ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL

SURVEILLANCE & MONITORING

External quality assessment of European SARS-CoV-2 and zoonotic influenza Bioinformatics (ESIB-EQA), 2024

ECDC SURVEILLANCE & MONITORING

External quality assessment of European SARS-CoV-2 and zoonotic influenza Bioinformatics (ESIB-EQA), 2024

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC) under the framework contract ECDC/2022/001 and specific contract No.11 (request for services RS/2023/DPR/26559), coordinated by Olov Svartström and Eeva Broberg, and produced by AURORAE Consortium.

Authors
Frédéric Lemoine, National Reference Center for Respiratory viruses, Institut Pasteur, Paris, France; Etienne Simon-Lorière, Evolutionary genomics of RNA viruses, Institut Pasteur, Paris, France; Ivo Van Walle, Dutch National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands.

Suggested citation: European Centre for Disease Prevention and Control. External quality assessment of European SARS-CoV-2 and zoonotic influenza Bioinformatics (ESIB-EQA), 2024. Stockholm: ECDC; 2025.

Stockholm, October 2025

ISBN 978-92-9498-804-1
doi: 10.2900/5914884
Catalogue number TQ-01-25-034-EN-N

Contents

Figures

Tables

Abbreviations

CDS	Coding sequence
COVID-19	Coronavirus disease 2019
EQA	External quality assessment
GR	Wuhan-1 global reference
GS	Comparator sequence
HRI	Highly reduced inhibition
IAV	Influenza A virus
INDEL	Insertion/Deletion
INFL1-4	Influenza components 1 to 4
IQR	Interquartile range
NI	Normal inhibition
NA	Influenza neuraminidase protein
PA	One of three influenza virus polymerase proteins
PB1	One of three influenza virus polymerase proteins
PB2	One of three influenza virus polymerase proteins
OCPs	Operational contact points
ORF	Open reading frame
RI	Reduced inhibition
RIVM	National Institute for Public Health and the Environment, Bilthoven, the Netherlands
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SARS1-4	SARS-CoV-2 components 1 to 4
SNP	Single nucleotide polymorphism

Executive summary

SARS-CoV-2 and influenza viruses are significant public health threats with varying impacts on global health. SARS-CoV-2, the virus behind the COVID-19 pandemic, exhibits high transmissibility and recurrent waves driven by new variants. Recent outbreaks of avian influenza in poultry and dairy cattle with cases of human infections highlight the importance of monitoring mutations and reassortment events in zoonotic influenza strains that may lead to increased human infections. Both SARS-CoV-2 and influenza challenge public health efforts due to their continuous evolution, necessitating ongoing surveillance, vaccine development, and strategic response measures.

Surveillance and laboratory networks

Surveillance for SARS-CoV-2 and influenza involves data collection, analysis, and interpretation to guide public health actions. The European Centre for Disease Prevention and Control sets surveillance objectives, including tracking disease trends, detecting outbreaks, evaluating prevention programs, identifying high-risk groups, and generating research hypotheses. External quality assessment (EQA) programmes support these objectives by ensuring laboratory accuracy and reliability.

ECDC coordinates the laboratory networks: The European COVID-19 reference laboratory network (ECOVID-LabNet) and the European Reference Laboratory Network for Human Influenza (ERLI-Net), both intended to enhance virological surveillance, laboratory capacity, and data sharing across Europe. These networks play a pivotal role in case detection, confirmation, and characterisation, contributing to effective public health responses.

Bioinformatics external quality assessment

The European SARS-CoV-2 and zoonotic influenza bioinformatics external quality assessment (ESIB-EQA) assesses the capability of laboratories in public health bioinformatics for SARS-CoV-2 and zoonotic influenza. It aims to provide troubleshooting advice and improve quality through training activities. This comprehensive approach to surveillance and quality assessment ensures preparedness for pandemics and major public health events, strengthening the overall capacity to manage SARS-CoV-2 and influenza effectively. This second iteration, conducted in 2024, focused on consensus sequence generation and quality control for both viruses, revealing several areas for improvement in laboratory practice.

Key findings

Key findings of this second iteration of the ESIB-EQA include:

- **Consensus sequence generation:** Issues included improper handling of co-infections, incorrect assessment of sequence quality, reliance on global reference sequences, and incomplete primer trimming. These issues were more prevalent in influenza virus analysis due to many possible references and software limitations.
- **Clustering and classification:** Laboratories often failed to consider indels in phylogenetic analyses and used outdated data definitions. For influenza, poor quality consensus sequences led to classification errors.
- **Mutation detection:** SARS-CoV-2 mutation detection showed no major issues, while influenza mutation analysis was challenging, mainly due to the lack of standardised datasets.

Feedback and participation

Feedback from participating laboratories indicated general satisfaction with the EQA's organisation and usefulness. Participation was robust among EU/EEA countries but limited in the Western Balkans and Türkiye.

