Title: High specificity and variable sensitivity in molecular detection of SARS-CoV-2 in European Expert Laboratories: External Quality Assessment, June – July 2020

During the ongoing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic about 23.7 million humans have been infected and 814,438 have died globally until the end of August (update WHO ref). As both specific medication and approved vaccines are not available yet, public health strategies focus on containment measures. Robust detection of acutely SARS-CoV-2 infected patients, typically done by real-time reverse transcription polymerase chain reaction (RT-PCR), is crucial to interrupt infection chains and thus to control the pandemic (1). An established tool to improve and support diagnostic quality is the conduction of external quality assessments (EQA) (2-4). In cooperation with the European Centre for Disease Prevention and Control (ECDC) an EQA of molecular detection of SARS-CoV-2 was organized for European expert laboratories. In April 2020 laboratories that were either part of the Emerging Viral Diseases-Expert Laboratory Network (EVD-LabNet) or of the European Reference Laboratory Network for Human Influenza (ERLI-Net) were invited to participate in the EQA.

The EQA panel was composed of 12 samples containing either SARS-CoV-2 at different concentrations to evaluate the sensitivity or samples containing other respiratory viruses to evaluate the specificity of SARS-CoV-2 testing of the participants (**Table 1**). SARS-CoV-2 samples were quantified using an E gene assay (5) and an in vitro-transcribed RNA standard. Accurate virus contents were confirmed by RdRp and N gene assays (5). Furthermore, two independent reference laboratories validated the EQA panel. Samples were diluted in FCS-free dulbecco's modified eagle's medium. All samples were heat-inactivated (65°C, 4h for SARS-CoV, SARS-CoV-2, MERS-CoV and *adenovirus*, 2h for all other viruses) and freeze-dried

before shipment on June 10th and June 11th. Successful inactivation was confirmed by absence of viral growth in three cell culture passages. Together with the EQA panels all participants received testing instructions, asking them to treat and test the EQA samples according to their routine molecular diagnostics workflow, but also including a recommendation on how to reconstitute the lyophilised samples.

Results

Laboratories in 29 EU/EEA countries (excluding Liechtenstein, no partner laboratory of EVD-LabNet), seven EU pre-accession countries and two other European countries were invited to participate in this EQA. 68 laboratories from 35 countries, i.e. 29 EU/EEA countries, five EU pre-accession countries and two other European countries, reported EQA results and technical details on testing within the scheduled timeframe (**Figure 1**). Of the 12 samples from the EQA panel, all samples were defined as core samples, except for sample no. 9, which contained SARS-CoV virus. Because some molecular SARS-CoV-2 assays cross-detect SARS-CoV, which is not circulating in human populations since 2004 (6), the SARS-CoV sample was excluded from statistical analyses in this study.

In the online result submission form, laboratories could select three result options: 'SARS-CoV-2 RNA detected', No SARS-CoV-2 RNA detected' or 'inconclusive'. Additionally, comments about the samples could be left. In total, laboratories selected the option 'inconclusive' for 3.0% of all samples. Most inconclusive results (56.5%) were reported for samples with the lowest SARS-CoV-2 concentration representing late, flattened, or unclear amplification curves. However, due to different test interpretations of similar test outcomes, results were scored not on the self-selected result declaration by the laboratories, but by using all information given for each sample, including the comment section. A sample was

scored positive, if at least one of the used SARS-CoV-2 tests gave a positive result. This was done for all core samples of the EQA.

39.7% of participating laboratories tested all samples correctly (**Figure 2A**), 30.9% reported correct results for 10/11 samples, 17.6% for 9/11 samples and 11.8% for 8/11 samples or less. False results were reported for 9.5% of all samples with 8.5% false-negative results reported for SARS-CoV-2 samples and 1.0% false-positive results reported for specificity controls (**Table 1**). As expected, the risk of false-negative tests increased significantly with lower SARS-CoV-2 concentrations ($p < 0.001$, Spearman correlation test) (**Figure 2 B**). Accordingly, Ct values of correctly tested SARS-CoV-2 samples increased with decreasing concentration. Median Ct values among all real-time RT-PCR assays were 25.9 (25.7-26.6, 95% CI) for sample 11 (2,500 cps/µL), 31.6 (31.1-32.1, 95% CI) for sample 12 (180 cps/µL), 34.0 (33.6-34.6, 95% CI) for sample 2 and 10 (12.5 cps/µL), and 36.4 (35.6-36.9, 95% CI) for sample 1 and 6 (2 cps/µL) (**Figure 2B**). 92.6% of all participants tested all specificity control samples correctly showing an overall good specificity in molecular SARS-CoV-2 diagnostics (**Figure 2A**). No included respiratory virus was more prone for false positive test results than others (**Table 1**).

