ Conducting animal surveillance and routine testing with a sensitive and specific SARS-CoV2 detection method is important in outbreak response and prevention. A joint statement from the United Nations Food and Agriculture Organization, World Organisation for Animal Health (WOAH), and

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

¹Corresponding author: Gregory H. Tyson, Center for Veterinary Medicine, U.S. Food and Drug Administration, 8401 Murkirk Rd, Laurel, MD 20708, USA. Gregory.Tyson@fda.hhs.gov

World Health Organization also noted the need to promote the monitoring of wildlife known to be potentially susceptible to SARS-CoV2 and reporting of confirmed animal cases of SARS-CoV2 to WOA, with these actions requiring a sensitive and accurate SARS-CoV2 test.

In response to the spread of SARS-CoV2 in animals, veterinary diagnostic laboratories (VDLs) receive animal specimens for detection of SARS-CoV2. Many VDLs also have Clinical Laboratory Improvement Amendments (<https://www.cdc.gov/clia/index.html>) certification and test human samples, meaning that they play a critical One Health role in assessing the impact of COVID-19 on both humans and animals. Among the available tests, detection of the viral RNA with reverse-transcription real-time PCR (RT-rtPCR) is the most widely used, sensitive, and specific detection method for SARS-CoV2.

The U.S. FDA Veterinary Laboratory Investigation and Response Network (Vet-LIRN) is a network of VDLs that investigates potential animal food- or drug-related issues.¹² In August 2020, an Inter-Laboratory Comparison Exercise Round 1 (ILC1) was conducted collaboratively by the FDA and other organizations to qualitatively and quantitatively evaluate the SARS-CoV2 RT-rtPCR detection methods used by VDLs⁵; ILC1 participants detected SARS-CoV2 RNA effectively in molecular transport medium (MTM) with the methods that they used routinely for testing clinical specimens. Two-thirds of the laboratories achieved nearly the theoretical optimum level of detection (LOD) of 3 copies.⁵ However, the viral RNA spiking levels used in ILC1 were not low enough to evaluate the method analytical sensitivity, specifically the LOD for each individual participant. Hence, ILC2 was designed to provide more challenging samples from which marginal detection results could be generated for statistical analysis.

New lineages of SARS-CoV2 were reported and quickly became dominant variants in different parts of the world after late 2020¹ (<https://covariants.org/>), including the Alpha (B.1.1.7) and Beta (B.1.351) variants.^{7,26} Additional variants of concern such as the Delta (B.1.617) and Omicron (B.1.1.529) variants emerged subsequently. These emerging variants carry various mutations throughout the viral genome, including on the spike (*S*), envelope (*E*), nucleocapsid (*N*), and/or *ORF1ab* genes. Most participants used assays detecting the *N* gene in their routine testing, specifically markers N1 and N2, whereas some laboratories used other gene markers, such as *ORF1ab*, *S*, or *E* genes. In ILC2, the variants designated as Alpha (B.1.1.7) and Beta (B.1.351) were used in addition to U.S.B.1, which was the most prevalent lineage in the United States at the end of 2020 and the beginning of 2021. The ILC2 samples were shipped to laboratories in June 2021, and we studied the detection of these variants with the various markers used by participants.

ILC2 was conducted collaboratively by the following: 1) the FDA Center for Veterinary Medicine Vet-LIRN (Laurel, MD, USA); 2) the Moffett Proficiency Testing

Laboratory (Bedford Park, IL, USA) located at the Institute for Food Safety and Health at the Illinois Institute of Technology (IIT/IFSH) and the FDA Division of Food Processing Science and Technology; 3) QuoData Quality and Statistics (Dresden, Germany); 4) Cornell University (Ithaca, NY, USA); 5) the U.S. Geological Survey (USGS) National Wildlife Center (Madison, WI, USA); 6) USDA National Veterinary Services Laboratories (NVSL), National Animal Health Laboratory Network (Ames, IA, USA); 7) the Integrated Consortium of Laboratory Networks (Washington, DC, USA); and 8) 45 participating laboratories. The study was a continuation of ILC1 with the following objectives: 1) to evaluate analytical sensitivity (i.e., LOD) of the methods used routinely by participating laboratories to detect SARS-CoV2 variant (B.1) RNA; 2) to evaluate the ability of the participants' methods to detect SARS-CoV2 variants Alpha (B.1.1.7) and Beta (B.1.351); and 3) to evaluate the methods' specificity by investigating cross-reactivity with the RNA of feline infectious peritonitis virus (FIPV; *Alphacoronavirus 1*), a non-SARS-CoV2 animal coronavirus. Our goal was to allow participating laboratories to assess their method performance, including strengths and limitations, to support the reliable diagnosis of COVID-19 in potentially infected animals and humans.

