Clinical Validation of a SARS-CoV-2 Real-Time Reverse Transcription PCR Assay Targeting the Nucleocapsid Gene

Jeffrey A. SoRelle, MD[1,*]
Ithiel Frame, MD, PhD[1,*]
Alejandra Falcon[2]
Jerin Jacob, MS[2]
Jennifer Wagenfuehr[2]
Midori Mitui[2]
Jason Y. Park, MD, PhD[1,2,3,**,†]
Laura Filkins, PhD[1,2**]

[1] Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas
[2] Department of Pathology and Laboratory Medicine, Children's Health System of Texas, Dallas, Texas
[3] McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, Texas

*These authors contributed equally to this work
**These authors contributed equally to this work

jeffrey.sorelle@utsouthwestern.edu
ithiel.frame@phhs.org
alejandra.falcon@childrens.com
jerin.jacob@childrens.com
jennifer.wagenfuehr@childrens.com
esperanza.mitui@childrens.com
jaspar@childrens.com
laura.filkins@utsouthwestern.edu

†Send correspondence to:
Jason Y. Park, MD, PhD
Department of Pathology
Children's Medical Center of Dallas
1935 Medical District Dr.
Dallas, TX 75235
Tel:214-456-1809
Email: jaspar@childrens.com

Word Count: 2004
Abstract: 222
Total Pages: 13
Figures: 0
Tables: 3
Supplemental Tables: 7

Key words: COVID-19, Coronavirus Disease 2019, SARS-CoV-2, Clinical Validation, EUA

Abstract

Background: Detection of SARS-CoV-2 viral RNA is important for the diagnosis and management of COVID-19.

Methods: We present a clinical validation of a RT-PCR assay for the SARS-CoV-2 nucleocapsid (N1) gene. Offboard lysis on an automated nucleic acid extraction system (EMAG®) was optimized with endemic Coronaviruses (OC43 and NL63). Genomic RNA and SARS-CoV-2 RNA in a recombinant viral protein coat (Accuplex) were used as control materials and compared for recovery from nucleic acid extraction.

Results: Nucleic acid extraction showed decreased recovery of endemic Coronavirus *in vitro transcribed* RNA (NL63) compared to attenuated virus (OC43). SARS-CoV-2 RNA (Accuplex) had more reliable recovery from extraction through amplification compared to genomic RNA. Recovery of genomic RNA was improved by combining lysis buffer with clinical matrix prior to adding RNA. The RT-PCR assay demonstrated 100% *in silico* sensitivity and specificity. The accuracy across samples was 100% (75 of 75). Precision studies showed 100% intra-run, inter-run, and inter-technologist concordance. The limit of detection was 264 copies per ml (estimated 5 copies per reaction; 35.56 mean Ct value).

Conclusions: This SARS-CoV-2 assay demonstrates appropriate characteristics for use under an emergency use authorization. Endemic Coronavirus controls were useful in optimizing the extraction procedure. In the absence of live or attenuated virus, recombinant virus in a protein coat is an appropriate control specimen type for assay validation during a pandemic.

Impact Statement

This is a clinical validation study of a molecular assay for SARS-CoV-2, the virus that causes the disease COVID-19. Laboratories providing or considering providing SARS-CoV-2 testing may find the results to be of interest. A unique aspect of this validation is the use of alternative validation materials. In this study we used a combination of endemic coronaviruses as well as commercial synthetic material to examine assay performance. This validation may be useful in the current COVID-19 outbreak as well as future viral outbreaks where molecular testing is needed.

Introduction

The global pandemic of COVID-19 (1) poses a diagnostic challenge that is best addressed by molecular diagnostic techniques. The COVID-19 pathogen, SARS-CoV-2 (2), is a single-stranded RNA Betacoronavirus with a 26 kilobase genome. The molecular detection of SARS-CoV-2 is based on targeting the viral genes (e.g., Orf1a/b, E, S, N genes) (3–7).

In the United States, the first clinical assay available for SARS-CoV-2 was developed by the Centers for Disease Control and Prevention (CDC) (3) under a US Food and Drug Administration (FDA) Emergency Use Authorization (EUA) on February 4, 2020 (8). The CDC assay initially targeted three regions of the viral nucleocapsid gene (N1, N2, N3) and the human RNase P (RP) gene as an internal control. Later changes to the original CDC EUA assay included removal of the N3 target and only required single detection of either the N1 or N2 target (9).

