Received: 21 April 2020 | Accepted: 5 May 2020

DOI: 10.1002/jmv.25988

SHORT COMMUNICATION

JOURNAL OF MEDICAL VIROLOGY WILEY

Comparison of SARS-CoV-2 detection from nasopharyngeal swab samples by the Roche cobas 6800 SARS-CoV-2 test and a laboratory-developed real-time RT-PCR test

Elisabet Pujadas[1] | Nnaemeka Ibeh[1] | Matthew M. Hernandez[2,3] | Aneta Waluszko[4] | Tatyana Sidorenko[4] | Vanessa Flores[5] | Biana Shiffrin[5] | Numthip Chiu[5] | Alicia Young-Francois[5] | Michael D. Nowak[1] | Alberto E. Paniz-Mondolfi[1] | Emilia M. Sordillo[1] | Carlos Cordon-Cardo[1] | Jane Houldsworth[1] | Melissa R. Gitman[1]

[1]Department of Pathology, Molecular, and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, New York

[2]The Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, New York

[3]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York

[4]Molecular Pathology Laboratory, Mount Sinai Health System, New York, New York

[5]Clinical Microbiology Laboratory, Mount Sinai Health System, New York, New York

Correspondence
Melissa R. Gitman, Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, 1425 Madison Ave, Rm 9-52C, New York, NY 10029.
Email: Melissa.gitman@mountsinai.org

Abstract

The urgent need to implement and rapidly expand testing for severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infection has led to the development of multiple assays. How these tests perform relative to one another is poorly understood. We evaluated the concordance between the Roche Diagnostics cobas 6800 SARS-CoV-2 test and a laboratory-developed test (LDT) real-time reverse transcription-polymerase chain reaction based on a modified Centers for Disease Control and Prevention protocol, for the detection of SARS-CoV-2 in samples submitted to the Clinical Laboratories of the Mount Sinai Health System. A total of 1006 nasopharyngeal swabs in universal transport medium from persons under investigation were tested for SARS-CoV-2 as part of routine clinical care using the cobas SARS-CoV-2 test with subsequent evaluation by the LDT. Cycle threshold values were analyzed and interpreted as either positive ("detected" or "presumptive positive"), negative (not detected), inconclusive, or invalid. Statistical analysis was performed using GraphPad Prism 8. The cobas SARS-CoV-2 test reported 706 positive and 300 negative results. The LDT reported 640 positive, 323 negative, 34 inconclusive, and 9 invalid results. When excluding inconclusive and invalid results, the overall percent agreement between the two platforms was 95.8%. Cohen's κ coefficient was 0.904 (95% confidence interval, 0.875-0.933), suggesting almost perfect agreement between both platforms. An overall discordance rate of 4.2% between the two systems may reflect differences in primer sequences, assay limit of detection, or other factors, highlighting the importance of comparing the performance of different testing platforms.

KEYWORDS

coronavirus, RNA extraction, SARS

Elisabet Pujadas and Nnaemeka Ibeh contributed equally to this work.

1 | INTRODUCTION

The 2019 novel coronavirus disease (COVID-19) pandemic has created the need for rapid and accurate diagnostic modalities. Reverse transcription-polymerase chain reaction (RT-PCR) has emerged as the primary mode of diagnosis of acute infection with the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) virus. Various groups have proposed different viral targets for the detection of the virus, including RNA-dependent RNA polymerase (RdRp), envelope (E), spike (S), open reading frame (ORF) 1a, and nucleocapsid (N). The Clinical Laboratories of the Mount Sinai Health System (MSHS Labs) perform testing for eight hospitals and their associated outpatient practices. In this study, we share our experience with two different testing platforms, the cobas 6800 SARS-CoV-2 test (Roche Molecular Systems, Branchburg, NJ) and a laboratory-developed test (LDT) real-time RT-PCR using the Centers for Disease Control and Prevention (CDC) 2019-nCoV primers and probes.[1]

2 | METHODS

Nasopharyngeal swabs (NPS) in universal transport medium from persons under investigation (PUIs) submitted to the MSHS Labs were tested as part of routine clinical care using the cobas SARS-CoV-2 test. A subset of 1006 samples collected between 17 March 2020 and 30 March 2020 were subsequently evaluated by the LDT for concordance. Both systems were previously validated against a panel of 30 NPS samples with known results from a reference laboratory (New York State Department of Health-Wadsworth Laboratories, New York City Department of Health and Mental Hygiene Public Health Laboratories, and Laboratory Corporation of America). NPS samples were transported at room temperature within 2 hours from the MSHS collecting location. The cobas SARS-CoV-2 test was set up upon receipt in the MSHS Labs. The samples for the LDT were then stored at 4°C for up to 3 days or at −70°C if longer.

For the cobas SAR-CoV-2 test, an 0.6 mL aliquot of each sample was loaded onto the Roche cobas 6800 where it was combined with the cobas SARS-CoV-2 master mix containing an internal RNA control, primers, and probes targeting the ORF1/a nonstructural region that is specific for SARS-CoV-2 (target 1), as well as the conserved, structural protein envelope E gene that is shared by the *Sarbecovirus* subgenus (target 2).