Materials and methods

Determination of SARS-CoV2 RNA concentration

Three synthetic SARS-CoV2 genomic RNA products were purchased from Twist Bioscience for the ILC2 study: B.1 (Twist Control 10), Alpha (B.1.1.7; Twist Control 15), and Beta (B.1.351; Twist Control 16). These single-stranded RNA controls are manufactured by *in vitro* transcription from 6 non-overlapping 5-kb synthetic gene fragments. According to the manufacturer, the synthetic RNAs cover 99.9% of the bases of the viral genomes that were predominant in the United States, including the U.S.B.1, United Kingdom (Alpha [B.1.1.7]), and South Africa (Beta [B.1.351]) variants, with GISAID names: USA/CA-PC101P/2020, England/205041766/2020, and South Africa/KRISP-EC-K005299/2020, respectively. Droplet digital RT-PCR (RT-ddPCR)-based quantification of these controls was performed by the Cornell University Genomics Facility (QX200 instrument; Bio-Rad). The CDC N1 primers and probe (IDT) were used for this analysis with the 1-step RT-ddPCR advanced kit for probes (Bio-Rad), on duplicate serial dilutions of the templates. The concentrations of the original Twist B.1, Alpha (B.1.1.7), and Beta (B.1.351) controls were determined by RT-ddPCR as 150,000, 345,000, and 300,000 copies/ μ L, respectively. Serial dilutions of the controls were then made in nucleic acid dilution solution from the VetMAX Xeno internal positive control DNA kit (Applied Biosystems) to levels of 2×10^5 to 2 copies/ μ L in 10-fold dilutions.

Table 1. Description of ILC2 samples and associated SARS-CoV2 RNA concentrations.

Sample	Sample description and RNA concentration
VM01	Blank MTM
VM11	
VM05	Blank + FIPV RNA ~10,000 copies/50 μ L MTM
VM10	
VM06	(B.1) 10 copies/50 μ L MTM
VM09	
VM08	(B.1) 50 copies/50 μ L MTM
VM16	
VM20	
VM07	(B.1) 100 copies/50 μ L MTM
VM19	
VM03	(B.1) 200 copies/50 μ L MTM
VM12	
VM04	(B.1) 500 copies/50 μ L MTM
VM18	
VM17	(B.1) 10,000 copies/50 μ L MTM
VM15	(Alpha, B.1.1.7) 500 copies/50 μ L MTM
VM13	(Alpha, B.1.1.7) 10,000 copies/50 μ L MTM
VM14	(Beta, B.1.351) 500 copies/50 μ L MTM
VM02	(Beta, B.1.351) 10,000 copies/50 μ L MTM

FIPV = feline infectious peritonitis virus. All samples are prepared in PrimeStore molecular transport medium (MTM).

FIPV RNA preparation

A cryopreserved suspension of the culture-adapted Black strain of FIPV was provided by Dr. Gary Whittaker at Cornell University. The culture was grown in Fcwf-4 CU cells as described previously.¹⁹ RNA was extracted and purified (MagMAX viral/pathogen kit; Thermo Fisher). Quantification by RT-ddPCR was performed as described above but with the P009 and P010 primers and P9/P10 probe,⁶ which targets the *N* gene of FIPV.

ILC2 sample stability and homogeneity studies

Acceptable homogeneity, stability, and targeted spike levels were verified in 3 studies. In study 1 at the Moffett Proficiency Testing Laboratory, samples were analyzed by 2 analysts in 2 trials on days 0, 7, 14, and 21 of storage at -80°C (Suppl. Table 1). Eighteen samples (S1–S18) for each set were prepared. S1–S16 were prepared by adding B.1 RNA to PrimeStore MTM (Longhorn Vaccines & Diagnostics) at levels of 0, 10, 20, 50, 100, or 1,000 copies/50 μ L. S17 and S18 were prepared by adding Alpha (B.1.1.7) and Beta (B.1.351) to MTM at 1,000 copies/50 μ L. The RNA in study 1 samples were isolated (Qiagen RNeasy mini kit; Qiagen) and subsequently analyzed (AgPath-ID one-step RT-PCR kit; Qiagen) using specific primers (i.e., N1 and N2 targeting 2 regions of the viral nucleocapsid gene) and probes for the virus *N* gene (IDT), according to

the CDC 2019-nCoV EUA kit method.³ RT-PCR was performed on the 7500 fast real-time PCR instrument (Applied Biosystems) with v.2.3 software.

