International Journal of Medical Microbiology 319 (2025) 151656

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

From crisis to routine – Standardization of SARS-CoV-2 genome detection by enhanced EQA schemes in a scientific pandemic network☆

Martin Kammel[a,b,*], Hans-Peter Grunert[b], Anika Zimmermann[b], Annemarie Martin[b], Vanessa Lindig[b], Samuel Samuleit[b], Ulf Dühring[b], Mechthild Adams-Bagusche[b], Dirk Sander[b], Hannah Zeichhardt[b], Christian Drosten[c], Victor M. Corman[c], Sandra Ciesek[d], Holger F. Rabenau[d], Martin Obermeier[e], Robert Ehret[e], Rolf Kaiser[f], Jim Huggett[g,h], Denise O'Sullivan[g], Peter M. Vallone[i], Megan H. Cleveland[i], Samreen Falak[j], Andreas Kummrow[j], Esmeralda Valiente[j], Rainer Macdonald[j], Mojca Milavec[k], Sabine Goseberg[a], Silke Kappler[a], Nathalie Weiss[a], Laura Vierbaum[a], Patricia Kaiser[a], Ingo Schellenberg[a,l], Heinz Zeichhardt[a,m,**]

[a] INSTAND e.V., Society for Promoting Quality Assurance in Medical Laboratories, Düsseldorf, North Rhine-Westphalia, Germany
[b] GBD Gesellschaft für Biotechnologische Diagnostik mbH, Berlin, Germany
[c] Institute of Virology, Charité - University Medicine Berlin, National Consultant Laboratory for Coronaviruses, German Centre for Infection Research, Berlin, Germany
[d] Institute for Medical Virology, University Hospital, Goethe University Frankfurt, Frankfurt, Hesse, Germany
[e] Medizinisches Infektiologiezentrum Berlin, Germany
[f] Institute of Virology, University of Cologne, Cologne, Germany
[g] National Measurement Laboratory, LGC, Teddington, Middlesex, United Kingdom
[h] Faculty of Health & Medical Science, School of Biosciences & Medicine, University of Surrey, Guildford, United Kingdom
[i] Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, United States
[j] Physikalisch-Technische Bundesanstalt, Berlin, Germany
[k] Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia
[l] Institute of Bioanalytical Sciences, Center of Life Sciences, Anhalt University of Applied Sciences, Bernburg, Saxony-Anhalt, Germany
[m] IQVD GmbH, Institut für Qualitätssicherung in der Virusdiagnostik, Berlin, Germany

ARTICLE INFO

Keywords:
COVID-19 pandemic
SARS-CoV-2
virus genome detection tests
reference materials
external quality assessment
laboratory medicine

ABSTRACT

In the beginning of 2020, the outbreak of the COVID-19 pandemic led to a crisis in which diagnostic methods for the genome detection of SARS-CoV-2 were urgently needed. Based on the very early publication of the basic principles for a diagnostic test for the genome detection of SARS-CoV-2, the first noncommercial laboratory-developed tests (LDTs) and commercial tests were introduced. As there was considerable uncertainty about the reliability and performance of different tests and different laboratories, INSTAND established external quality assessment (EQA) schemes for the detection of SARS-CoV-2 starting in April 2020. In close partnership in a scientific network, the EQA schemes were enhanced, especially the April, June and November 2020 terms. The enhancement included: (i) immediate provision of suitable virus including variants of concern at the beginning of the pandemic outbreak, (ii) short frequency of EQA schemes, (iii) concentration dependency of the testing and

Abbreviations: dPCR, digital polymerase chain reaction; DVV, German Association for the Control of Virus Diseases / Deutsche Vereinigung zur Bekämpfung der Viruskrankheiten e.V.; EQA, External quality assessment; GfV, Society of Virology / Gesellschaft für Virologie e.V.; LDT, Laboratory developed test system; MERS CoV, Middle East Respiratory Syndrome Coronavirus; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus 2; NAT, Nucleic acid testing; NMI, National metrology institute; PFU, Plaque forming unit; RM, Reference material; RT-dPCR, Reverse transcription dPCR; VOC, Variant of concern; WHO, World Health Organization.

☆ This article is dedicated to Prof. Dr. Martin Mielke, Head of Department 1 – Infectious Diseases at Robert Koch-Institut, Berlin, on the occasion of his retirement with sincere thanks for the many years of trustful cooperation.

* Correspondence to: GBD Gesellschaft für Biotechnologische Diagnostik mbH, Potsdamer Chaussee 80, Berlin 14129, Germany.
** Correspondence to: Institut für Qualitätssicherung in der Virusdiagnostik - IQVD GmbH, Potsdamer Chaussee 80, Berlin 14129, Germany.
E-mail addresses: Martin.Kammel@gbd-biotechdia.de (M. Kammel), Heinz.Zeichhardt@iqvd.de (H. Zeichhardt).

https://doi.org/10.1016/j.ijmm.2025.151656
Received 11 December 2024; Received in revised form 12 May 2025; Accepted 21 May 2025
Available online 24 May 2025

sensitivity check, achieved by using SARS-CoV-2-positive samples from a 10-fold dilution series of the same starting material, (iv) specificity check of the testing, achieved by using SARS-CoV-2-negative samples containing human coronaviruses or MERS CoV, (v) revealed samples for orientation on test performance during an ongoing or at the start of an EQA scheme using a pre-quantified SARS-CoV-2-positive EQA sample with a low viral RNA load of only 1 570 copies/mL assigned by digital PCR (dPCR) in June 2020 and (vi) quantified reference materials based on the experiences of the first two EQA schemes with dPCR-assigned values in copies/mL beginning in November 2020 for self-evaluation of the applied test system. This manuscript summarizes the results of a total of 13 EQA schemes for the detection of SARS-CoV-2 between April 2020 and June 2023 in which a total of 1 413 laboratories from 49 countries participated. The qualitative results for the detection of SARS-CoV-2-positive samples were between 95.8 % and 99.7 % correct positive, excluding extremely low concentration samples. For all SARS-CoV-2-negative EQA samples, the qualitative success rates ranged from 95.1 % to 99.4 % correct negative results. The widely varying values for the cycle threshold (Ct)/crossing point (Cq) reported for the different target genes and test systems were striking. A few laboratories reported quantitative results in copies/mL for several VOCs with an acceptable rate of over 93 % correct positive results in the majority of cases. The description of the enhanced EQA schemes for SARS-CoV-2 detection in terms of timing and scope can serve as a blueprint for the rapid development of a quality assessment of diagnostics for an emerging pathogen.

1. Introduction

After the first cases of a novel severe respiratory infectious disease emerged from Wuhan, China in December 2019, SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) was determined to be the causative agent of COVID-19. SARS-CoV-2 belongs to the beta coronaviruses, which also include SARS-CoV-1, Middle East Respiratory Syndrome (MERS) coronavirus and other human coronaviruses (Huang et al., 2020; Coronaviridae study group of the International Committee on Taxonomy of Viruses, 2020; Bhadoria et al., 2021).

SARS-CoV-2, which is mainly transmitted by droplets and aerosols, spread rapidly worldwide, leading WHO to declare the SARS-CoV-2 outbreak a pandemic on March 11, 2020 (Cucinotta and Vanelli, 2020). As of March 4, 2024, for COVID-19 the following cumulative numbers of confirmed cases and confirmed deaths were registered with the WHO since start of the pandemic: (i) worldwide: 774 699 366 confirmed cases; 7 033 430 confirmed deaths and (ii) out of these for Germany: about 38 400 000 confirmed cases and about 175 000 confirmed deaths (World Health Organization, 2024). Although it can be assumed that due to inadequacies in the testing and reporting systems of individual countries, the actual numbers are higher.

National health systems around the world faced the dilemma of building the infrastructure for preventive and clinical interventions in an extremely short period of time. In Germany, a National Pandemic Plan existed to prepare for an influenza pandemic, but the requirements had to be adapted to COVID-19 (Robert Koch-Institut, 2020a).

Since SARS-CoV-2 was an emerging virus, molecular diagnostics for genome detection of SARS-CoV-2 had to be launched immediately. This was indispensable for confirming positive patients and the detection of chains of infection.

While most diagnostic laboratories had established PCR systems for the detection of other respiratory pathogens, there was a lack of specific primers and probes for both commercial tests and laboratory-developed tests (LDTs) as well as the specific test conditions to detect the new virus at the beginning of the pandemic. By mid-January 2020 the first diagnostic tests for SARS-CoV-2 detection by PCR had already been established (Corman et al., 2020a). The first description of this test was made worldwide freely available by an international author consortium on the WHO homepage even before scientific publication (Corman et al., 2020b).

Once the capabilities for diagnostic detection of SARS-CoV-2 were established, it was necessary to introduce quality control measures to understand (i) the quality of measurements in the laboratories and (ii) the reliability of the commercial tests and noncommercial LDTs applied in the laboratory.

Shortly after the beginning of the pandemic, the first external quality assessment (EQA) schemes or proficiency testing programs for quality assurance of diagnostic methods for genome detection of SARS-CoV-2 were established by different organizations in different countries around the world, amongst others by American Proficiency Institute (API) (Edson et al., 2020); Austrian Association for Quality Assurance and Standardization of Medical and Diagnostic Tests (ÖQUASTA) (Goerzer et al., 2020, Buchta et al., 2021a); INSTAND - Society for Promoting Quality Assurance in Medical Laboratories (Zeichhardt and Kammel, 2020a); Korean Association of External Quality Assessment Service (KEQAS) (Sung et al., 2020); Quality Control for Molecular Diagnostics (QCMD) (Matheeussen et al., 2020); National Center for Clinical Laboratories in China (Wang et al., 2021); Department of Health Research and Indian Council of Medical Research (DHR-ICMR) (Kaur et al., 2022); Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP) (Lau et al., 2022).

In April 2020, INSTAND conducted one of the first global EQA schemes for the detection of the SARS-CoV-2 viral genome that included 463 participating laboratories from 36 countries. INSTAND is accredited according to DIN EN ISO/IEC 17043:2010 and serves as a reference institution of the German Medical Association and as a scientific society. INSTAND has run approximately 80 virological EQA schemes for many years in coordination with the Joint Diagnostic Commission of the German Association for the Control of Virus Diseases (DVV) and the Society of Virology (GfV). All virological EQA schemes are performed according to the Guideline of the German Medical Association on Quality Assurance of Laboratory Medical Examinations in Medical Laboratory Testing (Bundesärztekammer, 2023; Ahmad-Nejad et al., 2024) in an international scientific network.

This manuscript describes how, at the beginning of the COVID-19 pandemic, the quality control network quickly responded to initial challenges and adapted and improved. Beyond the typical approach in EQA programs, where the results are published after the completion of an EQA round, the EQA programs of April and June 2020 provided unblinded results for a subset of samples while the EQA programs were still running. This enabled participating laboratories to gain an immediate understanding of their capabilities and the performance of their test systems. Additionally, the frequency of the EQA schemes was significantly increased. In multiple schemes, a 10-fold dilution series of cultured inactivated SARS-CoV-2 was repeated, in parallel with individual samples revealed in the first EQA schemes. The combination of the "enhanced EQA schemes", which included independent reference materials, with quantitative values of the SARS-CoV-2 RNA load assigned using reverse transcription digital PCR (RT-dPCR), was extremely important, since the WHO International Standard and other well-characterized reference materials were not yet available. This enabled laboratories to monitor their tests for reliable routine diagnostics and, under certain conditions, even assisted clinical decisions for patient management. This evolving model of enhanced EQA schemes can serve as a model for future pandemics.

The results of a total of 13 EQA schemes for the detection of SARS-

CoV-2 are presented; these schemes were conducted between April 2020 and June 2023 with a total of 1 413 laboratories from 49 participating countries.

2. Materials and methods

2.1. Viruses and cells

For each of the EQA schemes, the currently circulating virus strains and variants of SARS-CoV-2, based on the epidemiological situation in Germany, were used for the preparation of the respective EQA samples (see Supplementary Table 1).

For the INSTAND EQA schemes, all SARS-CoV-2 virus strains and variants as well as the other coronaviruses were propagated in cell lines by the Coronavirus Consultant Laboratory at Charité - University Medicine Berlin, Germany, in a BSL 3 laboratory.

Cell monolayers of the following cell lines were used for growing of the different viruses at 37 °C using Dulbecco's Modified Eagle's Medium (D-MEM) supplemented with 10 % fetal bovine serum (FBS) at 5 % CO_2 for a different number of days depending on the virus isolate: Vero E6 cells (ATCC CRL-1586) for SARS-CoV-2 beta coronaviruses and variants (Vierbaum et al., 2022, 2023), Vero B4 cells (DZMZ ACC-33) for MERS-CoV beta coronavirus (Corman et al., 2021), Huh-7 cells (CVCL_0336) for the propagation of beta coronavirus hCoV OC43 (Corman et al., 2021) and alpha coronavirus hCoV 229E (Corman et al., 2021) and LLC-MK2 cells (ATCC CCL-7) for the propagation of alpha coronavirus hCoV NL63 (Vierbaum et al., 2023). See also Supplementary Table 1.

Inactivation of the propagated betacoronaviruses (SARS-CoV-2 and MERS-CoV) was performed either by heat inactivation (4 h, 60 °C) (Vierbaum et al., 2022) or chemically by 0.05 % beta-propiolactone (BPL) for 14 h at 4 °C followed by consecutive hydrolysis of BPL for 2 h at 37 °C (Vierbaum et al., 2023) (see Supplementary Table 1). This procedure reduced the plaque-forming units (PFU) of the virus-containing cell culture supernatants to 0 PFU. The results of inactivation were checked in double-blind passages on the respective cell monolayers before release of the virus suspensions for preparation of the EQA samples. The virus suspensions containing alpha coronaviruses were not inactivated.

Virus-negative cell culture lysates of MRC-5 cells and SARS-CoV-2-negative suspensions of different human coronaviruses, as well as MERS coronavirus, were used to test the specificity of the assay performance in the laboratories that participated in the EQA schemes.

The reference materials RM 1 and RM 2 were prepared from cell culture supernatants of Vero E6 cells (ATCC CRL-1586) that had been infected at a multiplicity of infection of 0.05 PFU/cell with SARS-CoV-2 (strain: BetaCoV/Munich/ChVir984/2020, GSAID: EPI_ISL_406862) and were subsequently heat inactivated (Vierbaum et al., 2022).

2.2. Preparation of EQA samples

The virus suspensions from cell culture supernatants and virus-negative cell culture lysates were diluted using a cell culture medium (Minimal Essential Medium, PAN-Biotech, Aidenbach, Germany), supplemented with non-essential amino acids (PanBioTech), HEPES buffer (PAN-Biotech) and fetal bovine serum (PAN-Biotech, gamma irradiated; 10–15 % v/v). Finally, 1.1 mL of the materials were aliquoted in screw cap micro tubes (2.0 mL; Sarstedt, Nümbrecht, Germany). Lyophilization was performed as described in Vierbaum et al., 2022. The sample vials were stored at < -20 °C until they were shipped to the laboratories at ambient temperature. Randomly selected vials of the EQA sample preparations were analyzed for the amount of virus, for homogeneity according to DIN EN ISO/IEC 17043:2010, and for stability during the period of the EQA survey in order to confirm the suitability of the control materials (Vierbaum et al., 2022, 2023).

2.3. Pre-study for EQA samples

Before distributing the EQA samples for any of the EQA schemes, randomly selected EQA sample sets were tested by up to ten INSTAND Expert Laboratories (see Acknowledgement), applying different genome detection methods to confirm the expected sample properties.

2.4. Determination of SARS-CoV-2 RNA loads by RT-dPCR

Reverse transcription digital PCR (RT-dPCR) assays were developed and applied to selected samples to confirm the expected values of SARS-CoV-2 RNA load. The methodological details for the characterization of EQA samples by RT-dPCR are described in detail for genome detection of SARS-CoV-2 (Vierbaum et al., 2022) as well as cytomegalovirus and HIV-1 (Milavec et al., 2022; Falak et al., 2022). The principles and application of dPCR in diagnostics are summarized by Huggett et al. (2015).

2.5. Performance of EQA schemes

The first SARS-CoV-2-positive samples were integrated into the pre-existing INSTAND EQA scheme for differentiation of human coronaviruses and MERS-CoV (program 340) in April 2020. This scheme was established in June 2015, and was set up in close cooperation with the National Consultant Laboratory for Coronaviruses at Charité – University Medicine Berlin, Institute of Virology, and the members of the ad hoc commission of DVV and GfV, University Hospital Frankfurt, Medizinisches Infektiologiezentrum Berlin (MIB), University Clinics of Cologne. Starting in March 2021, a separate EQA program for SARS-CoV-2 genome detection only was established (Program 409) to address the increasing number of emerging variants in the sample set.

From April 2020 until June 2023 INSTAND conducted 13 EQA schemes for SARS-CoV-2 virus genome detection. Participating laboratories received five to nine samples per EQA, depending on circulating SARS-CoV-2 variants at the time. Participants were instructed to reconstitute samples with 1.1 mL double-distilled water (sterile, pyrogen-free, PCR-grade) for 20 min at room temperature. Laboratories were instructed to apply their routine test method for genome detection of coronaviruses and were asked to report results INSTAND via the RV-Online platform (https://rv-online.instandev.de).

In the EQA schemes, all genome detection results for SAR-CoV-2 were recorded in the 'PCR/NAT' category, including isothermal methods or transcription-mediated amplification (TMA). Multiple results for each sample, obtained by different test methods, could be entered. Furthermore, the laboratories were asked to provide detailed information on the test method(s) used for each analysis, including the test kit manufacturer(s) and test kit(s) (Vierbaum et al., 2022).

2.6. Data evaluation

In total, 82 702 qualitative and 2 719 quantitative results from the 13 EQA schemes between April 2020 and June 2023, reported by 1 413 laboratories from 49 countries were included in the evaluation. (The corresponding EQA scheme reports are available upon request at instand@instand-ev.de). The evaluation was carried out in a test- and sample-dependent way according to the following definitions as described (Zeichhardt and Kammel, 2020a):

Correct result: A correct result is a result reported by a laboratory for a specific sample that is consistent with the qualitative target value or the quantitative target value interval (acceptance range).

Qualitative target value: The qualitative target value for a given sample is derived from the *consensus value* of all qualitative results reported by the participants, including the results reported by the INSTAND Expert Laboratories during the EQA scheme.

Quantitative target value: The quantitative target value for a given sample is derived from the *consensus value* of all quantitative results

reported by the participants. The consensus value is calculated as the *robust average* according to algorithm A/DIN ISO 13528/Annex C (DIN ISO 13528, 2020).

Quantitative target value interval: The target value interval for each SARS-CoV-2-positive sample is based on an interval of $-1.0\ \log_{10}$ to $+1.0\ \log_{10}$ of the target value (determined by the EQA scheme adviser based on all evaluated methods) according to the Guideline of the German Medical Association on Quality Assurance of Laboratory Medical Examinations in Medical Laboratory Testing (Bundesärztekammer, 2023; Ahmad-Nejad et al., 2024). When evaluating quantitative data, it is ensured that the coefficient of variation determined from all submitted results is substantially smaller than the specified target value interval.

Success rate: A success rate is depicted for each of the samples reflecting the portion of "correct" results (expressed as "percent" correct results and/or as "number of correct results per number of total results reported").

Data were analyzed using the INSTAND standard evaluation software (RV-ISI, Düsseldorf, Germany), Excel (Microsoft, Redmond, WA, USA). Basic statistical analyses were performed using JMP 16 from SAS Institute (Cary, North Carolina, USA).

3. Results

3.1. April 2020 - first EQA scheme with an interim report with three samples revealed before the submission deadline

The test samples for the first EQA scheme for genome detection of SARS-CoV-2, which was designated as an extra EQA scheme, were sent by INSTAND on 03 April 2020 to 488 laboratories from 36 countries. A total of 463 laboratories reported their test results.

This EQA sample set included a total of seven different samples with two vials per sample. Four samples (no. 340059, 304061, 340063, and 340064) were positive for inactivated SARS-CoV-2 (Table 1). These positive samples were part of a 10-fold dilution series from a supernatant of cells infected with SARS-CoV-2 to identify the proficiency of laboratories in detecting defined virus RNA load gradations.

In addition, a total of three samples that were negative for SARS-CoV-2 were used to provide information on specificity for SARS-CoV-2

Table 1
First INSTAND EQA Scheme (340) for Virus Genome Detection of SARS-CoV-2 in April 2020 – qualitative results.

