The versatility of external quality assessment for the surveillance of laboratory and in vitro diagnostic performance: SARS-CoV-2 viral genome detection in Austria

https://doi.org/10.1515/cclm-2021-0604
Received March 31, 2021; accepted June 21, 2021; published online June 29, 2021

Abstract

Objectives: External quality assessment (EQA) schemes provide information on individual and general analytical performance of participating laboratories and test systems. The aim of this study was to investigate the use and performance of SARS-CoV-2 virus genome detection systems in Austrian laboratories and their preparedness to face challenges associated with the pandemic.

Methods: Seven samples were selected to evaluate performance and estimate variability of reported results. Notably, a dilution series was included in the panel as a measure of reproducibility and sensitivity. Several performance criteria were evaluated for individual participants as well as in the cohort of all participants.

Results: A total of 109 laboratories participated and used 134 platforms, including 67 different combinations of extraction and PCR platforms and corresponding reagents. There were no false positives and 10 (1.2%) false negative results, including nine in the weakly positive sample ($C_t \sim 35.9$, $\sim 640$ copies/mL). Twenty (22%) laboratories reported results of mutation detection. Twenty-five (19%) test systems included amplification of human RNA as evidence of proper sampling. The overall linearity of $C_t$ values from individual test systems for the dilution series was good, but inter-assay variability was high. Both operator-related and systematic failures appear to have caused incorrect results.

Conclusions: Beyond providing certification for participating laboratories, EQA provides the opportunity for participants to evaluate their performance against others so that they may improve operating procedures and test systems. Well-selected EQA samples offer additional inferences to be made about assay sensitivity and reproducibility, which have practical applications.

Keywords: $C_t$; cycle threshold; EQA; external quality assessment; PCR; SARS-CoV-2.

Introduction

Regulation (EU) 2017/746 on In Vitro Diagnostic Medical Devices (IVDR) requires performance of in vitro diagnostics
Buchta et al.: Versatility of SARS-CoV-2 PCR external quality assessment

During a pandemic, the preparedness of laboratories to face arising challenges is of great interest. For example, it has been reported that several mutations in the SARS-CoV-2 genome are associated with reduced sensitivity of routine RT-qPCR systems [12–14]. To reduce the risk of positive samples escaping detection, the International Federation for Clinical Chemistry and Laboratory Medicine (IFCC) recommended the use of at least two SARS-CoV-2-specific gene targets in test systems [15]. An EQA can check to what extent this recommendation has been implemented. Another interesting point that EQA schemes can clarify is the preparedness of laboratories to interpret and report divergent or unexpected results (e.g., target gene “dropout”). This may include, for example, further examinations of putative SARS-CoV-2-positive samples for possible mutations or including amplification of human RNA as an internal control of sample quality [16].

The Austrian Association for Quality Assurance and Standardization (ÖQUASTA) is the national EQA provider in Austria and serves most of the relevant laboratories in the country. The EQA scheme “SARS-CoV-2 virus genome detection” was established in cooperation with the National Reference Laboratory for Respiratory Viruses at the Center for Virology of the Medical University Vienna, Austria, in early 2020, and two rounds have since been performed. For the third round (the first round in 2021), samples were selected so that the measurement results provided information in addition to detection of viral nucleic acid. Herein, we evaluated the current nationwide performance in SARS-CoV-2 genome detection, and we assess specific properties of laboratories and test systems using a specifically designed panel of sample materials.

Materials and methods

Organization of the EQA scheme

The “SARS-CoV-2 virus genome detection” scheme of ÖQUASTA and the Center for Virology of the Medical University Vienna corresponds to the general EQA scheme as summarized by the European Organization for External Quality Assurance Providers in Laboratory Medicine (EQALM) [17].