In studies 2 and 3 at Moffett, 2 sets of randomly chosen ILC2 samples (see their preparation below) were analyzed using the procedure described above: the first set (study 2) was analyzed prior to the shipment day, and the second set (study 3) was analyzed 2 d after the shipment day, when the ILC2 participants started to test their samples.

Qualitative data (Suppl. Tables 1A, 1C, 1D) indicated that all samples with RNA concentration ≥ 50 copies/50 μ L were detected; blank samples were not detected. The Ct values were subjected to quantitative analysis. There was no significant difference in the Ct values for the samples after 7, 14 and 21-d storage in study 1 (Suppl. Table 1B), indicating that the samples were stable for 21 d. The sample SD (s_{sample}) and the replicate measurements SD (s_r) were 0–1.13 and 0.11–1.14, respectively, when they were calculated based on the Ct values. The homogeneity and stability results demonstrated that the trial samples were deemed sufficiently homogeneous and stable, and the inoculation process was suitable to produce the targeted ILC2 samples.

ILC2 sample preparation and pre-shipment tests

The RNA was inoculated into MTM in bulk, and 150- μ L aliquots of each sample were transferred into 1.5-mL snap-top microfuge tubes, according to the sample composition in Table 1. All samples were stored at -80°C before shipping. To confirm successful inoculation before shipping, a set of ILC2 samples was tested as described in study 2 above. Another set of ILC2 samples was shipped to Cornell University to confirm the spiking levels by RT-ddPCR, using the procedures described above.

For the pre-shipment temperature trial, the packaging configuration was tested by packaging mock samples in dry ice in a shipping box according to the International Air Transport Association Dangerous Goods Regulations (<https://www.iata.org/en/publications/dgr/>). After holding the box for 72 h at room temperature, the sample containers were observed and qualitatively assessed as frozen or not. The assessment showed that the packaging configuration kept the primary sample containers frozen for 72 h.

ILC2 sample distribution

The final shipment samples were packaged (STP-320 UN 3373 Category B frozen insulated shipping system; Saf-T-Pak) according to the manufacturer-provided instructions and shipped (Priority Overnight; FedEx). We shipped 59 sets of blind-coded samples on dry ice to the 45 participating laboratories (14 laboratories requested a second set of samples to test 2 methods or to test by 2 analysts). Participants

were not aware of spike levels (i.e., analyte concentration) or number of replicates per spike level prepared by the organizers.

Sample analysis and data acquisition

Participants were instructed to use the SARS-CoV2 RNA extraction and detection methods that they used routinely in their laboratories. To facilitate statistical analysis, all analysts used the same input volume (50 μ L) for RNA extraction and reported their volumes of eluted RNA and PCR template to the ILC organizer. Before testing began, sample handling and result reporting instructions were discussed with the participants via 2 training sessions. To ensure confidentiality, each analyst was assigned an analyst identification number (AIN). Each analyst reported the results as detected (D), not detected (N), or inconclusive (I) for SARS-CoV2 viral RNA, according to their own laboratory protocols. Participants were also instructed to report Ct values for each PCR marker used, cutoff values, basic method information (e.g., PCR instrument model, extraction kit, and internal controls, as well as extraction, elution, and PCR input volumes), and any modifications to their methods. Detailed methods from each participating laboratory were kept confidential to maintain anonymity. We therefore refer to each assay target as “marker”, consistent with the terminology used in the ILC1 publication.²⁰

Qualitative and quantitative assessments

Rate of detection (ROD), the number of positive results divided by the total number of results, was calculated for all markers used (i.e., overall detection) and for N1 and N2 markers separately. Inconclusive results were classified as “not detected” for our assessment.

Analytical sensitivity (LOD) was calculated based on a probability of detection (POD) model. The complementary log-log regression model^{10,27} (i.e., the statistical model that corresponds to the Poisson assumption) was modified to take into consideration the analyst-specific actual copy numbers per well for calculating the marker- and analyst-specific POD curves. The POD curve was calculated based on ROD values obtained from the original (i.e., not rounded) Ct values. From these POD curves, the LOD95 values (the numbers of copies at which a POD of 95% is achieved) were derived. LOD95 was based on the effective volume (i.e., the part of the original sample volume used in the RT-PCR), which was calculated based on 3 volumes (i.e., extraction, elution, PCR input volumes) reported by participants to organizers (hereafter referred to as LOD95_{eff vol}). The adjustment was necessary for meaningful evaluation of method sensitivities for the individual laboratories.

The PCR amplification rate (i.e., efficiency) was calculated separately for each marker based on the nominal copy numbers (equivalent to dilution levels) and the submitted Ct values.

