

Christoph Buchta, Jeremy V. Camp, Jovana Jovanovic, Peter Chiba, Elisabeth Puchhammer-Stöckl, Maximilian Mayerhofer, Helga Plicka, Alexander Lercher, Alexandra M. Popa, Lukas Endler, Andreas Bergthaler, Wolfgang Huf, Bernhard Benka, Vincent Delatour, Mathias M. Müller, Andrea Griesmacher, Stephan W. Aberle and Irene Görzer*

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 ($C_t \sim 35.9$, ~ 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.

Christoph Buchta and Jeremy V. Camp contributed equally to this work.

*Corresponding author: Irene Görzer, Center for Virology, Medical University of Vienna, Vienna, Austria,
E-mail: irene.goerzer@meduniwien.ac.at

Christoph Buchta, Jovana Jovanović, Peter Chiba, Mathias M. Müller and Andrea Griesmacher, Austrian Association for Quality Assurance and Standardization of Medical and Diagnostic Tests (ÖQUASTA), Vienna, Austria

Jeremy V. Camp, Elisabeth Puchhammer-Stöckl and Stephan W. Aberle, Center for Virology, Medical University of Vienna, Vienna, Austria

Maximilian Mayerhofer and Helga Plicka, Armament and Defence Technology Agency, NBC & Environmental Protection Technology Division, Vienna, Austria

Alexander Lercher, Alexandra M. Popa, Lukas Endler and Andreas Bergthaler, CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria

Wolfgang Huf, Karl Landsteiner Institute for Clinical Risk Management, Vienna, Austria. <https://orcid.org/0000-0002-1479-2972>

Bernhard Benka, Federal Ministry of Social Affairs, Health, Care and Consumer Protection, Vienna, Austria

Vincent Delatour, Laboratoire National de Métrologie et d'Essais (LNE), Paris, France

Introduction

Regulation (EU) 2017/746 on In Vitro Diagnostic Medical Devices (IVDR) requires performance of *in vitro* diagnostics

(IVD) to be monitored by external quality assessment (EQA) schemes [1]. While the contribution of EQA schemes to post-market surveillance is not yet fully developed and standardized, EQA for medical laboratories has a tradition going back decades [2]. In such schemes, EQA providers distribute samples with known but undisclosed properties or content. Participating laboratories analyze these samples and report results to the EQA provider. Reported results are compared with assigned targets and scored as either matching (correct) or non-matching (false). Participants learn in individual reports how comparable their results are with those of their peers and gain additional educational effects from the general report [3]. Therefore, their motivation to participate in EQA is to have results assessed by an independent third party and to compare them with those obtained with both identical and alternative analytical systems. However, mere assessment of performance does not fully exploit the real potential of EQA schemes. The underutilized potential of such schemes lies in the aggregation of results, which provides an overview of the analytical performance and competencies of laboratories and test systems. This data is therefore useful for both post-market surveillance and epidemiological purposes.

Molecular diagnostics for SARS-CoV-2 infection from nasopharyngeal, nasal or oro-pharyngeal swabs were available in early 2020 and were quickly adopted by laboratories around the world. SARS-CoV-2-specific real time reverse transcription (RTq-) PCR is currently the gold standard for confirmation or exclusion of infection, monitoring COVID-19 therapy and infectious and epidemiologic follow-up testing [4, 5]. According to manufacturers' specifications, the reliability of such test results should be high. By and large, previously reported EQA results have confirmed that the various participating systems perform well for SARS-CoV-2 viral genome detection [6–11]. However, we believe an EQA scheme offers an opportunity to further probe the specific skills and competencies of participants by including samples with specific properties; and to our knowledge this approach has not yet been reported.

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-pharyngeal swab. The second negative sample, S5, consisted only of the diluent (0.9% NaCl solution).