Sample preparation and testing

The EQA test panel consisted of seven samples (900 μL each), five positive and two negative. The intended use of the individual samples with regard to respective challenges is summarized in Table 1. Positive samples S2 (nCoV-19/Austria/MUW_1359197/2021, GISAID: EPI_ISL_1020300) and S4 (nCoV-19/Austria/AGES804699/2021, GISAID: EPI_ISL_1020299) were obtained by 100-fold dilution of individual SARS-CoV-2 positive oro-nasopharyngeal swab samples with sterile sodium chloride solution (1.5 mL swab samples + 148.5 mL 0.9% NaCl). S4 contained a lineage B.1.1.7 variant virus [18] and was incubated at 56 °C for 30 min to reduce potential infectivity. Positive samples S3, S6, and S7 were obtained from cell culture supernatant of SARS-CoV-2 infected Vero cells (strain nCoV-19/Austria/CeMM0360/2020, GISAID: EPI_ISL_438123 [19]). S3 was a 10-fold dilution of S6, and S7 was a 100-fold dilution of S6. One negative sample, S1, was obtained by 100-fold dilution of a SARS-CoV-2 negative oro-nasopharyngeal swab. The second negative sample, S5, consisted only of the diluent (0.9% NaCl solution).
For quality control, all samples were stored at −80 °C and one panel was tested before shipment as follows: after thawing, samples were kept at ambient temperature for 2 days (anticipated maximum shipping time) before further storage at 4 °C for 4 days (maximal expected storage time before testing). Thereafter, samples were measured with an in-house SARS-CoV-2 molecular detection assay, and quantified by digital droplet PCR, as described previously [20]. Briefly, nucleic acid (NA) was extracted from a 200 μL aliquot of the original sample using the NucliSENS easyMAG platform (bioMérieux, France). The elution volume was 50 μL. For RT-qPCR, each 25 μL reaction mixture contained 12.5 μL of 2× reaction buffer, 0.4 μL of a 50 mM magnesium sulfate solution, 1 μL of SuperScript III RT/Platinum Taq Mix (SuperScript III One-Step RT-PCR System, Invitrogen, Germany), 0.5 μL each of LightMix® Modular SARS-CoV (COVID-19) E-gene and LightMix® Modular EAV RNA Extraction Control (TIB Molbiol, Berlin, Germany) and 10 μL of eluted NA as the template. Thermal cycling was performed using a Light Cycler 480 II (Roche, Vienna, Austria) under the following conditions: reverse transcription at 55 °C for 20 min; denaturation at 94 °C for 3 min; 45 amplification cycles of 94 °C for 15 s and 58 °C for 30 s (acquisition steps); a final incubation at 40 °C for 30 s. Analysis of the five positive samples was repeated three times and used as reference values (Table 2). Samples S1, S2, and S4 tested positive for beta actin as a marker for the presence of human genetic material.

**Table 1:** Sample selection for the SARS-CoV-2 virus genome detection EQA scheme: Samples S1–S6 and their intended use for respective challenges.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Sample</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Neg. swab</td>
<td>Pos. swab</td>
<td>Virus culture</td>
<td>Pos. swab</td>
<td>NaCl</td>
<td>Virus culture</td>
<td>Virus culture</td>
</tr>
<tr>
<td><strong>Individual and collective analytical performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytical sensitivity</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Analytical specificity</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Variability of SARS-CoV-2-specific C&lt;sub&gt;q&lt;/sub&gt; values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linearity of C&lt;sub&gt;q&lt;/sub&gt; values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detection of B.1.1.7 variant virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variability of C&lt;sub&gt;q&lt;/sub&gt; values in B.1.1.7 virus sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detection of &quot;invalid&quot; negative sample (no human RNA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variability of C&lt;sub&gt;q&lt;/sub&gt; values of human RNA</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences of C&lt;sub&gt;q&lt;/sub&gt; values in test systems employing more than one target gene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences in gene-specific C&lt;sub&gt;q&lt;/sub&gt; values</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Information about the community of participant laboratories**

<table>
<thead>
<tr>
<th>Different test systems in use</th>
<th>Number of SARS-CoV-2-specific gene targets employed in test systems</th>
<th>Proportion of test systems using a human housekeeping gene</th>
<th>Number of laboratories prepared to detect virus variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1: negative, diluted swab (human RNA positive). Sample 2: positive, diluted swab (human RNA positive), C&lt;sub&gt;q&lt;/sub&gt; ∼27.0, ∼252,000 copies/mL. Sample 3: positive, virus culture supernatant, 1:10 dilution of Sample 6 (human RNA negative), C&lt;sub&gt;q&lt;/sub&gt; ∼31.6, ∼22,300 copies/mL. Sample 4: positive, diluted swab, variant lineage B.1.1.7 (human RNA positive), C&lt;sub&gt;q&lt;/sub&gt; ∼26.3, ∼345,300 copies/mL. Sample 5: negative, sodium chloride solution (human RNA negative). Sample 6: positive, virus culture supernatant (human RNA negative), C&lt;sub&gt;q&lt;/sub&gt; ∼28.8, ∼96,100 copies/mL. Sample 7: positive; virus culture supernatant, 1:100 dilution of Sample 6 (human RNA negative); C&lt;sub&gt;q&lt;/sub&gt; ∼35.3, ∼640 copies/mL.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Reported results and test sample (S1–S7) specifications for the detection of SARS-CoV-2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin</td>
<td>Swab</td>
<td>Swab</td>
<td>Virus culture</td>
<td>Swab</td>
<td>NaCl</td>
<td>Virus culture</td>
<td>Virus culture</td>
</tr>
<tr>
<td>SARS-CoV-2</td>
<td>Negative</td>
<td>Positive</td>
<td>Positive</td>
<td>Swab</td>
<td>Negative</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Human RNA</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>C&lt;sub&gt;v&lt;/sub&gt; value (E-gene, mean; 95% C.I.)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&gt;45</td>
<td>27.2;</td>
<td>32.1;</td>
<td>26.7;</td>
<td>&gt;45</td>
<td>29.6;</td>
<td>35.9;</td>
</tr>
<tr>
<td>Viral RNA copies/mL</td>
<td>27.0–27.5</td>
<td>31.8–32.4</td>
<td>26.6–26.9</td>
<td>29.3–29.9</td>
<td>35.1–36.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reported results