Results

Overall detection results

Fifty-seven datasets submitted from 45 laboratories were collected and analyzed. The analysts submitted qualitative “overall detection” results (Table 2), Ct values for various markers (Suppl. Tables 2–6), Ct cutoff values (Table 3), and basic method information. The analysts targeted different markers for SARS-CoV2 RNA testing. The overall detection results and Ct cutoff values reported are based on criteria selected by individual analysts according to their internal protocols. Comparative evaluation of qualitative overall detection results, reported Ct values, and the Ct cutoff values revealed that individual analysts used different decision-making criteria during interpretation of their datasets (Table 2). Specifically, some analysts reported detected (D), whereas other analysts reported not detected (N) or inconclusive (I) when some, but not all, targeted markers were detected (i.e., 1 of 2 targeted markers was detected, or 3 of 4 targeted markers were detected).

The ROD was calculated and summarized at each spike level (Table 4). For the overall detection as well as for the 2 most common markers (N1 and N2), the ROD values increased with increasing copy numbers as expected, consistently achieving ROD values >95% at 500 copies/50 μ L. At 100 copies/50 μ L, overall detection was 85%. Lower copy levels were also included to help assess levels of detection, and at the lowest spike level of 10 copies/50 μ L, 26% of samples were still identified as positive for SARS-CoV2. One exception was observed for B.1 samples at 10,000 copies/50 μ L; the ROD for N2 marker was <100% (i.e., 97%). This was the result of one false-negative result submitted at this high level, which also affected the calculated value of LOD95_{eff vol} for the analyst (AIN 278).

We also sought to determine whether ROD varied for different SARS-CoV2 variants. For the Alpha (B.1.1.7) and Beta (B.1.351) variants, ROD values for N1, N2, and overall detection were nearly identical to the results from B.1, indicating that laboratories were able to detect SARS-CoV2 regardless of the variant. Among blank samples, there was only 1 false-positive (AIN 294), indicating a false-positive rate of <1% for all participants.

Analytical specificity of methods

The analytical specificity of the methods was evaluated by including the FIPV RNA as a confounder. Only 1 analyst (AIN 296) reported 1 false-positive for 1 of the 2 “Blank+FIPV” samples (Table 2). Importantly, this analyst detected the N2 marker, but not the N1 marker for this sample (Suppl. Table 2) and thus reported the sample as “detected” according to their internal protocols. Similarly, 2 other analysts (AINs 267, 286) detected the “Blank+FIPV” samples with only one of the markers (Suppl. Tables 2, 5); however, they reported the sample as “inconclusive” according to their

Table 3. Level of detection at 95% probability based on the effective volume (LOD95_{eff vol}) and PCR efficiency for each marker.

AIN	Effective volume, μL	Ct cutoff values	LOD95 _{eff vol} (efficiency, %)							RdRP	ORF1ab	S
			N1	N2	N	N3	E					
278	0.5	45	2.0† (210)	1.9† (106)								
276	0.5	40	3.1 (117)	4.6 (115)								
263	1.1	40	12.7 (114)	4.4 (106)								
293	2.5	40	14.1 (260)	6.4 (158)								
291	2.5	40	16.8 (193)	14.9 (155)								
285	2.5	45	3.3 (107)	3.3 (83)								
247	2.5	45							7.4 (140)			
250	2.5	40			26.7 (204)							
262	2.5	40			6.6* (99)							
294	2.5	40			4.9 (85)							
283	2.8	36	>556 (87)	106.7 (101)								
289	2.8	40	10.9 (106)	7.1 (104)								
241	2.8	39.99	10.9 (88)	3.7 (102)							18.2 (92)	
280	2.8	40	10.9 (97)	8.2 (90)								
246	2.8	40	13.2 (111)	3.7 (120)								
282	2.8	40	15.7 (101)	22.9 (79)								
267	2.8	40	17 (125)	6.4 (201)								
279	2.8	40	2.0† (122)	2.0† (108)								
265	2.8	40	2.0† (94)	4.6* (80)								
258	2.8	40	2.0† (96)	7.1 (93)								
245	2.8	45	22.9 (112)	3.7 (103)								
257	2.8	45	3.7 (106)	3.7 (92)								
256	2.8	40	3.7 (91)	2.0† (100)								
249	2.8	45	3.7 (92)		5.2 (132)							
274	2.8	45	34.1 (154)	10.9 (129)								
261	2.8	40	4.6* (101)	4.6* (111)								
284	2.8	40	4.6* (101)	4.6* (86)								
255	2.8	45	4.6* (95)	2.0† (67)		2.0† (99)						
242	2.8	40	41.8 (114)	8.2 (100)								
273	2.8	40	55.9 (107)	7.1 (108)								
252	2.8	40						3.7 (88)	4.6* (111)			
269	3.1	40	8 (112)	8 (134)								
275	3.9	45			6.5* (106)							
271	4.2	40	3 (104)	3 (91)								
244	4.2	40	3 (110)	3 (99)								
266	4.4	37	36.6 (136)	104 (84)								
254	5	40	12.8 (119)	9.3 (109)								
290	5	40	6.7 (102)	3.6 (108)								
268	5	45	6.7 (110)	8.3* (108)								
253	5	45	66.8 (72)	9.3 (107)								
259	5	40						3.6 (135)	57 (100)			
270	5	40						6.7 (102)	16 (93)			
286	5	45						6.7 (126)		2.4† (84)		
248	5.6	45			9.3* (138)							
287	5.8	38			9.7* (108)							
240	6.7	37			29.6 (94)					4.8 (115)	52.9 (88)	
295	7	40			3.2 (92)							
288	8	36						13.3* (103)	31.4 (101)			
292	8	36						13.3* (126)	5.7 (101)			
277	8.3	40	6 (117)	6 (99)								
296	8.9	40	6.4 (104)	3.7 (111)								
243	10	37			25.6 (107)					7.2 (121)	79.3 (85)	
251	10	37			56.4 (119)					7.2 (115)	>2,000 (65)	
260	10	37			91.6 (73)					122.7 (64)	>2,000 (59)	
264	10	37			82.5 (88)					56.4 (94)	>2,000 (70)	
272	10	39			44.4 (97)					13.3 (108)	>2,000 (93)	
281	11.1	40	18.5* (110)	8 (104)								