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

Challenge	Sample	S1 Neg. swab	S2 Pos. swab	S3 Virus culture	S4 Pos. swab	S5 NaCl	S6 Virus culture	S7 Virus culture
Individual and collective analytical performance								
Analytical sensitivity			x	x	x		x	x
Analytical specificity	x					x		
Variability of SARS-CoV-2-specific C_t values			x	x	x		x	x
Linearity of C_t values				x			x	x
Detection of B.1.1.7 variant virus					x			
Variability of C_t values in B.1.1.7 virus sample					x			
Detection of “invalid” negative sample (no human RNA)						x		
Variability of C_t values of human RNA	x		x		x			
Differences of C_t values in test systems employing more than one target gene			x	x	x		x	x
Differences in gene-specific C_t values			x	x	x		x	x
Information about the community of participant laboratories								
Different test systems in use								
Number of SARS-CoV-2-specific gene targets employed in test systems								
Proportion of test systems using a human housekeeping gene								
Number of laboratories prepared to detect virus variants								

Sample 1: negative, diluted swab (human RNA positive). Sample 2: positive, diluted swab (human RNA positive), $C_t \sim 27.0$, $\sim 252,000$ copies/mL. Sample 3: positive, virus culture supernatant, 1:10 dilution of Sample 6 (human RNA negative), $C_t \sim 31.6$, $\sim 22,300$ copies/mL. Sample 4: positive, diluted swab, variant lineage B.1.1.7 (human RNA positive), $C_t \sim 26.3$, $\sim 345,300$ copies/mL. Sample 5: negative, sodium chloride solution (human RNA negative). Sample 6: positive, virus culture supernatant (human RNA negative), $C_t \sim 28.8$, $\sim 96,100$ copies/mL. Sample 7: positive; virus culture supernatant, 1:100 dilution of Sample 6 (human RNA negative); $C_t \sim 35.3$, ~ 640 copies/mL.

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 $200\ \mu\text{L}$ aliquot of the original sample using the NucliSENS easyMAG platform (bioMérieux, France). The elution volume was $50\ \mu\text{L}$. For RT-qPCR, each $25\ \mu\text{L}$ reaction mixture contained $12.5\ \mu\text{L}$ of $2\times$ reaction buffer, $0.4\ \mu\text{L}$ of a $50\ \text{mM}$ magnesium sulfate solution, $1\ \mu\text{L}$ of SuperScript III

RT/Platinum Taq Mix (SuperScript III One-Step RT-PCR System, Invitrogen, Germany), $0.5\ \mu\text{L}$ each of LightMix[®] Modular SARS-CoV (COVID-19) E-gene and LightMix[®] Modular EAV RNA Extraction Control (TIB Molbiol, Berlin, Germany) and $10\ \mu\text{L}$ 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 2: Reported results and test sample (S1–S7) specifications for the detection of SARS-CoV-2.