| Total | 134 | 134 | 134 | 134 | 134 | 134 |
| Positive | 132 (99%) | 130 (97%) | 132 (99%) | 131 (98%) | 115 (86%) |
| Negative | 131 (98%) |
| Invalid | 3 (2%) | 2 (1%) | 3 (2%) | 2 (1%) | 24 (17%) |

Suspected or proven mutation

| Mutation | 1 | 20 | 1 |
| N501Y, B.1.1.7 | |
| B.1.351/P1 | 1 | 1 |

Reported C<sub>v</sub> values<sup>b</sup>

| E-gene | 27.4; 25.4–33.0 | 32.8; 30.5–40.0 | 28.6; 26.1–32.0 | 30.4; 28.5–35.0 | 36.2; 33.8–39.1 |
| N-gene | 28.3; 19.7–35.0 | 33.6; 25.2–39.1 | 29.3; 20.8–35.0 | 31.4; 25.1–37.0 | 37.2; 29.5–42.5 |
| S-gene | 24.4; 20.6–30.0 | 29.7; 25.8–35.7 | 26.2; 23.5–30.7 | 27.6; 24.3–32.9 | 33.6; 29.8–38.8 |
| RdRP-gene | 3.1; 15 | 3.5; 15 | 2.5; 12 | 2.9; 15 | 3.3; 11 |
| ORF1ab-gene | 27.9; 19.3–32.2 | 33.8; 29.8–38.4 | 29.5; 25.2–34.9 | 31.2; 26.6–34.8 | 37.5; 32.8–40.2 |
| Other<sup>c</sup> | 27.3; 24.1–32.1 | 32.2; 25.4–36.9 | 28.3; 25.1–33.9 | 29.8; 25.4–36.6 | 35.9; 25.4–41.4 |

Human RNA targets

| Beta actin | 24.6; 23.6–25.8 | 31.3; 29.4–32.8 | 37.0; 30.7–37.1 | 29.5; 22.7–31.1 | 33.4; 30.0–36.8 | 36.0; 34.2–37.2 |
| Beta globin | 31.0–31.0 | 31.0–31.0 | 31.0–31.0 | 31.0–31.0 | 31.0–31.0 | 31.0–31.0 |
| RNase P | 27.0; 22.0–38.6 | 32.8; 26.5–38.2 | 35.3; 30.7–39.1 | 31.1; 26.3–39.6 | 35.2; 25.3–39.9 | 37.7 |
| 0.0; 2 | 1.4; 2 | 1.4; 2 |

<sup>a</sup>Cross-reactive binding of primers targeting human beta actin, beta globin, and RNase P to Chlorocebus aethiops present in the Vero cell culture supernatant. <sup>b</sup>Target Ct values provided by the National Reference Laboratory. <sup>c</sup>Reported Ct values are summarized as the mean; range; standard deviation; and total stratified by target and sample. <sup>d</sup>“Other” included results obtained from two or more gene targets and therefore not shown separately.

Sample dispatch and instructions for participants

Samples were dispatched to participant laboratories by overnight delivery at ambient temperature conditions. The instructions provided with the samples recommended that the material either be tested immediately or, alternatively, be stored at 4 °C for a maximum of 4 days after arrival until analysis and to conduct tests in the same manner as for routine samples. Individual laboratories that wanted to participate...
with more than one protocol had to register each test system separately and received independently shipped sample sets for each of them.