Extraction method and extraction protocols can affect the overall diagnostic performance (2). Participants were thus asked for details about extraction methods, extraction volumes and elution volumes. Within this EQA, the diagnostic performance among laboratories conducting manual nucleic acid (NA) extraction (92.4% correct tests) and laboratories conducting automated extraction (89.8% correct tests) did not differ significantly ($p = 0.350$, Mann–Whitney U test). The EQA performance was neither corelated to the usage of extraction kits ($p = 0.938$, Kruskal–Wallis test) and ranged between 87.9% and 93.8% correct

results (**Table 2**). Most extraction kits and protocols increase the concentration of pathogen NA to improve diagnostic sensitivity. Within this study, increased concentrating of NA during extraction was not correlated with better diagnostic sensitivity ($p = 0.898$, Spearman correlation test). While NA extraction is a critical step during molecular diagnostics, different protocols and kits provide the same efficiency.

Apart from NA extraction, the specific detection of NA by real-time PCR is crucial in molecular diagnostics and the diagnostic performance may vary among different assays (7). In total, 26 different commercial assays and six different in-house assays were used by the EQA participants. Among those assays that were performed by at least 5 EQA participants, correct results ranged between 80% and 92% (**Table 3**). The diagnostic performance was not significantly correlated with PCR assays in general ($p = 0.525$, Kruskal–Wallis test). However, results were significantly better with the best performing test (IP RdRP gene) compared to "Others" ($p = 0.042$, Wilcoxon rank sum test), Corman N assay ($p = 0.008$, Wilcoxon rank sum test), and the Viasure ORF1 assay ($p = 0.011$, Wilcoxon rank sum test). The performance among commercial tests (84.6% correct results) and inhouse tests (85.7% correct results) was highly similar ($p = 0.969$, Wilcoxon rank sum test). As transcription in coronaviruses typically varies among different genomic and sub-genomic regions, target sites of the applied assays may affect the diagnostic sensitivity (8). Correct results did not differ significantly among different target sites of the real-time RT-PCRs ($p = 0.852$, Kruskal–Wallis test), being 84.9% for the E gene, 85.2% for the N gene, 82.5% for the ORF1 gene, 86.2% for the RdRp gene and 83.7% for the S gene (**Figure 2C**). Accordingly, the target region of SARS-CoV-2 real-time RT-PCR assays is no marker for good diagnostic performance.

Beyond technical details the quality of laboratory infrastructures may affect the diagnostic capacity. It was thus analysed whether the EQA performance was corelated with socio-

economic factors represented by the human development index (HDI). The performance in this EQA was clearly not corelated to the HDI ($p = 0.814$, Spearman correlation test) (**Figure 2D**) suggesting that the diagnostic capacity of laboratories cannot be assumed by national prosperity indicators.

Discussion

The overall performance of participating laboratories within this EQA was good. Most inconclusive or false results were reported for low concentrated SARS-CoV-2 samples. Although some laboratories should optimize workflows to achieve a better sensitivity, slightly reduced sensitivity is probably less dramatic regarding SARS-CoV-2 containment as infectious SARS-CoV-2 patients have commonly very high viral titres (9, 10). Low concentrations in clinical samples are often seen at a later infection stage that coincides with seroconversion (11). The inclusion of serological tests such as enzyme-linked immunosorbent assays to diagnostic workflows may thus be suitable to further reduce the risk of false-negative tests.

11.8% of the participating laboratories showed problematic EQA performances including five laboratories which had false-positive results. Notably, the laboratories reporting false positive results applied both extraction kits and RT-PCR tests that were also used by other laboratories. False-positive test results could thus be a consequence of contamination during sample handling and extraction or LOT-specific contamination of RT-PCR kits or oligonucleotides which has recently been reported (12). In both cases laboratories should adapt workflows to ensure good specificity which in general was excellent among participants. Adaptations might include pre-validation of new oligonucleotide batches and careful revision of protocols.

