National Institute for Public Health
and the Environment
*Ministry of Health, Welfare and Sport*

# External Quality Assessment of laboratories Performing SARS-CoV-2 Diagnostics for the Dutch Population, November 2021

# Colophon



John Sluimer[1]
Gabriel Goderski[1]
Chantal Herrebrugh[1]
Sharon van den Brink[1]
Lisa Wijsman[1]
Chantal Reusken[1]
Marion Koopmans[2]
Richard Molenkamp[2]
Adam Meijer[1]
Dirk Eggink[1]
Lance Presser[1]

1. National Institute For Public Health and The Environment (RIVM), Centre for Infectious Diseases Research, Diagnostics and Laboratory Surveillance, Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands.
2. Erasmus Medical Centre (Erasmus MC), Department Viroscience, Dr. Molewaterplein 40, 3015 GD Rotterdam

Corresponding author
John Sluimer opschalingslabs@rivm.nl

This investigation was performed by order, and for the account, of the Ministry of Health, Welfare and Sport, project number V/190002/01/PR.

This is a publication of the:
**National Institute for Public Health and the Environment**
P.O. Box 1 | 3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

# Contents



Version 2 31-1-2022: errata compared to version 1:
Page 8: Changed number (54.0% > 31.7%)
Page 9: Changed number (LEQA4_CoV20-06 > LEQA4_CoV21-09)
Page 11: Changed header of column in Figure 1 (hCoV-NL63 > hCoV-OC43)

# Summary

**Background**

Since January 2020 different workflows for molecular diagnostics of SARS-CoV-2 were implemented and checked for performance using specificity and sensitivity panels distributed by the National Institute for Public Health and the Environment (RIVM). A National External Quality Assessment (EQA) (Landelijk EQA; LEQA) program has been developed and implemented by the Dutch COVID-19 WHO reference laboratories at RIVM and Erasmus Medical Centre (Erasmus MC) and the Dutch Ministry of Health, Welfare and Sport ('Dienst Testen'). Main goal is to regularly check the performance of the network of COVID-19 molecular diagnostic labs simultaneously. This is also important because of a heterogeneity in used specificity and sensitivity panels for checking SARS-CoV-2 Nucleic Acid Amplification Tests (NAAT) performance over time and the fact that these quality checks were implemented only when workflows changed or laboratories were added to the network.

Since the start of this program three rounds of LEQA have been performed. This report describes the results of the fourth round of the LEQA program.

**Objective**

The goal of this LEQA round is to assess the quality of the Dutch SARS-CoV-2 molecular diagnostics field, using a panel that consists of 10 simulated clinical specimens, containing heat inactivated SARS-CoV-2, including variants of concern (VOC) SARS-CoV-2 B.1.617.2 (Delta variant) and SARS-CoV-2 B.1.1.7 (Alpha variant), or other respiratory viruses or their genetic material, or no virus. Each of the laboratories was asked to conduct molecular detection of SARS-CoV-2 according to their workflows used for SARS-CoV-2 diagnostics. Of these 10 specimens, 9 were considered core specimen with one additional educational specimen.

**Materials and Methods**

In October 2021 the LEQA panel was produced and pre-tested at the RIVM and Erasmus MC. After obtaining the correct results per specimen all laboratories performing SARS-CoV-2 diagnostics in the Dutch network were contacted and notified of the distribution of the panel in the second week of October 2021. Laboratories were asked to report their results via an online form. A number of workflows, especially the molecular point of care (mPOCT) ones, use expensive cartridges or pouches of which laboratories only receive a limited number every week. Therefore, laboratories that wanted to test these workflows with a limited panel were asked to indicate this. These laboratories were given the option to test a limited panel consisting of specimens 1, 2, 5, 8 and 9.

Workflows were given a score of 9 for 100% correct results for the 9 core specimens. The score was reduced by 1 point per specimen for a false positive/negative result and 0.5 points for a result reported as "Inconclusive". When a workflow tested the reduced panel (containing specimen 1, 2, 5, 8 and 9), 2 points for a false positive/negative result and 1 point for a result reported as "Inconclusive" were subtracted from the maximum score.

**Results**

Out of 158 workflows reported by 78 participating laboratories, 121 (76.58%) scored a 100% correct score for all 9 core specimens (9 points) and thus met all criteria set for reliable SARS-CoV-2 diagnostics, 23 (14.56%) scored between 8-8.5 points, making it likely that only minor adjustments need to be made to meet all criteria and 14 (8.86%) workflows scored below 8 points. A score below 8 points indicates that improvements need to be made for a workflow to be reliable for SARS-CoV-2 diagnostics in clinical diagnostic settings and surveillance. The laboratories sending in datasets of workflows scoring below 8 points have been contacted. In total 3 workflows were excluded from

further analyses for not matching the criteria set for participation in LEQA4 (e.g. too few specimens tested, or incorrect submission of data).
For the core specimens that do not contain SARS-CoV-2 only two false positive or inconclusive results were reported, confirming specificity of the vast majority of tested workflows. For the 7 SARS-CoV-2 containing core specimens, 55 false negative results and 10 inconclusive results were reported. These negative results corresponded to the highest dilution (lowest viral load) specimens considered core specimens. Despite the wide variety of kits, equipment and enzymes that are used in the different implemented workflows nationwide, the influence of these variations on the quality of molecular diagnostics for SARS-CoV-2 was limited.

**Conclusions**

Overall the workflows used for SARS-CoV-2 diagnostics perform very well and laboratories using them provide a reliable network. A small number of workflows could be further optimized to achieve full potential. The Dutch SARS-CoV-2 diagnostics laboratory network performs on a very high level with the vast majority of workflows detecting the core SARS-CoV-2 containing specimens correctly. The wide variety of kits, equipment and enzymes used in the Dutch SARS-CoV-2 diagnostic field do not affect adversely the quality of diagnostics. Instead, it allows for great flexibility during times of shortages in supplies and likely improves the capacity to detect possible future variants of SARS-CoV-2. In addition it prevents a collapse of the network if a particular assay fails due to evolution of the virus.

# 1. Introduction

Since January 2020 a wide variety of workflows for molecular diagnostics of COVID-19 were implemented and checked for performance using specificity and sensitivity panels distributed by the National Institute for Public Health and the Environment (RIVM). Although panels have been largely similar in viral load components for checking SARS-CoV-2 Nucleic Acid Amplification Tests (NAAT) performance, they initially contained SARS-CoV-1 RNA, later replaced by SARS-CoV-2 RNA, followed by SARS-CoV-2 whole heat inactivated virus particles, depending on when materials became available. Because of this heterogeneity in the past, the fact that patchy quality checks were implemented only when workflows changed or laboratories were added to the network, and because it is important to regularly check the performance of the COVID-19 molecular diagnostic lab network, the COVID-19 WHO reference laboratories at RIVM and Erasmus Medical Centre (Erasmus MC) and the Dutch Ministry of Health, Welfare and Sport ('Dienst Testen') have developed a National External Quality Assessment (EQA) (Landelijk EQA; LEQA) program.

At this moment this program consists of four rounds of EQA over the past one and a half year. In the second week of October 2021 the fourth round of EQA panels was distributed to all laboratories performing SARS-CoV-2 diagnostics on clinical specimens derived from Dutch patients. This panel consisted of 10 simulated clinical specimens that contained either heat inactivated SARS-CoV-2 (SARS-CoV-2 B.1.617.2 (Delta variant) or SARS-CoV-2 B.1.1.7 (Alpha variant)) or other respiratory viruses or their genetic material or no virus. Each of the laboratories was asked to conduct molecular detection of SARS-CoV-2 on this panel according to their workflows normally used for SARS-CoV-2 diagnostics. All data had to be reported back to the RIVM using an online reporting form.

# 2 Materials and methods

## 2.1 Approach

The LEQA panel was produced at the RIVM in September 2021 and pretested at the RIVM and Erasmus MC. Both centers obtained similar results. A month later, the laboratories performing SARS-CoV-2 diagnostics in the Dutch network were contacted and notified of the distribution of the panel in the second week of October 2021. All laboratories were asked to report their findings via an online form using Formdesk software (Wassenaar, The Netherlands) to allow for a more streamlined method of data collection. Laboratories were given time until the 31$^{st}$ of October to report their obtained results. After the 31$^{st}$ of October, laboratories that had not yet reported results were contacted and given one week grace time for reporting, after which the submission was closed on the 7$^{th}$ of November. A number of workflows, especially the molecular point of care (mPOCT) ones, use expensive cartridges or pouches of which laboratories only receive a limited number every week. Therefore laboratories that wanted to test these workflows with a limited panel were given the option to do so. The reduced panel consisted of specimens 1, 2, 5, 8 and 9 (Table 1), covering three different concentrations of SARS-CoV-2, a SARS-CoV-2 and influenza A(H3N2) double positive specimen and a no virus control.

## 2.2 Contents of LEQA4 panel

The LEQA4 panel consisted of 10 simulated clinical specimens (1ml) containing either whole infectious human respiratory seasonal viruses, or heat-inactivated SARS-CoV-2 viruses or no virus. Two specimens contained both SARS-CoV-2 and influenza virus. SARS-CoV-2 was isolated from clinical specimens on VERO E6 cells and heat-inactivated by heat treatment at 60 °C for two hours. The number of detectable copies of SARS-CoV-2 positive strand RNA in the stocks of SARS-CoV-2 was back-calculated from determination of the copy number after extraction of RNA by digital SARS-CoV-2 E-gene and RdRP-gene PCR. Because the viruses were not purified from the supernatant, the whole virus preparation contains in addition to genomic RNA, intermediate replication negative strand genomic

RNA and subgenomic E-gene RNA that contribute to detection in routine one-step RT-qPCR for SARS-CoV-2 RNA. Virus dilutions were made in MEM with Hanks' salts. HEp2 cells were added to the dilution at a concentration of 10.000 cells per ml panel specimen to simulate a clinical specimen. The 10 specimens included in the panel contained the following viruses: SARS-CoV-2 Delta (variant B.1.617.2;hCoV-19/Netherlands/NH-RIVM-27142/2021) in six concentrations, SARS-CoV-2 Alpha (variant B.1.1.7; 20B/501Y.V1; hCoV-19/Netherlands/UT-RIVM-12844/2021), hCoV-OC43 (ATCC), influenza A(H3N2)(A/Netherlands/10002/2019) and a specimen without any virus. Seasonal hCoV-OC43 and A(H3N2) viruses were not inactivated. In Table 1 all specimens are listed together with the expected target specific Cp values obtained at RIVM with routinely used diagnostic RT-qPCRs for the respective pathogens and the expected conclusion for SARS-CoV-2 detection in the specimens. The digital copies of RdRP-gene and E-gene are also listed in Table 1 for the SARS-CoV-2 containing specimens.