Data collection and assessment

Results of the qualitative interpretation (positive, negative, invalid) and test systems used were to be reported to the EQA provider via a web portal, email, fax or on paper in regular mail within 12 days. Participants were asked to report target gene(s) and C_t values for both virus-specific genes and, if applicable, for human RNA used by their test systems as a quality control, along with qualitative results. A sample was classified as “Invalid” if the results for SARS-CoV-2-specific gene targets and human RNA were negative. Only qualitative results were officially assessed for certification in this round. Assignment of qualitative targets for samples followed variant (c) “from knowledge of the origin of preparation of the proficiency item(s)” in chapter 11.3.1 of the applicable standard ISO 13528:2015, “Statistical methods for use in proficiency testing by interlaboratory comparison” [21]. Reported results were collected and evaluated in customized QuoStat software (QuoData, Dresden, Germany). Individual results were judged as either correct (=matching the target) or incorrect (=not matching the target) and in addition to individual reports, data for all test systems were made available in a general report.

Statistical analysis of reported C_t values included reporting basic summary statistics (mean, ranges, standard deviations and coefficients of variation) stratified by sample and/or gene target. Mann–Whitney U tests were used to compare C_t values between targets for a given sample, using Type I Error rate of alpha = 0.05.

Results

Information on participants

Results from 109 laboratories with a total of 134 individual test systems were available for evaluation. The test systems included combinations of nucleic acid extraction platforms and reagents and RT-qPCR platforms and reagents. In total, 67 different combinations of extraction and detection systems were used (Supplementary Material, Figure S1). Sample volume used varied from 2.5 to 900 μL and maximum cycle count varied between 38 and 50, as reported by participants and not verified by the authors. Ninety-one laboratories participated with one test system/protocol, and 18 laboratories used more than one test system/protocol. In particular, 14 laboratories used two, two laboratories used three, and two individual laboratories used either four or five distinct test systems, respectively. Gene-specific C_t values were reported for a total of 122 registered test systems.

IFCC’s recommendation to use at least two different target genes in SARS-CoV-2 detection systems was met by 79 (65%) participant test systems (two target genes were reported for 68 (56%), and three targets were reported for 11 (9%) test systems). A total of 25 (19%) participant systems included testing for human RNA as an internal sampling control. Human RNase P was reported to be used for 15 systems, beta actin for six, beta globin for three, and VPS29 for one system. Furthermore, 20 (22%) laboratories reported that they routinely test SARS-CoV-2-positive samples for mutations (Table 2).

Analytical performance

A total of 124 out of 134 test systems (93%) reported correct results for all five positive and both negative samples (Table 2). A total of 1,006 individual C_t values were reported from 122 test systems targeting the following viral gene targets: E (250 C_t values); N (337); S (68); RdRP (112); ORFlab (105); or “other” undisclosed gene targets (128), which may have included a single result obtained from two or more gene targets. Summary statistics of C_t values were calculated for individual targets and “other” target genes (Table 2 and Figure 1).

First, we note that the reported individual C_t values from each laboratory corresponded well with viral loads determined by the reference laboratory RT-qPCR and ddPCR (Figure 1). Generally the mode of the reported C_t values (all targets/test systems combined) were within 0.2–3.3 C_t values of the target mean C_t values (S2 = ~252,000/mL, mean target C_t = 27.2, mode = 27; S3 = ~22,300/mL, mean target C_t = 32.1, mode = 34; S4 = ~345,300/mL, mean target C_t = 26.7, mode = 30; S6 = ~96,100/mL, mean target C_t = 29.6, mode = 31; and S7 = ~640/mL, mean target C_t = 35.9, mode = 37) (Table 2, Figure 1).

We calculated the mean differences of C_t values in three test systems that target more than one gene (Figure 2). For GeneXpert/Xpert Xpress SARS-COV-2, mean difference between C_t for the N- and E-gene was 2.6; for Roche Cobas 6800,8800/cobas SARS-CoV-2 it was 0.1 for the E- and the ORFlab-gene; for Simplexa Liaison MDS Realtime PCR/Simplexa COVID-19 Direct Reaction Mix it was 0.4 for the ORFlab- and the S-gene.

The three samples from the dilution series were used to assess the linearity of C_t values (Figure 3). While we did not perform statistical analysis of these data, we note that approximately linear results were obtained from many test systems, but some were obviously non-linear (see Figure 3).