Compared to the performance of diagnostic laboratories in a recently published SARS-CoV-2 EQA the sensitivity of participants in this EQA was lower (4). This may indicate that the diagnostic capability in specialized diagnostic laboratories is higher compared to public reference laboratories. Accordingly, national reference laboratories should get further support to increase the diagnostic capability. Nevertheless, it is not uncommon for laboratories to show suboptimal performances in EQAs for molecular diagnostics of recently emerged viruses, particularly regarding a limited sensitivity (2, 13, 14). Notably, the results indicate that the diagnostic performance can rather be increased by harmonized workflows than by the selection of specific extraction or real-time RT-PCR kits. This study provides additional information for reference and diagnostic laboratories to optimize diagnostic workflows.

The conduction of follow-up EQA will be essential to support laboratories to systematically improve molecular diagnostic capabilities to detect SARS-CoV-2. Supporting this, there was considerable interest for further EQA panels from EQA participants and other laboratories.

References

1. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health. 2020;8(4):e488-e96.
2. Fischer C, Pedroso C, Mendrone A, Jr., Bispo de Filippis AM, Vallinoto ACR, Ribeiro BM, et al. External Quality Assessment for Zika Virus Molecular Diagnostic Testing, Brazil. Emerg Infect Dis. 2018;24(5).
3. Reusken CB, Mogling R, Smit PW, Grunow R, Ippolito G, Di Caro A, et al. Status, quality and specific needs of Ebola virus diagnostic capacity and capability in laboratories of the two European preparedness laboratory networks EMERGE and EVD-LabNet. Euro Surveill. 2018;23(19).
4. Matheeussen V, Corman VM, Donoso Mantke O, McCulloch E, Lammens C, Goossens H, et al. International external quality assessment for SARS-CoV-2 molecular detection and survey on clinical laboratory preparedness during the COVID-19 pandemic, April/May 2020. Euro Surveill. 2020;25(27).

5. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020;25(3).
6. Peeri NC, Shrestha N, Rahman MS, Zaki R, Tan Z, Bibi S, et al. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? Int J Epidemiol. 2020.
7. Igloi Z, Leven M, Abdel-Karem Abou-Nouar Z, Weller B, Matheeussen V, Coppens J, et al. Comparison of commercial realtime reverse transcription PCR assays for the detection of SARS-CoV-2. J Clin Virol. 2020;129:104510.
8. Thi Nhu Thao T, Labroussaa F, Ebert N, V'Kovski P, Stalder H, Portmann J, et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature. 2020;582(7813):561-5.
9. La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, et al. Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards. Eur J Clin Microbiol Infect Dis. 2020;39(6):1059-61.
10. Kim SE, Jeong HS, Yu Y, Shin SU, Kim S, Oh TH, et al. Viral kinetics of SARS-CoV-2 in asymptomatic carriers and presymptomatic patients. Int J Infect Dis. 2020;95:441-3.
11. Wolfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Muller MA, et al. Virological assessment of hospitalized patients with COVID-2019. Nature. 2020;581(7809):465-9.
12. Mogling R, Meijer A, Berginc N, Bruisten S, Charrel R, Coutard B, et al. Delayed Laboratory Response to COVID-19 Caused by Molecular Diagnostic Contamination. Emerg Infect Dis. 2020;26(8):1944-6.
13. Niedrig M, Linke S, Zeller H, Drosten C. First international proficiency study on West Nile virus molecular detection. Clin Chem. 2006;52(10):1851-4.
14. Pas SD, Patel P, Reusken C, Domingo C, Corman VM, Drosten C, et al. First international external quality assessment of molecular diagnostics for Mers-CoV. J Clin Virol. 2015;69:81-5.

Tables

Table 1. EQA panel composition and correct test results of participating laboratories

Sample information			Correct results	
Sample ID	Virus[a]	Cps/µL	%	n
1	SARS-CoV-2	2	60%	41
2	SARS-CoV-2	12.5	99%	67
3	HMPV A, HMPV B, PIV1, PIV2, PIV3, PIV4, RSV	500 2,000 2,000 2,000 2,000 2,000 1,500	96%	65
4	MERS-CoV	5,000	97%	66
5	*Enterovirus*, *Rhinovirus*, *Influenzavirus A* H1N1 *Influenzavirus B* *Adenovirus*	2,000 1,000 2,000 500 2,000	100%	68
6	SARS-CoV-2	2	60%	41
7	HCoV-229E, HCoV-NL63	1,000 2,000	97%	66
8	HCoV OC43	2,000	98%	65
9	SARS-CoV	5,000	89%	59
10	SARS-CoV-2	12.5	90%	61
11	SARS-CoV-2	2,500	99%	67
12	SARS-CoV-2	180	99%	67