AIN=analyst identification number. Empty cells indicate that a particular marker was not used by an analyst for ILC2 sample testing. Effective volume is that part of the original sample volume used in the PCR by each analyst. It was calculated based on 3 volumes (e.g., extraction, elution, PCR input) used and reported by analysts. LOD95_{eff vol} is the number of copies in a PCR reaction at which a probability of detection of 95% is achieved based on the effective volume used.

* Upper limit of CI; there are no false-negative results.

† Value is not significantly lower than the theoretical optimum of 3 target copies.

Table 4. Rate of detection (ROD) calculated for all analysts based on their overall detection and for those who used N1 and N2 markers.

SARS-CoV2 RNA spiking level	No. of PCR replicates	ROD across analysts		
		Overall detection	N1	N2
No. of analysts included in the calculations		57	37	36
Blank	2	1%	0%	0%
Blank + 10,000 FIPV copies/50 μ L	2	1%	0%	3%
B.1 10 copies/50 μ L	2	26%	30%	35%
B.1 50 copies/50 μ L	3	65%	56%	70%
B.1 100 copies/50 μ L	2	85%	81%	92%
B.1 200 copies/50 μ L	2	91%	95%	96%
B.1 500 copies/50 μ L	2	96%	95%	100%
B.1 10,000 copies/50 μ L	1	98%	100%	97%
Alpha (B.1.1.7) 500 copies/50 μ L	1	96%	97%	100%
Alpha (B.1.1.7) 10,000 copies/50 μ L	1	100%	100%	100%
Beta (B.1.351) 500 copies/50 μ L	1	98%	97%	100%
Beta (B.1.351) 10,000 copies/50 μ L	1	100%	100%	100%

ROD values were calculated based on the original (i.e., not rounded) Ct values. Inconclusive results and results with Ct values higher than the analyst-specific cutoff were classified as “not detected” in this statistical summary.

internal protocols. Overall, ILC2 results revealed that the participants’ methods are specific to SARS-CoV2 and do not routinely yield false-positive results as a result of detection of FIPV RNA.

Analytical sensitivity and efficiency of methods

Analytical sensitivity was evaluated with $LOD_{95_{eff\ vol}}$ (Table 3), which was calculated based on the actual RNA copy number added to the PCR for each analyst (i.e., effective volume for each spiking level). The 0.5–11.1- μ L effective volumes tested varied by a factor of 22 among analysts (Table 3). The $LOD_{95_{eff\ vol}}$ values varied greatly among participants and the markers they used. Sensitivities for N1 and N2 markers were summarized (Table 5), including $LOD_{95_{eff\ vol}}$ values calculated for AINs 283 and 266, which may be considered outliers. Specifically, AINs 283 and 266 reported Ct values for samples at low spike levels (≤ 200 copies; Suppl. Table 2). However, those values were counted as not detected because they were higher than the Ct cutoff values established by the analysts. Hence, the methods used in these 2 laboratories are likely sufficiently sensitive to detect more samples at lower spike levels, but established Ct cutoff values were too stringent (i.e., unoptimized), resulting in more false-negative results (Table 2) affecting LODs in these 2 laboratories. Thus, $LOD_{95_{eff\ vol}}$ calculations with these 2 laboratories (AINs 283 and 266, for N1 and N2) are shown in Table 5; however, $LOD_{95_{eff\ vol}}$ with these 2 laboratories excluded are interpreted below.