Sample	S1	S2	S3	S4	S5	S6	S7
Sample specifications							
Origin	Swab	Swab	Virus culture	Swab	NaCl	Virus culture	Virus culture
SARS-CoV-2	Negative	Positive	Positive	Positive	Negative	Positive	Positive
Human RNA	Positive	Positive	Positive ^a	Positive	Negative	Positive ^a	Positive ^a
C _t value (E-gene, mean; 95% C.I.) ^b	>45	27.2; 27.0–27.5	32.1; 31.8–32.4	26.7; 26.6–26.9	>45	29.6; 29.3–29.9	35.9; 35.1–36.7
Viral RNA copies/mL	–	~252,000	~22,300	~345,300	–	~96,100	~640
Reported results							
Total	134	134	134	134	134	134	134
Positive		132 (99%)	130 (97%)	132 (99%)		131 (98%)	115 (86%)
Negative	131 (98%)		1 (1%)		110 (83%)		9 (7%)
Invalid	3 (2%)	2 (1%)	3 (2%)	2 (1%)	24 (17%)	3 (2%)	10 (7%)
Suspected or proven mutation							
Mutation detected			1	20			1
N501Y, B.1.1.7				19			
B.1.351/P1			1	1			1
Reported C_t values^c							
<u>Virus-specific targets</u>							
E-gene		27.4; 25.4–33.0 1.4; 50	32.8; 30.5–40.0 1.7; 49	28.6; 26.1–32.0 1.2; 50		30.4; 28.5–35.0 1.6; 50	36.2; 33.8–39.1 1.3; 42
N-gene		28.3; 19.7–35.0 2.7; 72	33.6; 25.2–39.1 2.8; 70	29.3; 20.8–35.0 2.6; 72		31.4; 25.1–37.0 2.8; 72	37.2; 29.5–42.5 2.9; 56
S-gene		24.4; 20.6–30.0 3.1; 15	29.7; 25.8–35.7 3.5; 15	26.2; 23.5–30.7 2.5; 12		27.6; 24.3–32.9 2.9; 15	33.6; 29.8–38.8 3.3; 11
RdRP-gene		27.9; 19.3–32.2 2.7; 24	33.8; 29.8–38.4 2.4; 24	29.5; 25.2–34.9 2.2; 24		31.2; 26.6–34.8 1.8; 24	37.5; 32.8–40.2 1.8; 16
ORF1ab-gene		27.0; 22.7–37.1 3.0; 22	31.5; 27.3–34.8 2.3; 22	27.9; 24.1–31.2 2.0; 22		29.4; 24.9–33.2 2.3; 22	33.9; 25.1–37.7 3.0; 17
Other ^d		27.3; 24.1–32.1 1.7; 28	32.2; 25.4–36.9 2.1; 29	28.3; 25.1–33.9 1.9; 28		29.8; 25.4–36.6 2.1; 28	35.9; 25.4–41.4 3.2; 25
<u>Human RNA targets</u>							
Beta actin	24.6; 23.6–25.8 1.0; 6	31.3; 29.4–32.8 1.4; 6	37.0; 37.0–37.1 0.1; 2	29.5; 22.7–31.1 1.5; 6		33.4; 30.0–36.8 4.8; 2	36.0; 34.2–37.2 1.6; 3
Beta globin	31; 31.0–31.0 0.0; 2	38.0; 37.0–39.0 3.0; 22	37.0 3.0; 22	36.0; 35.0–37.0 2.0; 22		37.0 2.3; 22	
RNase P	27.0; 22.0–38.6 4.3; 15	32.8; 26.5–38.2 3.5; 12	35.3; 30.7–39.1 3.2; 7	31.1; 26.3–39.6 3.9; 14		35.2; 25.3–39.9 5.0; 6	37.7

^aCross-reactive binding of primers targeting human beta actin, beta globin, and RNase P to *Chlorocebus aethiops* present in the Vero cell culture supernatant. ^bTarget Ct values provided by the National Reference Laboratory. ^cReported Ct values are summarized as the mean; range; standard deviation; and total stratified by target and sample. ^d“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