A	B	C	D	D-1	E
Sample no.	**SARS-CoV-2 status**	**Dilutions of SARS-CoV-2-positive samples / negative samples**	**Success rates of correct qualitative results (%)** (for all target gene regions and all test systems) [number of correct results per total number of reported results]	**Success rates of correct qualitative results (%)** **reduced by the number of incorrect result assignments (mix-ups) for samples no. 340064 and no. 340065** (for all target gene regions and all test systems) [number of correct results per total number of reported results]	***"Overall Ct/Cq medians"*** (for all target gene regions and all test systems) according to category 1 for Ct/Cq value consideration
340059 revealed in the interim evaluation	positive	1: 1 000	**99.7 %** [980/983]	—	**22.8**
340063	positive	1: 10 000	**98.8 %** [971/983]	—	**26.0**
340064 revealed in the interim evaluation	positive	1: 100 000	93.2 % [916/983]	**98.9 %** [915/925]	**29.5**
340061	positive	1: 1 000 000	**93.0 %** [914/983]	—	**32.4**
340060 revealed in the interim evaluation	negative	HCoV OC43 1: 2 500 specificity control	**97.8 %** [961/983]	—	**n.a.**
340065	negative	HCoV 229E 1: 2 500 specificity control	92.4 % [908/983]	**98.1 %** [907/925]	**n.a**
340062	negative	non-infected MRC−5 cells specificity control	**98.6 %** [969/983]	—	**n.a.**

Table 1: For each of the seven EQA samples from the first EQA scheme in April 2020, the sample numbers (**column A**), the status for SARS-CoV-2 (positive or negative) (**column B**), the dilutions of the SARS-CoV-2-positive and negative samples (**column C**), the success rates of correct qualitative results (number of correct qualitative results per reported results) (**column D**) as well as the *"overall Ct/Cq medians"* (generated by including the results for all target gene regions and all test systems; according to category 1 for Ct/Cq value consideration; see Section 3.1.2) (**column E**) are shown. The *"Overall Ct/Cq medians"* were considered only for the SARS-CoV-2 RNA-positive samples. The positive samples no. 340059, 340061, 340063 and 340064 represent consecutive dilution steps of a dilution series of the same supernatant of cells which were infected with SARS-CoV-2 virus. See also Report for EQA scheme April 2020 virus genome detection of SARS-CoV-2 (Zeichhardt and Kammel, 2020a).

Column D-1: For sample no. 340064 (SARS-CoV-2-positive, diluted 1: 100 000), the reduced success rate of only 93.2 % was essentially due to incorrect assignments of results ("mix-ups") for sample no. 340064 and sample no. 340065 (negative for SARS-CoV-2 and positive for HCoV 229E). The "mix-ups" for samples no. 340064 and 340065 concerned 24 laboratories with a total of 58 results per sample. For sample no. 340065 (SARS-CoV-2-negative, HCoV 229E RNA-positive, diluted 1: 2 500), the reduced success rate of only 92.4 % was essentially due to incorrect assignments of results ("mix-ups") for samples no.340065 and 340064 (positive for SARS-CoV-2). The "mix-ups" for samples no. 340065 and 340064 concerned 24 laboratories with a total of 58 results per sample.
n.a. = not applicable

testing (Table 1). The specificity controls were either positive for heterologous human coronaviruses, such as the beta coronavirus OC43 (sample no. 340060) or the alphacoronavirus 229E (sample no. 340065) or contained cell lysates from uninfected control cells (sample no. 340062).

This extra EQA program was unique in that most of the participating laboratories had no prior knowledge of the reliability of the commercial test systems or LDTs used in terms of sensitivity and specificity immediately after the pandemic outbreak.

During the performance of this EQA scheme, i.e. before the extended submission deadline for the results on April 28, 2020, INSTAND received urgent requests from participating laboratories to reveal the characteristics of at least some of the samples to be tested. The laboratories asked for help to improve their test procedure on short notice in case of possible incorrect measurements. In agreement with the Joint Diagnostic Committee of DVV and GfV, INSTAND therefore decided to reveal the sample properties of three of the seven samples during the ongoing EQA scheme.

Overall, the characteristics were revealed for the following sample set members (Table 1): two samples positive for SARS-CoV-2, sample no. 340059 (diluted 1: 1 000) and sample no. 340064 (diluted 1: 100 000) and one of the three control samples (sample no. 340060, negative for SARS-CoV-2 and positive for HCoV OC43). An interim evaluation report based on the results of 112 laboratories as of April 14, 2020, had been published at INSTAND, Düsseldorf (Zeichhardt and Kammel, 2020b).

3.1.1. *Qualitative results*

The interim evaluation of results yielded very satisfactory results for the qualitative results for the two SARS-CoV-2-positive samples: Regardless of the gene region tested, the SARS-CoV-2 genomic detection tests yielded 100 % correct results in each gene region for both positive samples.

The qualitative results of the interim evaluation for the specificity control no. 340060 (positive for HCoV OC43) gave correct negative results between 90.5 % and 100 % for the individual gene regions tested (Zeichhardt and Kammel, 2020b). It was not possible to assess whether the positive results were due to a specificity problem of the applied tests or to carryover of SARS-CoV-2 from one of the positive samples during sample handling in the laboratories.

The final report included the results for all seven samples of the sample set, including the results of the three samples "revealed" in the interim report (Zeichhardt and Kammel, 2020a). However, only the four samples that had not yet been uncovered were taken into account when issuing a certificate of successful participation.

Considering the final qualitative results, the analyses for the four SARS-CoV-2-positive EQA samples from the 10-fold dilution series showed correct positive results between 99.7 % for sample no. 340059 (diluted 1: 1 000) and 93.0 % for sample no. 340061 (diluted 1: 1 000 000), regardless of the gene region investigated (Table 1, columns D and D-1).

It should be noted that for the SARS-CoV-2-positive sample no. 340064 two different success rates of correct positive results (%) are stated (Table 1). The lower success rate of only 93.2 % (column D) was due to incorrect assignments of results ("mix-ups") for sample no. 340064 (positive for SARS-CoV-2) and sample no. 340065 (negative for SARS-CoV-2 and positive for HCoV 229E) (Table 1). The "mix-ups" of samples no. 340064 and 340065 concerned a group of 24 laboratories with a total of 58 results per sample.

If the mix-up results were excluded from the results for the SARS-CoV-2-positive sample no. 340064, the corrected success rate for this sample was 98.9 % correct positive results (915/925) (column D-1).

This correction also changed the success rate for correct negative results for sample no. 340065 (Table 1, column D-1).

When differentiating the success rates for the qualitative genome detection for the four positive samples described above from the 10-fold dilution series according to the target gene regions tested, some differences for the rates of correct positive results were apparent for the different gene regions targeted by the different test systems (Fig. 1a). Overall, the success rates for correct positive results were between 85.4 % and 100 % for the specified target gene regions. Even with the highest diluted sample no. 340061 (robust average 25 978 copies/mL; see Supplementary Table 1), the success rates for the gene regions E, N, ORF1a and S were over 90 % correct positive results (Fig. 1a).

At the time of the first EQA scheme, most laboratories were already using commercial test systems. Of the 3 927 results for the four SARS-CoV-2-positive samples (related to all test systems and gene regions examined), 14 % (549/3 927 results) were performed using LDTs.

As Fig. 2a shows, these LDTs, which were produced out of necessity in the laboratories themselves at the beginning of the pandemic, yielded almost as good success rates for the qualitative results as the commercial tests when testing the four positive samples from the 10-fold dilution series (virus RNA loads between 17 071 604 copies/mL and 25 978 copies/mL (robust average values from all reported quantitative results for each of the samples); see Supplementary Table 1).

As shown in Table 1, the specificity of the qualitative genome tests in the laboratories was checked both with samples containing non-infected cell lysates and with samples containing supernatants of cells infected with human coronaviruses. Taking into account the previously described observation that a group of 24 laboratories had mixed-up the results for samples no. 340064 and 364065 (see legend to Table 1), the genome detection tests for SARS-CoV-2 for the three SARS-CoV-2-negative samples no. 340060 (HCoV OC43), no. 340065 (HCoV 229E), and no. 340062 (uninfected MRC-5 cells) gave predominantly correct negative results (97.8–98.6 % correct qualitative results) regardless of the gene region tested.

3.1.2. *Results in Ct/Cq values*

In the following considerations, the reported Ct/Cq values for a defined sample are considered at three different levels as:

Ct/Cq median - category 1: *"Overall Ct/Cq median"* - For a specific sample, an *overall Ct/Cq median* is generated by including the results for all target gene regions and all test systems;

Ct/Cq median - category 2: *"Target gene-specific Ct/Cq median"* - For a specific target gene region of a specific sample, a Ct/Cq median is generated from the reported results of all applied test systems;

Ct/Cq median - category 3: *"Assay-specific Ct/Cq median"* - For a defined test system, a Ct/Cq median is generated from the reported results for a specific target gene region of a specific sample.

When considering differences between distinct Ct/Cq values, it is important to remember that, in theory, a polymerase chain reaction (assuming 100 % reaction efficiency) shows a 3.32 amplification cycle difference of the Ct/Cq values for a 10-fold difference in concentration of the nucleic acid in the starting material. It can be concluded that a difference of 9.96 amplification cycles in Ct/Cq values for two individual results of the same test in the same run corresponds to a theoretical concentration difference of three orders of magnitude or a factor of 1000.

For each of the four SARS-CoV-2-positive samples of the EQA scheme in April 2020, the *"overall Ct/Cq median"* is shown in Table 1, column E. Considering this *"overall Ct/Cq median"* allowed only generalized conclusions, so it should be noted that the respective difference(s) between the *"overall Ct/Cq medians"* for the four samples, which were derived from a 10-fold dilution series, was remarkably close to the theoretically expected value of 3.32 for a 10-fold difference in concentration (or a multiple thereof) between the individual samples.

For example, the difference between the *"overall Ct/Cq medians"* for sample pair no. 340059 and no. 340063 (10-fold concentration difference) was 3.2 and for sample pair no. 340059 and no. 340064 (100-fold concentration difference) was 6.7. For sample pair no. 340059 and no. 340061 (1 000-fold concentration difference), the difference of the *"overall Ct/Cq medians"* was 9.6 (9.96 expected).

However, a more detailed examination of the Ct/Cq values for the

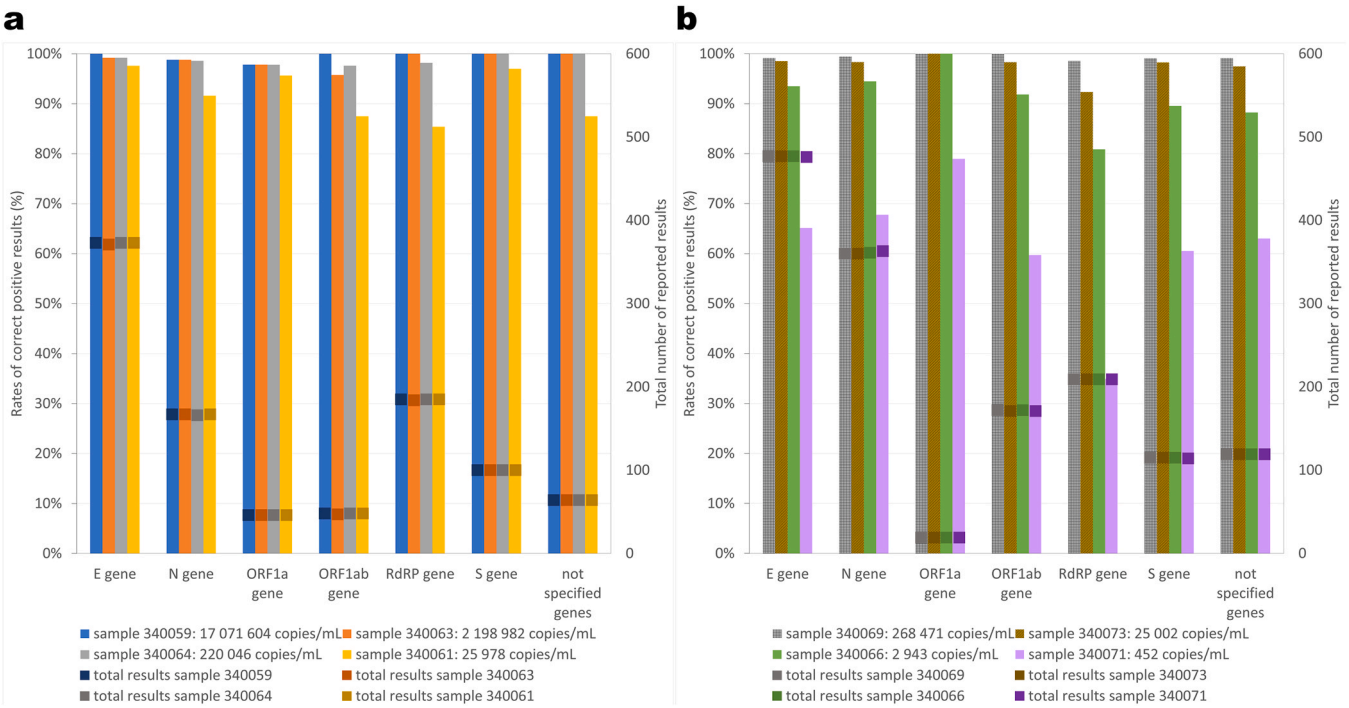


Fig. 1. Differentiation of correct positive results according to target gene regions for the qualitative SARS-CoV-2 detection in the EQA schemes April 2020 and June 2020. The success rates in percent for the results for the specified target gene regions are shown for all SARS-CoV-2-positive samples from the EQA schemes in April 2020 (Fig. 1**a**) and June 2020 (Fig. 1**b**). The non-VOC virus BetaCoV/Munich/ChVir984/2020 was used for all samples. Sample no. 340064 (April 2020) and sample no. 340069 (June 2020) contained virus RNA loads in the same order of magnitude. Samples no. 340061 (April 2020) and no. 340073 (June 2020) were also in the same order of magnitude. In order to better assess the reliability of the rates of correct results (left Y-axis), the total number of reported results is shown in the diagram on the right Y-axis (see squares in the respective column). **Columns in different colors:** The rates of correct positive results (%) for each sample are differentiated according to target gene regions: E gene, N gene, ORF1a, ORF1ab, RdRP gene, S gene and not specified genes. **Squares in the respective column**: The total number of reported results for each sample are differentiated according to the above target gene regions. **EQA scheme April 2020** (Fig. 1a) - **Blue columns**: sample no. 340059; robust average 17 071 604 copies/mL; diluted 1: 1 000. **Orange columns**: sample no. 340063; robust average 2 198 982 copies/mL; diluted 1: 10 000. **Grey columns**: sample no. 340064; robust average 220 046 copies/mL; diluted 1: 100 000. **Yellow columns**: sample no. 340061; robust average 25 978 copies/mL; diluted 1: 1 000 000. The quantitative values for each of the samples derive from the robust average values calculated from all reported quantitative results (Zeichhardt and Kammel, 2020a). **EQA scheme June 2020** (Fig. 1b) - **Grey columns with grid**: sample no. 340069; robust average 268 471 copies/mL; diluted 1: 50 000. **Yellow columns with stripes**: sample no. 340073; robust average 25 002 copies/mL; diluted 1: 500 000. **Green columns**: sample no. 340066; robust average 2 943 copies/mL; 1: 5 000 000 diluted. **Purple columns**: sample no. 340071; robust average 452 copies/mL; diluted 1: 50 000 000. The quantitative values for each of the samples derive from the robust average values calculated from all reported quantitative results (Zeichhardt and Kammel, 2020c).

individual SARS-CoV-2-positive samples showed a high degree of variation. As an example, Fig. 3a shows the Ct/Cq values for sample no. 340061 (SARS-CoV-2 diluted 1: 1 000 000) for each manufacturer and test name applied for the detection of the N gene region.

A total of 132 results were reported for the measured Ct/Cq values for the N gene region. The *"target gene-specific Ct/Cq median"* was 33.3 (Table 2, Fig. 3a).

Strikingly for the N gene region, the *"assay-specific Ct/Cq medians"* between the different test systems applied differed by more than 10 cycles (see Table 2, column E and columns C to D). A similar strong variation of the *"assay-specific Ct/Cq medians"* between the different test systems - as observed for the analyses of the N gene of sample no. 340061 - was also noted for the other target gene regions of SARS-CoV-2 (Table 2).

Even within a single group of participants who performed analyses with a single test system from one manufacturer specific for the N gene region, there were strong variations for the reported Ct/Cq values of up to 10 cycles, in the extreme 16 cycles (see red arrow in Fig. 3a).

Similarly, as for the SARS-CoV-2-positive sample no. 340061 in Table 2, column E, the difference values, i.e. the differences between the minimum and maximum *"assay-specific Ct/Cq medians"* for the respective target gene regions tested, were also very high for the other positive samples (no. 340059, no. 340063, and no. 340064). Consistent with the pattern for difference values for sample no. 340061 shown in Table 2, column E, the other three SARS-CoV-2-positive samples also showed that the difference values for the *"assay-specific Ct/Cq medians"* for the N gene, E gene, S gene, RdRP gene, ORF 1ab and not specified target gene regions ranged from 6.4 to 16.5. For all four positive samples, it was apparent that the difference values for the *"assay-specific Ct/Cq medians"* for the different test systems for ORF 1a only varied between 1.9 and 2.8. For details see Zeichhardt and Kammel, 2020a.

Further review of the Ct/Cq values reported by a single group of participants who performed analyses with a single test system from one manufacturer for the respective gene regions for samples no. 340059, no. 340063, and no. 340064, revealed that as with sample no. 340061, large variations in the reported Ct/Cq values of up to 10 cycles or more occurred already within one defined test system. For details see Zeichhardt and Kammel, 2020a.

3.1.3. *Quantitative results*

At the time of this extra EQA scheme in April 2020, different participants reported quantitative values in copies/mL for the four differently concentrated SARS-CoV-2-positive EQA samples (Zeichhardt and Kammel, 2020a). Obviously, some laboratories already had the possibility of quantification with their own internal standard preparations.

Table 3 shows the robust average values in copies/mL calculated from the quantitative values reported by the laboratories for each of the four SARS-CoV-2-positive samples (independent of the SARS-CoV-2 gene regions addressed by the respective test).

Looking at the quantitative result reports of the 21 laboratories

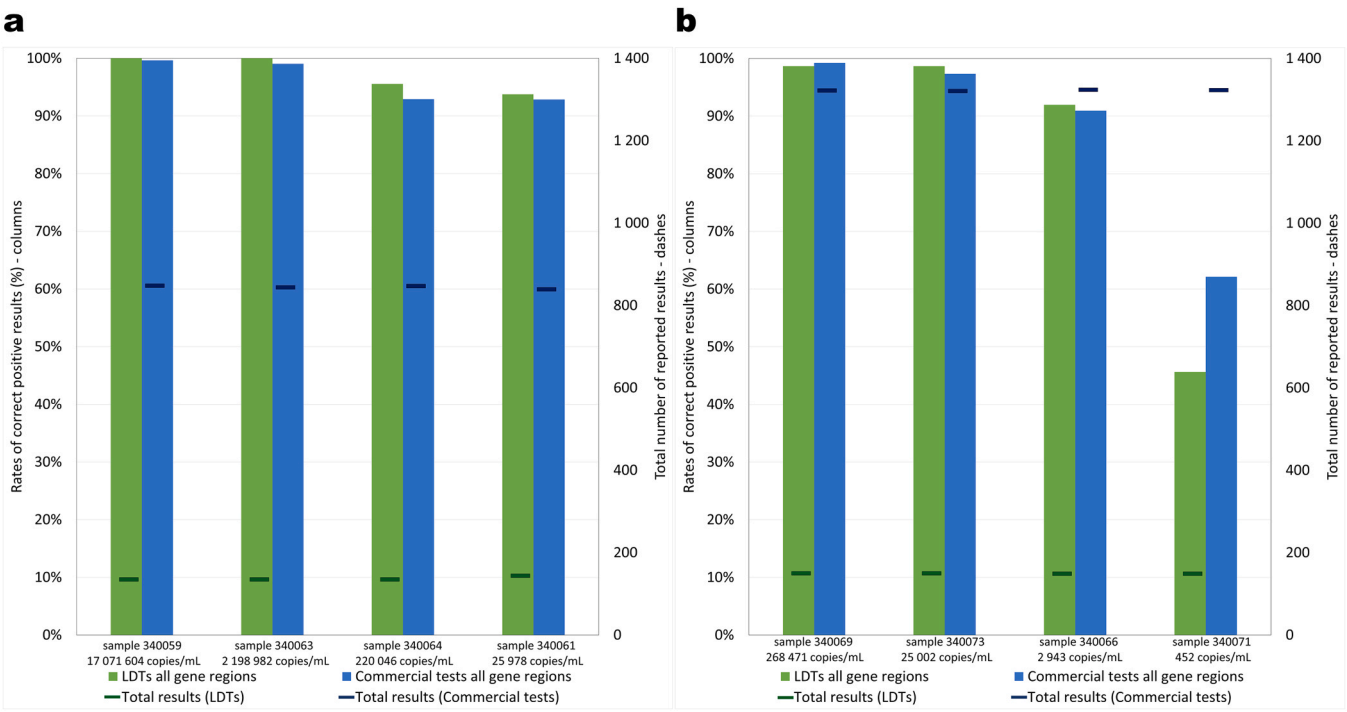


Fig. 2. Differentiation of correct positive results according to laboratory developed test systems (LDTs) and commercial test systems for the qualitative SARS-CoV-2 detection in the EQA schemes April 2020 and June 2020. The success rates in percent for the results differentiated according to LDTs (**green columns**) and commercial tests systems (**blue columns**) are shown for all SARS-CoV-2-positive samples from the EQA schemes in April 2020 (Fig. 2a) and June 2020 (Fig. 2b). The non-VOC virus BetaCoV/Munich/ChVir984/2020 was used for all samples. In order to better assess the reliability of the rates of correct results (left Y-axis), the total number of reported results is shown in the diagram on the right Y-axis (see dashes in the respective column). **Left Y-axis for the columns in green and blue:** Rates of correct positive results (%) for each sample analyzed. **Right Y-axis for the dashes in the respective columns**: The total number of reported results for each sample analyzed. **EQA scheme April 2020** (Fig. 2a) - success rates for qualitative genome detection of the SARS-CoV-2-positive sample no. 340059 (robust average 17 071 604 copies/mL, diluted 1: 1 000); sample no. 340063 (robust average 2 198 982 copies/mL, diluted 1: 10 000); sample no. 340064 (robust average 220 046 copies/mL, diluted 1: 100 000); sample no. 340061 (robust average 25 978 copies/mL, diluted 1: 1 000 000) (Zeichhardt and Kammel, 2020a). **EQA scheme June 2020** (Fig. 2b) - success rates for qualitative genome detection of the SARS-CoV-2-positive sample no. 340069 (robust average 268 471 copies/mL, diluted 1: 50 000); sample no. 340073 (robust average 25 002 copies/mL, diluted 1: 500 000); sample no. 340066 (robust average 2 943 copies/mL, diluted 1: 5 000 000); sample no. 340071 (robust average 452 copies/mL, diluted 1: 50 000 000) (Zeichhardt and Kammel, 2020c).