*Table 1: Composition of LEQA3 together with the target specific expected Cp values[1] based on the in-house assay(s) of the RIVM.*

| Panel coding | Virus[2] | Number of copies E gene target/ml specimen, determined with dPCR[3] | Number of copies RdRP gene target/ml specimen, determined with dPCR[3] | Target specific Cp[4] | E-gene (Sarbeco) Cp | RdRP-gene (SARS-CoV-2) Cp | Conclusion SARS-CoV-2 |
|---|---|---|---|---|---|---|---|
| LEQA4_CoV21-01 | SARS-CoV-2 Delta (d3) | 1010 | 788 | n/a | 34.8 (4/4) | 36.5 (4/4) | POSITIVE |
| LEQA4_CoV21-02[5] | Influenza virus A(H3N2) + SARS-CoV-2 Delta (1) | 505000 | 394000 | 34.8 (4/4) | 27.8 (4/4) | 28.7 (4/4) | POSITIVE |
| LEQA4_CoV21-03 | SARS-CoV-2 Alpha | 33900 | 23200 | n/a | 29.8 (4/4) | 30.5 (4/4) | POSITIVE |
| LEQA4_CoV21-04 | SARS-CoV-2 Delta (d1) | 101000 | 78800 | n/a | 29.1 (4/4) | 30.1 (4/4) | POSITIVE |
| LEQA4_CoV21-05[5] | Influenza virus A(H3N2) + SARS-CoV-2 Delta (2) | 5050 | 3940 | 28.1 (4/4) | 33.4 (4/4) | 35.4 (4/4) | POSITIVE |
| LEQA4_CoV21-06 | SARS-CoV-2 Delta (d3) | 1010 | 788 | n/a | 34.3 (4/4) | 36.5 (3/4) | POSITIVE |
| LEQA4_CoV21-07 | hCoV-OC43 | n/a | n/a | 29.8 (4/4) | n/a | n/a | Negative |
| LEQA4_CoV21-08 | No virus | n/a | n/a | n/a | n/a | n/a | Negative |
| LEQA4_CoV21-09[6] | SARS-CoV-2 Delta (d4) | 101 | 78.8 | n/a | 35.0 (1/4) | n/a | Weakly POSITIVE |
| LEQA4_CoV21-10 | SARS-CoV-2 Delta (d2) | 10100 | 7880 | n/a | 32.4 (4/4) | 33.8 (4/4) | POSITIVE |

[1] The expected Cp values shown in this table are based on RT-qPCR tests performed on the panel specimens using the routinely used RIVM in-house assays. The in-house real-time RT-qPCRs have been performed using the following reagents and volumes: ThermoFisher TaqMan® Fast Virus 1-Step Master Mix after extraction of 200 µl specimen on Roche MagNA Pure 96 instrument with Roche MagNA Pure 96 DNA and Viral NA Small Volume Kit, elution in 50 µl and 5 µl extract per RT-qPCR reaction on Roche LightCycler 480 mark I or II. Extractions and subsequent RT-qPCRs were performed in 4-fold; after the average Cp value between brackets () the number of times found positive is shown. SARS-CoV-2 E-gene Sarbeco specific primers and probes are those published by Corman *et al*. 2020; the RdRP primers and probes are modified from those published by Corman *et al*. (2020) to become SARS-CoV-2 specific and similar in LOD95 compared to the E-gene Rt-qPCR.

[2] d1, d2, d3 and d4 indicate that these specimens are a 1:10 dilution series. Dilution d1 has the highest concentration of SARS-CoV-2 and d4 the lowest. SARS-CoV-2 is heat inactivated.

[3] dPCR has been performed on + strand genomic RNA for RdRP-gene and E-gene. The one-step E-gene and RdRP-gene diagnostic RT-qPCR also detects - strand replicative form genomic RNA and the one-step E gene RT-qPCR in addition also detects subgenomic messengers, which probably increases the actual number of target templates for the diagnostic RT-qPCR in the specimen after extraction.

[4] For hCoV-OC43 N-gene; for influenza virus M-gene; n/a = not applicable.

[5] Double positive specimen for both SARS-CoV-2 Delta variant and influenza A(H3N2) virus. One specimen (1) contained a high copy number of SARS-CoV-2 and low of A(H3N2)y, whereas the other specimen (2) contained a low copy number of SARS-CoV-2 and high of A(H3N2) to be able to detect possible interference. Detection of influenza virus in these specimens is educational.

[6] Educational specimen: repeats of this specimen may have the E-gene and/or RdRP-gene negative; only 31.7% of reported workflows having tested this specimen reported this specimen positive for SARS-CoV-2

## 2.3 Scoring the workflows

The performance of each reported workflow was evaluated after which they were scored on a scale from 0 to 9, with 9 being the best grade. This scoring system was implemented based on the results with the core specimens present in the panel. All specimens except LEQA4_CoV21-09 (educational load of SARS-CoV-2) were deemed core specimens (specimens with clinically relevant amounts of virus or no virus). The laboratories were given the option to evaluate specimens with the following scores: Positive, Negative or Inconclusive. The term "Inconclusive" was used to indicate uncertainty concerning the test result(s) obtained.

As the panel consisted of 9 core specimens, each workflow started with a total of 9 points. For each wrongly determined core specimen (being positive for a specimen containing no SARS-CoV-2 or vice versa) 1 point was deducted (out of 9). When a core specimen was scored with an "Inconclusive" result, 0.5 point was deducted from the final mark of the workflow.

For some workflows (e.g. molecular point-of-care test (mPOCT) workflows) an option was given to test a smaller subset of specimens to enable making a limited statement about the sensitivity of detection of SARS-CoV-2 with the mPOCT used. These workflows only had to test LEQA4_CoV21-1, LEQA4_CoV21-02, LEQA4_CoV21-06, LEQA4_CoV21-08 and LEQA4_CoV21-09. The workflows testing the reduced panel were also graded according to a scale from 0 – 9 points. For each wrongly determined core specimen (being positive for a specimen containing no SARS-CoV-2 or vice versa) 2 points were deducted (out of 9). When a specimen was scored with an "Inconclusive" result, 1 point was deducted from the final mark of the workflow.

A workflow scoring 9 out of 9 passed all criteria set for SARS-CoV-2 diagnostics in terms of sensitivity and specificity deemed necessary for SARS-CoV-2 diagnostics in accordance with the set requirements for new workflows and laboratories. [1] Workflows scoring 8 or 8.5 out of 9 are valuable for SARS-CoV-2 diagnostics, but need adjustments in order to perform as desired. Adjustments depend on the type of result, e.g. an "Inconclusive" result for low viral load LEQA4_CoV21-03 or LEQA4_CoV21-08 specimens is less severe than detection of SARS-CoV-2 targets in specimens which do not contain SARS-CoV-2 (false positive). Any workflow scoring below 8 out of 9 points needs serious adjustments in order to be fit for SARS-CoV-2 diagnostics.

# 3. Results

## 3.1 Aggregated overview

Eighty-four laboratories were contacted with the announcement of panel distribution for this fourth EQA round. Seventy-eight (92.9%) of these laboratories reported their findings for a total of 161 workflows. Of these workflows, 3 workflows were excluded from the analysis due to not following instructions for LEQA4, dataset being submitted twice or due to poor performance of an experimental workflow. Data of the remaining 158 workflows were included for further analyses for this report. The workflow results reported for each panel specimen are summarized in Table 2. The panel scores obtained per laboratory and by number of workflows used are summarized in Table 3. Despite not all workflows obtained fully correct results with the core specimens (Table 2), nearly all laboratories (74/78; 94.9%) used at least one workflow for which a score of 8 to 9 was obtained (Table 3)**.** Using a heatmap, Figure 1 shows a summary of the overall results obtained by all workflows. In the subsequent chapters a more detailed insight in the results and their background is presented.

*Table 2. Aggregated overview of workflow conclusions by LEQA4 panel specimen.*

| Panel specimen | Content | № of workflows with test result reported (n=158)[1] | SARS-CoV-2 detection workflow conclusion | | | |
|---|---|---|---|---|---|---|
| | | | № Positive | № Indeterminate, equivocal or inconclusive | № Negative | Errors |
| LEQA4_CoV21-08 | No virus | 158 | 2 (1.27%) | 0 | 156 (98.73%) | False positive result (n=2) |
| LEQA4_CoV21-07 | hCoV-OC43 | 145 | 0 | 0 | 145 (100%) | None |
| LEQA4_CoV21-04 | SARS-CoV-2 Delta (d1) | 145 | 144 (99.31%) | 0 | 1 (0.69%) | False negative result (n=1) |
| LEQA4_CoV21-10 | SARS-CoV-2 Delta (d2) | 145 | 141 (97.24%) | 1 (0.69%) | 3 (2.07%) | False negative result (n=3); False inconclusive result (n=1) |
| LEQA4_CoV21-01 | SARS-CoV-2 Delta (d3) | 158 | 132 (83.54%) | 2 (1.27%) | 24 (15.19%) | False negative result (n=24); False inconclusive result (n=2) |
| LEQA4_CoV21-06 | SARS-CoV-2 Delta (d3) | 145 | 123 (84.83%) | 3 (2.07%) | 19 (13.10%) | False negative result (n=19); False inconclusive result (n=3) |
| LEQA4_CoV21-09 | SARS-CoV-2 Delta (d4) | 158 | 50 (31.65%) | 4 (2.53%) | 104 (65.82%) | Not applicable, educational specimen |
| LEQA4_CoV21-03 | SARS-CoV-2 Alpha | 145 | 144 (99.31%) | 0 | 1 (0.69%) | False negative result (n=1) |
| LEQA4_CoV21-02 | Influenzavirus A(H3N2) + SARS-CoV-2 Delta (1) | 158 | 158 (100%) | 0 | 0 | None |
| LEQA4_CoV21-05 | Influenzavirus A(H3N2) + SARS-CoV-2 Delta (2) | 158 | 151 (95.57%) | 0 | 7 (4.43%) | False negative result (n=7) |

[1] 13 of the 158 workflows tested were mPOCT testing specimens 1, 2, 6, 8 and 9 only. Therefore the number of workflows with test result per specimen is 145 for specimen 3, 4, 5, 7 and 10.