All test systems that included an internal control correctly reported the presence/absence of human RNA in all samples except for one participant reporting a false negative result. The sensitivity for human RNA (inferred from C_t values) varied between samples (coefficient of variation of C_t values for a given sample <16%). We found no
Sample S2 (C<sub>t</sub> ~27.2, ~252000 copies/ml)

Sample S3 (C<sub>t</sub> ~32.1, ~22300 copies/ml)

Sample S4 (C<sub>t</sub> ~26.7, ~345300 copies/ml)

Sample S6 (C<sub>t</sub> ~29.6, ~96100 copies/ml)

Sample S7 (C<sub>t</sub> ~35.9, ~640 copies/ml)

C<sub>t</sub> 27; n=42

C<sub>t</sub> 34; n=35

C<sub>t</sub> 30; n=45

C<sub>t</sub> 31; n=43

C<sub>t</sub> 37; n=27
differences in $C_t$ values between targets for the same sample (Mann–Whitney U tests, p>0.05) (Table 2 and Figure 1).

The sample containing B.1.1.7 variant virus (S4) was reported positive by all participants, but only 20 (22%) reported evidence of a mutation. Of these, 19 participants specifically reported S4 as either positive for the substitution in the spike protein at site N501Y and/or the deletions at amino acid positions H69/V70 or identified the sample as a lineage B.1.1.7 virus. One participant indicated they suspected variant lineage B.1.351 or P1 in Samples 3, 4, and 7. We noted that the S-gene target PCR of the test system of two participants ($2\times$ TaqPath) did not detect the lineage B.1.1.7.

**Discussion**

Evaluation of results of this round revealed good analytical performance of participants. There were no false positives and only a few false negatives, which were mostly in the sample with the highest mean $C_t$ value (“weakly positive”). As this was the third EQA round, we note the improvement in the false negative ratio over the two previous rounds: in the first and the current third round the weakly positive samples had roughly similar $C_t$ values (38 and 36, respectively) and yet the rate of false negatives fell from 40.3 to 6.7% (per 67 and 134 reported tests, respectively) [22].
In some cases, we could deduce the possible cause for the false negative results. One participant reported false “negative” for S3 and S7 using a test system that requires 400 µL of primary sample, but the participant reported that only 30 µL was used. Each Ct value reported by this participant was 3–4 cycles higher than other participants’ results, and therefore it is likely that low template led to a false negative result for two samples. Two participants reported the interpretation “negative” for S7, and yet reported Ct values between 36.8 and 39.5. We assume these were interpreted “negative” according to an internal laboratory cut-off value. For another five, no cause for the false negative result could be identified based on the data provided.

IFCC’s recommendation to use at least two different target genes in SARS-CoV-2 detection systems was met by only about two thirds of participant test systems. We were able to analyze three test systems that utilize two gene targets for differences in Ct values between the targets. These differences may be useful for laboratory quality control, for example, if the difference deviates beyond defined limits there may be indirect evidence of possible mutations in the target sequences. We note that one test system had a uniformly large difference in Ct values between the targets for the same sample (mean 2.6), the differences in targets between the other two test systems were similar (Figure 2). This may have implications for the use of Ct values as a qualitative assessment of patient status.

Figure 3: Linearity of Ct values in S6, S3, and S7 of the most frequently used gene targets for a dilution series derived from a virus isolate grown in cell culture.

Lines connect the Ct values reported by each participant from a dilution series of nCoV-19/Austria/CeMM0360/2020 [EPI_ISL_438123] (S6 undiluted, S3 = 10-fold dilution, S7 = 100-fold dilution) from various test systems for five viral gene targets (E, N, RdRP, and ORF1b) or “other” for systems which calculate Ct based on two gene targets.

In some cases, we could deduce the possible cause for the false negative results. One participant reported false “negative” for S3 and S7 using a test system that requires 400 µL of primary sample, but the participant reported that only 30 µL was used. Each Ct value reported by this participant was 3–4 cycles higher than other participants’ results, and therefore it is likely that low template led to a false negative result for two samples. Two participants reported the interpretation “negative” for S7, and yet reported Ct values between 36.8 and 39.5. We assume these were interpreted “negative” according to an internal laboratory cut-off value. For another five, no cause for the false negative result could be identified based on the data provided.

IFCC’s recommendation to use at least two different target genes in SARS-CoV-2 detection systems was met by only about two thirds of participant test systems. We were able to analyze three test systems that utilize two gene targets for differences in Ct values between the targets. These differences may be useful for laboratory quality control, for example, if the difference deviates beyond defined limits there may be indirect evidence of possible mutations in the target sequences. We note that one test system had a uniformly large difference in Ct values between the targets for the same sample (mean 2.6), the differences in targets between the other two test systems were similar (Figure 2). This may have implications for the use of Ct values as a qualitative assessment of patient status.