For the LDT, viral RNA was extracted from 120 to 140 µL of the same sample using either the QIAamp Viral RNA Mini Kit (Qiagen) on a QIAcube Connect (Qiagen) or the EZ1 DSP Virus Kit (Qiagen) on an EZ1 Advanced XL (Qiagen). Real-time RT-PCR was performed using the QuantiFast Pathogen RT-PCR Kit (Qiagen) in a LightCycler 480 II (Roche). The procedure otherwise followed the CDC protocol and used the same primers and probes as in the CDC 2019-nCoV Real-Time RT-PCR Diagnostic Panel.[1] These included N1, N2, and N3 probes that were selected from the regions of the virus nucleocapsid gene, with two that were specific for 2019-nCoV (N1 and N2), and also RNase P (RP) as an internal control. The limit of detection (LOD) for this assay was determined to be 1×10^3 viral RNA genome equivalents per mL.

The cycle threshold (C_t) values were reported by the cobas SARS-CoV-2 test as either "detected" (targets 1 and 2 detected), "presumptive positive" (target 1 not detected; target 2 detected), or "not detected." The C_t values were reported by the LDT as either "detected" ($N1_{Ct} < 38$ and $N2_{Ct} < 38$), "inconclusive" ($N1_{Ct} \geq 38$ and $N2_{Ct} < 38$, or $N1_{Ct} < 38$ and $N2_{Ct} \geq 38$), "not detected" ($N1_{Ct} \geq 38$, $N2_{Ct} \geq 38$, and $RP_{Ct} < 35$) or "invalid" ($N1_{Ct} \geq 38$, $N2_{Ct} \geq 38$, and $RP_{Ct} \geq 35$). In the absence of a gold standard, results between the two platforms were compared using overall percent agreement, with two-sided 95% score confidence intervals (CIs) for percent agreement calculated using Clopper-Pearson methodology.[2] To account for agreement due to chance, Cohen's κ coefficient was calculated.[2,3] Statistical analysis was performed using GraphPad Prism version 8 (GraphPad Software, La Jolla, CA).

3 | RESULTS

A total of 1006 NPS samples were tested by the cobas SAR-CoV-2 test and LDT. The cobas SARS-CoV-2 test reported 699 as "detected," 7 as "presumptive positive" (considered as detected) for a total of 706 "detected" and 300 as "not detected". The LDT reported 640 "detected", 323 "not detected", 34 inconclusive, and 9 invalids. After excluding invalid and inconclusive results, 639 of 678 (94.2%; 95% CI, 92.2-95.9) samples were in agreement between the "detected" results. Of the samples, 284 of 285 (99.6%; 95% CI, 98.1-99.9) were in agreement between the "not detected" results. This yielded an overall agreement in 923 of 963 (95.8%; 95% CI, 94.4-97.0) samples. A κ coefficient was calculated between the definitive results from the two platforms and was found to be 0.904 (95% CI, 0.875-0.933) suggesting almost perfect agreement between both platforms (Table 1).

4 | DISCUSSION

We compared the agreement between the cobas SARS-CoV-2 test (Roche Molecular Diagnostics) and an LDT based on a modified CDC

TABLE 1 Detected vs not detected results on the cobas SARS-COV-2 test and LDT

		Cobas SARS-CoV-2 test		
LDT		Detected	Not detected	Total
	Detected	639	1	640
	Not detected	39	284	323
	Total	678	285	963

- Overall percent agreement = 95.8%
- Positive percent agreement = 94.2%
- Cohen's κ coefficient: 0.904 (95% CI, 0.875-0.933)

Notes: Detected on cobas SARS-COV-2 test = targets 1 and 2 detected or target 1 not detected and target 2 detected.
Not detected on cobas SARS-COV-2 test = targets 1 and 2 not detected.
Detected on LDT = $N1_{Ct} < 38$ and $N2_{Ct} < 38$ and $RP_{Ct} < 35$.
Not detected = $N1_{Ct} \geq 38$, $N2_{Ct} \geq 38$ and $RP_{Ct} < 35$.
Abbreviations: CI, confidance interval; C_t, cycle threshold; LDT, laboratory-developed test real-time RT-PCR; RP, RNase P; SARS-COV-2; severe acute respiratory syndrome coronavirus-2.

protocol. There is currently no gold standard for the diagnosis of COVID-19; therefore, sensitivity and specificity could not be calculated. As an alternative, we calculated the overall agreement. We found a high degree of agreement between the two systems whether measured by overall percent agreement and by Cohen's κ. We found Cohen's κ coefficient to be 0.904 (95% CI, 0.875-0.933) indicating excellent agreement between the two tests.