*Table 3. Aggregated overview of scores for core specimens obtained by laboratories using various numbers of workflows.*

| № of workflows per lab | № of labs | № of workflows per lab with indicated score (No of labs) | | |
|---|---|---|---|---|
| | | Score 9 | Score 8 or 8.5 | Score < 8 |
| 6 | 2 | 3-5 (n=2) | 1-2 (n=2) | 1 (n=1) |
| 5 | 1 | 5 (n=1) | 0 | 0 |
| 4 | 9 | 1-4 (n=9) | 1-2 (n=3) | 1-2 (n=2) |
| 3 | 11 | 1-3 (n=6) | 1-3 (n=3) | 1-2 (n=3) |
| 2 | 17 | 1-2 (n=15) | 1 (n=5) | 1 (n=2) |
| 1 | 38 | 1 (n=29) | 1 (n=5) | 1 (n=4) |

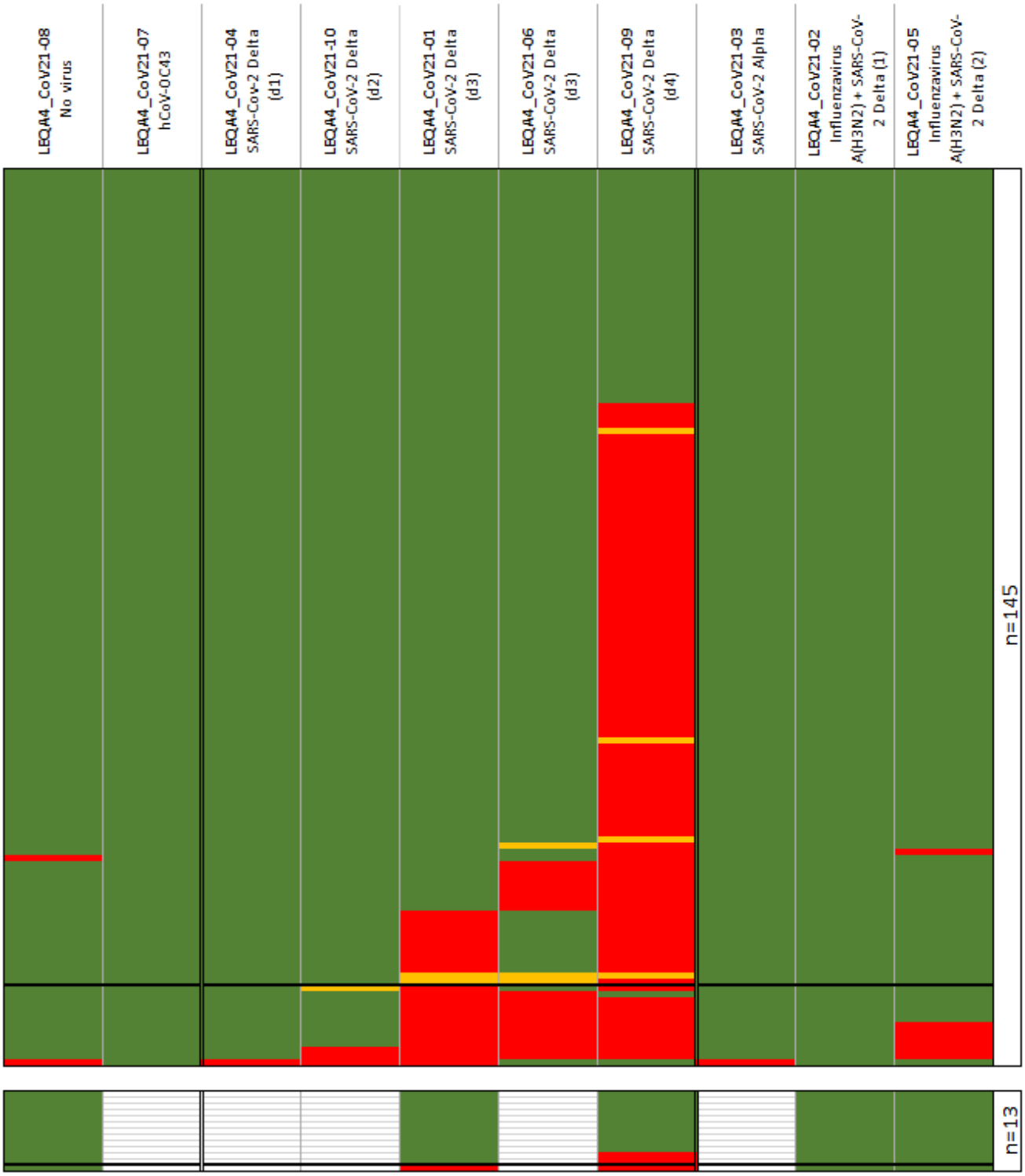



*Figure 1: A heatmap showing all final conclusions per sample per workflow submitted. Green blocks represent a correct result, red blocks indicate an incorrect result and orange blocks display inconclusive results. The workflows displayed in this heatmap are sorted by the obtained scores. The thick horizontal black line in the top segment of the figure separates all workflows scoring 8-9 points and all workflows scoring <8 points. When a box remains white, this sample was not tested by that particular workflow as it tested the reduced panel. All workflows testing the reduced panel are displayed in the lower segment of the figure (and are sorted by score). The thick horizontal black line in the lower segment of the figure separates all workflows scoring 8-9 points and the workflow scoring <8 points. All laboratories using a workflow scoring <8 points have been contacted. All samples are arranged similarly as in Table 2. The first two samples are SARS-CoV-2 negative samples, the next five are a SARS-CoV-2 Delta variant dilution series and the last three are various SARS-CoV-2 containing samples. As LEQA4_CoV21-09 is an educational sample with a very low viral load, many workflows failed to detect SARS-CoV-2 in this sample.*

## 3.2 Target genes used for RT-PCR or other NAAT

As the sensitivity of a workflow may also depend on the number of used target gene(s), for all workflows the target genes used were inventoried based on the online report form. Workflows used up to 4 target genes. Table 4 shows the scores obtained by all workflows testing LEQA4 sorted by the number target genes used. From the 158 workflows a total of 57 workflows used 1 target gene, 85 workflows used 2 target genes, 12 workflows used 3 target genes and 4 workflows used 4 target genes. Combinations of genes used by number of workflows are listed in Table 5. Some workflows using more than one target gene do not generate separate result for each independent gene but rather a composite conclusion.

*Table 4. Scores obtained by all workflows sorted by number of target genes used*

| № of target genes per workflow | № of workflows | № of workflows with indicated score | | |
|---|---|---|---|---|
| | | Score 9 | Score 8 or 8.5 | Score < 8 |
| 1 | 57 | 40 (70.1%) | 8 (14.04%) | 9 (15.79%) |
| 2 | 85 | 70 (82.35%) | 11 (12.94%) | 4 (4.71%) |
| 3 | 12 | 8 (66.67%) | 3 (25.00%) | 1 (8.33%) |
| 4 | 4 | 3 (75.00%) | 1 (25.00%) | 0 |

*Table 5. Overview of number and type of target genes used per reported workflow.*

| № target genes in workflow | Target gene(s) | № workflows |
|---|---|---|
| 1 | E-gene Sarbeco specific | 21 |
| | E-gene SARS-CoV-2 specific | 11 |
| | N-gene | 3 |
| | N1-gene | 1 |
| | NSP-2 | 1 |
| | ORF1a/b | 15 |
| | RdRP-gene | 5 |
| 2 | E-gene Sarbeco specific; N-gene | 4 |
| | E-gene Sarbeco specific; N1-gene | 7 |
| | E-gene Sarbeco specific; N2-gene | 33 |
| | E-gene Sarbeco specific; ORF1a/b | 3 |
| | E-gene Sarbeco specific; RdRP-gene | 3 |
| | E-gene Sarbeco specific; S-gene | 1 |
| | E-gene SARS-CoV-2 specific; N-gene | 1 |
| | E-gene SARS-CoV-2 specific; RdRP-gene | 1 |
| | M-gene; S-gene | 1 |
| | N-gene; ORF1a/b | 14 |
| | N-gene; RdRP-gene | 7 |
| | N1-gene; N2-gene | 7 |
| | N1-gene; ORF1a/b | 1 |
| | ORF1a/b; RdRP-gene | 2 |
| 3 | E-gene Sarbeco specific; N-gene; RdRP-gene | 3 |
| | E-gene Sarbeco specific; N-gene; S-gene | 1 |
| | E-gene SARS-CoV-2 specific; N-gene; RdRP-gene | 1 |
| | E-gene Sarbeco specific; RdRP-gene; S-gene | 1 |
| | N-gene; ORF1a/b; S-gene | 6 |
| 4 | E-gene Sarbeco specific; N-gene; RdRP-gene; S-gene | 4 |

## 3.3 Performance of the workflows

As described before, all workflows were graded using a point system from 0 (being the lowest grade) up to 9 (highest grade). One point was given for each of the core specimen in the complete panel. Two points for each of the core specimen was given when testing the reduced (mPOCT) panel. In total 121 workflows were given a '9', one workflow scored an '8.5', 22 workflows scored an '8', one workflow scored a "7.5", six workflows scored a '7', four workflows scored a '6', two workflows scored a '5' and one workflow scored a '4'. Figure 2 shows all grades given to the reported workflows.

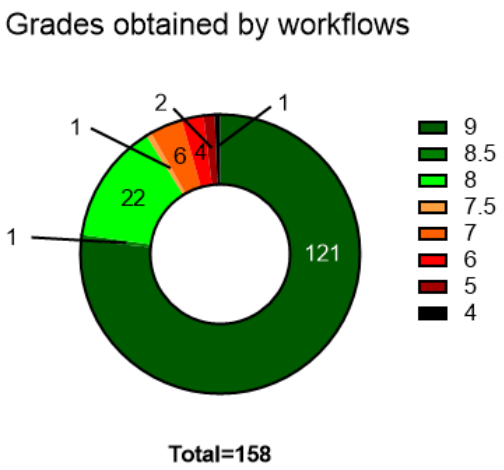



*Figure 2: All grades obtained by the reported workflows out of the maximum of 9 points (n=158).*

An overview containing the results obtained per target gene per specimen for workflows reporting Cp values is shown in Supplemental Figures 1 and 2. In these figures for each of the target genes used the Cp values are shown for each of the tested specimens.

The obtained scores per workflow are also linked to the used extraction kit or method, the performed PCR or other NAAT test and the used number of target genes in order to assess the effect of different techniques on the performance of workflows. An overview of these factors on the grade is shown in Supplemental Figure 8.

## 4. Discussion and conclusion

Here we describe the results of the fourth national LEQA panel. 78 laboratories, using in total 158 workflows have reported their findings. Out of the 158 workflows reported, 121/158 (76.58%) scored a maximum score for all 9 core specimens (9 points) and thus met all criteria set for reliable SARS-CoV-2 diagnostics. 23/158 (14.56%) scored between 8-8.5, making it likely that only minor adjustments need to be made to meet all criteria and 14/158 (8.86%) workflows scored a 7.5 or below. When scoring below 8 points, it is an indication that major improvements need to be made for these workflows to be reliable for SARS-CoV-2 diagnostics in clinical diagnostic settings and surveillance. All laboratories with workflows scoring below 8 points have been contacted and have been given recommendations how to improve their performance.