(Fig. 4a), it was noticeable that - with two exceptions - the laboratories were able to satisfactorily determine the concentration gradation for the four SARS-CoV-2-positive samples diluted in 10-fold steps (each of the parallel shifted lines represents an individual laboratory with its reported quantitative values for the four samples, related to a single gene region). However, it was found that the respective quantitative values in copies/mL for a given sample varied from laboratory to laboratory, so that up to 30 % of the reported quantitative values were outside the assessment range (target value interval of +/- 1 $\log_{10}$ of the target value) (Fig. 4a, Table 3/column D).

Separate consideration should be given to results obtained by digital PCR (dPCR). The dPCR is a calibration-independent genome detection method that provides quantitative results for viral nucleic acid load in copies/mL (Huggett et al., 2015; Whale et al., 2018; Falak et al., 2022, 2025; Milavec et al., 2022). Fig. 4a shows that almost all quantitative dPCR results (highlighted with green lines) were in agreement. This was true for the detections of the N, E and ORF1 gene regions. Only for the detection of RdRP as a target region, was under-quantification observed (lower green line in Fig. 4a).

3.2. June 2020 – second EQA scheme with a revealed sample with a predetermined SARS-CoV-2 RNA load with one sample revealed before EQA scheme

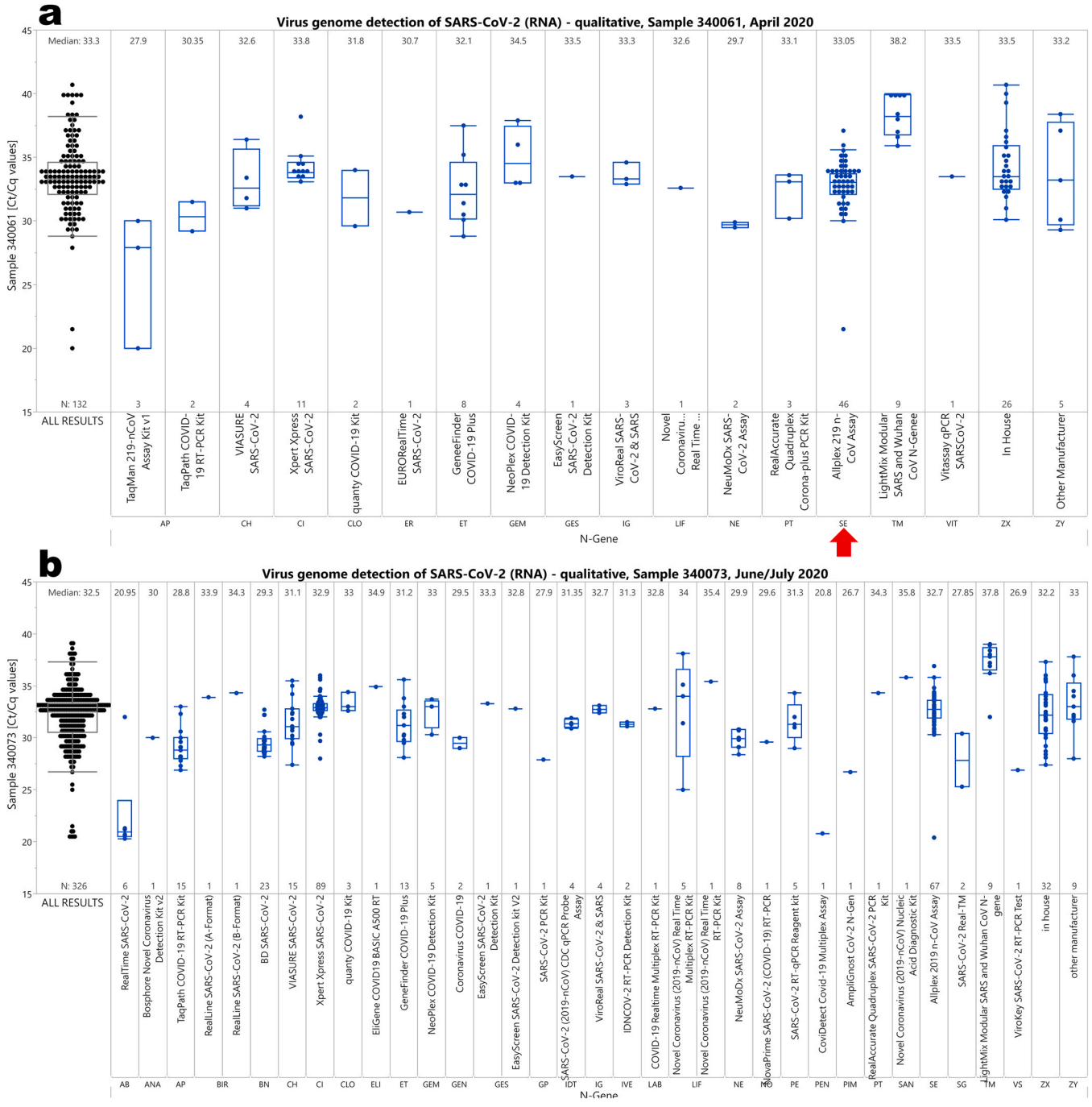
In order to provide EQA scheme participants with further assistance in assessing their testing performance with the test systems used at a short distance from the extra EQA scheme in April 2020, the second EQA scheme for the genome detection of SARS-CoV-2 was already carried out in June 2020 with nine samples in total (four samples positive for SARS CoV-2 in different dilutions and five samples negative for SARS CoV-2) (HZeichhardt and Kammel, 2020c). The June 2020 sample set included four SARS-CoV-2-positive samples from a 10-fold dilution series of the same B.1 non-VOC applied in April 2020 and five samples negative for SARS-CoV-2 (see Supplementary Table 1).

In this second EQA scheme, 636 laboratories from 40 countries reported their results to INSTAND, compared to 463 laboratories in the first EQA scheme in April 2020. This shows the increased need for quality assurance for molecular SARS-CoV-2 diagnostics.

Compared to the April 2020 EQA scheme (Tables 1 and 3), SARS-CoV-2 was provided in June 2020 in higher dilutions to check the performance of the laboratories with the used assays, also for low concentration samples. This mainly concerned sample no. 340071 with a dilution of the SARS-CoV-2-positive material of 1: 50 000 000 ("challenge sample", without evaluation for the certificate of participation) (see Supplementary Table 2 and Table 5).

A special feature of this EQA scheme was that one of the four SARS-CoV-2-positive samples, sample no. 340066 (diluted 1: 5 000 000), was provided as a "revealed sample" to the participants for orientation (without evaluation for the certificate of participation).

For this sample no. 340066, three National Metrology Institutes (NMI) from the UK, USA and Germany determined the virus RNA load in copies/mL by dPCR for the N gene region of SARS-CoV-2. The evaluation of the five analyses resulted in a combined value of 1 570 ± 360 copies/mL (Table 4).



(caption on next page)

Fig. 3. Reported Ct/Cq values for the detection of the N gene of SARS-CoV-2 in the EQA schemes April 2020 and June 2020. The Ct/Cq values for sample no. 340061, EQA scheme April 2020 (Fig. 3a), and sample no. 340073, EQA scheme June 2020 (Fig. 3b), are presented for the detection of N gene of SARS-CoV-2. **Box plots with black dots:** The box plots with the black dots summarize all Ct/Cq values reported by the EQA participants with their test systems used for the respective sample for the detection of the N gene. The median indicated in each of the box plots with the black dots represents the corresponding *"target gene-specific Ct/Cq median for N gene"* (generated for the specific target gene region from the reported results of all applied test systems; category 2 for Ct/Cq value consideration; see Section 3.1.2). **Box plots with blue dots:** The results are differentiated by test manufacturer and test name. Displayed are the reported Ct/Cq values of the individual test systems for N gene detection (represented by blue dots). In addition, for each manufacturer's test system, the corresponding *"assay-specific Ct/Cq median"* (generated for a defined test system from the reported results for a specific target gene region; category 3 for Ct/Cq value consideration) is given. The box for each specified test system includes its interquartile range of 50 % of the data points bounded by the 75 % quartile and 25 % quartile. In addition, the median is plotted for all data points. The whiskers delimit 75 % of the data points. **EQA scheme April 2020, sample no. 340061** (Fig. 3a): The *"target gene-specific Ct/Cq median for N gene"* (according to category 2; see Section 3.1.2) was 33.3 (robust average calculated of all reported quantitative results for this sample – 25 978 copies/mL; Supplementary Table 1) (Zeichhardt and Kammel, 2020a). An example of the range of variation in submitted assay-specific Ct/Cq values within a single group of participants who performed analyses using a single manufacturer's N-gene-specific test system is shown with a variation of up to 16 Ct/Cq values marked with a red arrow. It must be taken into account that the extreme results of the submitted Ct/Cq values may be influenced by strongly deviating incorrect analyses and/or reporting errors. **EQA scheme June 2020, sample no. 340073** (Fig. 3b): The *"target gene-specific Ct/Cq median for N gene"* (according to category 2) was 32.5 (robust average calculated of all quantitative results for this sample – 25 002 copies/mL; Supplementary Table 1) (Zeichhardt and Kammel, 2020c). Abbreviations for test manufacturers: 1DR: 1Drop Diagnostics; 3DM: 3DMed Corporation; AB: Abbott; ALF: Alifax; ANA: Analtolia Geneworks; ANI: Anicon; AP: Applied Biosystems; AST: Astra Biotech; BGI: BGI Genomics; BIR: Bioron GmhH; BMX: Biomaxima; BN: Becton Dickinson; BR: BioRad; CH: Certest Biotec; CI: Cepheid; CLO: Clonit; ELI: Elisabeth Pharmacon; ER: Euro Immun; ET: ELITech; FRI: Friz Biochem; GEM: Genematrix; GEN: Genesig; GES: Genetic Signatures; GP: Geneproof; HA: Hain Lifescience; IBT: Intron Biotechnology; IDT: Integrated DNA Technologies; IG: Ingenetix; IVE: IDvet Innovative Diagnostics; LAB: Labsystems Diagnostics; LGE: Labgenomics; LIF: Liferiver; LU: Luminex; MAC: Maccura Biotechnology; MBS: Mole Bioscience; MOB: Mobidiag; NE: Neumodx; NO: Novatec; PE: Perkin Elmer; PEN: Pentabase; PIM: Priv. Inst. f. Immunol. u. Mol.genetik; PT: Pathofinder; QG: Qiagen; RO: Roche Diagnostics; SAN: Sansure Biotech; SE: Seegene; SG: Sacace Biotechnologies; SGT: Solgent; SO: Adaltis; TFS: Thermo Fisher Scientific; TM: TIB Molbiol; VIT: Vitaassay; VS: Vela Diagnostics; ZX: Laboratory developed tests/LDT; ZY: Other manufacturer; ZYB: Zybio.

Table 2
Ct/Cq medians calculated from the reported Ct/Cq values for sample no. 340061 (April 2020) differentiated by target gene regions.

A	B	C	D	E	F
Target gene region	**"*Target gene-specific Ct/Cq median*"** (for all test systems) according to category 2 for Ct/Cq value consideration	**Minimum value for the "*assay-specific Ct/Cq median*"** (for the specific gene region) according to category 3 for Ct/Cq value consideration	**Maximum value for the "*assay-specific Ct/Cq median*"** (for the specific gene region) according to category 3 for Ct/Cq value consideration	**Difference value between D and C** (for the specific target gene region)	**Total number of reported Ct/Cq values**
N gene	33.3	27.9	38.2	10.3	132
E gene	32.1	29.4	35.5	6.1	336
S gene	31.5	26.1	36.3	10.2	91
RdRP gene	33.5	20.5	36.1	15.6	146
ORF 1a	31.7	31.7	34.5	2.8	40
ORF 1ab	31.5	29.2	35.2	6.0	38
not specified target gene regions	31.9	21.6	34.6	13.0	50

Overview for the different Ct/Cq medians for sample no. 340061 (SARS-CoV-2 RNA-positive, diluted 1: 1 000 000) of the EQA scheme of April 2020 differentiated by the reported results for each target gene region. For this sample the *"overall Ct/Cq median"* was 32.4 (generated by including the results for all target gene regions and all test systems; category 1 for Ct/Cq value consideration; see Section 3.1.2) (see Table 1). Shown are the different target gene regions (**column A**), the *"target gene-specific Ct/Cq medians"* (generated for the specific target gene region from the reported results of all applied test systems; category 2 for Ct/Cq value consideration) (**column B**), the corresponding minimum and maximum *"assay-specific Ct/Cq medians"* (generated for a defined test system from the reported results for a specific target gene region; category 3 for Ct/Cq value consideration) (**columns C and D**), the difference values between the minimum and maximum *"assay-specific Ct/Cq medians"* described in columns C and D (**column E**) and the number of analyses for each of the target gene regions (**column F**) (Zeichhardt and Kammel, 2020a). It must be taken into account that the extreme results of the submitted Ct/Cq values may be influenced by strongly deviating incorrect analyses and/or reporting errors.

Particularly in the early phase of the pandemic, the "revealed sample" no. 340066 with the dPCR assigned quantitative value of 1 570 copies/mL allowed laboratories to mirror the Ct/Cq values measured with their tests on a sample with an assigned virus RNA load in copies/mL. This sample was not considered for certification. Nevertheless, the participants were asked to report their results for this revealed sample.

3.2.1. Qualitative results

Two of the four SARS-CoV-2-positive samples were considered for granting certificates. Sample no. 340069 (diluted 1: 50 000) yielded 98.3 % (1460/1486 correct positive qualitative analyses) in the EQA scheme. Sample no. 340073 (diluted 1: 500 000) showed 96.5 % (1434/1486) correct positive qualitative analyses (Supplementary Table 2).

For the "revealed sample" no. 340066 with 1 570 copies/mL (not considered for the certificate) the detection rate was only 90.2 % (1341/1486) correct results which might be due to the reduced SARS-CoV-2 RNA load.

The sample with the lowest viral load, "challenge sample" no. 340071 (diluted 1: 50 000 000), expectedly achieved only 59.9 % (890/1486) correct positive qualitative analyses.

Differentiation of the reported qualitative results for all four SARS-CoV-2-positive samples of the 10-fold dilution series according to the targeted gene regions gave more insight into the concentration dependence of the results (Fig. 1b). It is noticeable for samples with SARS-CoV-2 RNA loads of approx. 25 000 copies/mL and lower that test systems that used RdRP as the target gene region showed lower success rates for correct positive results than the test systems for the other target gene regions. For the "challenge" sample no. 340071 with a robust average of

Table 3
First INSTAND EQA Scheme (340) for Virus Genome Detection of SARS-CoV-2 - April 2020 – quantitative results for the SARS-CoV-2 RNA-positive samples.

A	B	C	D
Sample no.	**Dilutions of SARS-CoV-2-positive samples**	**Robust average values (target values) for all reported quantitative results** (copies/mL) [target value interval of +/- 1.0 log_{10} of the target value for consideration as correct results]	**Success rates of results in the target value interval of +/- 1.0 log_{10} of the target value** [number of correct results per total number of reported results]
340059 revealed in the interim evaluation	1: 1 000	**17 071 604** [1 707 160 – 170 716 040]	**76.5 %** [26/34]
340063	1: 10 000	**2 198 982** [219 898 – 21 989 820]	**73.5 %** [25/34]
340064 revealed in the interim evaluation	1: 100 000	**220 046** [22 005 – 2 200 460]	**70.6 %** [24/34]
340061	1: 1 000 000	**25 978** [2 598 – 259 780]	**73.5 %** [25/34]

The table summarizes the quantitative results reported by 21 laboratories for the four SARS-CoV-2 RNA-positive samples of the first INSTAND EQA scheme for virus genome detection of SARS-CoV-2 in April 2020 (34 quantitative results per sample). Shown are the sample numbers (**column A**), the dilutions of the SARS-CoV-2 RNA-positive materials (**column B**), the robust average values (target values) for all reported quantitative results in copies/mL and the target value interval of +/- 1.0 log_{10} of the target value for consideration as correct results (**column C**) and the success rates of results in the target value interval of +/- 1.0 log_{10} of the target value including the number of correct results per total number of reported results (**column D**) (Zeichhardt and Kammel, 2020a).

452 copies/mL, the rates for correct positive results for the different target gene regions were only between 34.9 % (RdRP target gene region) and 78.9 % (ORF1a target gene region). While satisfactory success rates between 90 % and 100 % for all different target gene regions were achieved for the higher concentrated samples, sample no. 340069 (diluted 1: 50 000, 268 471 copies/mL) and sample no. 340073 (diluted 1: 500 000, 25 002 copies/mL). For comparison, see also Fig. 1a with the results from the EQA scheme in April 2020.

In this second EQA scheme, of the 5 888 reported results for the four SARS-CoV-2-positive samples (related to all test systems and gene regions examined), 10,2 % (598/5 888 results) were performed using LDTs (Fig. 2b). This means that the number of reported results for LDTs remained almost the same in the first two EQA schemes, while the number of reported results for the commercial test systems increased significantly from 3 378–5 290 (Figs. 2a and 2b). Looking at the three samples in the 10-fold dilution series with SARS-CoV-2 RNA loads between 268 471 copies/mL and 2 943 copies/mL for the EQA scheme June 2020, the LDTs showed good success rates of 91.9–98.7 %, as did the commercial test systems with 90.9–99.2 % for the correct positive results (Fig. 2b). For the "challenge" sample no. 340071 with the lowest SARS-CoV-2 RNA load of 452 copies/mL, the success rates were very low, i.e. 45.6 % correct positive results for LDTs and 62.1 % correct positive results for the commercial test systems.

In order to get insight into the performance of the laboratories in regard to the specificity of the applied test systems, the EQA scheme in June 2020 contained five samples negative for SARS-CoV-2. These samples consisted of either MERS CoV (inactivated), hCoV 229E, hCoV NL63, hCoV OC43 or uninfected cells (see Supplementary Table 1). In addition, the samples could also be used for virus differentiation. The success rates for all five control samples ranged satisfactorily between 97.1 % and 98.1 % of correct negative results for SARS-CoV-2 (Zeichhardt and Kammel, 2020c), as already described for the specificity controls in the EQA scheme April 2020 (see Table 1).

3.2.2. Results in Ct/Cq values

In the second EQA scheme in June 2020, up to 1 486 qualitative results with up to 1 306 results in Ct/Cq values were reported for each of the four SARS-2-positive samples from the 10-fold dilution series for the individual gene regions (e.g. 326 results in Ct/Cq values for the N gene region of sample no. 340073; see Fig. 3b).