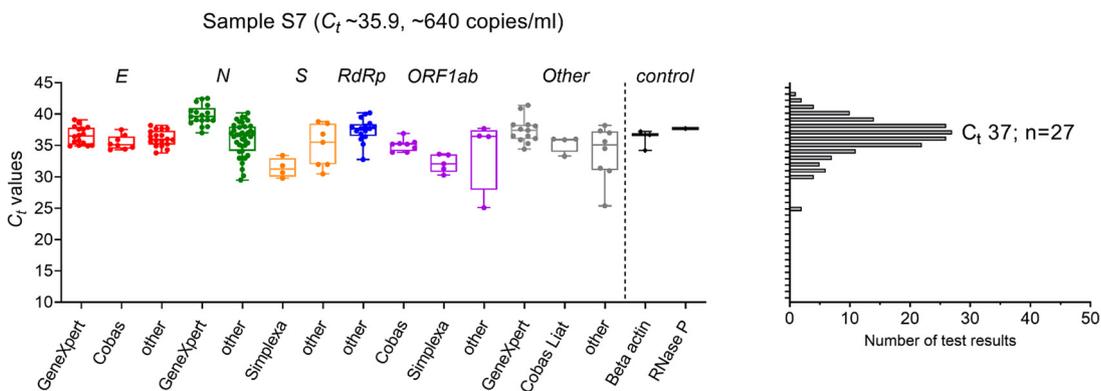
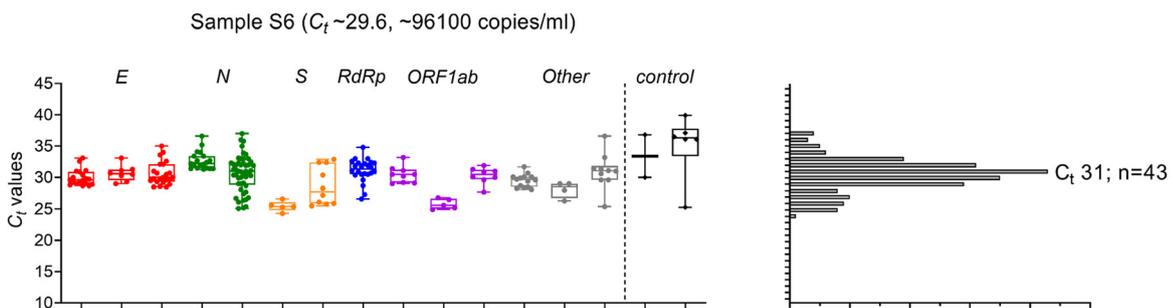
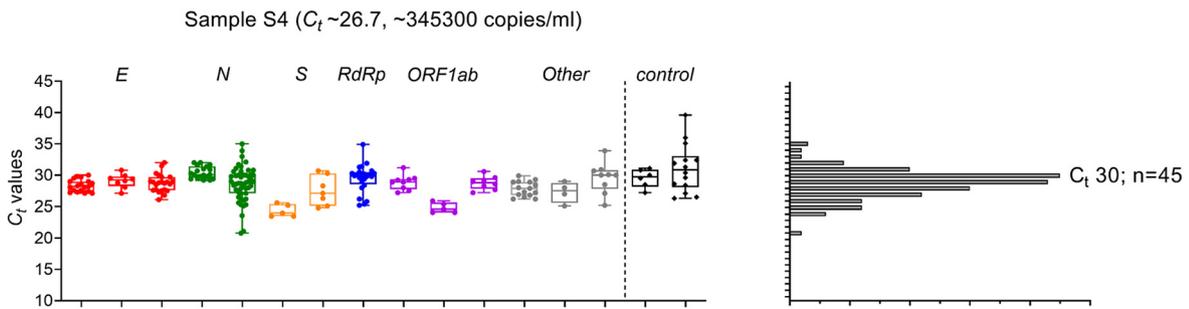
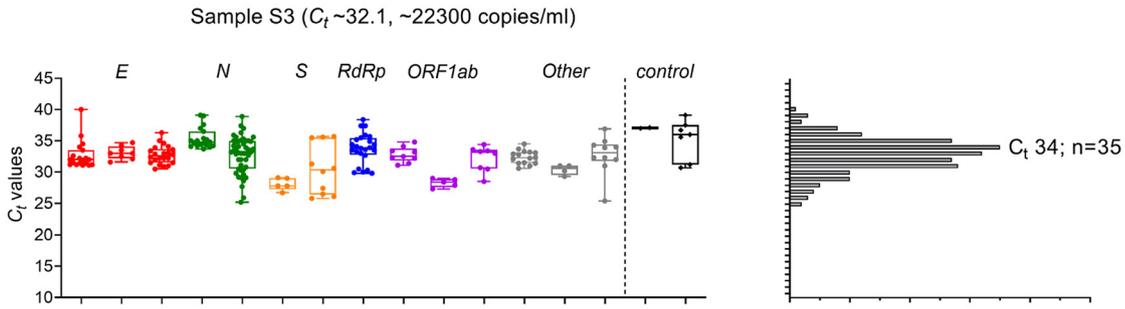
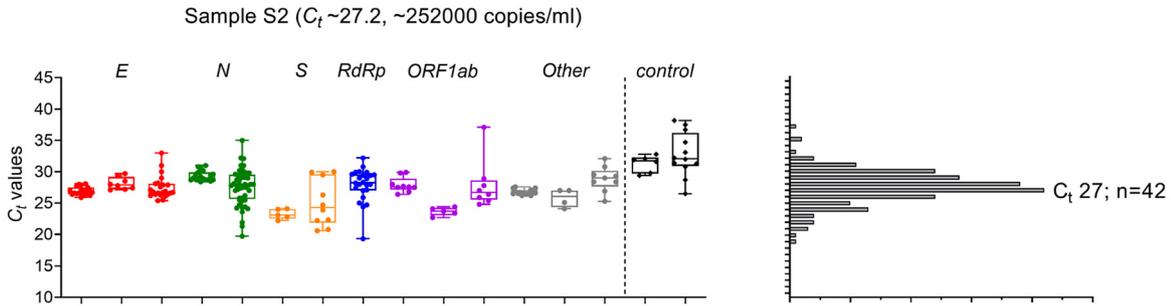
A total of 124 out of 134 test systems (93%) reported correct results for all five positive and both negative samples (Table 2). A total of 1,006 individual C_t values were reported from 122 test systems targeting the following viral gene targets: *E* (250 C_t values); *N* (337); *S* (68); *RdRP* (112); *ORF1ab* (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 ($S_2 = \sim 252,000/\text{mL}$, mean target $C_t = 27.2$, mode = 27; $S_3 = \sim 22,300/\text{mL}$, mean target $C_t = 32.1$, mode = 34; $S_4 = \sim 345,300/\text{mL}$, mean target $C_t = 26.7$, mode = 30; $S_6 = \sim 96,100/\text{mL}$, mean target $C_t = 29.6$, mode = 31; and $S_7 = \sim 640/\text{mL}$, 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 *ORF1ab*-gene; for Simplexa Liaison MDS Realtime PCR/Simplexa COVID-19 Direct Reaction Mix it was 0.4 for the *ORF1ab*- 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