Conclusion and recommendations

Issues from the first iteration of the ESIB-EQA persisted, particularly in consensus sequence generation and phylogenetic clustering. Improvements in consensus sequence generation, particularly for influenza, are needed. These include using complete reference sequences, properly masking primers, and addressing defective genomes. Clearer guidelines for mutation analysis and classification criteria are also recommended to enhance laboratory performance and public health response.

1 Introduction

SARS-CoV-2 and influenza viruses are pathogens of major public health impact [1-4]. These viruses pose a significant threat to global health, but their impact varies in terms of transmission, severity, and response measures.

SARS-CoV-2, the virus responsible for the COVID-19 pandemic, has demonstrated high transmissibility, leading to widespread outbreaks and a significant burden on healthcare systems worldwide. As of yet, there is no evidence of a seasonal pattern for SARS-CoV-2. Instead, new variants replace circulating ones [5] and their public health risk needs to be evaluated individually. Influenza viruses are responsible for seasonal outbreaks and, in some cases, severe epidemics.

While the impact of influenza and SARS-CoV-2 on populations are mitigated by vaccine development efforts, the continuous evolution of both viruses poses challenges for effective prevention and control efforts. Zoonotic influenza should be monitored, as sporadic transmission events to humans do occur. In 2024, large outbreaks of avian influenza in dairy cattle and poultry in North America have led to several cases of exposed humans becoming infected [6]. Genomic analysis of isolated viruses has revealed concerning mutations, such as those associated with mammalian adaptation and enhanced replication. Such mutations and reassortment events may increase the risk of infections in or among humans.

Understanding the public health impact of both SARS-CoV-2 and influenza is crucial for informing policy decisions, resource allocation, and the development of effective strategies to mitigate their respective impacts on population health. In this context, surveillance plays a vital role in monitoring the spread and impact of SARS-CoV-2 and influenza [7].

Surveillance involves the collection, analysis, and interpretation of data to inform public health actions. The European Centre for Disease Prevention and Control recognises the importance of disease surveillance and has set specific surveillance objectives to address these pathogens [8]:

- Continuously track and analyse trends in communicable diseases across Member States;
- Detect and monitor outbreaks of communicable diseases occurring across multiple countries, including SARS-Cov-2 and Influenza;
- Contribute to the evaluation and monitoring of prevention and control programs;
- Identify high-risk population groups; and
- Generate hypotheses and identify research needs.

External quality assessment (EQA) programmes are designed to assess and improve the performance of laboratories involved in disease surveillance. These programmes have specific objectives that align with and contribute to ECDC's surveillance objectives. The primary objectives of EQAs include evaluating the accuracy and reliability of laboratory data analysis methods and identifying areas for improvement. By ensuring standardised and high-quality laboratory testing, EQA programmes aim to optimise laboratory performance and ultimately enhance overall surveillance efforts.

As part of this strategy, ECDC concluded a framework contract through public procurement that aimed to provide laboratory support for surveillance, preparedness and response to COVID-19 and influenza, and to specifically strengthen the capacity for genomic epidemiology and public health bioinformatics through interdisciplinary trainings. All these activities are crucial for response during a pandemic or unexpected major public health events. The geographic scope of these activities covers European Union/European Economic Area (EU/EEA) countries, the Western Balkans (Albania, Kosovo, Montenegro, Serbia, North Macedonia, Bosnia and Herzegovina) and Türkiye.

To effectively monitor and respond to the public health impact of SARS-CoV-2 and influenza, a robust laboratory network is crucial. ECDC coordinates the European COVID-19 reference laboratory network (ECOVID-LabNet) and the European Reference Laboratory Network for Human Influenza (ERLI-Net), both aiming to carry out virological surveillance, strengthen laboratory capacity, and foster data sharing – for example, through conducting EQAs [9,10]. These networks comprise laboratories across Europe, which play a pivotal role in supporting disease surveillance by contributing to the detection, confirmation, and characterisation of cases. Through their coordinated efforts, these laboratory networks strengthen the surveillance infrastructure and enable the generation of accurate and reliable data for effective public health response.

The main objectives of this bioinformatics EQA were:

- To assess the capability of participating laboratories for public health bioinformatics for SARS-CoV-2 and zoonotic influenza, and
- To provide troubleshooting advice where needed and possible (larger needs for improving quality are expected to be fulfilled through other training activities).

This bioinformatics EQA consisted of different components designed to each independently test one particular type of analysis for SARS-CoV-2; for influenza virus, the focus was on strains of zoonotic origin, and here a scenario-based approach was adopted whereby the same samples (reads) were used for components that were equivalent to those applied for SARS-CoV-2 (Table 1). The influenza virus components were managed by Institut Pasteur and the SARS-CoV-2 components by the Dutch National Institute for Public Health and the Environment (RIVM).

This second iteration of the bioinformatics EQA was conducted in 2024; the first was executed one year earlier in 2023. For influenza virus, the 2024 EQA was set up using scenarios specifically for viruses of zoonotic origin. For SARS-CoV-2, this EQA was set up in the same way as in the 2023 EQA [11].