As mentioned above, $LOD_{95_{eff\ vol}}$ values from participants were combined for each marker and their median, minimum, and maximum $LOD_{95_{eff\ vol}}$ values for comparative evaluation (Table 5). The $LOD_{95_{eff\ vol}}$ values were 2.0–66.8 for the N1

marker, with a median of 6.7. This corresponded to a factor of ~ 30 between the lowest and highest values; although some analysts detected every copy of SARS-CoV2 RNA, others detected $< 10\%$ of the copies. For the N2 marker, the $LOD_{95_{eff\ vol}}$ values were 1.9–22.9 with the median of 4.6 copies, indicating that methods based on the N2 marker were the most sensitive in ILC2. An ideal $LOD_{95_{eff\ vol}}$ is calculated to be 3 copies per PCR reaction based on a hypothetical POD curve.²⁷ Given random variation, $LOD_{95_{eff\ vol}}$ values < 3 may be observed. For the N1 marker, 30 of 36 (83%) analysts had an $LOD_{95_{eff\ vol}}$ value ≤ 20 copies, and 20 of 37 (54%) analysts had an $LOD_{95_{eff\ vol}}$ value not statistically significantly > 3 (meaning that the $LOD_{95_{eff\ vol}}$ is within the margin of error of the best possible value). For the N2 marker, 33 of 34 (97%) analysts had an $LOD_{95_{eff\ vol}}$ value ≤ 20 copies, and 27 of 36 (75%) analysts had an $LOD_{95_{eff\ vol}}$ value not statistically significantly > 3 . For the E and N3 markers, all analysts (100%) had $LOD_{95_{eff\ vol}}$ values at ≤ 20 copies. For the N, ORF1ab, and RdRP markers, only half of the analysts had $LOD_{95_{eff\ vol}}$ near or equal to the best possible LOD—theoretical minimum of 3 copies. For the S marker, none of the analysts had $LOD_{95_{eff\ vol}}$ near or equal to the best possible value of 3 copies, and only 1 of 7 (14%) analysts had an $LOD_{95_{eff\ vol}}$ of ≤ 20 copies.

Calculated efficiency varied greatly among participants and markers that they used, as well (Table 3). In general, a PCR efficiency of 100% indicates that the target sequence of interest doubles during each cycle. If the Ct values change < 3.3 cycles between 10-fold dilutions of the PCR template, efficiency values were $> 100\%$. Some of the calculated efficiencies are indeed $> 100\%$, which could be interpreted as an indication of problems in the amplification process. On the other hand, it should be noted that the calculation of efficiencies is associated with considerable statistical uncertainty as a result of unavoidable

Table 5. Comparative summary of sensitivity (number of false negative and LOD95_{eff vol}) of each marker.

Marker	Total no. of analysts	No. of analysts without false-negative results	LOD95 _{eff vol}			No. of analysts with LOD95 _{eff vol} at ≤20 copies
			Median	Min.	Max.	
N1	37	4	8.0	2.0	>556	30/37 (81%)
N1*	36	4	6.7	2.0	66.8	30/36 (83%)
N2	36	4	5.3	1.9	106.7	33/36 (92%)
N2†	34	4	4.6	1.9	22.9	33/34 (97%)
E	6	2	6.7	3.6	13.3	6/6 (100%)
N	14	4	17.7	3.2	91.6	7/14 (50%)
N3	1	0	—	—	—	1/1 (100%)
ORF1ab	7	0	7.2	2.4	122.7	5/7 (71%)
RdRP	6	1	11.7	4.6	57.0	4/6 (67%)
S	7	0	52.9	18.2	79.3	1/7 (14%)

Dash (—) indicates not applicable.

* Level of detection (LOD)_{eff vol} value of >556 copies by analyst identification number (AIN) 283 was excluded from the summary because this value was considered an outlier given unoptimized (i.e., too stringent) Ct cutoff values used by the analyst and substantially affecting the summary of LODs for N1 marker.

† LOD at 95% probability based on the effective volume (LOD95_{eff vol}) values of 106 copies (AIN 283) and 104 copies (AIN 266) were excluded from calculations for the same reason described above.

random fluctuations in the Ct values. This is especially true when, as in the present case, the underlying dilution levels differ by only a few orders of magnitude.

Discussion

The ILC2 provides insight into the performance of methods to detect SARS-CoV2 RNA in MTM. ILC2 demonstrated that most participants have relatively sensitive and specific methods to detect 3 SARS-CoV2 variants.