For the seven SARS-CoV-2 containing core specimens, 55 false negative results and 10 inconclusive results were reported. Two false positive results were reported for specimens without SARS-CoV-2. Apparently, sensitivity is a bigger issue than specificity for the workflows used for SARS-CoV-2 diagnostics. When considering all workflows each laboratory has access to and has reported LEQA4 results for, four laboratories have workflows which do not have a grade of 8 or above. Compared to LEQA1-3, the overall scores obtained for LEQA4 seems to have decreased slightly. This is likely due to the slightly lower viral loads present in the SARS-CoV-2 dilution series (d1 – d4; see Table 1). [2-4] Hopefully the laboratories scoring less than 9/9 points for their workflow(s) will try to adapt them in order to increase their performance where possible.

Throughout the reported workflows lots of different target genes and combinations of them are reported, but the E-gene as target gene as either a Sarbeco specific or SARS-CoV-2 specific target is most prevalent (95 out of 158 workflows).

When comparing all workflows, there does not appear to be a correlation between the number of target genes used and the score obtained by the workflow. The average scores obtained are 8.42 (n=57), 8.75 (n=70), 8.50 (n=12) and 8.75 (n=4) out of 9 for workflows containing one, two, three and four target genes, respectively. The median score was 9 for all four types of workflows. Despite this apparent difference in sensitivity/specificity for workflows based on number of target genes, an argument in favor of multiple target genes is that it decreases the chance of false negative test results when testing (novel) SARS-CoV-2 variants. Mutations in (novel) SARS-CoV-2 variants might cause an target gene dropout in assays. Also other mutations are able to render primer/probe combinations less effective in detecting SARS-CoV-2. [5,6] It is possible that other novel SARS-CoV-2 variants cause target gene dropouts as primer/probe sets are not optimized for that particular new genetic variant. If an assay were to use only one target gene, there is a bigger chance on a false negative result than in assays with multiple target genes in case of new variants that affect primer or probe target sites. A downside of using multiple target genes is an increased uncertainty in final conclusion when two (or more) target genes show different results. The latter will especially be the case for specimen with low viral loads.

Despite target genes generating a Cp value in some workflows, the target gene result was deemed negative for SARS-CoV-2 presence, possibly related to a laboratory specific determined cut off value or other criterium, e.g. shape of the amplification curve.

For all LEQA4 specimens tested with a workflow providing Cp values as an output, a broad range of Cp values for the individual SARS-CoV-2 containing specimens has been reported. The biggest range of reported Cp values was found for LEQA4_CoV20-03 for target gene ORF1a/b. The average Cp value for ORF1a/b was 27.35 (SD 4.97) and the median Cp value was 28.67 (range Cp 13.24 – 33.17). This indicates a wide spread of Cp values reported for the same specimen, generated by the large variety of workflows used. Despite this wide range of Cp values for the same specimen by different workflows,

it did not seem to affect the sensitivity of the workflows when looking at the reported diagnostic result (positive vs negative). This finding indicates that comparing Cp values between workflows and laboratories is not possible without prior calibration using a standard.

A wide array of varying in-house and kit-based SARS-CoV-2 workflows have been reported. Compared to the 2009 influenza pandemic, the Dutch clinical diagnostic field for respiratory diagnostics shows a divergent pattern in use of kits, reagents and equipment. [7] A more divergent use of kits, reagents and equipment can be quite useful in a laboratory network as a shortage of any of these can be compensated by switching to different equipment or when certain workflows are less capable of detecting new strains of SARS-CoV-2. This is highly relevant with the rising level of infections with novel variants of SARS-CoV-2.

Although part of the workflows assessed need some work in order for them to perform as desired. For these workflows we hope that the laboratories involved will look into possible improvements on their workflow(s). All in all the Dutch SARS-CoV-2 diagnostics laboratory network appears to perform on a very high level providing accurate SARS-CoV-2 molecular detection services.

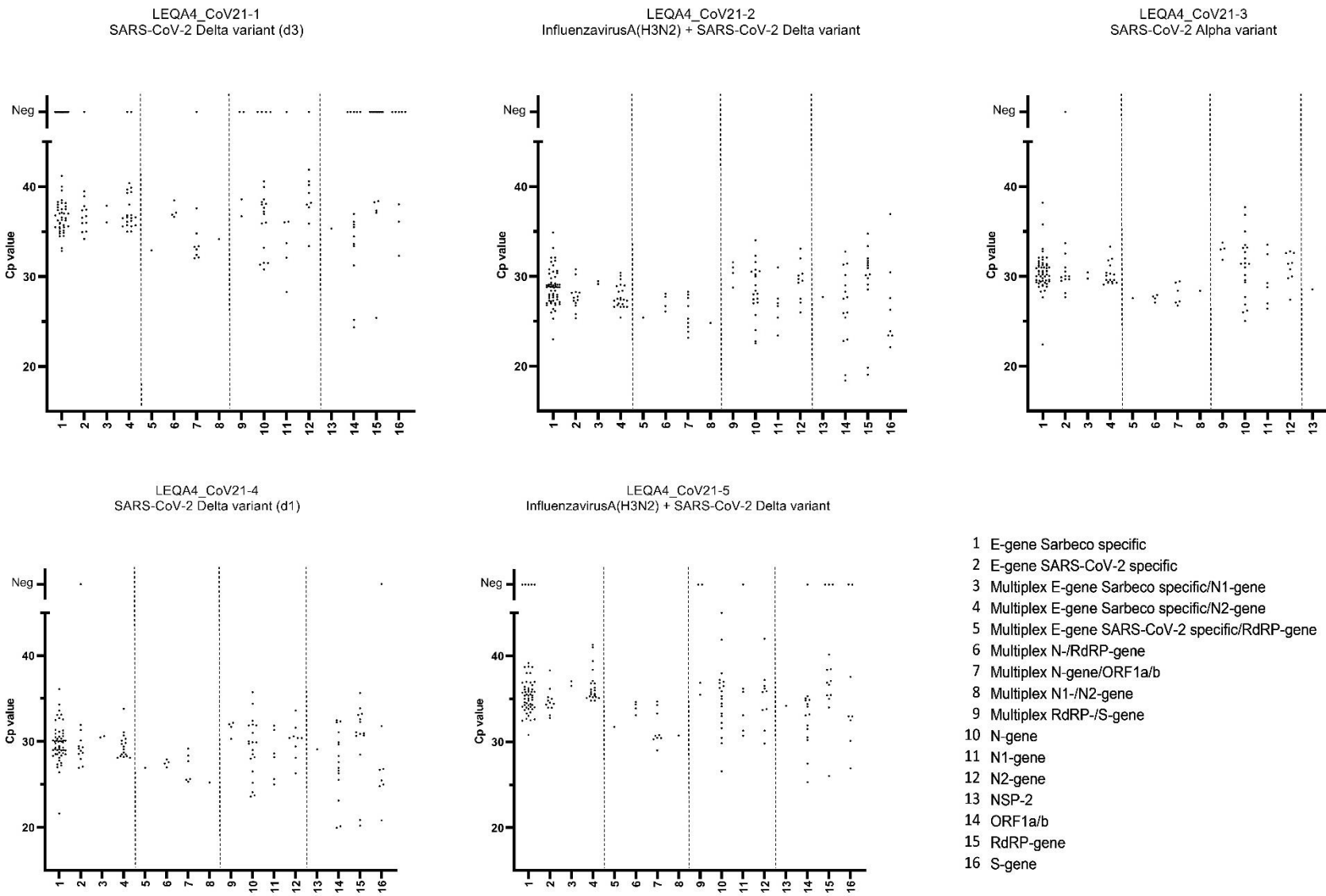
# 5. References


1. RIVM. Aanvullende informatie diagnostiek COVID-19. 2021 [cited 2021 28-6-2021]; 24: Available from: https://lci.rivm.nl/covid-19/bijlage/aanvullend
2. J. Sluimer *et al*., "External Quality Assessment of laboratories Performing SARS-CoV-2 Diagnostics for the Dutch Population, November 2020", RIVM report. Available from: https://www.rivm.nl/documenten/external-quality-assessment-of-laboratories-performing-sars-cov-2-diagnostics-for-dutch
3. J. Sluimer *et al.*, "External Quality Assessment of laboratories Performing SARS-CoV-2 Diagnostics for the Dutch Population, February 2021", RIVM report. Available from: https://www.rivm.nl/documenten/external-quality-assessment-of-laboratories-performing-sars-cov-2-diagnostics-for-0
4. J. Sluimer et al., "External Quality Assessment of laboratories Performing SARS-CoV-2 Diagnostics for the Dutch Population, May 2021", RIVM report. Available from: https://www.rivm.nl/documenten/eqa-of-laboratories-performing-sars-cov-2-diagnostics-for-dutch-may-2021
5. Kidd, M., et al., *S-Variant SARS-CoV-2 Lineage B1.1.7 Is Associated With Significantly Higher Viral Load in Specimens Tested by TaqPath Polymerase Chain Reaction.* J Infect Dis, 2021. **223**(10): p. 1666-1670.
6. WHO. Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern [cited 2021 12-12-2021]; 1: Available form: https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern
7. A. Meijer *et al*., "Preparing the outbreak assistance laboratory network in the Netherlands for the detection of the influenza virus A(H1N1) variant", J. Clin. Virol., vol. 45, no. 3, pp. 179–184, 2009.