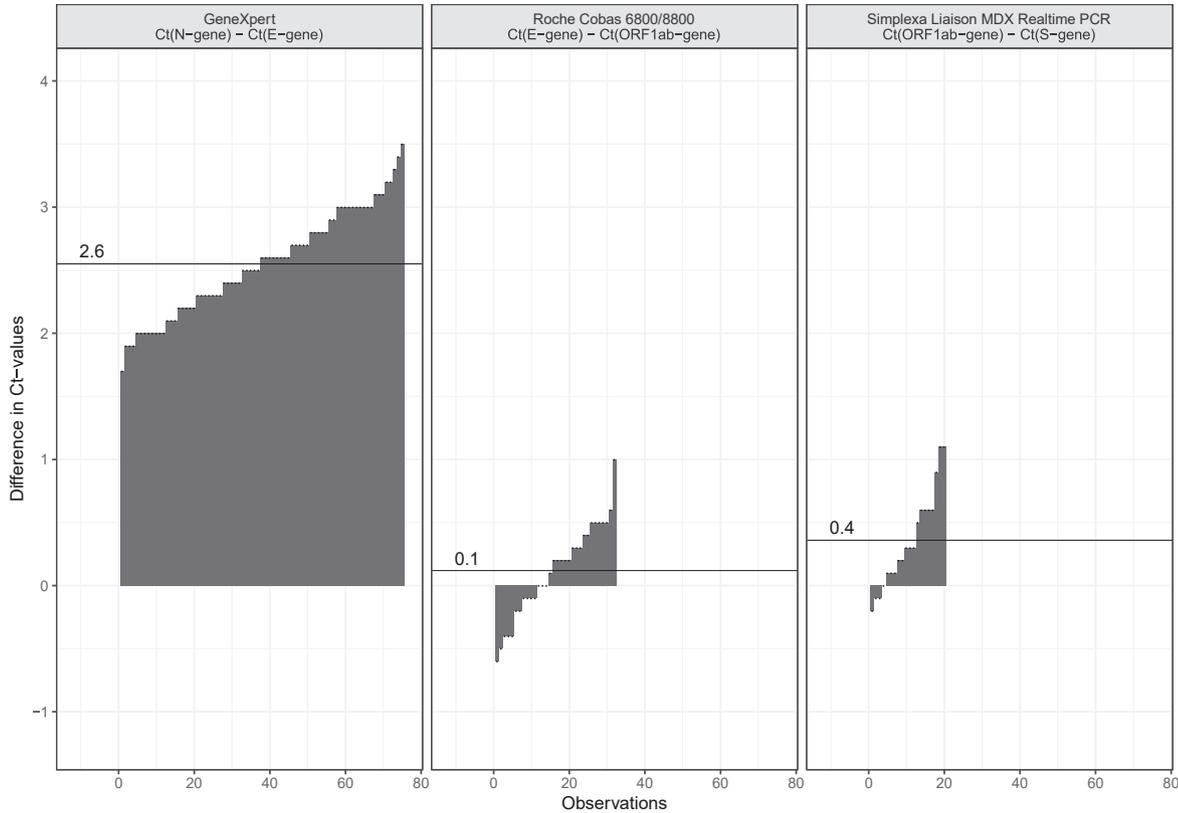


Figure 2: Differences of C_t values in test systems employing more than one gene target.

Barplots show the difference in C_t values for each reported pair, ordered from least to greatest difference, within the following test systems: GeneXpert, $C_t N$ -gene – $C_t E$ -gene; Roche Cobas, $C_t E$ -gene – $C_t ORF1ab$ -gene; Simplexa Liaison, $C_t ORF1ab$ -gene – $C_t S$ -gene. The horizontal lines display the mean difference, indicated in the text annotation.

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× 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].

Figure 1: SARS-CoV-2-specific and human RNA C_t values obtained for each sample and the indicated target genes.

Boxplots of five test samples (S2 and S4 are patient-derived swabs; S6, S3, and S7 are a 1:1, 1:10, 1:100 dilution series, respectively, of virus culture supernatant) in the left column show 50% inter-quartile range and median within the box for various test systems (x -axis) targeting specific viral gene targets (subheading in italics). The plot titles indicate the sample number, the reference C_t value, and the template copy number as determined by digital droplet PCR. “Other” = results obtained from two or more gene targets in one channel and therefore not shown separately. The histogram in right column displays the distribution of reported C_t values (y -axis is the same as in left column boxplots) for each target, indicating the mode and total sample number.