We did detect a discordance rate of 4.2% between the cobas SAR2-CoV-2 test and the LDT. This is consistent with other reports that have found tests targeting the N gene to have a higher LOD than the E gene.[4] We note that of the samples detected by the cobas SARS-CoV-2 test and inconclusive on LDT, the majority (22/25) would have been classified as detected if the criteria were modified to consider detection of N1 sufficient. It is worth considering whether these criteria may be excessively strict, especially given that all samples with inconclusive results that were N1 positive were resulted as detected by the cobas SARS-CoV-2 test. Conversely, it has been reported that the N2 target may be prone to false positive results[5]; detection of the N2 target alone was found in only a small minority of the inconclusive results (3/25) from the LDT test.

The category of samples resulting as "presumptive positive" on the cobas SARS-CoV-2 test but "not detected" on the LDT present an interesting dilemma in the absence of a gold standard test, particularly in light of known SARS-CoV-2 carriage by asymptomatic persons, and anecdotal clinical reports of patients with COVID-19-like presentations despite negative tests for SARS-CoV-2. At this time, the possibility that these are false positive results, arising either from spurious amplification or from detection of a closely related virus, cannot be excluded. Another possibility is that these samples contain low levels of virus, near the limit of detection for the cobas SARS-CoV-2 tests but below the threshold of detection for the LDT.

Similarly, samples with "detected" results by the cobas SARS-CoV-2 test that yielded "not detected" results by the LDT may have been below the limit of detection for the LDT. Examination of C_t values in discordant samples revealed a wide range of results, suggesting that while low viral loads in some samples may be contributory, they do not account for all the observed differences. An alternative explanation is that differences in the regions of the virus targeted by the two assays play an important role, as has been explored by others.[5,6] Interpretation is further complicated by the external limitation of inconsistent availability of access to essential reagents including polymerases, buffers or controls due to COVID-19 related shortages, necessitating alternative steps in the LDT procedure. An additional consideration given the sequential workflow for testing would be the longer interval between sample collection and performance of LDT. Although stored at 4°C or frozen at −70°C, some sample degradation during refrigeration or the freeze-thaw process may have occurred, although after review, the time elapsed between performances of the cobas SARS-CoV-2 test versus the LDT did not appear to be related to the discrepancy rate. Of note, Poljak et al[6] reported a similar experience with the cobas 6800 SARS-CoV-2 test as compared to an LDT targeting E ans RdRp coronavirus genes as developed at Charité Hospital, Berlin, Germany, finding great concordance (98.1%-99.6% agreement) between the two assays. In conjunction with the high prevalence of SARS-CoV-2 in New York City and that testing was limited to symptomatic patients meeting criteria for PUI at the time the NPS samples included in our study were collected, it is more likely that these samples represent true positives.

This study has several important limitations. The first is that due to the very high volume of clinical samples requiring testing and limited availability of reagents, it was not feasible to test all samples received by our laboratories on both platforms. A second limitation is that clinical understanding of COVID-19 infection is still evolving, so that currently there is no clinical or laboratory gold standard to serve as an absolute reference for comparison. Further studies will be needed to evaluate the performance of diagnostic tests for SARS-CoV-2 infection as this information becomes available, but is beyond the scope of the current paper.

In conclusion, we observed overall excellent agreement between the two tests methods, although our results suggest that cobas SARS-CoV-2 test may have a lower limit of detection than the LDT based on a modified CDC assay. Further studies are needed to compare testing platforms, along with correlation with clinical data to guide testing decisions.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Viviana Simon for here assistance in the development of the LDT.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Elisabet Pujadas http://orcid.org/0000-0003-4795-7945
Melissa R. Gitman http://orcid.org/0000-0002-6960-0763

REFERENCES

1. https://www.fda.gov/emergency-preparedness-and-response/mcm-legal-regulatory-and-policy-framework/emergency-use-authorization. Accessed 15 April 2020.
2. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. https://doi.org/10.2307/2F2529310
3. Guidance for Industry and FDA Staff Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-guidance-reporting-results-studies-evaluating-diagnostic-tests-guidance-industry-and-fda. Accessed 5 January 2020.
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(3):2000045.
5. Vogels C, Brito A, Wyllie A, et al. Analytical sensitivity and efficiency comparisons of SARS-COV-2 qRT-PCR assays. *medRxiv*. 2020. https://doi.org/10.1101/2020.03.30.20048108

6. Poljak M, Korva M, Knap Gašper N, et al. Clinical evaluation of the cobas SARS-CoV-2 test and a diagnostic platform switch during 48 hours in the midst of the COVID-19 pandemic [published online ahead of print April 10, 2020]. *J Clin Microbiol*. 2020. https://doi.org/10.1128/JCM.00599-20

How to cite this article: Pujadas E, Ibeh N, Hernandez MM, et al. Comparison of SARS-CoV-2 detection from nasopharyngeal swab samples by the Roche cobas 6800 SARS-CoV-2 test and a laboratory-developed real-time RT-PCR test. *J Med Virol*. 2020;92:1695–1698. https://doi.org/10.1002/jmv.25988