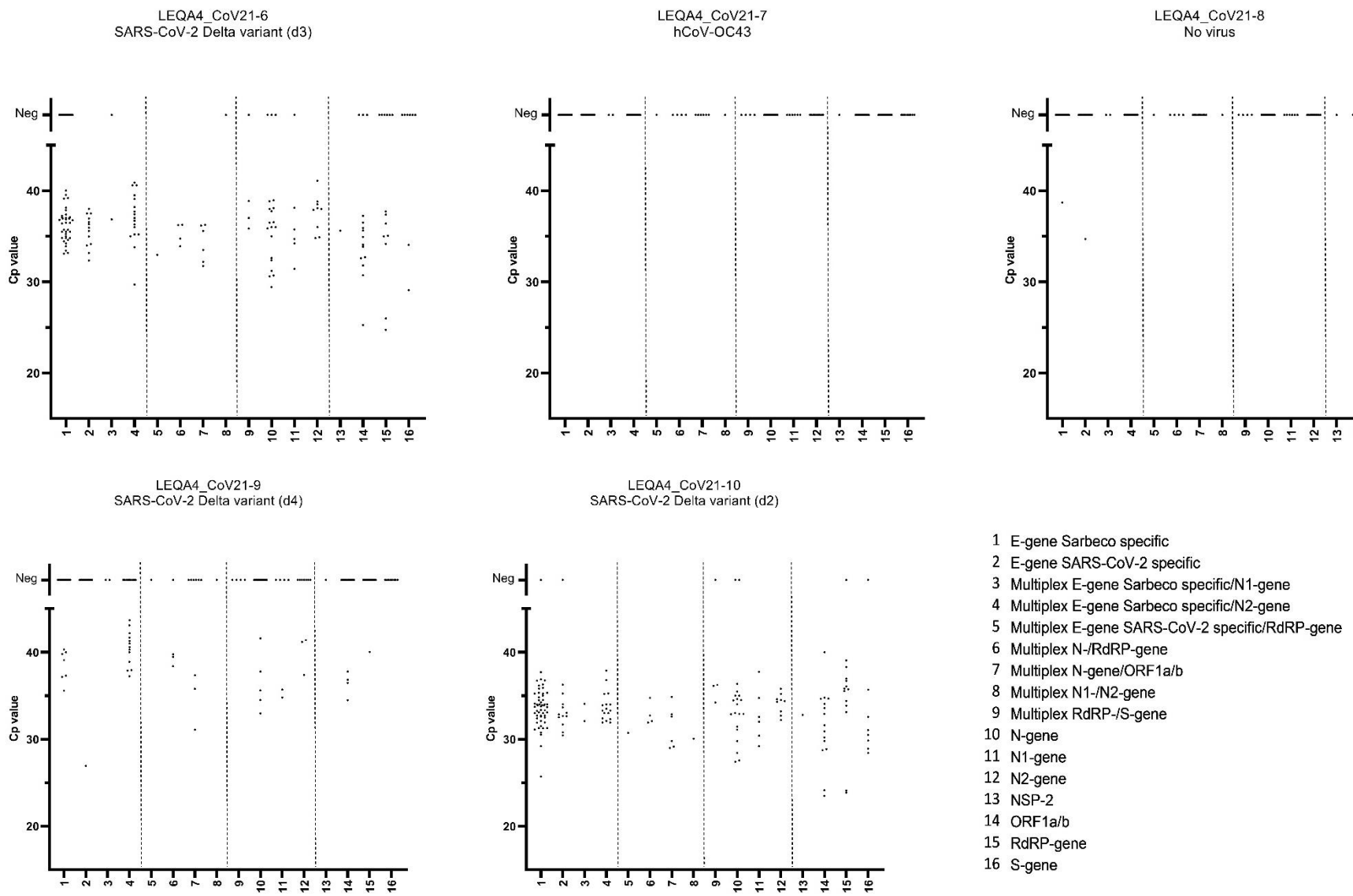
# 6. Supplemental material

## 6.1 Cp values obtained per panel specimen per target gene

In total 16 different target genes were reported in the submitted workflows. When looking at the Cp values obtained, it seems that using the same target genes can result in big variations even within the same specimen. Besides the generally big spread of Cp values, it is notable that in the SARS-CoV-2 Alpha variant containing specimen (LEQA4_CoV21-03) there is a relatively high number of negative results for the S-gene as target gene. This is not unexpected as mutations in SARS-CoV-2 variant B.1.1.7/20B/501Y.V1 can cause an S-gene dropout in certain assays. [5, 6] In this LEQA round it happened for the following kits: Applied Biosystems, TaqPath™ COVID-19 Multiplex Diagnostic Solution (CE-IVD) and Applied Biosystems, TaqPath™ COVID-19 RT-PCR Kit). In Supplemental Figures 1 and 2 the Cp values found for each target gene per specimen are shown.



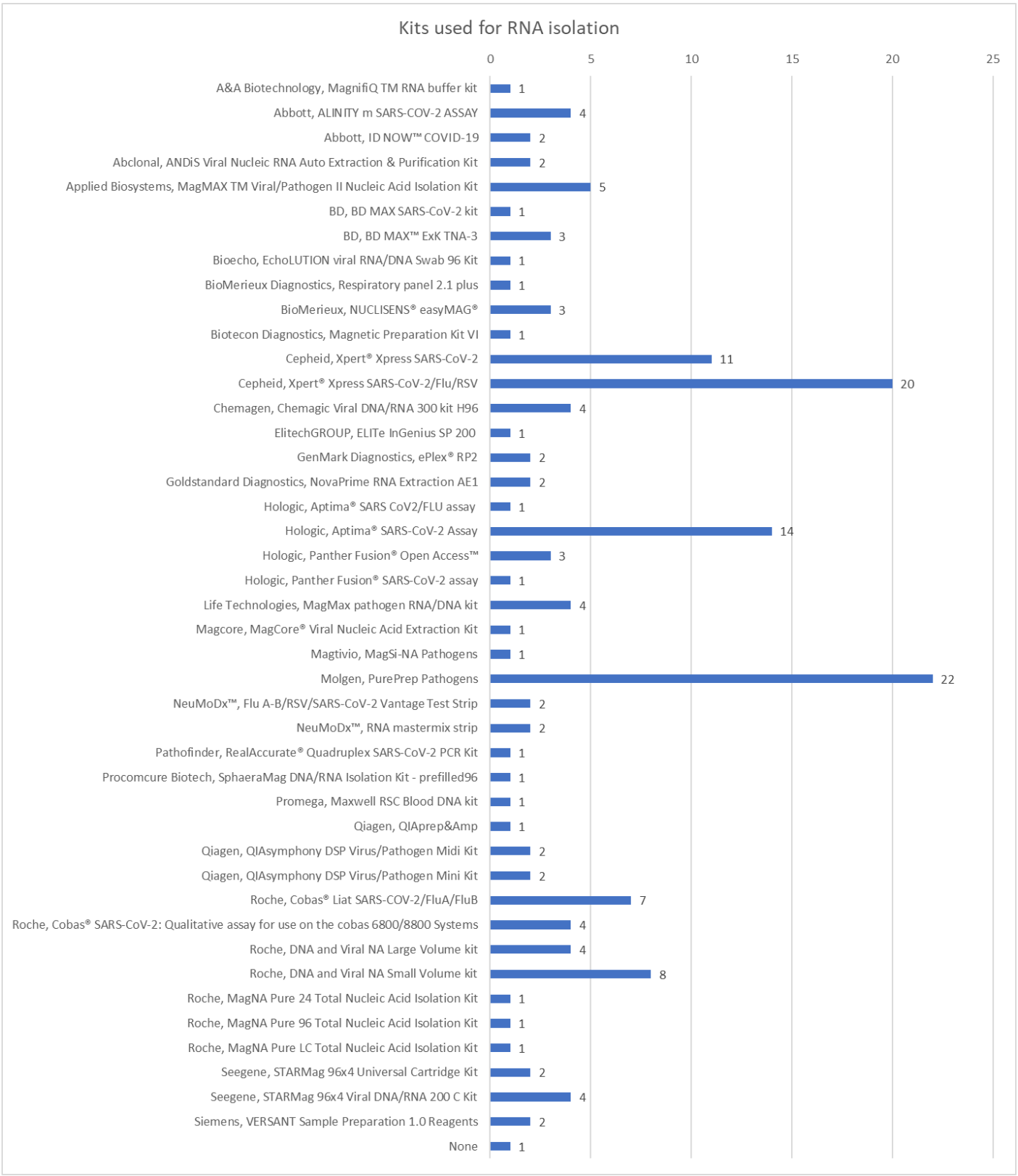

*Supplemental Figure 1: All Cp values obtained by Cp value reporting workflows for specimens LEQA4_CoV21-01 to LEQA4_CoV21-05. In total 16 different target genes were reported, some of which were multiplex target genes. Cp values obtained from workflows using multiple individually reported target genes area all reported separately from each other. The legend displayed shows which target genes were used. Negative results are clustered separately above the Y-axis.*



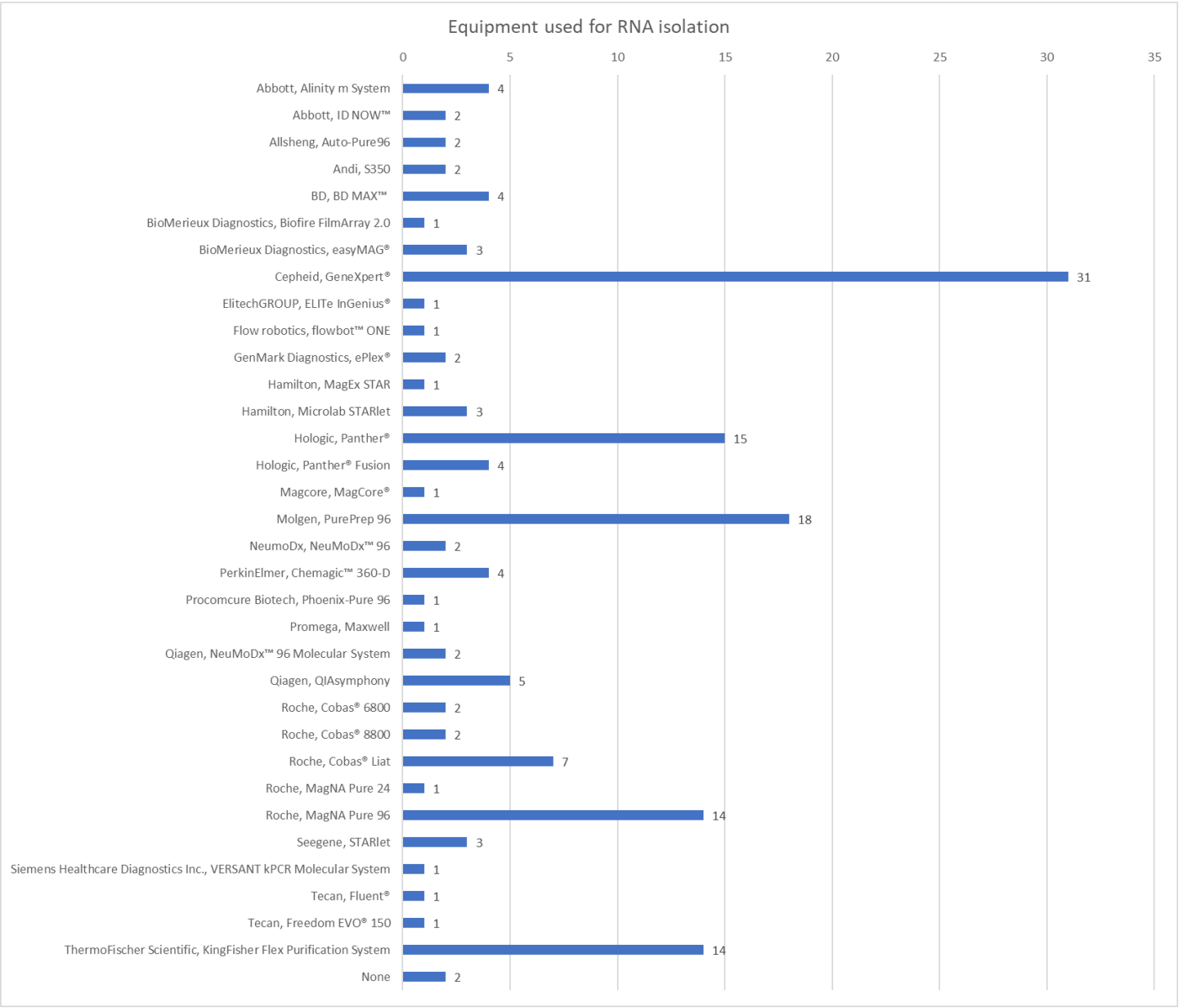

*Supplemental Figure 2: All Cp values obtained by Cp value reporting workflows for specimens LEQA4_CoV21-06 to LEQA4_CoV21-10. In total 16 different target genes were reported, some of which were multiplex target genes. Cp values obtained from workflows using multiple individually reported target genes area all reported separately from each other. The legend displayed shows which target genes were used. Negative results are clustered separately above the Y-axis.*