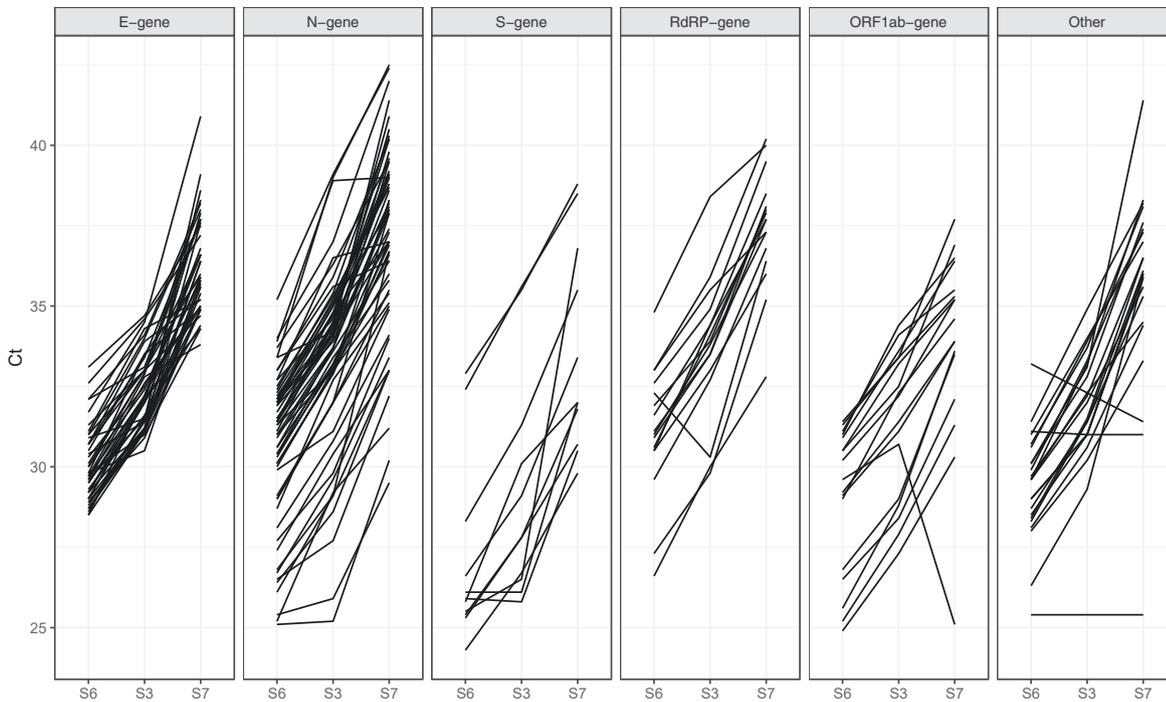


Figure 3: Linearity of C_t values in S6, S3, and S7 of the most frequently used gene targets for a dilution series derived from a virus isolate grown in cell culture.

Lines connect the C_t 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, S, RdRP, and ORF1b) or “other” for systems which calculate C_t 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 C_t 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 C_t 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 C_t 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 C_t 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 C_t values as a qualitative assessment of patient status.

Reporting of C_t 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 C_t values. First, we observed the well-documented property of RT-qPCR, that primer-probe sets specific for different genomic targets produce different C_t 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 C_t values reported for the same target, in the same sample, by different participants was on average approximately 10 C_t units. However, the variability of C_t 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 C_t 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 C_t according to template concentration, but some were notably non-linear (Figure 3).

Possible reasons for the high variability revealed by individual C_t 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 C_t 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 C_t 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. 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

1. The European Parliament and the Council of the European Union. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. *Off J Eur Union* 2017;L 117:176–331.
2. Belk WP, Sunderman FW. A survey of the accuracy of chemical analyses in clinical laboratories. *Am J Clin Pathol* 1947;17:853–61.
3. Libeer JC, Baadenhuijsen H, Fraser CG, Petersen PH, Ricós C, Stöckl D, et al. Characterization and classification of external quality assessment schemes (EQA) according to objectives such as evaluation of method and participant bias and standard deviation. External Quality Assessment (EQA) Working Group A on Analytical Goals in Laboratory Medicine. *Eur J Clin Chem Clin Biochem* 1996; 34:665–78.
4. Magleby R, Westblade LF, Trzebucki A, Simon MS, Rajan M, Park J, et al. Impact of SARS-CoV-2 viral load on risk of intubation and mortality among hospitalized patients with coronavirus disease 2019. *Clin Infect Dis* 2020;30:ciaa851.
5. Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Euro Surveill* 2020;25:2001483.