The analysts participating in ILC2 were instructed to follow their routine SARS-CoV2 detection procedure; hence, the testing methods varied between laboratories. The laboratories' Ct cutoff values also varied, which resulted in some inconsistency in interpretations among analysts. For a study such as ILC2 in which various extraction and detection methods were involved, applying a universal Ct cutoff value for each marker to provide a score for each individual analyst is not realistic, even though it would simplify result interpretation. Method information provided by each analyst allowed the ILC2 organizers to statistically identify possible correlations with result variability. Analyst-specific results broken down by Ct values for different markers and with extraction and PCR methods annotated were summarized and provided to the analysts in an ILC2 report. These data are not shown here to protect the confidentiality of participants, particularly those who were using methods unique to their laboratory.

Most of the participating laboratories used the CDC N1 and N2 assays to detect SARS-CoV2 RNA. Research has reported one marker as more sensitive than the other or vice versa.^{22,29} In ILC1,⁵ differences in ROD values between the 2 markers were very minor. In ILC2, the ROD values were higher for N2 than for N1, especially for 50–100 copies/50 µL for the B.1 variant. The LOD95_{eff vol} values generated for the N1 and N2 markers indicated that 81% and 92% of ILC2 analysts, respectively,

demonstrated an LOD95_{eff vol} value of ≤20 copies. Other markers, such as E, N, N3, ORF1ab, RdRP, or S were used in the non-CDC assays. The number of analysts who used these markers, the LOD95_{eff vol} values, and the sensitivity for each marker are also summarized in Table 5. All 6 analysts using the E marker were able to achieve an LOD95_{eff vol} value of ≤20 copies; those using the S marker demonstrated the lowest sensitivity with only 1 of 7 analysts meeting this threshold. Importantly, all analysts using the S gene marker also had N and ORF1ab markers as part of their multiplex PCR assays.

The relationship between the effective volumes used and the observed LOD95_{eff vol} values was not significant. Although we cannot rule out that high effective volumes make it more difficult to detect all copies, the observed LOD95_{eff vol} values also depend on other factors. These factors may include extraction kit, PCR kit, reagents, model, or type of equipment, such as PCR machine, centrifuges, and pipettes, level of analyst's experience, current quality control system in laboratory, multiplex versus singleplex approaches, and others. Information on some of these factors was provided in the confidential report for the participants. The variability in Ct values observed for particular analysts is also informative and can point to potential issues with methods that may affect success of analysts in future exercises.

ILC2 demonstrated a successful collaboration involving government agencies, universities, and private industry. By using a larger range of SARS-CoV2 RNA spiking levels, including lower concentrations, we evaluated the LOD95_{eff vol} values of the methods used by participating analysts more accurately than in ILC1. The laboratories were able to detect Alpha and Beta variants of SARS-CoV2 with their current methods. The method specificity was confirmed by using FIPV RNA as a confounder and reached over 99%.

ILC2 was a success in meeting the stated objectives. The exercise allowed organizers not only to characterize

important parameters of participants' method performance (e.g., analytical sensitivity and specificity, efficiency, and suitability for different variants) but also allowed participants to compare their performance to each other. Organizers processed the submitted data using various statistical approaches that allowed them to not only identify possible weaknesses and strengths of methods used but to offer suggestions on improving participants' performance in the future. Specifically, an important finding of the ILC2 is that individual analysts used different decision-making criteria during interpretation of similar datasets. This indicates a need for laboratories to review data from ILC2 and potentially reassess their decision-making criteria during interpretation of Ct values when using multiple markers. Our ILC2 study also indicates that the false-negative rate and sensitivity of some methods can be improved if Ct cut-off values used are re-evaluated (e.g., on a test that a too-stringent Ct cutoff value was originally used) and optimized by analysts accordingly.

In the current era of rapidly developing methodology and lack of international standards, participation in ILCs such as our study is very beneficial. Further exercises will be important to ensure continued detection of emerging SARS-CoV2 variants. In contrast to other types of proficiency testing exercises that only aim to assess which results are correct or incorrect, our ILC revealed much more about the methods used and assisted participants in continuous efforts to improve their performance.

Acknowledgments

We acknowledge the diligence and hard work of the laboratory scientists who developed SARS-CoV2 assays rapidly for animals and participated in our collaborative study. We acknowledge the following individuals for technical assistance and administrative support: Angelica Jones, Olgica Ceric, Ellen Hart, and Dave Rotstein from FDA; Joseph Flint from Cornell University; 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 this facilitation. We also thank Robert Newkirk from the FDA CFSAN Proficiency Team for sample shipment assistance.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the official policy of the Department of Health and Human Services, the U.S. Food and Drug Administration, or the U.S. Government. The use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

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


ILC2 was funded by the FDA Vet-LIRN program, and laboratories were not charged to participate in this exercise. Vet-LIRN laboratories may have also used infrastructure grant funds provided via Vet-LIRN funding opportunity PAR-17-141 to cover the cost of supplies.