The assay described in this study has almost all of the components described under the EUA for the CDC assay; however, we use a non-regulated PCR instrument (i.e., ABI 7500), which the FDA determined is not equivalent to a regulated PCR instrument (i.e., ABI 7500 Fast Dx). Any modification to the CDC assay, including the use of a non-regulated PCR instrument, required a new FDA EUA application. In our validation, we used the FDA EUA guidance to determine the target sensitivity and specificity by *in silico* and wet bench analyses. Several challenges arose for validating this molecular virology assay, including a lack of reference materials, a changing regulatory landscape, and an unstable supply chain. In our assay development, we used endemic coronaviruses as surrogates for the extraction efficiency of SARS-CoV-2. In addition, we examined two types of positive control materials (genomic RNA and recombinant viral protein encapsulated RNA).

Materials and Methods

RNA Extraction

Under a Class II biosafety cabinet, 300 µL of swabbed nasopharyngeal (NP) samples inoculated into universal viral transport medium (VTM; BD, 220529) or spiked control material in VTM were transferred to conical tubes containing 2 mL Nuclisens® Lysis Buffer (bioMérieux; Durham, NC) for lysis/virus inactivation (10 min.) before extraction. Total nucleic acids from pooled or individual residual NP collections and controls were obtained using the EMAG® Nucleic Acid Extraction System (bioMérieux) with an offboard lysis protocol and the following volume parameters: 300 µL input, 50 µL magnetic silica, and 80 µL output/elution.

Control material and Patient Specimens

Control material for endemic coronavirus strains was obtained from Exact Diagnostics (Fort Worth, TX; respiratory panel pooled control including human coronavirus NL63 [*in vitro* transcribed RNA] and human coronavirus OC43 [whole inactivated virus]). These strains were detected with SYBR green RT-PCR as described previously (10,11). SARS-CoV-2 genomic RNA was from University of Texas Medical Branch (6.00E+07 copies per µL; Galveston, TX). Recombinant Sindbis virus containing SARS-CoV-2 RNA was obtained as a commercial control material (Accuplex™ Reference Material; Cat No. 0505-0126; SeraCare®, Milford, MA). Residual de-identified patient samples from a public health laboratory with a SARS-CoV-2 FDA EUA assay included both positive (n=5) and negative (n=5) samples. The transport medium of these samples was presumed to be a formulation of VTM, though the exact manufacturer is unknown. Each of these patient samples underwent freeze-thaw cycles at least twice prior to our extraction and PCR. Positive patient specimens, when indicated, were diluted in VTM before nucleic acid extraction. Twenty-five residual patient NP swab specimens in VTM collected prior to December 2019 were tested as SARS-COV-2 negative samples.

Real-time polymerase chain reaction

Reverse transcription, real-time PCR was performed using the SARS-CoV-2 N1 (2019-nCoV_N1) and human RNase P (RP) primer/probe mixes (IDT, Cat No. 10006606) (Table S1) following the CDC protocol (3) using TaqPath 1-Step RT-qPCR Master Mix, CG (ThermoFisher, Cat No. A15299) and 5µL of extracted nucleic acid in a final reaction volume of 20µL. Our initial evaluation of the N1 and N2 targets suggested similar performance with slightly lower Ct values for N1. Upon notification that the FDA was permitting single detection of N1 or N2 for the CDC EUA, we chose to pursue the N1 target. Unlike the CDC EUA protocol that uses the ABI 7500 FAST Dx, the assay was validated on an ABI 7500. The cycling conditions are as follows: 25°C (2 min), 50°C (15 min), and 95°C (2 min), then amplification for 40 cycles (95°C 3 sec, 55°C 30 sec) with fluorescence measured at 55°C. The NL63 and OC43 strains were detected with SYBR green RT-PCR as described previously (10,11). In brief, extracted RNA was random primed for first strand cDNA synthesis (PreSeq RNA QC, ArcherDX; Boulder, CO) and then PCR amplified (KAPA SYBR FAST qPCR, Roche; Wilmington, MA). PCR program was 1 cycle at 95°C for 5 min followed by 40 cycles (95°C 30 sec, 49°C 30 sec, 60°C 45 sec).

***In silico* sensitivity (Inclusivity) and specificity (Cross-Reactivity)**

For *in silico* sensitivity, Forward (N1-F), Reverse (N1-R) and Probe (N1-TaqMan) sequences were queried in the National Center for Biotechnology Information (NCBI) BLASTN search across all Betacoronaviruses (7864 sequences on the 2020/03/17 NCBI update). For *in silico* specificity, Forward (N1-F), Reverse (N1-R) and Probe (N1-TaqMan) sequences were queried in NCBI BLASTN search against public domain nucleotide sequences. The database search parameters were as follows: 1) Standard databases were used 2) Organism was specified and the taxonomy ID was recorded (NCBI taxid), 3) Program Selection was selected to optimize for "Highly similar sequences" (Megablast), 4) Sequences with the highest E-values were

selected and the total number of aligned bases across the whole primer or probe length were used to calculate percent homology.