As shown in Supplementary Table 2 and Fig. 3b, the second EQA scheme in June 2020 confirmed the observation of the first EQA scheme when looking at Ct/Cq values (for comparison, see Table 1 and Fig. 3a):

(i) For the four SARS-CoV-2-positive samples from the 10-fold dilution series, the *"overall Ct/Cq medians"* showed that the expected Ct/Cq differences of the dilution series (theoretically expected value 3.32 for a 10-fold gradation) were well reproduced. The difference in the respective *"overall Ct/Cq medians"* for the sample pair no. 340069 and no. 340073 (concentration difference of 1: 10) was 3.3 and for the sample pair no. 340069 and no. 340066 (concentration difference of 1: 100) was 6.2. For the sample pair no. 340069 and no. 340071 (concentration difference of 1: 1 000), the difference of the *"overall Ct/Cq medians"* was only 8.6 (9.9 expected), which could be due to the high dilution of sample no. 340071 (1: 50 000 000).

(ii) The wide assay and target gene-specific variation of Ct/Cq values was confirmed for the SARS-CoV-2-positive samples in this second EQA scheme (Fig. 3b). Again, the results for SARS-CoV-2 detection in the N gene serve as an example when looking at the *"assay-specific Ct/Cq medians"*. As with the results for sample no. 340061 in the EQA scheme April 2020 with differences between the *"assay-specific Ct/Cq medians"* by ten (up to 16) (see Fig. 3a), these differences were similar (up to 17.0) for sample no. 340073 in the EQA scheme June 2020 (Fig. 3b). It could not be decided whether this remaining high difference compared to the EQA scheme April 2020 could be explained by the respective larger number of participants and the introduction of new test system applied in the EQA scheme June 2020.

Fig. 3b shows for the N gene region of sample no. 340073 a *"target gene-specific Ct/Cq median for N gene"* of 32.5.

3.2.3. Quantitative results

In total 32 laboratories reported 44 quantitative values determined by qPCR and dPCR in copies/mL for each of the four positive samples (Zeichhardt and Kammel, 2020c).

The reported values in copies/mL for each of these samples yielded robust average values (regardless of the SARS-CoV-2 gene regions addressed by each assay) with satisfying expected values for 10-fold gradation (Table 5).

For quantitative consideration, two of the four SARS-CoV-2-positive samples in June 2020 were of particular interest.

For the revealed sample no. 340066, which had a value of 1 570

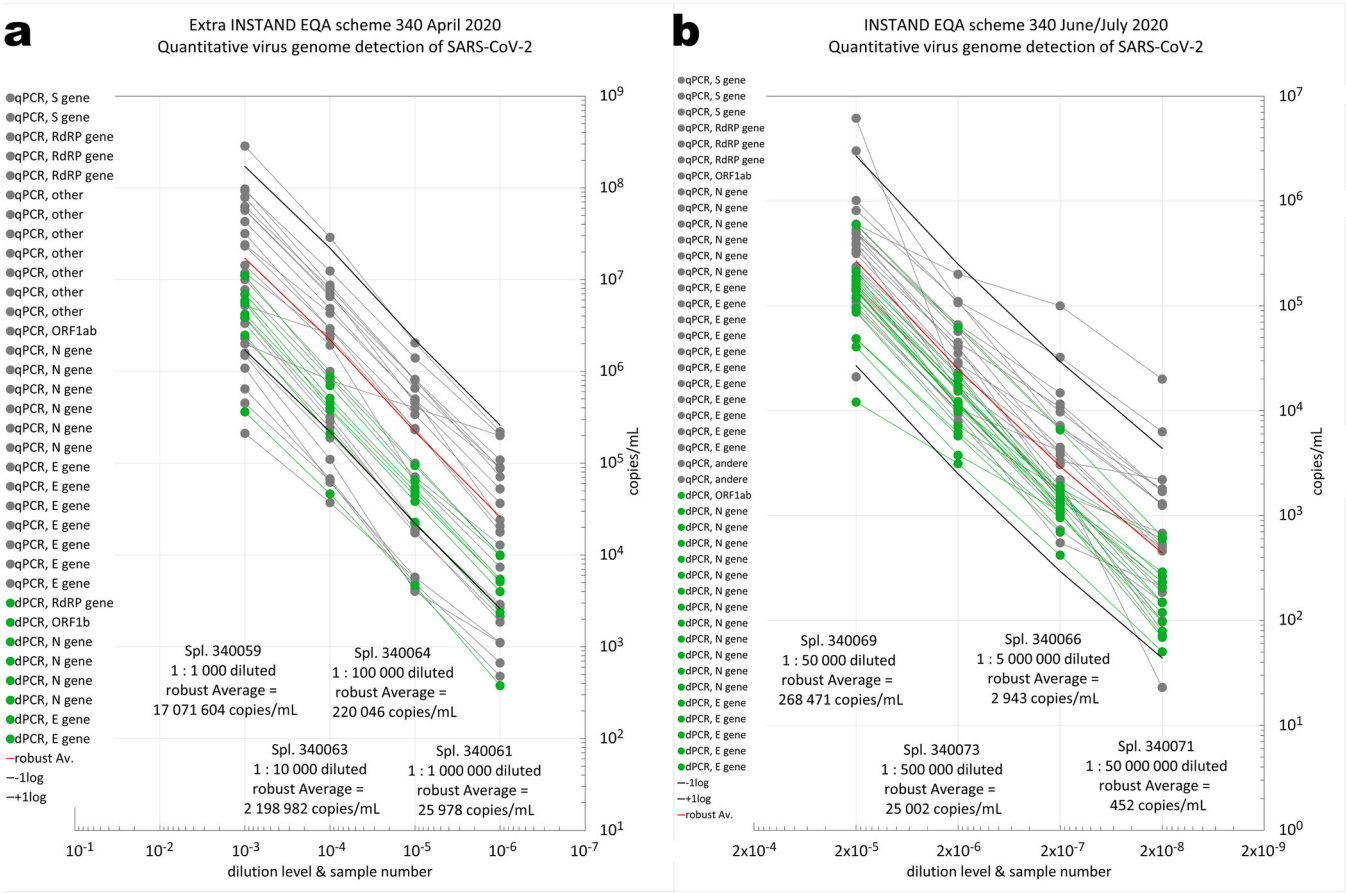


Fig. 4. Quantitative results for the SARS-CoV-2 samples in the EQA schemes. The samples no. 340059, no. 340061, no. 340063 and no. 340064 (Extra EQA scheme April 2020; Fig. 4a) and the samples no. 340066, no. 340069, no. 340071 and no. 340073 (EQA scheme June 2020; Fig. 4b) derive from two independent ten-fold dilution series of the same initial SARS-CoV-2 variant BetaCoV/Munich/ChVir984/2020 (Supplementary Table 1). Each line represents the results for the respective sample from an individual laboratory. **Grey lines and dots** represent results from quantitative PCR (qPCR), **green lines and dots** represent results from digital PCR (dPCR). The **red line** marks the robust average values as a "consensus line" from all quantitative results reported by all laboratories. The **black lines** indicate the assessment range (target value interval of +/- 1 $\log_{10}$ of the target value). In April 2020, 18 laboratories performed analyses by qPCR and 3 laboratories analyses by dPCR (Zeichhardt and Kammel, 2020a). In June 2020, 32 laboratories reported quantitative results (Zeichhardt and Kammel, 2020c). Of these, one laboratory performed studies with both qPCR and dPCR. Furthermore, 22 laboratories performed qPCR alone and 9 laboratories performed dPCR alone.

copies/mL assigned by dPCR, it is noteworthy that the robust average value of the 32 laboratories having applied different test systems for quantification was close, namely 2 943 copies/mL. Despite this good agreement of the robust average value with the assigned dPCR value, it should be noted that the values for the individual determinations by qPCR showed a high variation (550 – 100 000 copies/mL) (Fig. 4b). A total of 36 of 44 reported results were obtained with LDTs (data not shown).

Sample no. 340071, which contained the highest dilution of SARS-CoV-2 (diluted 1: 50 000 000) and was declared a "challenge" sample in the evaluation of qualitative results because of the reduced success rate (59.9 % correct qualitative results, see Supplementary Table 2), yielded 452 copies/mL as a robust average value.

For all SARS-CoV-2-positive samples in the EQA scheme June 2020, as with the EQA scheme April 2020, the quantitative values in copies/mL for the respective SARS-CoV-2-positive samples varied up to three orders of magnitude (Fig. 4b).

3.3. *Reference materials for self-assessment of the test systems used*

For the first two EQA schemes, April 2020 and June 2020, both the reported results in Ct/Cq values and the quantitative results in copies/mL showed a high variation of the measured values depending on the individual laboratory, applied test system and SARS-CoV-2 target gene. Especially the results of the second EQA scheme with the revealed sample no. 340066, which was pre-quantified by dPCR at 1570 copies/mL, clearly showed that no consistent Ct/Cq value could be assigned to this virus RNA load. In order to improve this situation, there was a need for reference materials against which laboratories could anchor their individually determined results in order to assess the characteristics of the tests applied.

In mid-2020, there was no WHO International Standard for SARS-CoV-2 RNA with a predetermined virus RNA load available, so two independent reference materials (RM 1 and RM 2) were established in Germany in collaboration between the Robert Koch-Institut (RKI), the National Coronavirus Consultant Laboratory at the Institute of Virology of the Charité – University Medicine Berlin, INSTAND, GBD and members of the Joint Diagnostic Commission of DVV and GfV (Vierbaum et al., 2022).

Standardization of virus RNA load determination or estimation of Ct/Cq values using reference materials with assigned values in copies/mL for SARS-CoV-2 RNA was one of the prerequisites to use SARS-CoV-2 RNA load in patient materials as a surrogate for assessing the transmission risk that a patient may pose. This was then associated with clinical decisions for patients to be discharged from isolation (Vierbaum et al., 2022). It was recognized from the beginning that pre-analytical factors during sampling have a great influence on the quality of the test material (in general nasopharyngeal or nasal swabs) and thus on the

Table 4
Assignment of a quantitative value in copies/mL for sample no. 340066 by dPCR.

A		B	C
NMI and dPCR target gene region		**SARS-CoV-2 RNA load determined by dPCR (± standard uncertainty)** [copies/mL]	**Consensus value SARS-CoV-2 RNA load determined by dPCR ±expanded uncertainty (95 % CI)** [copies/mL]
NMI1	**N2**	1 421 (± 175)	1 570 (± 360)
NMI2	**China N**	1 547 (± 89)	
NMI2	**N1**	1 437 (± 92)	
NMI2	**N2**	1 916 (± 65)	
NMI3	**China N**	1 333 (± 373)	

Table 4 summarizes the quantitative results which were independently obtained by the three National Metrology Institutes (NMIs) by application of dPCR for sample no. 340066 (diluted 1: 5 000 000). Shown are the target gene regions used for the different dPCR assays (**column A**), the reported SARS-CoV-2 RNA load in copies/mL (**column B**) and the consensus value for the SARS-CoV-2 RNA load in copies /mL calculated from the 5 independent dPCR measurements (**column C**). The dPCR assigned quantitative value of 1 570 ± 360 copies/mL is a consensus value from the analyses of three NMIs: National Metrology Laboratory for Chemical and Biological Measurements, LGC, Teddington, UK; - Physikalisch-Technische Bundesanstalt, AG 8.32 Cell and Molecular Biological Measurement Methods, Berlin, Germany; National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, U.S.A.
The dPCR analyses of the three NMIs were performed on basis of dPCR quantifications with N1/N2 assays (Lu et al., 2020) and China N assay (Chinese Centre for Disease Control and Prevention, 2020), together with the consensus value estimated using the DerSimonian-Laird approach.

obtained measured values in Ct/Cq values. It was, therefore, recommended that such statements on patient management, based on Ct/Cq values and anchored to quantified reference material, would only be appropriate in the case of professionally reliable sampling during follow-up examinations.

Against this background, the question was raised as to which criteria should be used to decide at which virus RNA load an approximate statement about the infectivity of a patient can be made. For this purpose, it was deduced from literature data that the probability of virus cultivation – as biological marker for infectiousness – is low for diagnostic samples with a SARS-CoV-2 RNA load below ~10^6 to ~10^7 copies/mL (conservatively estimated at about 20 %), which applies in particular to diagnostic samples taken after the onset of symptoms. Under these conditions, virus RNA load has been used as a surrogate for patient management in terms of transmission risk or as a criterion for discharging patients from isolation. (von Kleist et al., 2020; Arons et al., 2020; Perera et al., 2020; van Kampen et al., 2021; Woelfel et al., 2020a, 2020b).

The above-mentioned cooperation partners decided that a sample set with two tubes each of reference material RM 1 (with approx. 10^7 copies/mL) and reference material RM 2 (with approx. 10^6 copies/mL) should be produced and provided free of charge for the German laboratories.

The virus RNA load of sample no. 340066 (1 570 copies/mL), which had been quantified by dPCR by the three NMIs described above for the EQA scheme in June 2020 (see above Section 3.2), was the basis for setting the reference materials RM 1 and RM 2 at 10^7 copies/mL and 10^6 copies/mL, respectively. RM 1 and RM 2 were prepared by GBD. The virus RNA loads of RM 1 and RM 2 were quantified by (i) qPCR on synthetic RNA standards by the National Consultant Laboratory for Coronaviruses, Berlin, and Robert Koch-Institut and (ii) dPCR by the above mentioned three National Metrology Institutes.

This set of reference materials was offered to German EQA participants starting in November 2020 and was intended to serve as an anchor

Table 5
INSTAND EQA Scheme (340) for Virus Genome Detection of SARS-CoV-2 - June 2020 – quantitative results for the SARS-CoV-2 RNA-positive samples.

A	B	C	D
Sample no.	**Dilutions of SARS-CoV−2-positive samples**	**Robust average values (target values) for all reported quantitative results** (copies/mL) [target value interval of + /- 1.0 $\log_{10}$ of the target value for consideration as correct results]	**Success rates of results in the target value interval of + /- 1.0 $\log_{10}$ of the target value** [number of correct results per total number of reported results]
340069	1: 50 000	**268 471** [26 847 – 2 684 710]	**90.9 %** [40/44]
340073	1: 500 000	**25 002** [2 500 – 250 020]	**97.7 %** [43/44]
340066 revealed sample	1: 5 000 000	**Not evaluated** 2 943 [294 – 29 430]	**Not evaluated** 95.5 % [42/44]
340071 challenge sample	1: 50 000 000	**Not evaluated** 452 [45 – 4 520]	**Not evaluated** 81.8 % [36/44]

The table summarizes the 44 quantitative results of 32 laboratories for each of the SARS-CoV-2 RNA-positive samples of the INSTAND EQA scheme for virus genome detection of SARS-CoV-2 in June 2020. Shown are the sample numbers (**column A**), the dilutions of the SARS-CoV-2 RNA-positive materials (**column B**), the robust average values (target values) for all reported quantitative results in copies/mL and the interval +/- 1.0 $\log_{10}$ of the target value for consideration as correct results (**column C**) and the success rates of results in the target value interval of +/- 1.0 $\log_{10}$ of the target value including the number of correct results per total number of reported results (**column D**) (Zeichhardt and Kammel, 2020c).

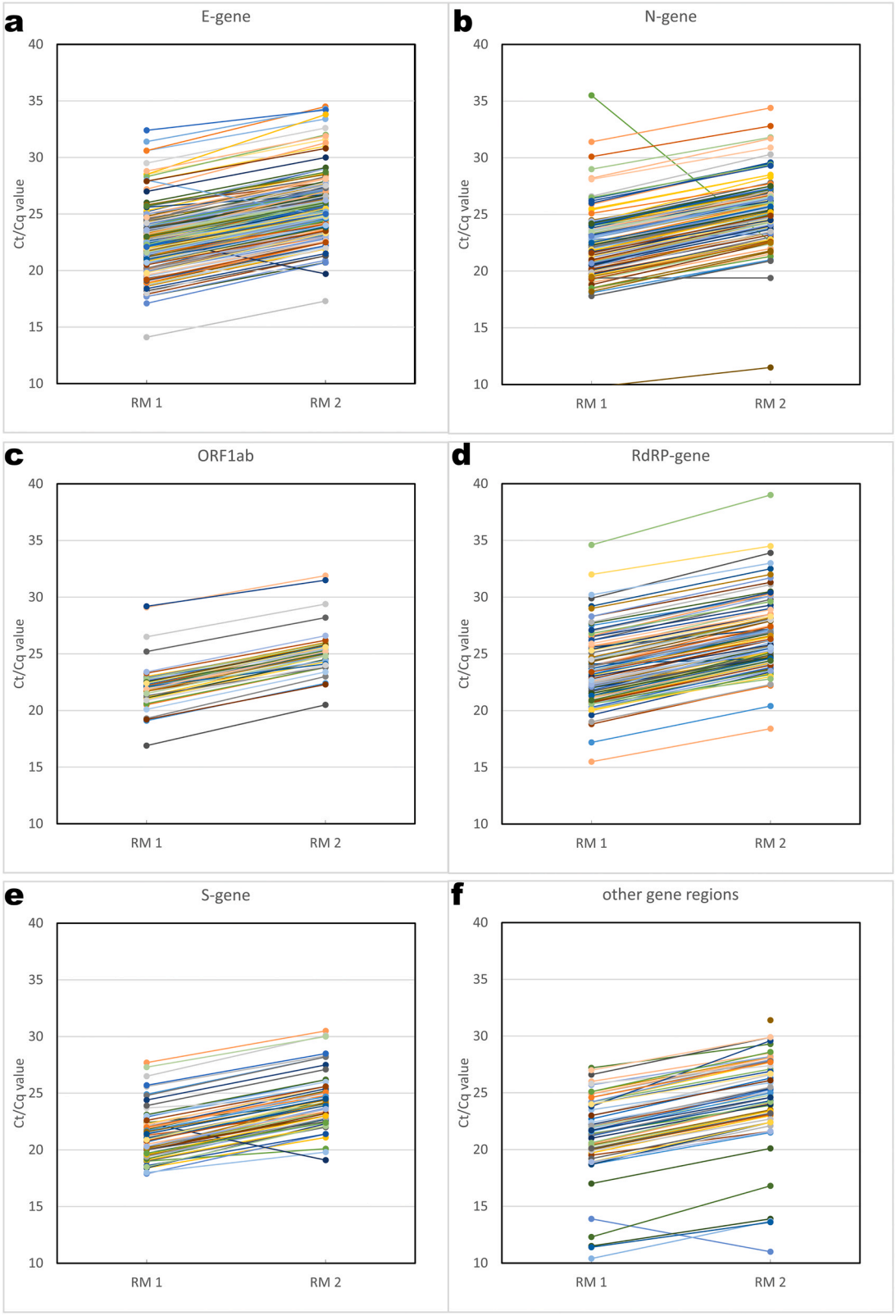


Fig. 5. Results for quantitative reference materials RM 1 and RM 2 differentiated by target gene region from 234 laboratories. Each line represents the results of one laboratory for the respective selected test for the respective target gene region (as of January 15, 2021) for E gene (**a**), N gene (**b**), ORF1ab (**c**), RdRP gene (**d**), S gene (**e**), and other gene regions (**f**) (Zeichhardt and Kammel, 2021).

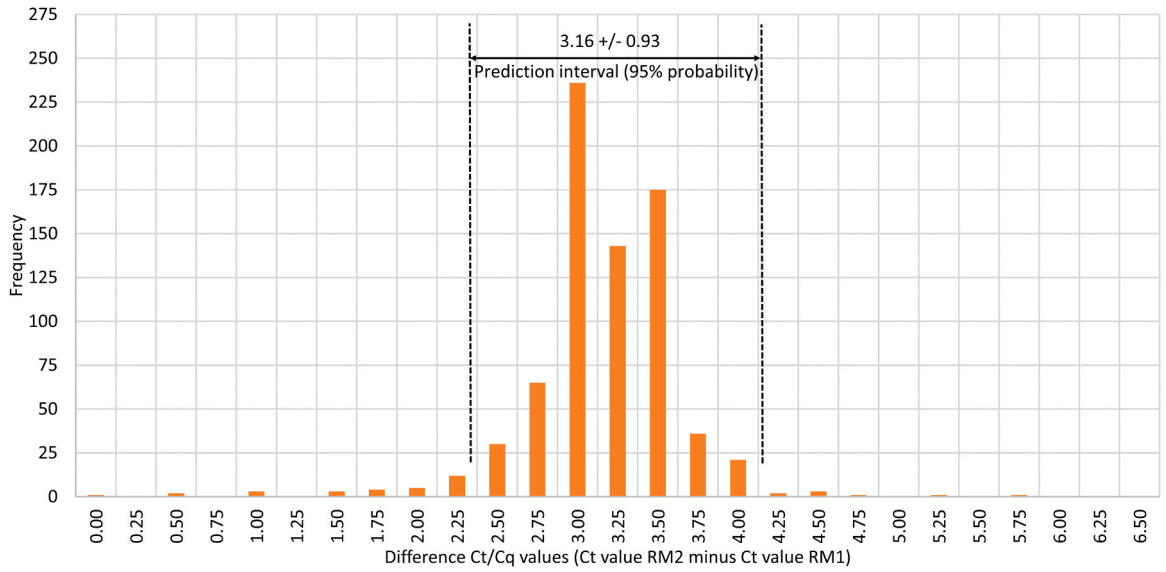


Fig. 6. Histogram of the calculated differences of the submitted Ct/Cq values for RM 2 and RM 1 (as of January 15, 2021). The average Ct/Cq value difference was 3.16 with a prediction interval of 0.93 (95 % probability) (Zeichhardt and Kammel, 2021). Not shown are Ct/Cq differences from 5 analyses with sample mix-ups.

for estimating the Ct/Cq values determined with different test systems targeting different gene regions. The reference materials were sent to 305 German EQA scheme participants in three shipments (03 and 17 November 2020 and 15 January 2021).