## 6.2 Used equipment, kits and reagents

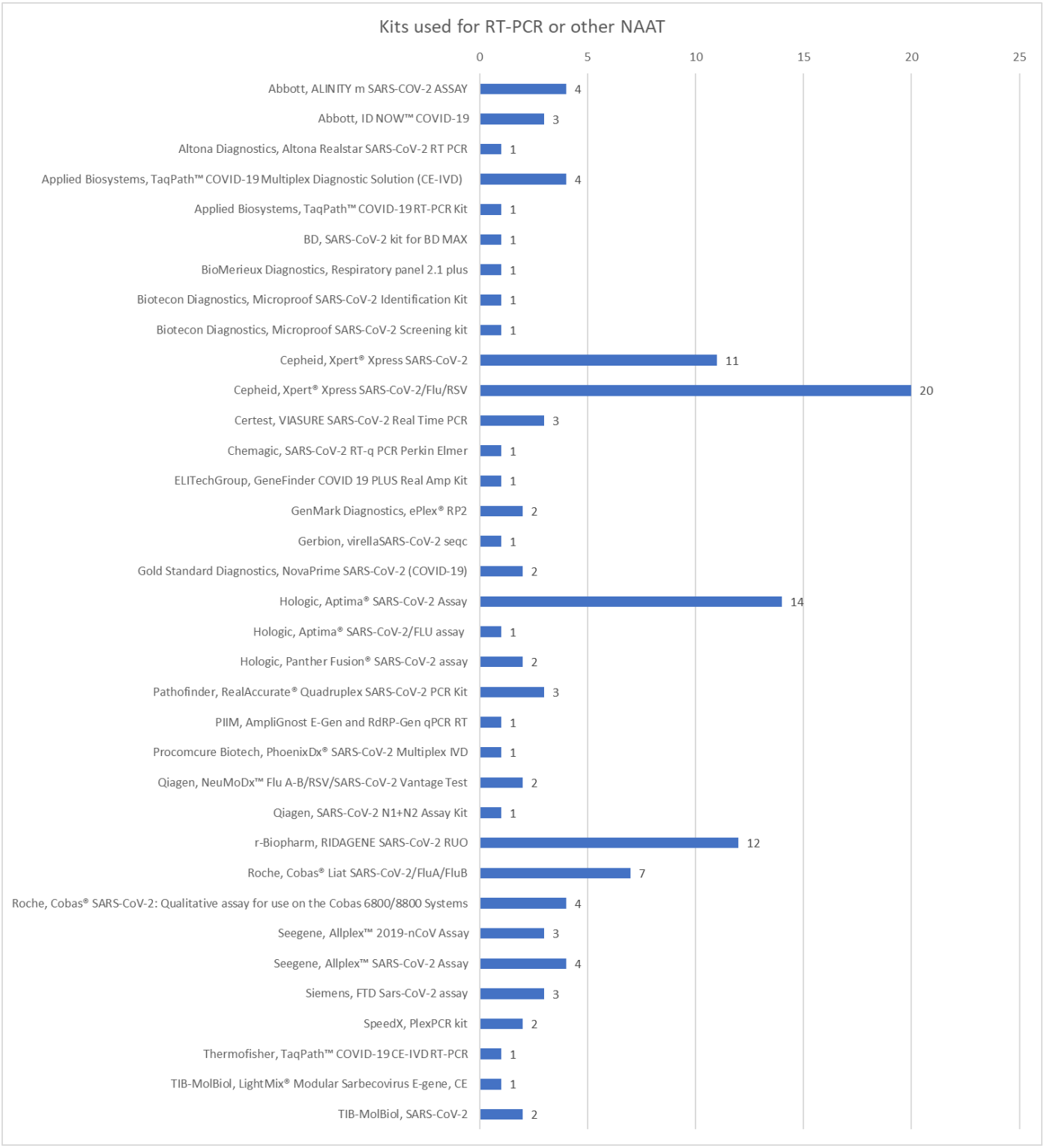
Some of the factors that may determine the performance of the workflows are the used kits, equipment and/or separate enzymes used for extraction and amplification implemented in SARS-CoV-2 diagnostics for the Dutch population. Therefore for each workflow these details were inventoried. Supplemental Figure 3 shows the kits used for RNA/total NA isolation, Supplemental Figure 4 shows the RNA isolation equipment, Supplemental Figure 5 shows the kits used for the RT-PCR or other NAAT reaction, Supplemental Figure 6 shows the separate enzymes used for the in-house RT-PCR or other NAAT reaction and Supplemental Figure 7 shows the equipment used for the RT-PCR or other NAAT reaction. In several occasions the kit used for extraction and for RT-qPCR or other NAAT has the same name because these are all-in-one kits.



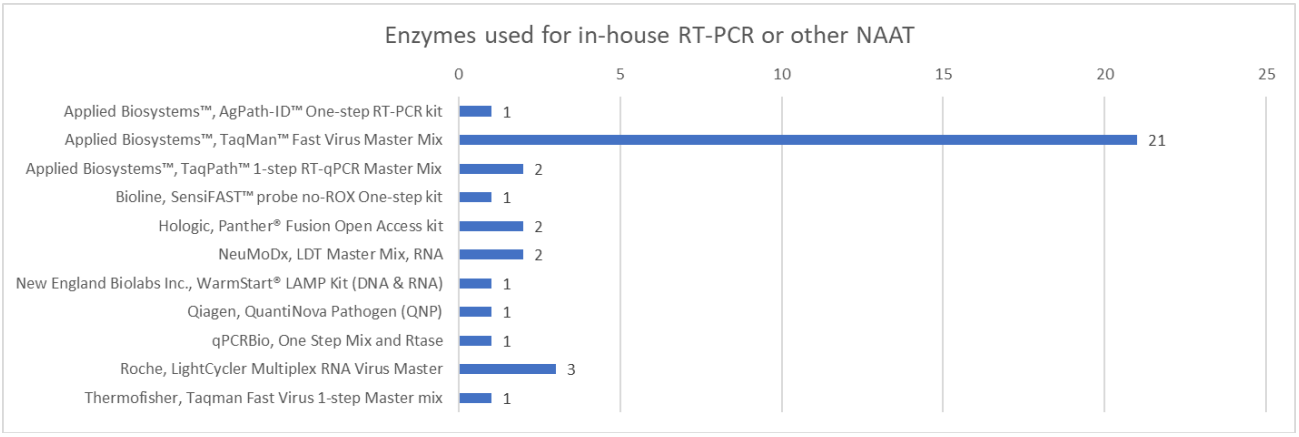

*Supplemental Figure 3: The RNA isolation kits used by workflows testing for SARS-CoV-2 together with the number of workflows per kit (n=158)*



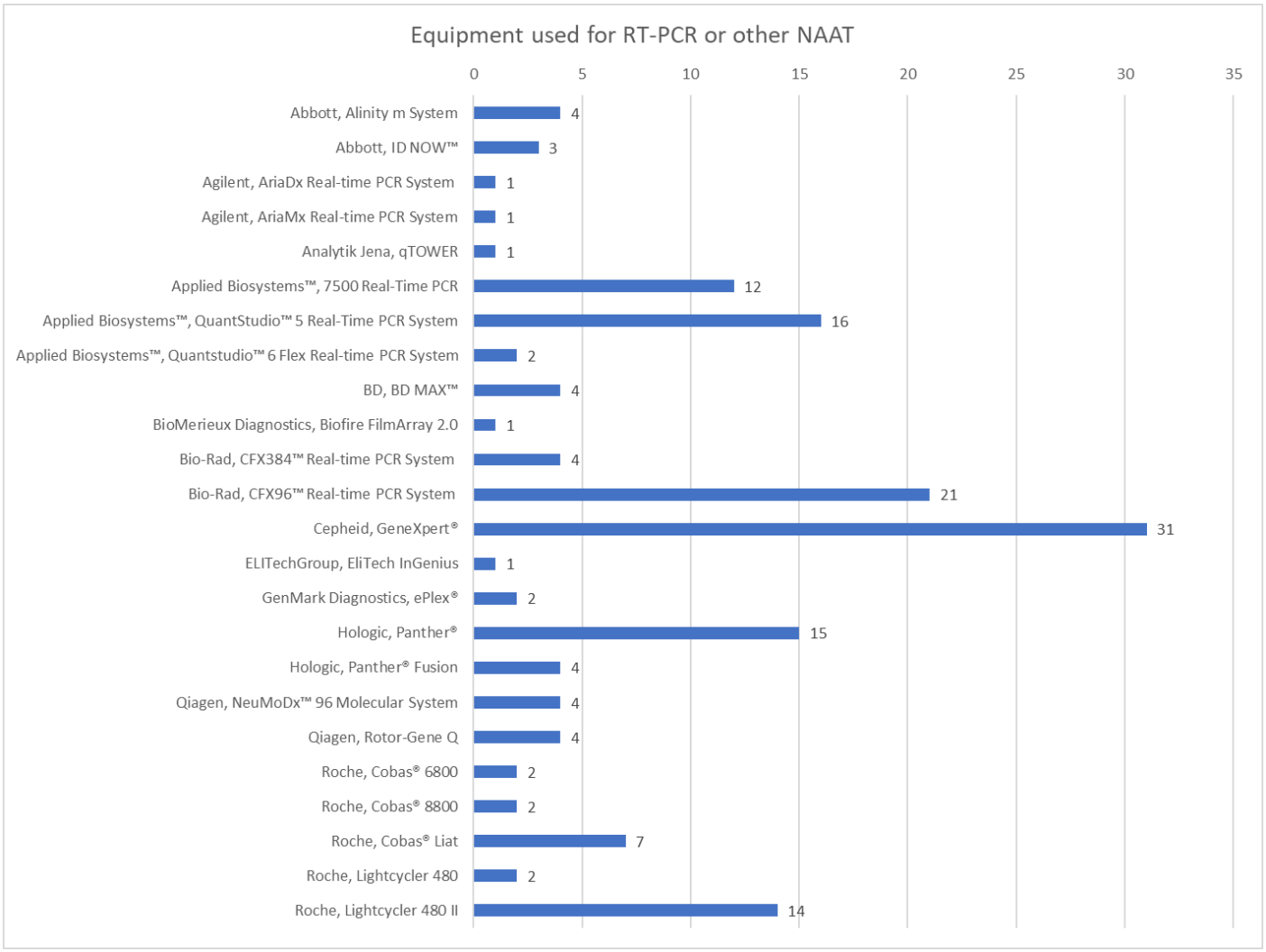

*Supplemental Figure 4: The RNA isolation equipment used by workflows testing for SARS-CoV-2 together with the number of workflows per machine (n=158)*