Analytic specificity

Specificity studies were drawn from previously collected and frozen patient NP specimens positive for microorganisms other than SARS-CoV-2 and supplemented with pooled NP specimens spiked with cultured organisms (50 µL 0.5 MacFarland bacterial or fungal isolates into 250 µL NP) or QC material (HKU1, 229E, NL63, OC43, Rhinovirus type 1A, and Parainfluenza virus 4, ZeptoMetrix, Buffalo, NY). For bacterial or fungal isolate spiking, the following isolates were grown on blood agar, chocolate agar, or Sabouraud dextrose agar plates overnight at 30°C (for *Candida albicans*) or 35-37°C (all others), 5% CO_2: *Haemophilus influenzae* ATCC 49247, *Streptococcus pneumoniae* ATCC 49619, *Streptococcus pyogenes* ATCC 19615, *Candida albicans* ATCC 10231, *Pseudomonas aeruginosa* ATCC 27853, *Staphylococcus aureus* ATCC 29213, *Staphylococcus epidermidis* ATCC 12228, and *Streptococcus salivarius* clinical isolate.

Accuracy, limit of detection, and precision studies

Accuracy was determined using a combination of positive clinical specimens at various dilutions (range of undiluted to 1:20 dilution; n=36 samples derived from 5 positive patient specimens; undiluted Ct values for the 5 positive patients ranged from 25.39 to 35.56), Accuplex encapsulated SARS-CoV-2 (n=9), and negative clinical specimens (n=30). Ct values greater than 40 for N1 were considered negative and values ≤40 Ct were considered positive. Limit of detection was tested with recombinant encapsulated SARS-CoV-2 and genomic SARS-CoV-2 RNA. For precision analysis, two materials were used: positive sample (SeraCare SARS-CoV-2

RNA at 1040 copies per ml) and negative sample (NP). These samples were tested in triplicate within the same assay run (intra-assay precision) and were also examined as single sample analysis across six different assay runs (inter-assay precision). Three technologists performed nucleic acid extraction and two technologists performed nucleic acid amplification procedures in four different paired combinations throughout the six assay runs. The reference method was a public health laboratory performing the CDC EUA assay.

Statistics

Excel 2016 (Microsoft Corporation, Redmond, WA) was used for calculations. A descriptive statistic of percentage was used. Probability values (p-values) were not used for hypothesis testing.

Results

Extraction Optimization

Optimization of the extraction protocol was performed using a pooled respiratory panel control material (Exact Diagnostics) that included endemic coronavirus OC43 whole inactivated virus and endemic coronavirus NL63 *in vitro* transcribed (IVT) RNA in known concentrations. Nucleic acid recoveries of endemic coronavirus OC43 and NL63 in NP matrix, VTM and undiluted were compared with differing input, magnetic silica, and elution volumes. Ct values were lowest when control material was directly added to lysis buffer (Table S2, samples 1-3). In the presence of NP or VTM matrix, there was an increase in Ct values for detecting NL63, but not OC43 (Table S2, samples 1, 4, 7). Despite a 5-log greater concentration, detection of NL63 was similar or worse than detection of OC43 when in NP or VTM matrix (Table S2, samples 4-9). These data suggest a loss of IVT RNA compared to whole virus when present in matrix prior to a lysis step.

In Silico Sensitivity (Inclusivity) and Specificity (Cross-Reactivity)

An *in silico* analysis of sensitivity (inclusivity) and specificity (cross-reactivity) for the primers and probes used for SARS-CoV-2 assay validation was performed. BLASTN search across Betacoronaviruses (7864 sequences on 03/17/2020) identified all SARS-CoV-2 genomes (100%, representative homologies for 32 isolates in Table S3). When combining primer and probe sequences, there was no significant homology to high priority pathogens or organisms as defined by the FDA EUA (Table S4).

Assayed Specificity (Cross-Reactivity)

Analytical cross reactivity of primers and probes was assessed in clinical specimens positive for or spiked with common respiratory pathogens or microbiota (n=25) and negative patient specimens (n=5) (Table S5). All reactions were valid and none of the specimens were amplified by the N1 target (0 of 30).