The German laboratories that had received the two reference materials RM 1 and RM 2 were asked to report their measurement results to INSTAND analogous to EQA survey results. By January 15, 2021, an interim analysis of the results of the first two shipments was performed and communicated to participating laboratories in an INSTAND notice in order to give the laboratories an early orientation (Zeichhardt and Kammel, 2021). This interim analysis included Ct/Cq results from a total of 234 laboratories (754 results per reference material for the test systems used).

The reported Ct/Cq values are shown in Fig. 5a to f for each target gene (E gene, N gene, ORF1ab, RdRP gene, S gene, other gene region) separately for RM 1 and RM 2. Each line represents the reported Ct/Cq values of an individual laboratory for RM 1 and RM2. The analyses of nearly all laboratories reflected well the 10-fold concentration difference between the two reference materials. However, the parallel shifted lines of the Ct/Cq value measurements of the individual laboratories showed a strong variation of the Ct/Cq values of the individual test systems. For example, the parallel lines for the E gene show Ct/Cq differences of up to 18.3 Ct/Cq levels.

As already observed in the first two EQA programs, the reported Ct/Cq values showed a high variation depending on the results of the performing laboratory and the test system used (Fig. 5). This large variation in Ct/Cq values did not allow differentiation in the reported results by the target gene regions.

In some cases, the results of individual laboratories for the sample pair did not follow the expected line progression, which was either due to sample interchanges and/or mis-entries in the protocol or due to incorrect measurements.

For each participating laboratory, the difference between the submitted Ct/Cq values for RM 1 and RM 2 was calculated and shown according to their frequency in Fig. 6. For the calculation, all 750 out of 754 results (each for RM 1 and RM 2) were included regardless of the test system used and the individual target gene regions.

The average value of the calculated Ct/Cq difference values ("Ct/Cq value RM 2 minus Ct/Cq value RM 1") was 3.16, which is very close to the theoretically expected Ct/Cq difference value of 3.32. This reflected well the 10-fold dilution between the different concentrations of RM 1 and RM 2.

A detailed analysis of the results reported for the reference materials RM 1 and RM 2 has been published by Vierbaum et al., 2022.

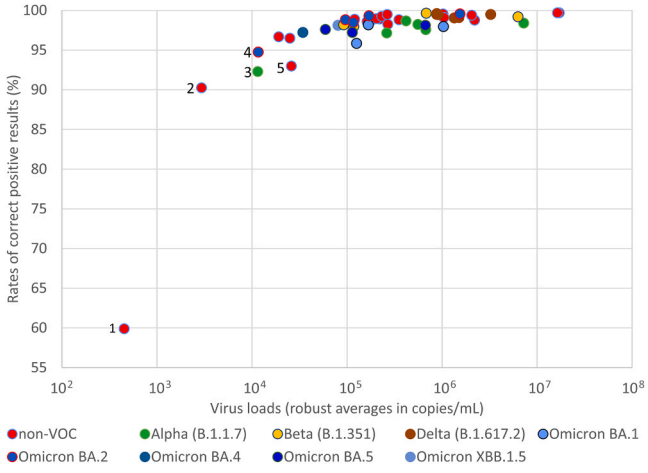


Fig. 7. Rates of correct positive results differentiated by SARS-CoV-2 variants, virus RNA loads and time-dependent use in the EQA schemes. The data refer to the 13 EQA schemes from April 2020 to June 2023. **The different virus variants are color coded.** For a given sample the diagram shows the rate of correct positive results in percent according to the respective qualitative target value (see Materials and Methods, Data evaluation). For each sample the rate of correct positive results is plotted against the corresponding quantitative target value for the SARS-CoV-2 RNA load derived from the consensus value from all quantitative results (based on the robust average in copies/mL). Numbered data points: 1 = sample no. 340071/June 2020/diluted 1: 50 000 000; 2 = sample no. 340066/June 2020/diluted 1: 5 000 000; 3 = sample no. 409–230613–04/June 2023/ diluted 1: 2 500 000; 4 = sample no. 409033/September 2022/ diluted 1: 2 000 000; 5 = sample no. 340061/ April 2020/ diluted 1: 1 000 000.

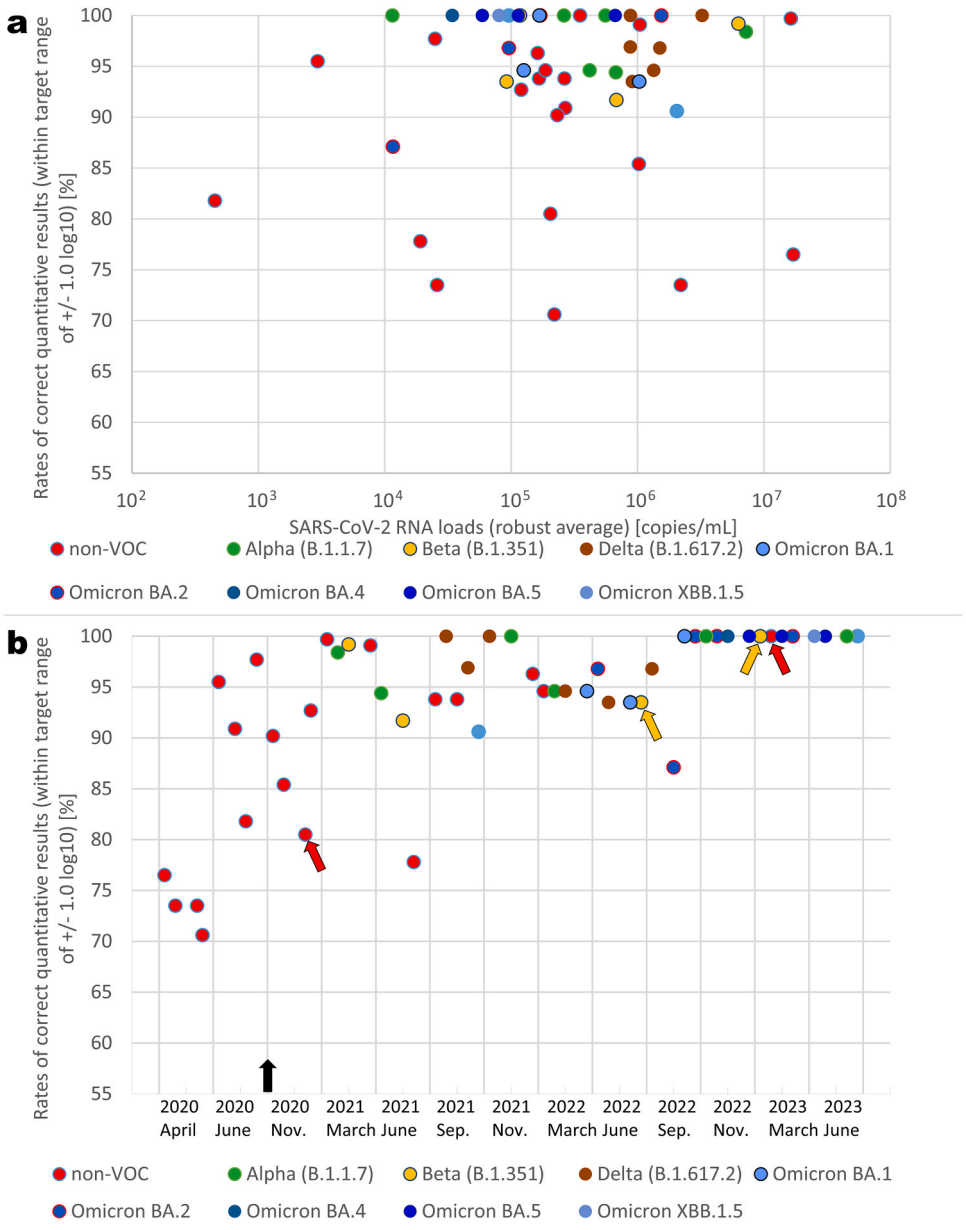


Fig. 8. Rates of correct quantitative results differentiated by SARS-CoV-2 variants, virus RNA loads and time-dependent use in the EQA schemes. The data refer to the 13 EQA schemes from April 2020 to June 2023 differentiated according to the SARS-CoV-2 RNA loads (copies/mL) (Fig. 8a) and date of EQA scheme (Fig. 8b). The different virus variants are color coded. The diagrams show the rates of correct quantitative results in percent for each sample according to the respective quantitative target intervals (see Materials and Methods, Data evaluation). Fig. 8a**:** For each sample the rate of correct quantitative results (within a target value interval of +/- 1.0 $\log_{10}$ of the target value) is plotted against the corresponding quantitative target value for the SARS-CoV-2 RNA load derived from the consensus value from all quantitative results (based on the robust average in copies/mL). Fig. 8b**:** The individual data points shown in Fig. 8a are assigned to the respective EQA term. In detail, for each sample the rate of correct quantitative results (within a target value interval of +/- 1.0 $\log_{10}$ of the target value) is plotted against the time-dependent use in the different EQA schemes. The **black arrow** represents the start of shipments of the two reference materials (RM) containing SARS-CoV-2 at two different RNA loads (RM 1 at approximately 10^7 copies/mL) and RM 2 at approximately 10^6 copies/mL). Samples were shipped on November 03 and 17, 2020, and January 15, 2021 (see Section 3.3). The **red arrows** mark two identical EQA samples, each containing the same non-VOC virus (BetaCoV/Baden-Württemberg/1/ChVir1577/2020 isolateBER; diluted 1: 50 000) from the same preparation with the same virus RNA load but with different coding, and used on two different EQA dates [sample no. 340079 in November 2020 (204 795 copies/mL robust average) and sample no. 409–230314–02 in March 2023 (351 642 copies/mL robust average)]. The **yellow arrows** mark two identical EQA samples, each containing the same VOC Beta B.1.351 (BetaCoV/South Africa/ChVir22131/2020; diluted 1: 350 000) from the same preparation with the same virus RNA load but with different coding, and used on two different EQA dates [sample no. 409030 in June 2022 (92 064 copies/mL robust average) and sample no. 409–230314–01 in March 2023 (117 581 copies/mL robust average)].

3.4. Longitudinal view of EQA scheme results in the period April 2020 to June 2023 including results with variants adapted to the epidemiological situation

The number of EQA scheme participants showed a wide range for the period of April 2020 to June 2023. At the beginning of the pandemic in April 2020, 463 laboratories reported results. In June 2020, the number of laboratories having submitted results reached a peak of 636 laboratories. From November 2020 to June 2023, reporting laboratories declined from 571 (November 2020) to 260 (June 2023).

A total of 13 EQA schemes for genome detection of SARS-CoV-2 with 52 SARS-CoV-2-positive samples were performed between April 2020 and June 2023 (see Supplementary Table 1). For the first three EQA dates (April, June and November 2020), three initial isolates from Bavaria, Berlin and Baden-Wuerttemberg were used in different concentrations. From March 2021, current variants of concern (VOC) were included to reflect the current epidemiological situation. A total of eight newly emerged variants of concern (VOC) were analyzed until the EQA scheme in June 2023.

A total of 21 SARS-CoV-2-negative samples were tested. Of these, nine samples contained MERS CoV or the human coronaviruses CoV 229E or CoV NL63 or CoV OC43. A total of 12 samples contained lysates of non-infected control cells, MRC-5 cells.

3.4.1. Qualitative results

For 46 of the 52 SARS-CoV-2-positive EQA samples, the success rates for qualitative detection of SARS-CoV-2 ranged from 95.8 % to 99.7 % correct positive results (Fig. 7). Five samples containing highly diluted SARS-CoV-2 (dilutions between 1: 1 000 000 and 1: 50 000 000) showed reduced success rates between 59.9 % and 94.8 % correct positive results. Sample no. 340071 with 59.9 % success rate was already described above as a "challenge" sample in the June 2020 EQA scheme (robust average 452 copies/mL). The other four high dilution samples had robust averages only between 2 943 and 25 978 copies/mL, which may have accounted for the reduced success rates.

The result of one sample in the EQA scheme November 2020 (sample no. 340081) will be considered separately. This sample with only 82.9 % correct results was used as another "challenge" sample because it comprised synthetic SARS-CoV-2 RNA fragments that did not contain all diagnostically relevant gene targets.

For all 21 SARS-CoV-2-negative EQA samples, success rates for qualitative detection ranged from 95.1 % to 99.4 % correct negative results. For three of these SARS-CoV-2-negative samples, success rates were between 95.1 % and 95.3 % correct negative results. It is unclear whether the incorrect negative results were due to (i) laboratory handling errors such as sample mix-up, protocol entry errors, incorrect measurements or carry-over of SARS-CoV-2 from samples with higher SARS-CoV-2 RNA loads or (ii) the test systems used not recognizing certain virus variants.

Fig. 7 provides information on the rates of correct positive results with respect to the different virus variants used, differentiated by SARS-CoV-2 RNA load. For a given sample the virus RNA load in copies/mL refers to the individual target value derived from the consensus value from all quantitative results (based on the robust average). Each data point represents between 532 and 1486 results considering all target gene regions.

With regard to the rate of correct positive results, it was found that for SARS-CoV-2 RNA loads higher than 100 000 copies/mL, the respective correct positive rates ranged from 95.8 % to 99.7 % irrespective of the virus variant tested. For SARS-CoV-2 RNA loads in the range between 10 000 and 100 000 copies/mL (robust averages), there was a decrease in the corresponding rates of correct positive results, with the majority of rates still above 95 %. Again, this was shown to be very similar for all virus variants examined. Three EQA samples showed rates of correct positive results between 92 % and 95 % (Fig. 7, data points labelled with numbers 3, 4 and 5). For the two samples with lower SARS-CoV-2 RNA loads, less than 10 000 copies/mL, the rates of correct positive results decreased notably. As already described above for the EQA scheme June 2020, the sample with a rate of 90.2 % correct positive results and a robust average of 2 943 copies/mL was the revealed sample no. 340066 (data point number 2), which had been assigned a SARS-CoV-2 RNA load of 1 570 copies/mL by dPCR (Section 3.2, Table 4). The sample with a correct positive rate of only 59.9 % and a SARS-CoV-2 RNA load of 452 copies/mL (robust average) was used in the EQA scheme June 2020 as a highly diluted "challenge" sample (sample no. 340071, diluted 1: 50 000 000; data point number 1) as described above (Section 3.2.1).

3.4.2. Results in Ct/Cq values

For the longitudinal analysis of the results in Ct/Cq values, sample no. 409–230613–01 from the June 2023 EQA scheme is selected as a representative sample (Zeichhardt and Kammel, 2023a). This sample no. 409–230613–01 contained the VOC Omicron XBB 1.5 (B.1.1.529) and was diluted by 1: 250 000 resulting in a robust average value of 80 360 copies/mL. For this sample the *"overall Ct/Cq median"* was 29.9 based on 404 analyses (data not shown).

The Ct/Cq values from the N gene analyses are considered in Supplementary Figure 1. A total of 107 results were reported for the measured Ct/Cq values for the N gene region. The *"target gene-specific Ct/Cq median for N gene"* was 29.9 (see Supplementary Table 3, Supplementary Figure 1). As already done for sample no. 340061 of the EQA scheme April 2020 (Table 2, Fig. 3a), for sample no. 409–230613–01 the differentiation of the *"assay-specific Ct/Cq median"* for N gene analyses yielded a variation of up to 13 Ct/Cq values between the different test systems (see Supplementary Table 3, column E and Supplementary Figure 1).

For the other target gene regions, as observed in the analyses of the N gene, the differences between the minimum and maximum *"assay-specific Ct/Cq medians"* were also variable, ranging from 3.8 to 24.1 cycles (see Supplementary Table 3, column E). It must be taken into account that the extreme results of the submitted Ct/Cq values may be influenced by strongly deviating incorrect analyses and/or reporting errors.

Again, even within a single group of participants who performed analyses with a test system from one manufacturer specific for the N gene region, there was a variation in reported Ct/Cq values of up to 6.0 cycles (see Supplementary Figure 1).

3.4.3. Quantitative results

Fig. 8 shows a longitudinal analysis of the quantitative results in copies/mL reported in EQA schemes from April 2020 to June 2023 for each of the 51 samples positive for SARS-CoV-2 RNA with the respective SARS-CoV-2 variants (see Supplementary Table 1). The virus variants shown in Fig. 8 have the same color coding as in Fig. 7.

In detail, Fig. 8a presents the success rates, i.e. the rates of correct quantitative results reported (in percent). Shown are the rates of correct quantitative results for each data point based on the individually reported results in copies/mL (within a target range of +/- 1.0 $\log_{10}$) in the corresponding EQA sample and expressed as a percentage (Y axis). The target values on basis of robust average values in copies/mL (with a target value interval of +/- 1.0 $\log_{10}$ of the target value) obtained for each EQA sample are depicted (X axis). Even though the individual data points are based on only 25 to 44 reports of quantitative results each considering all target gene regions, it is noticeable that those points representing the various VOCs were in the majority of cases at an acceptable rate of correct positive results higher than 93 %. In contrast, the data points for the non-VOC virus (data points coded in red) represented predominantly lower rates of correct positive results (71 %-93 %). This also occurred at SARS-CoV-2 RNA loads higher than 1 000 000 copies/mL (robust average).

The success rates of correct quantitative results shown for the individual samples in Fig. 8a are assigned to the individual EQA schemes between April 2020 and June 2023 in ascending chronological order in

Fig. 8b, whereby the color coding for the virus variants is retained. It was found that a considerable number of the data points with reduced rates of correct quantitative results (71 %-93 %) were for the EQA samples with non-VOC viruses (highlighted in red) tested in the first three EQA schemes in April, June, and November 2020.

As mentioned above, the reduced rates of correct quantitative results were not related to particularly low SARS-CoV-2 RNA loads when considering the 12 samples with the non-VOC viruses during this period. Three samples are representative of this observation (see also Supplementary Table 1): (i) EQA term April 2020: Sample no. 340059 with a rate of 70.5 % correct quantitative results and a high virus RNA load of 17 071 604 copies/mL (robust average); (ii) EQA term June 2020: Sample no. 340066 with a rate of 95.5 % correct quantitative results and a low virus RNA load of 2 943 copies/mL (robust average); (iii) EQA term November 2020: Sample no. 340079 with a rate of 80.5 % correct quantitative results and an intermediate virus RNA load of 204 795 copies/mL (robust average).

From the EQA term in March 2021 to June 2023 term, the 39 subsequently analyzed SARS-CoV-2 positive samples showed significantly improved rates of correct positive results (Fig. 8b). 25 of the 39 samples had rates of correct positive results between > 95 %-100 %. It should be emphasized that these mostly satisfactory results also and primarily concerned the samples tested with different VOCs/lineages, which were regularly updated in the EQA schemes from the March 2021 onwards depending on the epidemiological situation.

Considering the time-dependent observation, sample pairs containing the same virus variant from the same preparation with the same virus RNA load but with different coding at different EQA terms allowed to demonstrate quality improvements in quantitative virus RNA load determination over time.

One example in Fig. 8b (two red arrows) is the pair of samples with the same non-VOC variant (BetaCoV/Baden-Württemberg/1/ChVir1577/2020 isolate BER; diluted 1: 50 000), where the rate of correct quantitative results improved from 80.5 % [sample no. 340079 in November 2020 (204 795 copies/mL robust average)] to 100 % [sample no. 409–230314–02 in March 2023 (351 642 copies/mL robust average)].

A second example in Fig. 8b (two yellow arrows) for improvement in the rate of correct quantitative results was observed for a pair of samples containing the Beta VOC B.1.351 (diluted 1: 350 000). While sample no. 409030 in June 2022 (92 064 copies/mL robust average) showed a rate of correct quantitative results of 93.5 %, the twin sample no. 409–230314–01 in March 2023 (117 581 copies/mL robust average) revealed a rate of 100 %.

In the context of the longitudinal considerations, it is important to discuss to what extent an improvement in diagnostic results might be related to the introduction of the reference materials for SARS-CoV-2 in November 2020 (see Section 3.3), an WHO International Standard or other standards.