[a] *Enterovirus*, *rhinovirus*, *respiratory syncytial virus* (RSV), adenovirus, *Influenza A virus* H1N1, *Influenza B virus*, *middle east respiratory syndrome coronavirus* (MERS-CoV), *severe acute respiratory syndrome coronavirus* (SARS-CoV) and *severe acute respiratory syndrome*

coronavirus 2 (SARS-CoV-2) BetaCoV/Munich/ChVir984/2020 were grown on Vero cells. *Parainfluenza 1* (PIV1), *parainfluenza 2* (PIV2), *parainfluenza 3* (PIV3), *parainfluenza 4* (PIV4), *human coronavirus 229E* (hCoV-229E), and *human coronavirus OC43* (hCoV-OC43) were grown on CaCo-2 cells. *Human metapneumovirus A* (hMPV A), *human metapneumovirus B* (hMPV B), and *human coronavirus NL63* (hCoV-NL63) were grown on LLC-MK2 cells.

Table 2. Comparison of different extraction methods in EQA performance

Extraction method	# Tested samples	Correct results
Other method	286	89.4%
Roche Cobas Omni	55	89.1%
EZ1 Virus Mini Kit v2.0	32	93.8%
MagMAX Viral/Pathogen Nucleic Acid Isolation Kit	33	90.9%
MagNA Pure Compact Nucleic Acid Isolation Kit I	33	87.9%
MagNA Pure 96 DNA and Viral NA SV Kit	44	90.9%
BioMérieux NucliSENS easyMAG	88	89.8%
QIAamp Viral RNA Mini Kit	143	92.3%
QIAamp Viral RNA Mini QIAcube Kit	54	92.6%

None of the extraction methods performed significantly better than "Other methods" considering two-sided Yates corrected chi square.

Table 3. Performance among different real-time RT-PCR assays. None of the assays performed significantly better than "Others" considering two-sided Yates corrected chi square. Sample 9 was excluded for calculation of the EQA performances.

RT-PCR assay		Genome target		No. of submitted EQA results	Core samples												Performance (core samples)		
					SARS-CoV-2 [copies/ml]						hCoV-229E hCoV-NL63	hCoV-OC43	MERS-CoV	Respiratory panel A	Respiratory panel B	SARS-CoV			
Name	Method type	Gene	Specificity		2,50E+06	1,80E+05	1,25E+04	1,25E+04	2,00E+03	2,20E+03							Total correct	False pos.	False neg.
					Sample 11	Sample 12	Sample 2	Sample 10	Sample 1	Sample 6	Sample 7	Sample 8	Sample 4	Sample 3	Sample 5	Sample 9			
Corman *et al.*	in house	E	Sarbeco	31	100%	100%	97%	87%	35%	39%	97%	100%	100%	94%	100%	100%	88%	2%	24%
Corman *et al.*	in house	RdRp	SARS-CoV-2	11	100%	100%	100%	73%	18%	36%	100%	100%	100%	100%	100%	90%	84%	0%	29%
CDC	in house	N	SARS-CoV-2	10	100%	100%	90%	70%	60%	70%	100%	100%	100%	100%	100%	100%	90%	0%	18%
IP	in house	RdRp	SARS-CoV-2	10	100%	100%	100%	90%	70%	50%	100%	100%	100%	100%	100%	88%	92%	0%	18%
Allplex 2019-nCoV Assay	commercial	E	Sarbeco	9	100%	89%	89%	56%	44%	56%	100%	100%	100%	100%	100%	67%	85%	0%	28%
Allplex 2019-nCoV Assay	commercial	N	SARS-CoV-2	9	100%	100%	78%	89%	56%	56%	100%	100%	89%	100%	100%	100%	88%	2%	20%
Allplex 2019-nCoV Assay	commercial	RdRp	SARS-CoV-2	9	100%	89%	78%	67%	44%	33%	100%	100%	100%	100%	100%	100%	83%	0%	31%
RealStar SARS-CoV-2 RT-PCR Kit 1.0	commercial	E	Sarbeco	5	100%	100%	100%	100%	40%	40%	100%	100%	100%	100%	100%	100%	89%	0%	20%
RealStar SARS-CoV-2 RT-PCR Kit 1.0	commercial	S	SARS-CoV-2	5	100%	100%	100%	100%	20%	40%	100%	100%	100%	100%	100%	100%	87%	0%	23%
Cobas SARS-CoV-2 Test	commercial	E	Sarbeco	5	80%	100%	100%	60%	40%	40%	100%	100%	100%	100%	100%	80%	84%	0%	30%
Cobas SARS-CoV-2 Test	commercial	ORF1ab	SARS-CoV-2	5	80%	100%	100%	60%	80%	40%	100%	100%	100%	100%	100%	100%	87%	0%	23%
Corman *et al.*	in house	N	Sarbeco	5	100%	100%	100%	80%	0%	0%	100%	100%	100%	100%	100%	100%	85%	0%	22%
Viasure SARS-CoV-2 Real Time PCR Detection Kit	commercial	N	SARS-CoV-2	5	100%	100%	100%	100%	0%	40%	100%	100%	100%	80%	100%	80%	84%	4%	27%
Viasure SARS-CoV-2 Real Time PCR Detection Kit	commercial	ORF1ab	SARS-CoV-2	5	100%	80%	100%	100%	0%	0%	100%	100%	100%	100%	100%	100%	80%	0%	37%
Others	both	various	n.a.	38	97%	92%	79%	71%	37%	39%	100%	97%	100%	95%	100%	n.a.	82%	6%	31%