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>

Gregory H. Tyson  <https://orcid.org/0000-0002-2729-5035>

Supplemental material

Supplemental material for this article is available online.

References

- Badua CLDC, et al. Genomic and proteomic mutation landscapes of SARS-CoV-2. *J Med Virol* 2021;93:1702–1721.
- Centers for Disease Control and Prevention (CDC). Animals and COVID-19. CDC, 2022 Apr 27. <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/animals.html>
- Centers for Disease Control and Prevention. CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel: for emergency use only: instructions for use. CDC, 2021 Jul 21. CDC-006-00019, revision 07. <https://www.fda.gov/media/134922/download>
- Coleman CM, et al. Emergence of the Middle East respiratory syndrome coronavirus. *PLoS Pathog* 2013;9:e1003595.
- Deng K, et al. Interlaboratory comparison of SARS-CoV2 molecular detection assays in use by U.S. veterinary diagnostic laboratories. *J Vet Diagn Invest* 2021;33:1039–1051.
- Dye C, et al. Evaluation of real-time RT-PCR for the quantification of FCoV shedding in the faeces of domestic cats. *J Feline Med Surg* 2008;10:167–174.
- Galloway SE. Emergence of SARS-CoV-2 B.1.1.7 lineage—United States, December 29, 2020–January 12, 2021. *MMWR* 2021;70:95–99.
- Hamer SA, et al. Natural SARS-CoV-2 infections, including virus isolation, among serially tested cats and dogs in households with confirmed human COVID-19 cases in Texas, USA. *bioRxiv* 2020:2020.12.08.416339. Preprint.
- He S, et al. Backward transmission of COVID-19 from humans to animals may propagate reinfections and induce vaccine failure. *Environ Chem Lett* 2021;19:763–768.
- 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>
- Ip HS, et al. An opportunistic survey reveals an unexpected coronavirus diversity hotspot in North America. *Viruses* 2021;13:2016.
- Kaneene JB, et al. Changes in multidrug resistance of enteric bacteria following an intervention to reduce antimicrobial resistance in dairy calves. *J Clin Microbiol* 2009;47:4109–4112.

13. Ksiazek TG, et al. A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* 2003;348:1953–1966.
14. Kumar R, et al. COVID-19 and domestic animals: exploring the species barrier crossing, zoonotic and reverse zoonotic transmission of SARS-CoV-2. *Curr Pharm Des* 2021;27:1194–1201.
15. Lee AC, et al. Oral SARS-CoV-2 inoculation establishes subclinical respiratory infection with virus shedding in golden Syrian hamsters. *Cell Rep Med* 2020;1:100121.
16. McAloose D, et al. From people to *Panthera*: natural SARS-CoV-2 infection in tigers and lions at the Bronx Zoo. *mBio* 2020;11:e02220-20.
17. Meisner J, et al. Household transmission of SARS-CoV-2 from humans to pets in Washington and Idaho: burden and risk factors. *bioRxiv* 2022:2021.04.24.440952. Preprint.
18. Newman A, et al. First reported cases of SARS-CoV-2 infection in companion animals—New York, March–April 2020. *MMWR* 2020;69:710–713.
19. O'Brien A, et al. Characterizing replication kinetics and plaque production of type I feline infectious peritonitis virus in three feline cell lines. *Virology* 2018;525:1–9.
20. Oude Munnink BB, et al. Jumping back and forth: anthropozoonotic and zoonotic transmission of SARS-CoV-2 on mink farms. *bioRxiv* 2020:2020.09.01.277152. Preprint.
21. 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.
22. Perchetti GA, et al. Validation of SARS-CoV-2 detection across multiple specimen types. *J Clin Virol* 2020;128:104438.
23. 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.
24. 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.
25. Segalés J, et al. Detection of SARS-CoV-2 in a cat owned by a COVID-19-affected patient in Spain. *Proc Natl Acad Sci U S A* 2020;117:24790–24793.
26. Tegally H, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* 2021;592:438–443.
27. Uhlig S, et al. Validation of qualitative PCR methods on the basis of mathematical-statistical modelling of the probability of detection. *Accredit Qual Assur* 2015;20:75–83.
28. 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/animal_health/one_health/downloads/sars-cov2-in-animals.pdf
29. Vogels CBF, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. *Nat Microbiol* 2020;5:1299–1305.
30. Wacharapluesadee S, et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat Commun* 2021;12:972.