*Supplemental Figure 5: The RT-qPCR or other NAAT kits used by workflows testing for SARS-CoV-2 together with the number of workflows per kit (n=122). Not all workflows use 'ready to use' kits for their RT-qPCR or other NAAT, so the total N is not equal to the number of workflows tested. For each kit the used target genes are listed. Workflows using separate enzymes and primers and probe are listed in Supplemental Figure 6.*



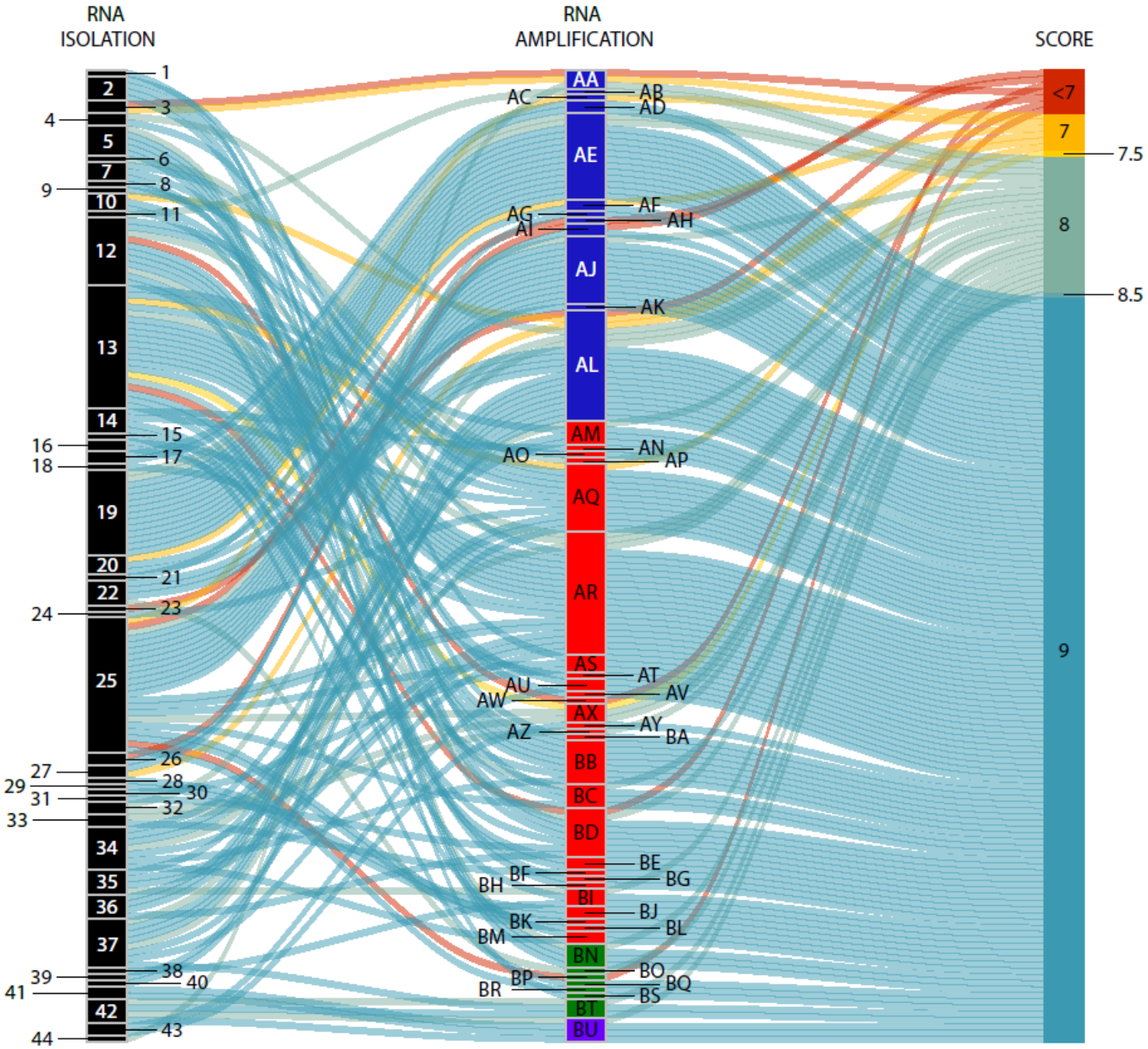

*Supplemental Figure 7: The enzymes used for performing RT-PCR or other NAAT by workflows testing for SARS-CoV-2 together with the number of workflows per enzyme (n=36). Not all workflows use separate enzymes for their RT-PCR or other NAAT, so the total N is not equal to the number of workflows tested. Workflows using complete 'ready to use' kits for SARS-CoV-2 detection are listed in Supplemental Figure 5.*




*Supplemental Figure 6: The RT-PCR or other NAAT equipment used by workflows testing for SARS-CoV-2 together with the number of workflows per machine (n=158)*

## 6.3 Effect of RNA isolation and NAAT reaction on score obtained by workflow

In order to see how well certain combinations of RNA isolation and amplification techniques perform together, an analysis was performed in which these factors were combined with their final score. These data are summarized in Supplemental Figure 8. If for example one type of RNA isolation technique has poor compatibility with a specific RNA amplification technique, you would expect a big band of poor scores flowing through both categories.




*Supplemental Figure 8: A flow diagram showing all workflows reported to have tested the LEQA4 panel with extraction method, PCR test, the number of target genes used and the final score achieved by each workflow. In the alluvial plot PCR tests using 1 target gene are depicted in blue, PCR tests using 2 target genes are shown in red, PCR tests using 3 target genes are shown in green and PCR tests using 4 target genes are shown in purple. Color of trails per workflow are based on the grade obtained for LEQA4. All workflows receiving grades below 7 are grouped in <7.*

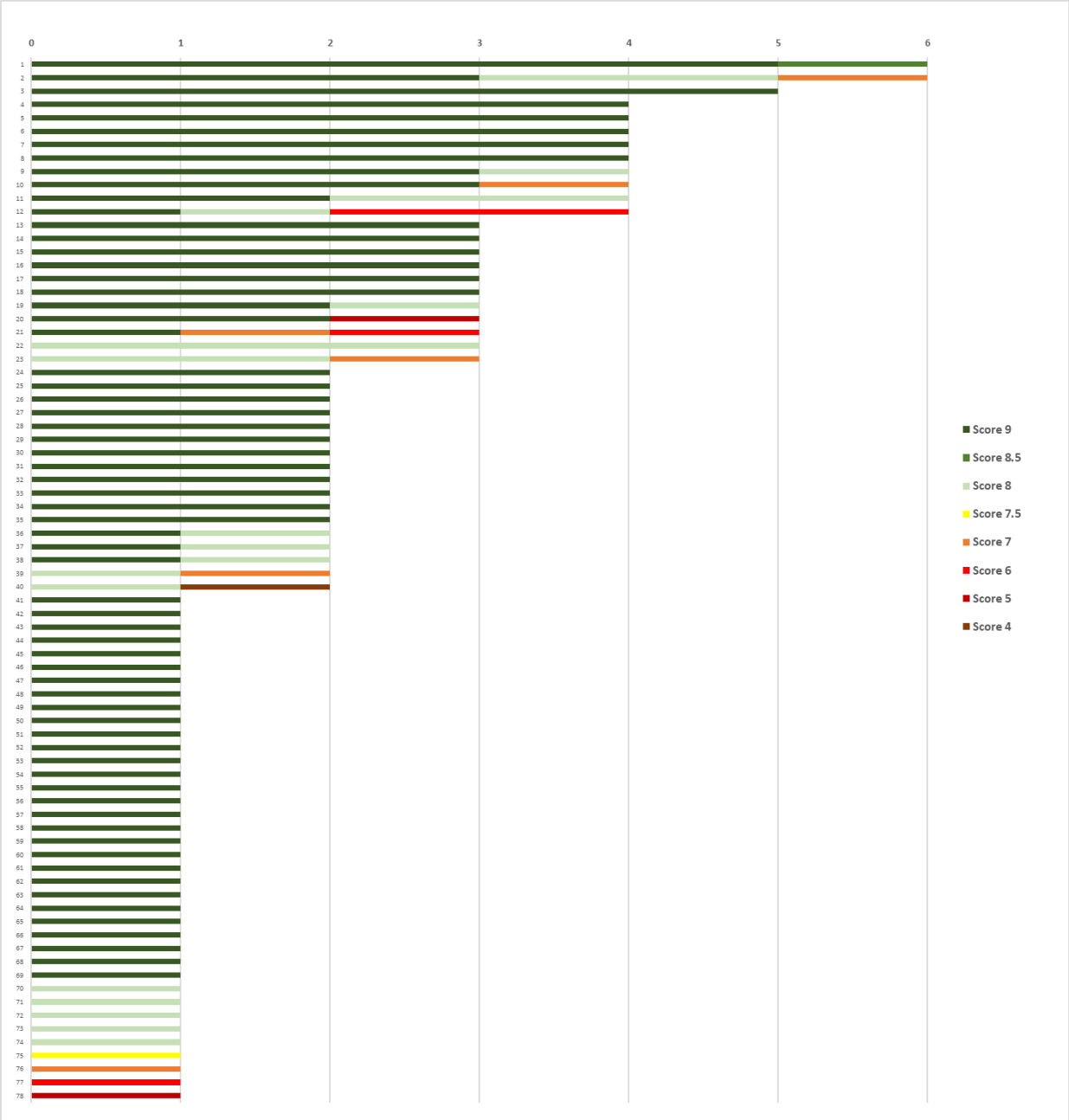
| RNA isolation | |
|---|---|
| 1 | A&A Biotechnology, MagnifiQ TM RNA buffer kit |
| 2 | Abbott, ALINITY m SARS-COV-2 ASSAY |
| 3 | Abbott, ID NOW™ COVID-19 |
| 4 | Abclonal, ANDiS Viral Nucleic RNA Auto Extraction & Purification Kit |
| 5 | Applied Biosystems, MagMAX TM Viral/Pathogen II Nucleic Acid Isolation Kit |
| 6 | BD, BD MAX SARS-CoV-2 kit |
| 7 | BD, BD MAX™ ExK TNA-3 |
| 8 | Bioecho, EchoLUTION viral RNA/DNA Swab 96 Kit |
| 9 | BioMerieux Diagnostics, Respiratory panel 2.1 plus |
| 10 | BioMerieux, NUCLISENS® easyMAG® |
| 11 | Biotecon Diagnostics, Magnetic Preparation Kit VI |
| 12 | Cepheid, Xpert® Xpress SARS-CoV-2 |
| 13 | Cepheid, Xpert® Xpress SARS-CoV-2/Flu/RSV |
| 14 | Chemagen, Chemagic Viral DNA/RNA 300 kit H96 |
| 15 | ElitechGROUP, ELITe InGenius SP 200 |
| 16 | GenMark Diagnostics, ePlex® RP2 |
| 17 | Goldstandard Diagnostics, NovaPrime RNA Extraction AE1 |
| 18 | Hologic, Aptima® SARS CoV2/FLU assay |
| 19 | Hologic, Aptima® SARS-CoV-2 Assay |
| 20 | Hologic, Panther Fusion® Open Access™ |
| 21 | Hologic, Panther Fusion® SARS-CoV-2 assay |
| 22 | Life Technologies, MagMax pathogen RNA/DNA kit |
| 23 | Magcore, MagCore® Viral Nucleic Acid Extraction Kit |
| 24 | Magtivio, MagSi-NA Pathogens |
| 25 | Molgen, PurePrep Pathogens |
| 26 | NeuMoDx™, Flu A-B/RSV/SARS-CoV-2 Vantage Test Strip |
| 27 | NeuMoDx™, RNA mastermix strip |
| 28 | Pathofinder, RealAccurate® Quadruplex SARS-CoV-2 PCR Kit |
| 29 | Procomcure Biotech, SphaeraMag DNA/RNA Isolation Kit - prefilled96 |
| 30 | Promega, Maxwell RSC Blood DNA kit |
| 31 | Qiagen, QIAprep&Amp |
| 32 | Qiagen, QIAsymphony DSP Virus/Pathogen Midi Kit |
| 33 | Qiagen, QIAsymphony DSP Virus/Pathogen Mini Kit |
| 34 | Roche, Cobas® Liat SARS-COV-2/FluA/FluB |
| 35 | Roche, Cobas® SARS-CoV-2: Qualitative assay for use on the cobas 6800/8800 Systems |
| 36 | Roche, DNA and Viral NA Large Volume kit |
| 37 | Roche, DNA and Viral NA Small Volume kit |
| 38 | Roche, MagNA Pure 24 Total Nucleic Acid Isolation Kit |
| 39 | Roche, MagNA Pure 96 Total Nucleic Acid Isolation Kit |
| 40 | Roche, MagNA Pure LC Total Nucleic Acid Isolation Kit |
| 41 | Seegene, STARMag 96x4 Universal Cartridge Kit |
| 42 | Seegene, STARMag 96x4 Viral DNA/RNA 200 C Kit |
| 43 | Siemens, VERSANT Sample Preparation 1.0 Reagents |
| 44 | None |