Figures

Figure 1. Overview of countries of the participating laboratories in SARS-CoV-2 molecular EQA, June/July 2020.

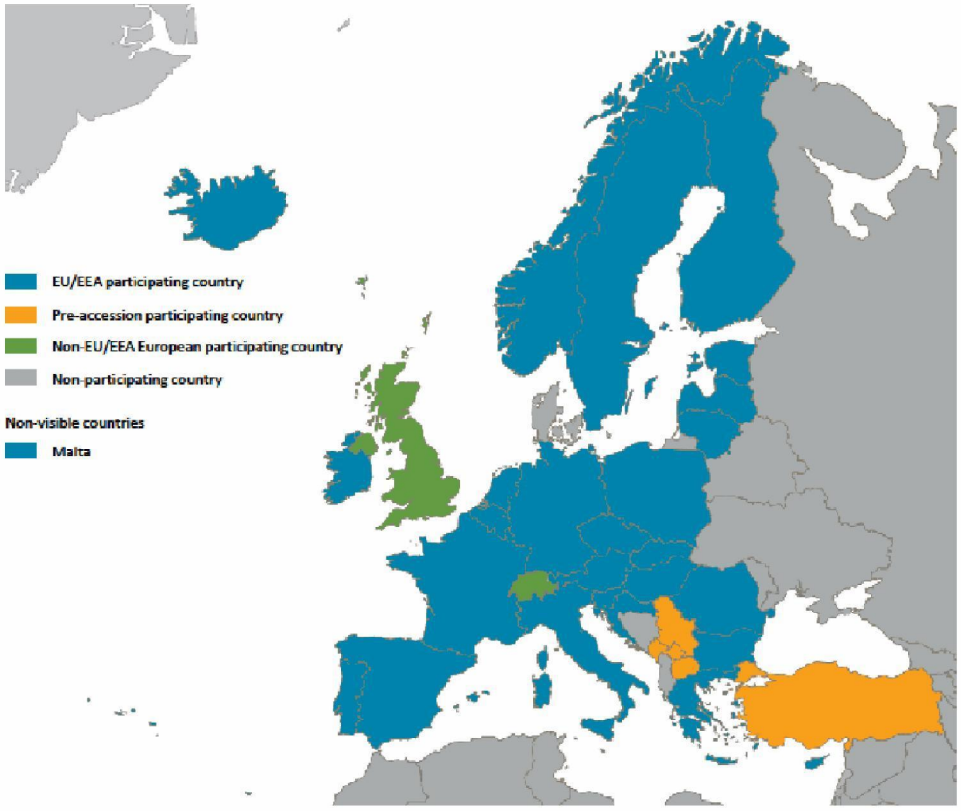

Figure 2. EQA performance among laboratories and in relation to specific factors.

A, the number of correct, false negative and false positive samples per laboratory. B, reported Ct values per SARS-CoV-2 positive samples considering all results reported including multiple tests conducted by one participant. Median Ct values are indicated by bars, quartiles by boxes, and interquartile range by whiskers. Single events are indicated by grey (positive) and purple (negative) dots. C, EQA performance depending on different real-time RT-PCR targets. For boxplot explanation see B. D, EQA performance depending on the countries human development index (HDI) of EQA participants. Results are visualized binned to protect the anonymity of the participants. For boxplot explanation see B.