6. Matheussen V, Corman VM, Donoso Mantke O, McCulloch E, Lammens C, Goossens H, et al. International external quality assessment for SARS-CoV-2 molecular detection and survey on clinical laboratory preparedness during the COVID-19 pandemic, April/May 2020. *Euro Surveill* 2020;25:2001223.
7. Sung H, Han M, Yoo C, Lee S, Chung Y, Park J, et al. Nationwide external quality assessment of SARS-CoV-2 molecular testing, South Korea. *Emerg Infect Dis* 2020;26:2353–60.
8. Wang Z, Chen Y, Yang J, Han Y, Shi J, Zhan S, et al. External quality assessment for molecular detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in clinical laboratories. *J Mol Diagn* 2021;23:19–28.
9. Edson DC, Casey DL, Harmer SE, Downes FP. Identification of SARS-CoV-2 in a proficiency testing program. *Am J Clin Pathol* 2020;154:475–8.
10. Fischer C, Mögling R, Melidou A, Kühne A, Oliveira-Filho EF, Wolff T, et al. Variable sensitivity in molecular detection of SARS-CoV-2 in European expert laboratories: external quality assessment, June - July 2020. *J Clin Microbiol* 2021;59:e02676–20.
11. Carroll A, McNamara E. Comparison and correlation of commercial SARS-CoV-2 real-time PCR assays, Ireland, June 2020. *Euro Surveill* 2021;26:2002079.
12. Vanaerschot M, Mann SA, Webber JT, Kamm J, Bell SM, Bell J, et al. Identification of a polymorphism in the N gene of SARS-CoV-2 that adversely impacts detection by reverse transcription-PCR. *J Clin Microbiol* 2020;59:e02369–20.
13. Ziegler K, Steininger P, Ziegler R, Steinmann J, Korn K, Ensser A. SARS-CoV-2 samples may escape detection because of a single point mutation in the N gene. *Euro Surveill* 2020;25:2001650.
14. Kim S, Lee JH, Lee S, Shim S, Nguyen TT, Hwang J, et al. The progression of SARS coronavirus 2 (SARS-CoV2): mutation in the receptor binding domain of spike gene. *Immune Netw* 2020;20:e41.
15. Bohn MK, Mancini N, Loh TP, Wang CB, Grimmler M, Gramegna M, et al. IFCC interim guidelines on molecular testing of SARS-CoV-2 infection. *Clin Chem Lab Med* 2020;58:1993–2000.
16. Basso D, Aita A, Navaglia F, Franchin E, Fioretto P, Moz S, et al. SARS-CoV-2 RNA identification in nasopharyngeal swabs: issues in pre-analytics. *Clin Chem Lab Med* 2020;58:1579–86.
17. Stavelin A, Albe X, Meijer P, Sarkany E, MacKenzie F. An overview of the European Organization for external quality assurance providers in laboratory medicine (EQALM). *Biochem Med* 2017;27:30–6.
18. Virological, discussion forum for analysis and interpretation of virus molecular evolution and epidemiology. Available from: <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> [Accessed 22 Feb 2021].
19. Popa A, Genger JW, Nicholson MD, Penz T, Schmid D, Aberle SW, et al. Genomic epidemiology of superspreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. *Sci Transl Med* 2020;12:eabe2555.
20. Buchta C, Görzer I, Chiba P, Camp JV, Holzmann H, Puchhammer-Stöckl E, et al. Variability of cycle threshold values in an external quality assessment scheme for detection of the SARS-CoV-2 virus genome by RT-PCR. *Clin Chem Lab Med* 2020;59:987–94.
21. International Organization for Standardization. Statistical methods for use in proficiency testing by interlaboratory comparison (ISO 13528:2015). Geneva: International Organization for Standardization (ISO); 2015.
22. Görzer I, Buchta C, Chiba P, Benka B, Camp JV, Holzmann H, et al. First results of a national external quality assessment scheme for the detection of SARS-CoV-2 genome sequences. *J Clin Virol* 2020;129:104537.
23. Austrian Federal Ministry for Social Affairs, Healthcare, Care and Consumer Protection. Available from: https://www.sozialministerium.at/dam/jcr:531f7e21-0f53-4180-b214-848e19668b52/Empfehlung%20zur%20Entlassung%20von%20COVID-19-Fällen%20aus%20der%20Absonderung_final_230720.pdf [Accessed 22 Feb 2021].
24. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. *Clin Biochem Rev* 2007;28:139–47.
25. Miller WG, Budd J, Greenberg N, Weykamp C, Althaus H, Schimmel H, et al. IFCC Working Group Recommendations for correction of bias caused by noncommutability of a certified reference material used in the calibration hierarchy of an end-user measurement procedure. *Clin Chem* 2020;66:769–78.

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