Limit of Detection

Recombinant virus with SARS-CoV-2 RNA was used to determine the limit of detection of the assay by a two-fold dilution series. A preliminary dilution in triplicate was performed; when all triplicate samples were detected, then an extended replicate series of 20 samples was examined. We found 100% positive rate at 264 copies per mL (Table 1). This was substantially lower than we were able to achieve using genomic RNA spiked into NP samples (Table 2), which only consistently recovered at $24x10^6$ copies per mL (mean Ct=37.07) when tested in triplicate. However, when NP was combined with lysis buffer before spiking the RNA, recovery of RNA was improved to 750 copies per mL (Table 2). The three-fold difference in limit of detection between virus in a recombinant protein coat compared to genomic RNA suggests loss in recovery of genomic RNA or RNA quantitation differences.

Accuracy Study

Residual SARS-CoV-2 positive patient specimens were tested undiluted or diluted in VTM. These positive samples were diluted because of limited availability of positive samples at the time of assay development. A total of 45 positive samples were used (36 derived from positive patient samples and 9 positive commercial reagents (Accuplex)). Thirty negative samples included residual clinical specimens confirmed negative for SARS-CoV-2 (n=5),

residual clinical samples positive for respiratory pathogens (n=21, originally tested before December 2019), microorganism isolates (n=2), and reference quality control material (n=2) spiked into pooled residual NP collection matrix in VTM (total of 30 negative specimens). All specimens for accuracy were tested in a blinded manner. All 75 specimens were concordant (Table 3 and Table S6).

Reproducibility Studies

All runs were performed using aliquots of the same control material. For inter-assay reproducibility, the positive and negative controls were run across six assay runs and yielded Ct values with CVs of 1.65% and 1.02% for N1 and RP targets, respectively. For the intra-assay reproducibility, the positive and negative control samples were run in triplicate within a single run. Intra-assay reproducibility yielded CVs of 1.11% and 1.10% for N1 and RP respectively. The imprecision for the inter-technologist precision across four paired technologist combinations (one technologist for extraction and one technologist for RT-PCR analysis) was 1.49% and 0.62% for N1 and RP respectively. Concordance was 100%, and the CV was <2% and standard deviation was <0.5 Ct (Table S7).

Conclusions

In this study, we developed and validated a real-time PCR molecular assay to measure RNA from SARS-CoV-2 virus. The limit of detection of SARS-CoV-2 was 264 copies per mL for viral protein encapsulated RNA and 750 copies per mL for genomic RNA. *In silico* and analytical specificity studies showed no cross-reactivity with common respiratory pathogens. The test was validated with a total of 75 accuracy samples (45 samples positive for SARS-CoV-2 and 30 samples negative for SARS-CoV-2). The samples positive for SARS-CoV-2 RNA had Ct values ranging from 25.39 to 35.56.

A major hurdle to validation was the lack of access to SARS-CoV-2 live or inactivated virus. Purified genomic RNA was available but demonstrated variable efficiency in extraction recovery. Endemic coronaviruses were used as surrogates to optimize the extraction process. In addition, combining lysis buffer with NP specimens before spiking non-enveloped RNA improved recovery probably by decreasing RNA degradation. Improved RNA stability by spiking into matrix combined with buffer has been previously reported (12). This pre-extraction, offboard lysis protocol also has an advantage of improving the safety for laboratory testing personnel because it can be performed in a biosafety cabinet.

The challenge of limited control material or patient specimens may arise again in future infectious disease outbreaks. The quickest specimens to be available in the COVID-19 outbreak were *in vitro* transcribed RNA and genomic RNA. These RNAs were helpful to optimize post-extraction assay characteristics, but they showed poor extraction characteristics. In future outbreaks, production and widespread distribution of viral RNA within recombinant protein coat would improve the speed and reliability of molecular assay validation.

In summary, we validated a modified version of the CDC assay under the FDA's Emergency Use Authorization with optimizations in offboard lysis and the use of SARS-CoV-2 RNA in a recombinant viral protein coat. This assay may be used in high complexity labs for the diagnosis of SARS-CoV-2 in a high-throughput setting.

Acknowledgments

We would like to acknowledge the kind gift of *in vitro* transcribed SARS-CoV-2 viral genes by Dr. Nathan Grubaugh (Yale University) (13,14) and Dr. John Schoggins (University of Texas Southwestern Medical Center) which have been used as positive control materials.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: J.Y. Park, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: J.Y. Park, Miraca Holdings; Baylor Genetics, Fujirebio and SRL Labs are subsidiaries.
Stock Ownership: J. SoRelle, Myriad Genetics
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: None declared.

Role of Sponsor: No sponsor was declared.