4. Discussion

This report summarizes the results of 13 EQA schemes for SARS-CoV-2 genome detection in the period from April 2020 to June 2023 and provides an overview of the proficiency of the participating laboratories, suitability of test materials and the performance of PCR/NAT based test systems used over the entire course of the COVID-19 pandemic. The procedure for establishing and conducting these new molecular EQA schemes was based on the experience gained from the INSTAND virological EQA schemes since 1988. SARS-CoV-2 posed enormous challenges for the newly established diagnostics. At the beginning of the pandemic, a rapidly increasing number of analyses for the genome detection of SARS-CoV-2 had to be carried out, with an initially limited number of commercial test systems and LDTs. The performance and reliability of these diagnostics was initially poorly understood.

At the beginning of the pandemic, it was necessary to **enhance the external quality control** to provide adequate assistance to the laboratories, so that they would be able to test the performance and reliability of their newly established test systems. In close coordination with the cooperation partners and experts in the here described national and international scientific network, the following focal points were considered for the successful introduction and permanent performance of **enhanced EQA schemes** for the genome detection of SARS-CoV-2.

Immediate provision of suitable virus right at the beginning of the pandemic outbreak*:* Due to the exceptionally fast provision of inactivated SARS-CoV-2 by the National Consultant Laboratory for Coronaviruses at Charité - University Medicine Berlin, it was possible to conduct a first international EQA scheme for the genome detection of SARS-CoV-2 on the basis of a pre-existing EQA program for the differentiation of coronaviruses in April 2020.

Short frequency of EQA schemes: A short time interval between the EQA schemes - every three months at the beginning of the pandemic – was essential so that the laboratories could monitor their test systems at the most up-to-date level. This approach, together with increasing the number of EQA samples (up to nine samples in June 2020), also ensured that the currently circulating virus variants could be represented in the EQA schemes and revealed whether newly emerging virus variants were sufficiently well recognized by the test systems. It should be emphasized that the frequent EQA schemes with their reports served to provide up-to-date post-marketing surveillance at all pandemic stages, not only for laboratories and manufacturers, but also for the public health service and the scientific community.

Concentration dependency of the testing for sensitivity check: Based on the experience from earlier EQA schemes, samples with viruses from 10-fold dilution series were recurrently used to check the sensitivity and linearity of the test systems and critical concentration range for detection when performing the various test systems for virus genome detection (see Supplementary Table 1).

Specificity check of the testing: From the outset, the EQA schemes included samples containing other coronaviruses, such as MERS or various human coronaviruses, in addition to non-infected cells as negative samples, to be able to check the specificity of the test procedure and the test systems (see Supplementary Table 1).

Revealed samples for orientation on test performance: The properties of EQA scheme samples were revealed *during an ongoing or at the start of an EQA scheme* so that laboratories could anchor their results to the predetermined properties at an early stage. In the first EQA scheme in April 2020, a total of three out of seven of the EQA samples were already revealed two weeks before the submission deadline for the results (see Table 1 and Section 3.1). The second EQA scheme in June 2020 contained one sample that was revealed and had a *dPCR assigned quantitative value* from the very beginning (see Section 3.2).

Reference materials for independent verification: Reference samples with assigned SARS-CoV-2 RNA loads were made available outside of the EQA scheme starting November 2020 to provide laboratories with quantitative guidance in interpreting results that were necessary for clinical decisions during the height of the pandemic (see Section 3.3; Vierbaum et al., 2022).

Note: The steps described to enhance the EQA schemes are marked in bold in the following discussion.

4.1. Reliability of qualitative results

When looking at the results of the first EQA program for qualitative genome detection of SARS-CoV-2, which was enhanced by immediate **provision of suitable virus** provided by the NRZ **at the beginning of the pandemic outbreak** in April 2020, it was shown that the success rates for both the SARS-CoV-2-positive samples from a 10-fold dilution series (**for concentration dependency and sensitivity check**) and the SARS-CoV-2-negative samples (**for specificity check**) were already satisfactory and ranged between 93.0 % and 99.7 % correct results (Table 1). Compared to other respiratory viruses, the respective first

INSTAND EQA schemes for virus genome detection showed correspondingly satisfactory success rates: for human coronaviruses including the MERS coronavirus between 96.9 % and 100 % (Zeichhardt and Donoso Mantke, 2015), for influenza A and B viruses rates between 90.9 % and 100 % (Zeichhardt and Habermehl, 2000) and for respiratory syncytial virus (RSV) rates between 97.8 % and 100 % (Zeichhardt and Donoso Mantke, 2014).

A comparison of the satisfying success rates for the qualitative genome detection of SARS-CoV-2 of the INSTAND EQA scheme of April 2020 with the success rates in EQA schemes from other EQA scheme providers showed similar results. In an international EQA study in April/May 2020, five SARS-CoV-2-positive and three SARS-CoV-2-negative samples in different concentrations were determined by 365 laboratories with success rates of 86.0 %–99.2 % correct qualitative results (Matheeussen et al., 2020). In the first EQA scheme in Austria in May 2020, 52 laboratories tested sample sets with three SARS-CoV-2-positive samples in different concentrations and one negative sample each (Görzer et al., 2020). In this study the success rate for the correct positive results was summarized for all three positive samples. Overall, 60 % of laboratories were able to correctly identify all three SARS-CoV-2-positive samples, while 100 % of laboratories correctly identified the negative sample. An overview of the EQA programs described by various authors for the genome detection of SARS-CoV-2 can be found in Buchta et al., 2023a. In the course of the pandemic, it has become increasingly clear that EQA is a good tool for monitoring the performance of testing (Buchta et al., 2022, 2023b; Mercer et al., 2022).

The results were appreciated by medical colleagues, but also attracted interest from groups far beyond the scientific community. People unfamiliar with the subject publicly expressed the suspicion that the 1.4 % false-positive results reported for sample 340062 in Table 1 generally reflect poor specificity of the test systems used and that genome detection for SARS-CoV-2 is unreliable. It should be clarified that the results evaluated as “false” for this sample in the sense of an overall medical evaluation of the results are not generally to be evaluated as “false-positive” and in no way allow a statement to be made about the specificity and sensitivity of the tests used in the examination of SARS-CoV-2-negative samples. A detailed examination of the results for this sample showed that clearly false-positive results were in fact only reported by 2 of the 284 German laboratories (Spannagl et al., 2020). The remaining falsely evaluated results were due to borderline results or incomplete data reporting.

Against this background that the misinterpretation of the EQA test results by persons outside the field led to incorrect general conclusions about the performance of the laboratories with the test systems used, the Robert Koch-Institut, as the German public health authority, showed great interest in an interpretation of the test specificity and sensitivity data obtained in the EQA scheme in order to be able to assess the testing capability of the laboratories in Germany. For details see Robert Koch-Institut, 2020b.

An ad hoc group of the Joint Diagnostic Commission of DVV and GfV in collaboration with the Consultant Laboratory for Coronaviruses, the Robert Koch-Institut and INSTAND drafted a general statement on the significance of EQA test results (Spannagl et al., 2020).

Based on the knowledge gained from the first EQA scheme in April 2020, it was essential to perform the subsequent EQA schemes in **short frequency**, so that the next EQA scheme was already run in June 2020. In the second INSTAND EQA scheme, SARS-CoV-2 RNA-positive samples of a **10-fold dilution series** were again used for enhancing the EQA scheme in order to **check the concentration dependencies and sensitivities** of the applied test systems, however, this time with higher virus dilutions than in the first EQA scheme (see Supplementary Table 2 and Table 1). Highly diluted samples containing non-VOC virus (see Supplementary Table 1) were used to find out at which SARS-CoV-2 RNA load the genome detection lead to difficulties in the laboratories performing the tests. At the two highest virus dilutions of 1: 5 000 000 (sample no. 340066; 2 943 copies/mL) and 1: 50 000 000 (no. 340071; 452 copies/mL) reduced success rates of 90.2 % and 59.9 %, respectively, of correct positive results occurred (Supplementary Table 2). For the lowest concentration sample, no. 340071, a differentiated analysis of the results by manufacturers and test names showed that the reduced numbers of correct positive results were not primarily attributable to distinct test systems. Rather, it was found that individual laboratories using different test systems had increasing difficulties with qualitative virus detection as the virus RNA load decreased (Zeichhardt and Kammel, 2020c). This demonstrates that the limits of detection of the test systems used were essentially within the range of 100 to 1 000 copies/mL as later required by the WHO (World Health Organization, 2021).

The majority of the applied SARS-CoV-2 virus genome detection tests were multi-target RT-qPCRs, in which the amplicons are generated in many cases in a one-step RT-dPCR reaction. Among the first published SARS-CoV-2 multiplex RT-PCR assays were tests using primer/probe sequences in gene regions such as the E gene and/or N gene for pan-sarbeco screening and other gene regions such as RdRP genes for SARS-CoV-2 confirmation (Corman et al., 2020a, 2020b; Centers for Disease Control and Prevention, 2020; FIND, 2020). In addition, sequences in the S gene, ORF1a and ORF1ab were used as target gene regions. In the EQA schemes of April and June 2020, results for the E gene region were reported most frequently, followed by RdRP gene, N gene and, from June 2020, also ORF1ab gene region (Fig. 1a and b). The number of participants reporting results for the S gene and ORF1a was lower. The rates of reported correct positive results were satisfactory for all target gene regions. Only highly diluted samples with SARS-CoV-2 RNA loads of 2 943 copies/mL or most clearly at 452 copies/mL showed lower success rates in some cases, indicating that the detection limit was reached at SARS-CoV-2 RNA load of 452 copies/mL for some of the test systems or test runs (Fig. 1b).

When the rates for correct positive results were differentiated according to LDTs and commercial tests, the LDTs show equally good or even slightly better success rates than the commercial test systems (Fig. 2a and b). It was remarkable that in the first EQA scheme in April 2020, 14 % of the reported results came from LDTs and 86 % from commercial test systems (see Section 3.1.1). In contrast, in June 2023 the proportion of results based on LDT analyses was strongly reduced. Of the 2 128 results for the four SARS-CoV-2-positive samples (related to all test systems and gene regions examined), only 2.1 % (44/2 128 results) were performed using LDTs.

This reflected the importance of LDTs for the development of detection systems as quickly as possible at the beginning of an epidemic or pandemic to detect a newly emerging pathogen. Later in the pandemic, laboratories certainly increasingly replaced LDTs by commercial test systems, as the efforts required for continued performance monitoring of LDTs were obviously significantly higher for a laboratory than for the verification of a commercial test system.

Looking in the longitudinal view at the concentration dependency of testing, the success rates for correct positive results for all 52 SARS-CoV-2-positive samples plotted against the respective SARS-CoV-2 RNA load (robust average values in copies/mL) (see Fig. 7, see also Section 3.4.1), showed a concentration dependency with decreasing success rates for correct positive results at SARS-CoV-2 RNA loads below 100 000 copies/mL. The qualitative SARS-CoV-2 genome detection was already sufficient (90 % correct positive results) or better at the beginning of the pandemic, provided that the virus RNA loads were not below 10 000 copies/mL. In contrast, virus RNA loads above 100 000 copies/mL consistently showed success rates of over 95 %. It should be emphasized that this applies to non-VOC viruses as well as to the various VOCs tested (Fig. 7).

4.2. Assessment of the strongly varying Ct/Cq values

When looking at the Ct/Cq values reported together with the qualitative results, an extremely heterogeneous and complex picture emerged

across all EQA scheme terms. Very large variations in Ct/Cq values of 10 cycle thresholds and more were observed across all levels of analysis. This was evident for the results for (i) the respective *"target gene-specific Ct/Cq medians* as shown in Table 2 and Supplementary Table 3, (ii) the *"assay-specific Ct/Cq medians"* as shown in Fig. 3a and b and Supplementary Figure 1 for the N gene region and (iii) the individual Ct/Cq values reported by a single group of participants who performed analyses with a single test system from one manufacturer again shown as an example for the N gene region (marked in Fig. 3a with a red arrow).

This wide variation in Ct/Cq values was actually not really surprising, as each and every PCR system is known to depend on a large number of different parameters, such as extraction procedure, selection of target gene sequences, one-step or two-step reactions, temperature profiles, number and length of cycles, primers and probes, etc. This is also evident in other recent INSTAND EQA schemes for qualitative genome detection of viruses such as hepatitis A virus, measles virus, rubella virus or West Nile virus (Zeichhardt and Kammel, 2023b, 2023c, 2023d, 2023e), where the reported Ct/Cq values for the respective virus detection also showed a wide range of variation depending on the test systems used and the selected target gene regions in the different laboratories. Additionally, a strong variation in Ct/Cq values in the genome detection of SARS-CoV-2 has been described elsewhere (Görzer et al., 2020, Buchta et al., 2021b; Evans et al., 2022).

In the context of Ct/Cq value variation, the element of enhancing the EQA schemes was **checking the concentration dependency of testing** by applying EQA samples from **defined 10-fold dilution series** of the same source virus was of great importance both for understanding the variation of the Ct/Cq median values in the 3 different categories (see Section 3.1.2) and for assessing the reliability of the diagnostics at the level of the individual laboratory.

Looking at first at the *"overall Ct/Cq medians"* as median values of the results of all participants, the 10-fold dilution levels were well reflected in both EQA schemes of April 2020 (Table 1) and June 2020 (see Supplementary Table 2). For sample pairs with a 10-fold difference in concentration between each other, a Ct/Cq difference value of 3.32 could theoretically be expected. Satisfactorily, nearly all of the Ct/Cq differences value for the respective dilution series in the two EQA schemes April and June 2020 were between 2.9 and 3.5. In the EQA scheme June 2020, the most highly diluted sample no. 340071 (1: 50 000 000; 452 copies/mL), however, revealed a difference to the 10-fold higher concentrated sample no. 340066 (1: 5 000 000; 2 943 copies/mL) of only 2.4 (see Supplementary Table 2). This was not unexpected, as some test systems were probably already at the edge or outside their linear measurement range at the low concentration of sample no. 340071.

A special feature for the laboratories participating in these EQA schemes was that they were able to check their own Ct/Cq values for the samples from the described 10-fold dilution series to verify the concentration-dependent linearity of their own test system in comparison to the results of the other EQA scheme participants with the respective tests applied. If a linearity reflecting the 10-fold dilution was not given, the implementation of the test system with the individual parameters should be urgently checked.

An important tool for providing laboratories with **orientation on their test performance** early in the pandemic was **revealing the properties of individual samples** during an ongoing or at the start of an EQA scheme. This was particularly important at the beginning of the pandemic, when laboratories were still very uncertain about the performance of their tests. For this reason, the characteristics of two SARS-COV-2 RNA-positive samples from a 10-fold dilution series were already revealed **during the ongoing EQA scheme** in April 2020. The interim evaluation as of 14 April 2020 showed a very strong variation of Ct/Cq values for the two SARS-CoV-2-positive samples 340059 and 340064 up to 24 for the analyzed target genes (for details see Zeichhardt and Kammel, 2020b).

In order to gain insights into the SARS-CoV-2 content of samples, one of the samples in the June 2020 EQA scheme was revealed with regard to its qualitative and quantitative properties at the start of the EQA scheme. This was a great advantage before reference materials of different virus RNA loads (see Sections 3.2 and 3.3; Vierbaum et al. 2022), a WHO International Standard or reference materials certified to DIN EN ISO 15194:2009 or other standard materials were available.

One example for revealing a sample for orientation on test performance **at the start of an EQA scheme** was the prequantified sample no. 340066, which was **assigned a value of 1 570 copies/mL by dPCR** (consensus value of five dPCR analyses from three independent NMIs). This sample was provided to the laboratories to enable them for the first time in the pandemic to orient their measured Ct/Cq values with a given virus RNA load during the ongoing EQA scheme (see Section 3.2, Table 4).

Right at the beginning of the pandemic, as the number of severe cases and fatalities drastically increased, there was an urgent need from the medical community for a way to make a medical statement about the infectivity of diagnostic samples obtained and the contagiousness of patients, despite the widely varying Ct/Cq values from the diagnostic results. Compared to systemic infections with HIV or hepatitis B and C viruses, where serum or plasma is obtained by phlebotomy, it had been recognized in SARS-CoV-2 infections that the pre-analytical influences on the diagnostic results could be very strong when nasopharyngeal swabs are taken as stated in Vierbaum et al. (2022). It is known that the diagnostic results of respiratory infections are highly dependent on the quality of the swab obtained (e.g. depending on the personnel, the collection method and the type of swab) (Lippi et al., 2020; see also ISO/TS, 5798:2022).

Given all this, it was desired to have a surrogate for assessing the risk of virus transmission from one patient to another based on determining Ct/Cq levels in a nasopharyngeal swab by a quantitative estimate of SARS-CoV-2 RNA load. The question of patient management, when a patient could be discharged from isolation, was of eminent importance as the many clinics were completely overloaded during the pandemic. To reduce the aforementioned pre-analytical influences on the diagnostic result, it was suggested to perform such determinations only with swabs from consecutive samplings by the same sender with knowledge of the time-dependent patient history to be able to make statements about the changes in virus RNA load.

To fulfill the requirements described above, the **reference materials** RM 1 and RM 2 were developed and established **for independent verification** in collaboration between the Robert Koch-Institut (RKI), the National Coronavirus Consultant Laboratory at the Institute of Virology of the Charité – University Medicine Berlin, INSTAND, GBD and members of the Joint Diagnostic Commission of DVV and GfV (see Section 3.3; Vierbaum et al., 2022). The insights gained from **the two enhanced EQA schemes** in April and June 2020 were the basis for determining the quantitative properties of the reference materials RM 1 and RM 2.

The concentrations of the two reference materials RM 1 and RM 2 for the SARS-CoV-2 RNA load were set at virus RNA loads of $\sim 10^6$ and $\sim 10^7$ copies/mL, respectively. It was assumed that the probability of transmission was less than 20 % for virus RNA loads below $\sim 10^6$ copies/mL (see Section 3.3; cited in Vierbaum et al., 2022). With a virus RNA load of more than $\sim 10^7$ copies/mL, it was expected that SARS-CoV-2 is very likely to be transmitted to other patients.

Figs. 5 and 6 impressively show that the vast majority of laboratories that reported their Ct/Cq values to INSTAND recognized the 10-fold gradation of the concentration of the two reference materials RM 1 and RM 2 reflected by an average Ct/Cq difference value between RM 1 and RM 2 of 3.16 +/- 0.93 (prediction interval, 95 % probability). Parallel shifted lines for the Ct/Cq values were found for all investigated gene regions, but again with Ct/Cq value deviations from laboratory to laboratory of 10 values and more. Despite these deviations, the positive conclusion from the use of the two reference materials is that this approach has made it possible to anchor Ct/Cq values to a given viral

RNA load in copies/mL. However, this requires that each laboratory determines this correlation between Ct/Cq values and viral RNA load individually for the test system used for each of the targeted gene regions. It should be noted that for different laboratories there is no general validity for the relationship found between Ct/Cq and a given virus RNA load, even when using the same test system targeting the same gene region. Conversely, it must be clearly stated that it is not appropriate to choose a single Ct/Cq value - such as the value of 30 - to make a statement about the transmissibility of SARS-CoV-2. In this context, the RKI provided revised "Instructions for testing patients for SARS-CoV-2 / Hinweise zur Testung von Patientinnen und Patienten auf SARS-CoV-2" on its online portal (Robert Koch-Institut, 2023). With regard to the interlaboratory variance of Ct/Cq values it is proposed that "… a conversion of Ct/Cq values into virus RNA loads (RNA copies per sample volume) by calibration using a standardized viral RNA preparation" such as the above-described reference materials should be aimed for.

For an extended use of reference materials for test calibration, at least three samples with different concentrations in 10-fold gradation should be included in a set of reference materials in the future, so that the linearity of the test systems can be better checked independently of the EQA schemes. A prerequisite would be that the set of reference materials is based on defined viral RNA load values assigned by dPCR.

4.3. Quantitative consideration

Some laboratories already reported quantitative results in copies/mL in the EQA scheme April 2020 and thus before the introduction of the described reference materials RM 1 and RM 2 and the WHO International Standard for SARS-CoV-2 RNA (NIBSC, 2021). The number of laboratories reporting quantitative results in the EQA schemes increased from 21 laboratories in April 2020 to 32 laboratories in June 2020 and declined to 14 laboratories in June 2023 (Zeichhardt and Kammel, 2020a, 2020c, 2023a).