Reporting of Ct values of positive results is still requested by medical authorities and may offer additional qualitative assessment of a virus-positive result [23]. However, our results clearly demonstrate that caution should be taken when comparing and interpreting Ct values. First, we observed the well-documented property of RT-qPCR, that primer-probe sets specific for different genomic targets produce different Ct values for the same sample. This is likely related to reaction efficiency and/or the relative abundance RNA template for the different viral genes. Second, the range of Ct values reported for the same target, in the same sample, by different participants was on average approximately 10 Ct units. However, the variability of Ct values was less than in the previous round of this EQA scheme [20]. Finally, in the previous EQA scheme, the data seemed to suggest that laboratories were at least consistently over- or under-estimating the mean Ct value for a given sample. We therefore tested this hypothesis with a
dilution series of a virus culture supernatant. The results revealed that many test systems reported approximately linear changes in Ct, according to template concentration, but some were notably non-linear (Figure 3).

Possible reasons for the high variability revealed by individual Ct values in general, and departures from expected efficiency in the dilution series, include differences in assay performance and/or in EQA materials characteristics. Commutability of an EQA material describes its ability to properly mimic the behavior of real patient samples that have not been modified [24, 25]. Although a formal commutability analysis could not be performed, the fact that similar results were obtained in both swab and virus culture samples suggests that assays are likely not greatly impacted by matrix effects, at least for the materials involved in this study. Although the amplification of human RNA as an internal control (e.g., as evidence of proper sampling via swab) was reported by some labs, it does not appear to be very widespread. However, if laboratories do not supervise correct sampling, including such an internal control may be helpful to identify invalid samples and prevent the reporting of false negative results. As mentioned, we noted variability in the volume of sample used for extraction, and believe this, combined with assay efficiency related to PCR reagents (primers, polymerases, buffer systems), may have had the greatest impact on the variability of individual Ct values.

In conclusion, this round of EQA was beneficial in simultaneously serving several interested parties. Participant laboratories have received feedback on the quality of their individual results, which had improved over prior rounds. Supervisory authorities can read information about the general analytical performance and laboratory preparedness; and here we note that, from a technical standpoint, considerations should be made concerning the use of Ct values to assess patient status. Finally, we demonstrated that EQA may be useful to IVD manufacturers (as well as diagnostic laboratories) to compare the performance of products in routine use.

Acknowledgments: We gratefully acknowledge all laboratories that participated in this study and made special efforts to report more data than was required to participate in this EQA round.

Research funding: None declared.

Author contributions: CB: Conceptualized, conducted and analyzed this EQA study, wrote and edited the manuscript draft. JVC: Conducted and analyzed this EQA study, visualized data, critically reviewed and edited the manuscript. JJ: Analyzed data and provided technical EQA support, critically reviewed manuscript. PC: Provided scientific advice, critically reviewed and edited the manuscript. EPS: Provided scientific advice, critically reviewed the manuscript. MM: Performed and analyzed ddPCR of samples, wrote the respective part of the manuscript. HP: Performed and analyzed ddPCR of samples, wrote the respective part of the manuscript. AB: Performed and analyzed sequencing of samples, wrote the respective part of the manuscript. AL: Performed and analyzed sequencing of samples, wrote the respective part of the manuscript. AMP: Performed and analyzed sequencing of samples, wrote the respective part of the manuscript. AMP: Performed and analyzed sequencing of samples, wrote the respective part of the manuscript. LE: Performed and analyzed sequencing of samples, wrote the respective part of the manuscript. WH: Performed statistical evaluation and visualized data, critically reviewed manuscript. BB: Critically reviewed and edited the manuscript. VD: Provided scientific advice regarding commutability, wrote the respective part of the manuscript. MMM: Provided scientific advice to the overall study, reviewed and edited the manuscript. AG: Provided scientific advice to the overall study, reviewed and edited the manuscript. SWA: Provided sample material, conceptualized, conducted and supervised this EQA study, provided scientific advice to the study, reviewed and edited the manuscript. IG: Conceptualized, conducted and analyzed this EQA study, visualized data, wrote and edited the manuscript draft. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

Informed consent: Not applicable.

Ethical approval: Not applicable.

References


Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/ccm-2021-0604).