References

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579:270–3.
2. Gorbalenya AE, Baker SC, Baric RS, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020;5:536–44.
3. CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel For Emergency Use Only Instructions for Use. Centers for Disease Control and Prevention, Atlanta, GA.
4. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020;25. doi: 10.2807/1560-7917.ES.2020.25.3.2000045
5. Coronavirus disease (COVID-19) technical guidance: Laboratory testing for 2019-nCoV in humans. World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance. (Accessed March 31, 2020)
6. Chan JF, Yip CC, To KK, et al. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-polymerase chain reaction assay validated in vitro and with clinical specimens. J Clin Microbiol. 2020 March 4. doi: 10.1128/jcm.00310-20.
7. Chu DKW, Pan Y, Cheng SMS, et al. Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. Clin Chem. 2020;66:549–55.
8. Policy for Diagnostic Tests for Coronavirus Disease-2019 during the Public Health Emergency. https://www.fda.gov/regulatory-information/search-fda-guidance-

documents/policy-diagnostic-tests-coronavirus-disease-2019-during-public-health-emergency. (Accessed March 31, 2020)

9. FAQs on Diagnostic Testing for SARS-CoV-2. https://www.fda.gov/medical-devices/emergency-situations-medical-devices/faqs-diagnostic-testing-sars-cov-2. (Accessed March 31, 2020)

10. Silva CS, Mullis LB. Human Respiratory Coronaviruses Detected In Patients with InfluenzaLike Illness in Arkansas, USA. Virol Mycol. OMICS Publishing Group; 2014;01.

11. Escutenaire S, Mohamed N, Isaksson M, et al. SYBR Green real-time reverse transcription-polymerase chain reaction assay for the generic detection of coronaviruses. Arch Virol. 2007;152:41–58.

12. Mitchell S, George KS, Rhoads D, Butler-Wu S, Dharmarha V, Miller M. Verification procedure for commercial tests with Emergency Use Authorization for the detection of SARS-CoV-2 RNA. https://asm.org/ASM/media/Protocol-Images/ASM_EUA_verification_040220_FINAL.pdf (Accessed March 31, 2020)

13. Vogels C, Brito AF, Wyllie AL et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 qRT-PCR assays. medRxiv. doi: https://doi.org/10.1101/2020.03.30.20048108 (Accessed April 16, 2020)

14. Vogels C, Fauver J, Ott I, Grubaugh N. Generation of SARS-CoV-2 RNA transcript standards for qRT-PCR detection assays. https://www.protocols.io/view/generation-of-sars-cov-2-rna-transcript-standards-bdv6i69e (Accessed April 16, 2020)

Table 1. Limit of Detection with Enveloped Virus Containing SARS-CoV-2

Virus Copies† per mL	Virus Copies per reaction	Total Replicates††	Positive Replicates	Positive Rate (%)	Mean Ct	1 SD
2113	40	20	19	95	31.66	1.29
1057	20	20	20	100	32.21	0.31
528	10	20	20	100	35.61	0.77
264	5	20	20	100	35.56	0.65
132	2.5	3	2	66	37.39	2.47

† Recombinant Sindbis virus containing SARS-CoV-2 RNA (Accuplex™ Reference Material). ††All concentrations were first performed in triplicate, only if all triplicate samples were detected was a set of 20 replicates examined. Ct - threshold cycle.

Table 2. LOD Determination Using Genomic SARS-CoV-2 RNA

	Virus Copies per mL	Virus Copies per reaction	Total Replicates†	Positive Replicates	Positive Rate (%)	Mean Ct
RNA spiked to NP prior to lysis buffer	$2.4x10^7$	450,000	3	3	100	37.07
	$1.2x10^7$	225,000	3	2	66	38.45
	$6x10^6$	112,500	3	1	33	38.19
	$3x10^6$	56,250	3	0	0	>40
NP added to lysis buffer prior to RNA	3000	56	3	3	100	31.34
	1500	28	3	3	100	32.62
	750	14	20	20	100	34.06
	375	7	20	15	75	37.23

†All concentrations were first performed in triplicate, only if all triplicate samples were detected was a set of 20 replicates examined. Ct - threshold cycle; NP- Nasopharyngeal matrix

Table 3. Accuracy

	Number Tested	Number Detected	Detected (%)
Clinical specimens* and Commercial Control**	45	45	100 (45 of 45)
Negative	30	0	0 (0 of 30)

*Clinical specimens were performed either undiluted, or diluted (1:2, 1:8, 1:10, 1:20). **Recombinant Sindbis virus containing SARS-CoV-2 RNA (Accuplex™ Reference Material).