As shown in Fig. 4 for the first two EQA schemes in April and June 2020, the quantitative results reported by the laboratories for the respective sample sets from the independent **ten-fold dilution** series showed very good concentration gradations and sufficient **sensitivities**. However, for each of the positive samples, huge variations were observed between the laboratories for the quantitative determinations. The quantitative results in copies/mL differed from laboratory to laboratory by factors between approx. 600 and approx. 1 400 for the individual samples in April 2020 (Fig. 4a). In both EQA terms, it was noticeable that the results obtained by dPCR (green lines in Fig. 4a and b) were (i) rather near to the respective red line, which marks the robust average values as the "consensus line" from all quantitative results reported by all laboratories and (ii) quite close to each other. In June 2020, the dPCR results of the individual laboratories were only separated by factors ranging from about 10 to 50 (Fig. 4b).

To determine the "rates of correct quantitative results", a target value interval of +/- 1.0 $\log_{10}$ around the previously mentioned consensus value, i.e. the robust average value for all reported results in copies/mL, was used. For this term, the expression "success rates of results in the target value interval +/- 1.0 $\log_{10}$ of the target value (number of correct results per total number of reported results)" was also used in this review (see Table 3). Such an evaluation criterion with a quantitative consensus value and a specified range of variation was based on the Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen ("Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations – Rili-BAEK", Bundesärztekammer, 2023; Ahmad-Nejad et al., 2024). A target value interval of +/- 1.0 $\log_{10}$ around the consensus value is regularly used for a number of quantitative INSTAND EQA schemes for virus genome detection, e.g. for varicella zoster virus.

An analysis of the robust average values for the 34 quantitative results reported for each of the four positive samples in April 2020, showed that the success rates for the various samples were only between 70.6 % and 76.5 % (see Section 3.1.3, Table 3, Fig. 8b). These rather low success rates may have been due to the lack of a WHO International Standard for SARS-CoV-2 RNA or other accepted reference materials. It can be assumed that some of the laboratories initially used different calibration materials, such as mostly synthetic genomic RNA fragments of SARS-CoV-2 from their own laboratory or from different manufacturers.

In the second EQA scheme in June 2020, in which only two of four samples from another 10-fold dilution series were evaluated (see Section 3.2.3, Table 5, Fig. 8b), 44 quantitative results were submitted for each of the samples. The quantitative results expressed as robust average values improved with success rates of 90.9 % and 97.7 % for the two evaluated samples.

The progression of the success rates was shown by the longitudinal analysis of the reported quantitative results of all 13 EQA schemes between April 2020 and June 2023 (see Section 3.4, Fig. 8a). Altogether 51 SARS-CoV-2 RNA-positive samples consisting of different SARS-CoV-2 virus isolates and circulating variants were analyzed during this period (see Supplementary Table 1). The number of reported quantitative results varied between 25 and 44 for the corresponding samples.

Fig. 8b clearly shows that from the EQA scheme November 2020 onwards, the success rates for the quantitative results improved continuously, regardless of which SARS-CoV-2 virus isolates including VOCs were used in the samples. This was the case for most of the test systems used, regardless of the targeted gene regions.

As of the March 2021 scheme, the success rates for all samples tested - with the exception of four samples - were between 93 % and 100 %. Between April 2020 and June 2023 the reduced success rates of less than 90 % (in some cases only between 70 % and 80 %), which were primarily observed for non-VOC viruses, were striking. This was mainly seen in the three early EQA schemes between April and November 2020, and can only partly be explained by the comparatively low virus RNA loads of 100 000 copies/mL (Fig. 8b). Since samples with virus RNA loads of between approx. 2 000 000 and 17 000 000 copies/mL also resulted in success rates of less than 80 % in this early phase (Fig. 8a), it can be assumed that the lack of sufficient expertise in the laboratories and the lack of suitable reference materials were responsible for this. In addition, as mentioned above, the described reference materials RM 1 and RM 2 and the WHO International Standard for SARS-CoV-2 RNA were only available from November 2020 (see black arrow in Fig. 8b).

The statement that the rates of correct positive results improved over time was further supported by the use of EQA samples that came from one and the same preparation and were therefore fully identical in terms of their source virus in dilution and matrix. Such identical samples were used with different coding on two different EQA dates. An example of this were the identical samples 340079 (November 2020) and sample no. 409–230314–02 (March 2023) (see Supplementary Table 1) with the non-VOC BetaCoV/Baden-Württemberg/1/ChVir1577/2020 isolate BER, which showed a strong improvement in the rates of correct positive results from 80.5 % to 100 % (Fig. 8b, see two red arrows). A second example was the two identical samples 409030 (June 2022) and sample no. 409–230314–01 (March 2023) with the VOC BetaCoV/South Africa/ChVir22131/2020. Here, the rates of correct positive results increased from 93.5 % to 100 % (Fig. 8b, see two yellow arrows).

In summary, the **enhanced EQA schemes** described here, consisting of a combination of high EQA scheme frequency at the beginning of the pandemic together with the use of individual revealed samples and independent reference materials with dPCR assigned quantitative values, have proven very successful in the course of the SARS-CoV-2 pandemic in order to obtain an accurate picture of both the proficiency in the laboratories and the performance of the test systems used. During the preparation of the manuscript, a special series of publications on the importance of EQA systems was published (Buchta et al., 2025a, 2025b, 2025c, 2025d, 2025e, 2025f). The holistic approach of EQA schemes described there supports the concept of enhanced EQA schemes that we have established. The importance of expert networks for consultation,

guidance, decision-making and recommendation processes to improve the quality of laboratory diagnostics, especially in the event of a pandemic, is emphasized in our study. It is stated that it is essential for the continuous improvement of the quality of laboratory diagnostics and for the immediate response capability in the event of a new pandemic that international networks based on trust and collaboration between experts from hospitals, laboratories (including blood donation organizations), industry, scientific societies, governmental and regulatory institutions, and national and international reference institutes (including European reference laboratories) work together (Mercer et al., 2022).

5. Conclusion

The lessons learned from the enhanced EQA schemes at the beginning of the COVID-19 crisis can be taken as a blueprint for preparedness for future pandemics:

a) A prerequisite to set up "enhanced EQA schemes" is to provide appropriate target sequences for primers and probes for the pathogen as basis for diagnostic molecular detection (test systems) as quickly as possible. Therefore, LDTs can be of great importance because they can be set up in a few days, if reagents are rapidly available. The test and work protocols should be freely accessible via the internet.
b) National and international networks of experts are needed to review and optimize the protocols, define target gene regions, collect, exchange and analyze clinical materials from patients, discuss the evaluation of test results and recommend which further measures are needed.
c) Immediate provision of a suitable pathogen right at the beginning of the pandemic outbreak is necessary for EQA sample production. The source pathogen used for the preparation of EQA samples may have to be inactivated according to its risk group to ensure national and international shipment.
d) EQA schemes have to be established quickly and must include adequate positive and negative samples, dilution series of the pathogen, currently circulating variants of the pathogen, sensitivity and specificity testing with genomically closely related pathogens to test the proficiency of the laboratories and performance of the applied test systems.
e) Inclusion of revealed EQA samples with assigned quantitative values of the pathogen during the ongoing EQA schemes gives laboratories guidance to check the performance of a defined test system in their laboratory and to compare it with other laboratories with the same or other test systems of other EQA participating laboratories.
f) The establishment of independent reference materials with assigned quantitative values by digital PCR (dPCR), preferably with the support of National Metrology institutes (NMIs), should be carried out rapidly.

This approach of “enhanced EQA schemes” serves to continuously ensure and improve the quality of diagnostics and is of central importance for public health, especially during a pandemic.

Funding

INSTAND e.V. was funded by the German Federal Ministry of Health with regard to the reference materials described in Section 3.3 (Vierbaum et al., 2022) on the basis of Chapter 1503 Title 68403 for the 2020 financial year.

CRediT authorship contribution statement

Martin Kammel: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hans-Peter Grunert:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Anika Zimmermann:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Annemarie Martin:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Vanessa Lindig:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Samuel Samuleit:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Ulf Dühring:** Writing – review & editing, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Mechthild Adams-Bagusche:** Writing – review & editing, Validation, Methodology, Investigation. **Dirk Sander:** Writing – review & editing, Validation, Methodology, Investigation. **Hannah Zeichhardt:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **Christian Drosten:** Writing – review & editing, Resources, Methodology, Investigation, Conceptualization. **Victor M. Corman:** Writing – review & editing, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Sandra Ciesek:** Writing – review & editing, Investigation. **Holger F. Rabenau:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Martin Obermeier:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Robert Ehret:** Validation, Investigation, Formal analysis. **Rolf Kaiser:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Jim Huggett:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Denise O’Sullivan:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Peter M. Vallone:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Megan H. Cleveland:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Samreen Falak:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Andreas Kummrow:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Esmeralda Valiente:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Rainer Macdonald:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Mojca Milavec:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Sabine Goseberg:** Writing – review & editing, Software, Data curation. **Silke Kappler:** Writing – review & editing, Software, Data curation. **Nathalie Weiss:** Writing – review & editing, Validation, Software, Formal analysis, Data curation. **Laura Vierbaum:** Writing – review & editing, Validation, Software, Formal analysis, Data curation. **Patricia Kaiser:** Writing – review & editing, Software, Formal analysis, Data curation. **Ingo Schellenberg:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Heinz Zeichhardt:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Acknowledgements

We gratefully acknowledge the contributions of the INSTAND Expert Laboratories:

Bernhard-Nocht-Institut, Nationales Referenzzentrum für tropische Infektionserreger, Abteilung für Virologie, WHO Collaborating Centre for Arbovirus and Haemorrhagic Fever Reference and Research, Hamburg: Prof. Dr. S. Günther, Prof. Dr. Dr. J. Schmidt-Chanasit, PD Dr. P. Emmerich

Charité - Universitätsmedizin Berlin, Campus Mitte, Institut für Virologie, Nationales Konsiliarlaboratorium für Coronaviren: Prof. Dr.

C. Drosten, Dr. V. M. Corman, Dr. D. Niemeyer;

Medizinische Hochschule Hannover, Institut für Virologie, Nationales Konsiliarlaboratorium für Adenoviren: Prof. Dr. T. Schulz, PD Dr. A. Heim, Dr. W. Puppe, Dr. C. Schmitt

National Measurement Laboratory for Chemical and Bio-Measurement, LGC, Teddington, UK: Dr. J. Huggett, Dr. D. OSullivan

Medizinisches Infektiologiezentrum, Berlin: Dr. M. Obermeier, Dr. R. Ehret

Philipps Universität Marburg, Institut für Virologie, Nationales Konsiliarlaboratorium für Filoviren: Prof. Dr. S. Becker, Dr. Ch. Keller, Dr. M. Eickmann

Robert Koch-Institut, Abt. Infektionskrankheiten, FG 17 Influenzaviren und weitere Viren des Respirationstraktes, Nationales Referenzzentrum für Influenza, Nationales Konsiliarlaboratorium für Respiratorische Syncytialviren (RSV), Parainfluenzaviren, Metapneumoviren, Berlin: Dr. R. Dürrwald, Dr. B. Biere, Dr. J. Reiche

Robert Koch-Institut, ZBS 1, Zentrum für biologische Gefahren und spezielle Pathogene/Hochpathogene Viren: Prof. Dr. A. Nitsche, Dr. J, Michel

Uniklinik Köln, Institut für Virologie, Nationales Referenzzentrum für Papillom- und Polyomaviren: Prof. Dr. F. Klein, Prof. Dr. U. Wieland, Dr. S. Silling, Dr. R. Kaiser, Dr. E. Heger, Dr. E. Knops

Universitätsklinikum Bonn, Institut für Virologie: Prof. Dr. H. Streeck, Prof. Dr. A.-M. Eis-Hübinger

Universitätsklinikum Düsseldorf, Institut für Virologie, Nationales Referenzzentrum für Hepatitis-C-Viren: Prof. Dr. J. Timm, Prof. Dr. O. Adams, Dr. N. Lübke

Universitätsklinikum Frankfurt, Institut für Medizinische Virologie, Frankfurt/Main: Prof. Dr. S. Ciesek, Prof. Dr. H. F. Rabenau, Prof. Dr. A. Berger

Universitätsklinikum Freiburg, Institut für Virologie, Nationales Konsiliarlaboratorium für HSV und VZV: Prof. Dr. H. Hengel, Dr. D. Huzly, Prof. Dr. M. Panning, Prof. Dr. M. Schwemmle

NIST Disclaimer

Points of view in this document are those of the co-authors (Vallone and Cleveland) and do not necessarily represent the official position or policies of the U.S. Department of Commerce. Certain commercial software, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ijmm.2025.151656.

References

Ahmad-Nejad, P., Bauersfeld, W., Baum, H., Behre, H.M., Burkhardt, R., Cassens, U., Ceglarek, U., Christmann, M., Cremers, J.-F., Diedrich, S., Döring, S., Gässler, N., Haase, Gerhard, Haselmann, V., Hofmann, J., Holdenrieder, S., Hübner, M.P., Hunfeld, K.-P., Huzly, D., Kohlschmidt, N., Köhn, F.-M., Kornak, U., Kreuzer, K.-A., Kunz, J., Lackner, K., Niendorf, S., Peetz, D., Petersmann, A., Pick, K.-H., Rabenau, H.F., Sack, U., Schächterle, C., Schaffer, S., Schneider, S., Schuppe, H.C., Seidl, C., Tönnies, H., Uhr, M., Ullmann, K., Volkmann, M., Weiss, N., Wellinghausen, N., Zeichhardt, H., App, U., Auch, D., Barion, J., Hiester, P., Kaiser, P., Klouche, M., Kolanowski-Albrecht, M., Macdonald, R., Malms-Fleschenberg, W., Michelsen, A., Schellenberg, I., Schiffner, R., Spannagl, M., Stosch, R., van Diepen, L., Wettmarshausen, S., Ziesing, S., Knabbe, C., Schoerner, C., Kliesch, S., Nauck, M., 2024. Revision of the "Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations – Rili-BAEK. J. Lab. Med. 48 (6), 263–306. https://doi.org/10.1515/labmed-2024-0131.

Arons, M.M., Hatfield, K.M., Reddy, S.C., Kimball, A., James, A., Jacobs, J.R., Taylor, J., Spicer, K., Bardossy, A., Oakley, L.P., Tanwar, S., Dyal, J.W., Harney, J., Chisty, Z., Bell, J.M., Methner, M., Paul, P., Carlson, C.M., Mc Laughlin, H.P., Thornburg, N., Tong, S., Tamin, A., Tao, Y., Uehara, A., Harcourt, J., Clark, S., Brostrom-Smith, C., Page, L.C., Kay, M., Lewis, J., Montgomery, P., Stone, N.D., Clark, T.A., Honein, M. A., Duchin, J.S., Jermigan, J.A., 2020. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N. Engl. J. Med. 382 (22), 2081–2090. https://doi.org/10.1056/NEJMoa2008457.

Bhadoria, P., Gupta, G., Agarwal, A., 2021. Viral Pandemics in the Past Two Decades: An Overview. J. Fam. Med. Prim. Care 10 (8), 2745–2750. https://doi.org/10.4103/jfmpc.jfmpc_2071_20.

Buchta, C., Benka, B., Delatour, V., Faé, I., Griesmacher, A., Hellbert, K., Huggett, J., Kaiser, P., Kammel, M., Kessler, A., Kessler, H., Müller, D., Rosendahl, J., Scheiblauer, H., Schweiger, C., Zeichhardt, H., Cobbaert, C., 2025f. Reference, calibration and referral laboratories – a look at current European provisions and beyond. Clin. Chem. Lab. Med. 63 (4), 656–669. https://doi.org/10.1515/cclm-2024-1066.

Buchta, C., Camp, J.V., Jovanovic, J., Chiba, P., Puchhammer-Stöckl, E., Mayerhofer, M., Plicka, H., Lercher, A., Popa, A.M., Endler, L., Bergthaler, A., Huf, W., Benka, B., Delatour, V., Müller, M.M., Griesmacher, A., Aberle, S.W., Görzer, I., 2021a. The versatility of external quality assessment for the surveillance of laboratory and in vitro diagnostic performance: SARS-CoV-2 viral genome detection in Austria. Clin. Chem. Lab. Med. 59 (10), 1735–1744. https://doi.org/10.1515/cclm-2021-0604.

Buchta, C., Camp, J.V., Jovanovic, J., Radler, U., Puchhammer-Stöckl, E., Benka, B., Huf, W., Müller, M.M., Griesmacher, A., Aberle, S.W., Görzer, I., 2021b. A look at the precision, sensitivity and specificity of SARS-CoV-2 RT-PCR assays through a dedicated external quality assessment round. Clin. Chem. Lab. Med. 60 (2), e34–e37. https://doi.org/10.1515/cclm-2021-1004.

Buchta, C., De la Salle, B., Marrington, R., Aburto Almonacid, A., Albarède, S., Badrick, T., Bullock, D., Cobbaert, C.M., Coucke, W., Delatour, V., Faria, A.P., Geilenkeuser, W.J., Griesmacher, A., Huggett, J.F., Ianovska, V., Kammel, M., Kessler, A., Körmöczi, G.F., Meijer, P., Miranda, A., Patel, D., Pezzati, P., Sandberg, S., Schennach, H., Schweiger, C.R., Schwenoha, K., Spannagl, M., Sung, H., Thelen, M., Weykamp, C., Zeichhardt, H., Restelli, V., Perrone, L.A., 2025e. Behind the scenes of EQA - characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part V – Benefits for stakeholders other than participants, 2025 Jan 6 Clin. Chem. Lab. Med.. https://doi.org/10.1515/cclm-2024-1293.

Buchta, C., De la Salle, B., Marrington, R., Albarède, S., Badrick, T., Bicker, W., Bietenbeck, A., Bullock, D., Delatour, V., Dequeker, E., Flasch, M., Geilenkeuser, W. J., van Hellemond, J., Huggett, J.F., Ianovska, V., Kaiser, P., Kammel, M., Kessler, A., Laudus, N., Luppa, P.B., Morandi, P.A., O'Connor, G., Panteghini, M., Pezzati, P., Rosendahl, J., Sandberg, S., Scheiblauer, H., Skitek, M., Spannagl, M., Stavelin, A., Thelen, M., Unterberger, U., Zeichhardt, H., Restelli, V., Perrone, L.A., 2025d. Behind the scenes of EQA-characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part IV – Benefits for participant laboratories, 2025 Jan 6 Clin. Chem. Lab. Med.. https://doi.org/10.1515/cclm-2024-1292.

Buchta, C., Marrington, R., De la Salle, B., Albarède, S., Albe, X., Badrick, T., Berghäll, H., Bullock, D., Cobbaert, C., Coucke, W., Delatour, V., Geilenkeuser, W., Griesmacher, A., Henriksen, G., Huggett, J., Juhos, I., Kammel, M., Luppa, P., Meijer, P., Pelanti, J., Pezzati, P., Sandberg, S., Spannagl, M., Thelen, M., Thomas, A., Zeichhardt, H., Restelli, V., Perrone, L., 2025c. Behind the scenes of EQA – characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part III – EQA samples, 2025 Jan 6 Clin. Chem. Lab. Med.. https://doi.org/10.1515/cclm-2024-1291.

Buchta, C., Marrington, R., De la Salle, B., Albarède, S., Badrick, T., Berghäll, H., Bullock, D., Coucke, W., Delatour, V., Geilenkeuser, W., Griesmacher, A., Henriksen, G., Huggett, J., Luppa, P., Pelanti, J., Pezzati, P., Sandberg, S., Spannagl, M., Thelen, M., Restelli, V., Perrone, L., 2025b. Behind the scenes of EQA – characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part II – EQA cycles, 2025 Jan 6 Clin. Chem. Lab. Med.. https://doi.org/10.1515/cclm-2024-1290.

Buchta, C., Marrington, R., De la Salle, B., Albarède, S., Badrick, T., Bietenbeck, A., Bullock, D., Cadamuro, J., Delatour, V., Dusinovic, E., Geilenkeuser, W., Gidske, G., Griesmacher, A., Haliassos, A., Holzhauser, D., Huggett, J., Karathanos, S., Pezzati, P., Sandberg, S., Sarkar, A., Solsvik, A., Spannagl, M., Thelen, M., Restelli, V., Perrone, L., 2025a. Behind the scenes of EQA – characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part I – EQA in general and EQA programs in particular, 2025 Jan 6 Clin. Chem. Lab. Med.. https://doi.org/10.1515/cclm-2024-1289.

Buchta, C., Müller, M.M., Griesmacher, A., 2022. The importance of external quality assessment data in evaluating SARS-CoV-2 virus genome detection assays. Lancet Microbe 3 (3), e168. https://doi.org/10.1016/s2666-5247(22)00003-9.

Buchta, C., Zeichhardt, H., Aberle, S.W., Camp, J.V., Görzer, I., Weseslindtner, L., Puchhammer-Stöckl, E., Huf, W., Benka, B., Allerberg, F., Mielke, M., Griesmacher, A., Müller, M.M., Schellenberg, I., Kammel, M., 2023a. Design of external quality assessment schemes and definition of the roles of their providers in future epidemics. Lancet Microbe 4 (7), e552–e562. https://doi.org/10.1016/S2666-5247(23)00072-1.