| RNA amplification | |
|---|---|
| AA | Abbott, ID NOW™ COVID-19 |
| AB | Biotecon Diagnostics, Microproof SARS-CoV-2 Identification Kit |
| AC | Biotecon Diagnostics, Microproof SARS-CoV-2 Screening kit |
| AD | Gold Standard Diagnostics, NovaPrime SARS-CoV-2 (COVID-19) |
| AE | Hologic, Aptima® SARS-CoV-2 Assay |
| AF | Hologic, Panther Fusion® SARS-CoV-2 assay |
| AG | N1-gene, Lu et al. 2020 |
| AH | ORF1a/b (unknown; LAMP assay) |
| AI | Qiagen, NeuMoDx™ Flu A-B/RSV/SARS-CoV-2 Vantage Test |
| AJ | r-Biopharm, RIDAGENE SARS-CoV-2 RUO |
| AK | RdRP, Corman et al. 2020 (adapted) |
| AL | Sarbeco E-gene, Corman et al. 2020 |
| AM | Abbott, ALINITY m SARS-COV-2 ASSAY |
| AN | Altona Diagnostics, Altona Realstar SARS-CoV-2 RT PCR |
| AO | BD, SARS-CoV-2 kit for BD MAX |
| AP | BioMerieux Diagnostics, Respiratory panel 2.1 plus |
| AQ | Cepheid, Xpert® Xpress SARS-CoV-2 |
| AR | Cepheid, Xpert® Xpress SARS-CoV-2/Flu/RSV |
| AS | Certest, VIASURE SARS-CoV-2 Real Time PCR |
| AT | Chemagic, SARS-CoV-2 RT-q PCR Perkin Elmer |
| AU | GenMark Diagnostics, ePlex® RP2 |
| AV | Hologic, Aptima® SARS-CoV-2/FLU assay |
| AW | N1-gene, Lu et al. (2020); ORF1a/b, Lu et al. (2020) |
| AX | Pathofinder, RealAccurate® Quadruplex SARS-CoV-2 PCR Kit |
| AY | Procomcure Biotech, PhoenixDx® SARS-CoV-2 Multiplex IVD |
| AZ | Qiagen, SARS-CoV-2 N1+N2 Assay Kit |
| BA | r-Biopharm, RIDAGENE SARS-CoV-2 RUO |
| BB | Roche, Cobas® Liat SARS-CoV-2/FluA/FluB |
| BC | Roche, Cobas® SARS-CoV-2: Qualitative assay for use on the Cobas 6800/8800 Systems |
| BD | Sarbeco E-gene, Corman et al. 2020; N1-gene, Lu et al. 2020 |
| BE | Sarbeco E-gene, Corman et al. 2020; N2-gene, Lu et al. 2020 |
| BF | Sarbeco E-gene, Corman et al. 2020; N-gene, Lu et al. 2020 |
| BG | Sarbeco E-gene, Corman et al. 2020; RdRP-gene, Corman et al. 2020 |
| BH | Sarbeco E-gene, Corman et al. 2020; RdRP-gene, Corman et al. 2020 (adapted) |
| BI | Siemens, FTD Sars-CoV-2 assay |
| BJ | SpeedX, PlexPCR kit |
| BK | Thermofisher, TaqPath™ COVID-19 CE-IVD RT-PCR |
| BL | TIB-MolBiol, LightMix® Modular Sarbecovirus E-gene, CE |
| BM | TIB-MolBiol, SARS-CoV-2 |
| BN | Applied Biosystems, TaqPath™ COVID-19 Multiplex Diagnostic Solution (CE-IVD) |
| BO | Applied Biosystems, TaqPath™ COVID-19 RT-PCR Kit |
| BP | E-gene (unknown); N-gene (unknown); S-gene (unknown) |
| BQ | ELITechGroup, GeneFinder COVID 19 PLUS Real Amp Kit |
| BR | Gerbion, virellaSARS-CoV-2 seqc |
| BS | PIIM, AmpliGnost E-Gen and RdRP-Gen qPCR RT |
| BT | Seegene, Allplex™ 2019-nCoV Assay |
| BU | Seegene, Allplex™ SARS-CoV-2 Assay |

*Supplemental figure 8 continued: Legend*

## 6.4 Scores obtained per laboratory

Here all obtained scores per workflow per laboratory are summarized. In total two laboratory reported data for 6 workflows, one laboratory reported data for 5 workflows, nine laboratory reported data for 4 workflows, eleven laboratories reported data for 3 workflows, seventeen laboratories reported data for 2 workflows and thirty-eight laboratories reported data for 1 workflow. There are two laboratories which only have scores of < 7 for all reported workflows. The obtained scores per workflow are sorted (anonymously) per laboratory and shown in Supplemental Figure 9. Of the 78 laboratories, 65 laboratories reported at least one workflow with fully correct results.




*Supplemental Figure 9: Grades obtained per workflow per laboratory (anonymized). For each of the laboratories the number of reported workflows is shown on the X-axis together with their accompanying grades. In total 78 laboratories sent in data of their workflows testing LEQA4. There are two laboratories which only have scores of < 7 for all reported workflows.*

## 6.5 Participating laboratories

All participating laboratories are listed below. We would like to thank colleagues from these laboratories for their participation in this round of LEQA for the Dutch SARS-CoV-2 diagnostics field.

| **Laboratory name** |
|---|
| Amsterdam UMC |
| Antoniusziekenhuis Nieuwegein |
| ArminLabs |
| Atalmedial |
| Canisius Wilhelmina Ziekenhuis |
| Catharina Ziekenhuis |
| CBSL Tergooi |
| Certe |
| Comicro B.V. |
| Deventer Ziekenhuis |
| Diagnostiek voor u |
| Diakonessenhuis Utrecht |
| Erasmus MC |
| Eurofins Genomics Europe Applied Genomics GmbH |
| Eurofins HVL |
| Eurofins Medische Microbiologie |
| Eurofins NMDL-LCPL |
| Fenelab Consortium - Merieux Nutrisciences |
| Fenelab Consortium - NofaLab |
| Fenelab Consortium - Normec Biobeheer |
| Fenelab Consortium - Nutreco - MasterLab |
| Fenelab Consortium - NutriControl |
| Fenelab Consortium - Nutrilab B.V. |
| Fenelab Consortium - SGS |
| Fenelab Consortium - Triskelion |
| Franciscus Gasthuis en Vlietland |
| Gelre Ziekenhuizen Apeldoorn |
| GGD Amsterdam |
| Groene Hart ziekenhuis |
| Haaglanden Medisch Centrum |
| Hagaziekenhuis |
| Howareyou diagnostics |
| IJssellandziekenhuis |
| Ikazia ziekenhuis |
| inBiome |
| Isala |
| Izore |
| Jeroen Bosch Ziekenhuis |
| LabMicTA |
| Labor Wisplinghoff |
| Laboratorium Healthcare Medical |
| Laurentius ziekenhuis |
| Leiden University Medical Center |
| Maasstadziekenhuis |
| Maastricht UMC+ |

| Laboratory name |
| --- |
| Meander Medisch Centrum |
| Medische Laboratoria Dr. Stein & Collegea |
| Microbe&Lab |
| Microvida |
| Microvida ETZ |
| Mozand B.V. |
| Noordwest Ziekenhuisgroep Alkmaar |
| Novogenia GmbH |
| OLVG Lab B.V. |
| Pure Medical |
| Radboudumc |
| Reinier Haga MDC |
| Rijnstate Velp |
| RLM |
| Royal GD |
| Saltro, locatie Hudsondreef |
| Saltro, locatie Mississippidreef |
| Sanquin, NSS lab |
| Star-SHL |
| Stichting PAMM |
| Streeklab Haarlem |
| Streekziekenhuis Koningin Beatrix Winterswijk |
| Synlab Heppignies |
| Synlab Jena Oncoscreen Coronalabor |
| Synlab Laboratoire Collard |
| Synlab Leverkusen |
| Synlab Trier |
| TLR International Laboratories |
| UMCU |
| VieCuri MC |
| Wageningen Bioveterinary Research |
| Ziekenhuis Gelderse Vallei |
| Zuyderland Medisch Centrum |