Buchta, C., Zeichhardt, H., Griesmacher, A., Schellenberg, I., Kammel, M., 2023b. Ignoring SARS-CoV-2 testing performance during COVID-19. Lancet Microbe 4 (5), e296. https://doi.org/10.1016/S2666-5247(23)00030-7.

Bundesärztekammer, 2023. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen ("Guideline of the German Medical Association", in German language), as of 30 May 2023, anounced in: Deutsches Ärzteblatt 120 (21-22):A-994 / B-858. ⟨https://doi.org/10.3238/arztebl.2023.rili_baek_QS_Labor⟩.

Centers for Disease Control and Prevention, 2020. CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. ⟨https://www.fda.gov/media/134922/download⟩ (assessed 18 March 2024).

Chinese Centre for Disease Control and Prevention. Specific primers and probes for detection of 2019 novel coronavirus [posted on 24 January 2020] ⟨https://www.who.int/docs/default-source/coronaviruse/whoinhouseassays.pdf?sfvrsn=de3a76aa_2⟩ (assessed 03 April 2025).

Corman, V.M., Bleicker, T., Brünink, S., Drosten, C., Landt, O., Koopmans, M., Zambon, M., 2020a. Diagnostic detection of 2019-nCoV by real-time RT-PCR, as of 17.01.2020. Available at: ⟨https://www.who.int/docs/default-source/coronaviruse/protocol-v2-1.pdf⟩ (assessed 03 April 2025).

Corman, V.M., Haage, V.C., Bleicker, T., Schmidt, M.L., Mühlemann, B., Zuchowski, M., Jo, W.K., Tscheak, P., Möncke-Buchner, E., Müller, M.A., Krumbholz, A., Drexler, J. F., Drosten, C., 2021. Comparison of seven commercial SARS-CoV-2 rapid point-of-care tests: a single-centre laboratory evaluation study. Lancet Microbe 2 (7). https://doi.org/10.1016/S2666-5247(21)00056-2.

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., Mulders, D.G.J.C., Haagmans, B.L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.L., Ellis, J., Zambon, M., Peiris, M., Goossens, H., Reusken, C., Koopmans, M.P.G., Drosten, C., 2020b. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eur. Surveill. 25 (3), 23–30. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045.

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5, 536–544. https://doi.org/10.1038/s41564-020-0695-z.

Cucinotta, D., Vanelli, M., 2020. WHO Declares COVID-19 a Pandemic. Acta Biomed. 91 (1), 157–160. https://doi.org/10.23750/abm.v91i1.9397.

DIN EN ISO 15194:2009. In-vitro-Diagnostika - Messung von Größen in Proben biologischen Ursprungs - Anforderungen an zertifizierte Referenzmaterialien und an den Inhalt der Begleitdokumentation (ISO 15194:2009) / In vitro diagnostic medical devices - Measurement of quantities in samples of biological origin - Requirements for certified reference materials and the content of supporting documentation (ISO 15194:2009).

DIN EN ISO/IEC 17043:2010. Konformitätsbewertung - Allgemeine Anforderungen an Eignungsprüfungen (ISO/IEC 17043:2010) / Conformity assessment - General requirements for proficiency testing (ISO/IEC 17043:2010).

DIN ISO 13528:2020. Statistische Verfahren für Eignungsprüfungen durch Ringversuche (ISO 13528:2015, korrigierte Fassung 2016-10-15) / Statistical methods for use in proficiency testing by interlaboratory comparison (ISO 13528:2015, Corrected version 2016-10-15).

Edson, D.C., Casey, D.L., Harmer, S.E., Downes, F.P., 2020. Identification of SARS-CoV-2 in a Proficiency Testing Program. Am. J. Clin. Pathol. 1–4. https://doi.org/10.1093/ajcp/aqaa128.

Evans, D., Cowen, S., Kammel, M., O'Sullivan, D.M., Stewart, G., Grunert, H.-P., Moran-Gilad, J., Verwilt, J., In, J., Vandesompele, J., Harris, K., Hong, K.H., Storey, N., Hingley-Wilson, S., Dühring, U., Bae, Y.-K., Foy, C.A., Braybrook, J., Zeichhardt, H., Huggett, J.F., 2022. The Dangers of Using Cq to Quantify Nucleic Acid in Biological Samples: A Lesson From COVID-19. Clin. Chem. 68 (1), 153–162. https://doi.org/10.1093/clinchem/hvab219.

Falak, S., Macdonald, R., Busby, E.J., O'Sullivan, D.M., Milavec, M., Plauth, A., Kammel, M., Zeichhardt, H., Grunert, H.-P., Kummrow, A., Huggett, J.F., 2022. An assessment of the reproducibility of reverse transcription digital PCR quantification of HIV-1. Methods 201, 34–40. https://doi.org/10.1016/j.ymeth.2021.03.006.

Falak, S., O'Sullivan, D.M., Cleveland, M.H., Cowen, S., Busby, E.J., Devonshire, A.S., Valiente, E., Jones, G.M., Kammel, M., Milavec, M., Vierbaum, L., Schellenberg, I., Zeichhardt, H., Kummrow, A., Vallone, P.M., Macdonald, R., Huggett, J.F., 2025. The Application of Digital PCR as a Reference Measurement Procedure to Support the Accuracy of Quality Assurance for Infectious Disease Molecular Diagnostic Testing. Clin. Chem. 71 (3), 378–386. https://doi.org/10.1093/clinchem/hvae187.

FIND, 2020. FIND evaluations of SARS-CoV-2 molecular tests. ⟨https://www.finddx.org/covid-19/find-evaluations-of-sars-cov-2-assays/find-evaluations-of-sars-cov-2-molecular-tests/⟩ (accessed 18 March 2024).

Goerzer, I., Buchta, C., Chib, P., Benk, B., Camp, J.V., Holzmann, H., Puchhammer-Stöckl, E., Aberle, S.W., 2020. First results of a national external quality assessment scheme for the detection of SARS-CoV-2 genome sequences. J. Clin. Virol. 129, 1–6. https://doi.org/10.1016/j.jcv.2020.104537.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wie, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395 (10223), 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

Huggett, J.F., Cowen, S., Foy, C.A., 2015. Consideration for digital PCR as an accurate molecular diagnostic tool. Clin. Chem. 61 (1), 79–88. https://doi.org/10.1373/clinchem.2014.221366.

ISO/TS 5798:2022. In vitro diagnostic test systems — Requirements and recommendations for detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by nucleic acid amplification methods.

Kaur, H., Mukhopadhyay, L., Aggarwal, N., Gupta, N., Narayan, J., Vijay, N., Gupta, S., Rana, S., Kaur, J., Kumar, V., Singh, H., 2022. Inter-laboratory testing as a strategy for external quality assessment for qualitative detection of SARS-CoV-2 by real-time RT-PCR testing in India. Indian J. Med. Res. 155, 86–90. https://doi.org/10.4103/ijmr.ijmr_2433_21.

Lau, K.A., Kaufer, A., Gray, J., Theis, T., Rawlinson, W.D., 2022. Proficiency testing for SARS-CoV-2 in assuring the quality and overall performance in viral RNA detection in clinical and public health laboratories. Pathology 54 (4), 472–478. https://doi.org/10.1016/j.pathol.2022.01.006.

Lippi, G., Simundic, A.-M., Plebani, M., 2020. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). Clin. Chem. Lab. Med. 58 (7), 1070–1076. https://doi.org/10.1515/cclm-2020-0285.

Lu, X., Wang, L., Sakthivel, S.K., Whitaker, B., Murray, J., Kamili, S., Lynch, B., Malapati, L., Burke, S.,A., Harcourt, J., Tamin, A., Thornburg, N.J., Villanueva, J.M., Lindstrom, S., 2020. US CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome Coronavirus 2. Emerg. Infect. Dis. 26 (8), 1654–1665. https://doi.org/10.3201/eid2608.201246.

Matheeussen, V., Corman, V.M., Donoso Mantke, O., McCulloch, E., Lammens, C., Goossens, H., Niemeyer, D., Wallace, P.S., Klapper, P., Niesters, H.G.M., Drosten, C., Ieven, M., 2020. International external quality assessment for SARSCoV-2 molecular detection and survey on clinical laboratory preparedness during the COVID-19 pandemic, April/May 2020. Eur. Surveill. 25 (27), 1–5. https://doi.org/10.2807/1560-7917.ES.2020.25.27.2001223.

Mercer, T., Almond, N., Crone, M.A., Chain, P.S.G., Deshpande, A., Eveliegh, D., Freemont, P., Fuchs, S., Garlick, R., Huggett, J., Kammel, M., Li, P.-E., Milavec, M., Marlowe, E.M., O'Sullivan, D.M., Page, M., Pestano, G.A., Suliman, S., Simen, B., Sninsky, J.J., Sopchak, L., Tato, C.M., Vallone, P.M., Vandesompele, J., White, T.J., Zeichhardt, H., Salit, M., 2022. The Coronavirus Standards Working Group's roadmap for improved population testing. Nat. Biotechnol. 40, 1563–1568. https://doi.org/10.1038/s41587-022-01538-1.

Milavec, M., Pavšič, J., Bogožalec Košir, A., Jones, G.M., O'Sullivan, D.M., Alison, S., Devonshire, A.S., Van Heuverswyn, F., Karczmarczyk, M., Neeb, J., Plauth, A., Corbisier, P., Schimmel, H., Kummrow, A., Neukammer, J., Carole, A., Foy, C.A., Kammel, M., Grunert, H.-P., Zeichhardt, H., Huggett, J.F., 2022. The performance of human cytomegalovirus digital PCR reference measurement procedure in seven external quality assessment schemes over four years. Methods 201, 65–73. https://doi.org/10.1016/j.ymeth.2021.03.016.

NIBSC, 2021. WHO International Standard First WHO International Standard for SARS-CoV-2 RNA NIBSC code: 20/146 Instructions for use (Version 3.0, Dated 11/02/2021). ⟨https://nibsc.org/documents/ifu/20-146.pdf⟩.

Perera, R.A.P.M., Tso, E., Tsang, O.T.Y., Tsang, D.N.C., Fung, K., Leung, Y.W.Y., Chin, A. W.H., Chu, D.K.W., Cheng, S.M.S., Poon, L.L.M., Chuang, V.W.M., Peiris, M., 2020. SARS-CoV-2 Virus Culture and Subgenomic RNA for Respiratory Specimens from Patients with Mild Coronavirus Disease. Emerg. Infect. Dis. 26 (11), 2701–2704. https://doi.org/10.3201/eid2611.203219.

Robert Koch-Institut, 2020a. Website: Ergänzung zum Nationalen Pandemieplan – COVID-19 – neuartige Coronaviruserkrankung (4.3.2020). ⟨https://www.rki.de/DE/Themen/Infektionskrankheiten/Infektionskrankheiten-A-Z/C/COVID-19-Pandemie/Strategie/Ergaenzung_Pandemieplan_Covid.html?nn= 16911042⟩ (assessed 30 April 2025).

Robert Koch-Institut, 2020b. Bericht zur Optimierung der Laborkapazitäten zum direkten und indirekten Nachweis von SARS-CoV-2 im Rahmen der Steuerung von Maßnahmen, Berlin 7.7.2020. Bericht zur Optimierung der Laborkapazitäten zum direkten und indirekten Nachweis von SARS-CoV-2 im Rahmen der Steuerung von Maßnahmen (assessed 17 March 2025).

Robert Koch-Institut, 2023. Website: Hinweise zur Testung von Patientinnen und Patienten auf SARS-CoV-2 - Quantitative Bezugsproben. ⟨https://www.rki.de/DE/Themen/Infektionskrankheiten/Infektionskrankheiten-A-Z/C/COVID-19/Diagnostik-Hinweise.html⟩ (assessed 30 April 2025).

Spannagl, M., Schellenberg, I., Hunfeld, K.-P., Zeichhardt, H., Kammel, M., Rabenau, H., Kaiser, R., Obermeier, M., Schmidt-Chanasit, J., Mielke, M., Drosten, C., Corman, V., 2020. Aussagekraft und Bedeutung von Ringversuchen zum Nachweis von SARS-CoV-2. Hier: INSTAND-Ringversuch zum „Virusgenom-Nachweis - SARS-CoV-2" im April 2020: Beurteilung der PCR-Ergebnisse [Significance and importance of round robin tests for verification from SARS-CoV-2. Here: INSTAND round robin test for "virus genome detection - SARS-CoV-2" in April 2020: Evaluation of the PCR results] [Internet]. Düsseldorf; 2020 Sep [cited 2023 May 3]. Available from: ⟨https://www.instand-ev.de/fileadmin/uploads/Aktuelles/Konsentierte_Stellungnahme_zum_INSTAND_RV_340_SARS-CoV-2_NAT_April_2020_bt.pdf⟩ (assessed 18 March 2024).

Sung, H., Han, M.-G., Yoo, C.-K., Lee, S.-W., Chung, Y.-S., Park, J.-S., Kim, M.-N., Lee, H., Hong, K.H., Seong, M.-W., Lee, K., Chun, S., Lee, W.G., Kwon, G.-C., Min, W.-K., 2020. Nationwide External Quality Assessment of SARS-CoV-2 Molecular Testing, South Korea. Emerg. Inf. Dis. 26 (10), 2353–2360. https://doi.org/10.3201/eid2610.202551.

van Kampen, J.J.A., van de Vijver, D.A.M.C., Fraaij, P.L.A., Haagmans, B.L., Lamers, M. M., Okba, N., van den Akker, J.P.C., Endeman, H., Gommers, D.A.M.P.J., Cortnelissen, J.J., Hoek, R.A.S., van der Eerden, M.M., Hesselink, D.A., Metselaar, H. J., Verbon, A., de Steenwinkel, J.E.M., Aron, G.I., van Gorp, E.C.M., van Boheemen, S., Voermans, J.C., Boucher, C.A.B., Molenkamp, R., Koopmans, M.P.G., Geurtsvankessel, C., van der Eijk, A.A., 2021. Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). Nat. Commun. 12 (267), 1–6. https://doi.org/10.1038/s41467-020-20568-4.

Vierbaum, L., Wojtalewicz, N., Grunert, H.P., Lindig, V., Duehring, U., Drosten, C., Corman, V., Niemeyer, D., Ciesek, S., Rabenau, H.F., Berger, A., Obermeier, M., Nitsche, A., Michel, J., Mielke, M., Huggett, J., O'Sullivan, D., Busby, E., Cowen, S., Vallone, P.M., Cleveland, M.H., Falak, S., Kummrow, A., Keller, T., Schellenberg, I., Zeichhardt, H., Kammel, M., 2022. RNA reference materials with defined viral RNA loads of SARS-CoV-2 - A useful tool towards a better PCR assay harmonization. PLOS ONE 17 (1), e0262656. https://doi.org/10.1371/journal.pone.0262656.

Vierbaum, L., Wojtalewicz, N., Grunert, H.P., Zimmermann, A., Scholz, A., Goseberg, S., Kaiser, P., Duehring, U., Drosten, C., Corman, V., Niemeyer, D., Rabenau, H.F., Obermeier, M., Nitsche, A., Michel, J., Puyskens, A., Huggett, J.F., O'Sullivan, D.M., Busby, E., Cowen, S., Vallone, P.M., Cleveland, M.H., Falak, S., Kummrow, A., Schellenberg, I., Zeichhardt, H., Kammel, M., 2023. Results of German external

quality assessment schemes for SARS-CoV-2 antigen detection. Nat. Sci. Rep. 13, 13206. https://doi.org/10.1038/s41598-023-40330-2.

von Kleist, M., Ruehe, B., Oh, D.J., Nitsche, A., Haas, W., Stoliaroff-Pépin, A., Eckmanns, T., Abu Sin, M., van der Toorn, W., Jenny, M., Mielke, M., Herzog, C., Wieler, L.H., 2020. Abwaegung der Dauer von Quarantaene und Isolierung bei COVID-19. Epidemiologisches Bulletin RKI (39/2020), 3-11. Available from: ⟨https://www.rki.de/DE/Content/Infekt/EpidBull/epid_bull_form.html⟩ (assessed 18 March 2024).

Wang, Z., Chen, Y., Yang, J., Han, Y., Shi, J., Zhan, S., Peng, R., Li, R., Zhang, R., Li, J., Zhang, R., 2021. External Quality Assessment for Molecular Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Clinical Laboratories. J. Mol. Diagn. 23 (1), 19–28. https://doi.org/10.1016/j.jmoldx.2020.10.008.

Whale, A.S., Jones, G.M., Pavšič, J., Dreo, T., Redshaw, N., Akyürek, S., Akgöz, M., Divieto, C., Sassi, M.P., He, H.-J., Cole, K.D., Bae, Y.-K., Park, S.-R., Deprez, L., Corbisier, P., Garrigou, S., Taly, V., Larios, R., Cowen, S., O'Sullivan, D., Bushell, C. A., Goenaga-Infante, H., Foy, C.A., Woolford, A.J., Parkes, H., Huggett, J.F., Devonshire, A.S., 2018. Assessment of digital PCR as a primary reference measurement procedure to support advances in precision medicine. Clin. Chem. 64 (9), 1296–1307. https://doi.org/10.1373/clinchem.2017.285478.

Woelfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Mueller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmann, R., Zwirglmaier, K., Drosten, C., Wendtner, C., 2020a. Virological assessment of hospitalized patients with COVID-2019. Nature 581 (7809), 465–469. https://doi.org/10.1038/s41586-020-2196-x.

Woelfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Mueller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmann, R., Zwirglmaier, K., Drosten, C., Wendtner, C., 2020b. Virological assessment of hospitalized patients with COVID-2019. Nature, Corrections and amendments, 588 E35. ⟨https://doi.org/10.1038/s41586-020-2984-3⟩ Correction to: Nature ⟨https://doi.org/10.1038/s41586-020-2196-x⟩.

World Health Organization, 2021. Technical Specifications for Selection of Essential In Vitro Diagnostics for SARS-COV-2. WHO reference number: WHO/2019-nCoV/Essential_IVDs/2021.1. Website: ⟨https://www.who.int/publications/i/item/WHO-2019-nCoV-Essential_IVDs-2021.1⟩ (assessed 18 March 2024).

World Health Organization, 2024. Website: WHO COVID-19 dashboard / Number of COVID-19 cases reported to WHO - cumulative total and Number of COVID-19 deaths reported to WHO - cumulative total, 06 March 2024. ⟨https://covid19.who.int/⟩ (assessed 18 March 2024).

Zeichhardt, H., Habermehl, K.-O., 2000. INSTAND Ringversuch zum Nachweis von Virusgenom und Antigen, Influenza A und B Viren, Vorauswertung, February 2000. EQA Report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Donoso Mantke, O., 2014. INSTAND Report on EQAS Group No. 359 Virus Immunology, Respiratory Syncytial Virus, June 2014. EQA Report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Donoso Mantke, O., 2015. INSTAND Report on EQAS Group No. 340 Virus Genome Detection – Coronaviruses (incl. MERS Coronavirus) June 2015. EQA report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2020a. INSTAND Report on Extra EQAS Group No. 340 Virus Genome Detection SARS-CoV-2 April 2020. EQA report RV-Online: ⟨https://www.instand-ev.de/System/rv-files/340%20EN%20SARS-CoV-2%20Genom%20April%202020%2020200502j.pdf⟩ (assessed 05 December 2024).

Zeichhardt, H., Kammel, M., 2020b. INSTAND Report on Extra EQAS Group No. 340 Virus Genome Detection SARS-CoV-2 April 2020 – Interim Evaluation. EQA report RV-Online. Report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2020c. INSTAND Report on EQAS Group No. 340 Virus Genome Detection Coronaviruses incl. SARS-CoV-2 June 2020. EQA report RV-Online. Report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2021. Begleitheft - Quantitative Bezugsproben zur Verbesserung der Vergleichbarkeit und Bewertung von Laborergebnissen zum Virusgenom-Nachweis von SARS-CoV-2 (Stand 24.03.2021), available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2023a. INSTAND Report on EQAS Group No. 340 Virus Genome Detection SARS-CoV-2 June 2023. EQA report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2023b. INSTAND Report on EQAS Group No. 377 Virus Genome Detection – Hepatitis A Virus June 2023. EQA Report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2023c. INSTAND Report on EQAS Group No. 386 Virus Genome Detection – Measles Virus June 2023. EQA Report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2023d. INSTAND Report on EQAS Group No. 389 Virus Genome Detection – Rubella Virus June 2023. EQA Report RV-Online, report available upon request at instand@instand-ev.de.

Zeichhardt, H., Kammel, M., 2023e. INSTAND Report on EQAS Group No. 391 Virus Genome Detection – West Nile Virus June 2023. EQA Report RV-Online, report available upon request at instand@instand-ev.de.