Table 1. Independent components of the European SARS-CoV-2 and zoonotic influenza bioinformatics external quality assessment that laboratories could choose to participate in, 2024

Name	Organism	Samples	Description
SARS1	SARS-CoV-2	20	Consensus sequence generation from complete amplicons based on Nanopore reads.
SARS2	SARS-CoV-2	20	Consensus sequence generation from fragmented amplicons based on Illumina reads.
SARS3	SARS-CoV-2	10	Clustering and classification of full-length genomes.
SARS4	SARS-CoV-2	10	Detection of particular amino-acid substitutions.
INFL1, INFL3, INFL4 (linked)	Influenza virus (zoonotic, in human)	10	Consensus sequence generation from complete amplicons or metagenomic sequencing based on Nanopore reads (INFL1). Clustering and classification of full-length genomes (INFL3). Prediction of reduced susceptibility to antivirals and/or host adaptation mutations (major and minor variants) (INFL4).
INFL2, INFL3, INFL4 (linked)	Influenza virus (zoonotic, in human)	10	Consensus sequence generation from fragmented amplicons or metagenomic sequencing based on Illumina reads (INFL2). Clustering and classification of full-length genomes (INFL3). Prediction of reduced susceptibility to antivirals and/or host adaptation mutations (major and minor variants) (INFL4).

2 Study design and methods

Participants and process

The invitation for the EQA was sent to ECDC's Operational Contact Points for COVID-19, for influenza and for bioinformatics, with ECDC's National Focal Points for Viral Respiratory Diseases and National Microbiology Focal Points (NMFPs) in copy. These contact points cover the 27 European Union Member States, as well as the three European Economic Area countries. In addition, ECDC's Country Cooperation Office contacted the Western Balkan countries and Türkiye.

For SARS-CoV-2, each component had its own dataset consisting of several samples, for which data could be downloaded by participants, processed, and the results provided back in an Excel template. Sequence data could be provided in Fasta format as well. For influenza virus, the exercises were divided in two parts, INFL1/3/4 and INFL2/3/4, each consisting of 10 samples constituting four virtual scenarios (see Annex 2, Supplementary text 1). INFL results were provided in an Excel template as well. In addition to the assessment of the particular type of analysis for which the component was designed, two general aspects were assessed as well for each component: quality control of the input data for the analysis, and the throughput time of the analysis. The methodology used for all these aspects is described in the sections below.

For each SARS-CoV-2 dataset, a comparator sequence was prepared for each sample. All individual differences versus the comparator sequence of good-quality samples were manually investigated to determine the likely cause for each difference. Where a difference was found to be due to an issue in the comparator sequence, the latter was adjusted and the assessment was re-run. In practice, this did not occur. The discrepancies were aggregated and compared with the methodology provided by the laboratory to try to identify concrete methodological issues such as inappropriate parameter settings. No criteria were applied to classify the laboratory as having passed or failed each EQA component. Instead, the focus was on providing useful information for improvement, by highlighting potential issues and potential causes.

The evaluation of the responses for each zoonotic influenza component/scenario was similar to that of SARS-CoV-2. Reads for 10 samples were created, each derived through simulation from an artificial virus genome sequence. For the presence of mutations, it should be noted that the effect of them in some of the virus samples submitted may be the subject of debate in the literature. Laboratories were not penalised for providing detections of additional mutations. We focused our reference samples on a subset of mutations.

After the period allotted to laboratories to provide their results, laboratories received an individual report with those results assessed and compared with those of other laboratories in an anonymised form. Laboratories were subsequently given the opportunity to discuss the report with the organisers, and finally feedback could be given through an optional, anonymous questionnaire.

Throughput time

The total throughput time needed to process the component was measured for educational purposes only. Throughput time was defined as the duration of time from the moment the password was provided to decrypt the dataset for a particular component until the moment the results for that component were submitted. An adjusted version of the throughput time was calculated to better reflect actual throughput time by subtracting non-working hours from the total throughput time. The latter included weekend days as a whole and outside 8:00–18:00 on working days. In addition, when passwords were initially provided and/or results were submitted for two or more components at the same time, the corresponding throughput time was generally excluded from statistics on adjusted throughput time, since it no longer represented the throughput time for a single component. This was only done for the SARS-CoV-2 components, since the influenza virus components were combined into scenarios as described further below.

Quality control

Quality control is an essential element of each step in a bioinformatic analysis. Often, performing the analysis in question on sequence reads with bad quality may lead to misleading results such as missing mutations or unjustified clusters. In addition, the results from processed sequence reads can also be of bad quality, either due to the input data or the algorithm used and would as such affect any downstream analyses. Each dataset therefore contained some clearly bad quality samples, and participants needed to indicate for each sample whether the quality was either 'OK' or 'BAD'.

For SARS-CoV-2, a distinction was made between overall and specific quality. Overall quality considered the entire genome and therefore the usability for analyses that take the entire genome as input, such as clustering or classification. Specific quality only considered a specific part of the genome that is used for a more targeted analysis, such as detection of mutations in a particular gene. The assessment of the differences in quality control versus the comparator sequence was performed identically for both types.

For the SARS1/2 components (Table 1), the overall quality control needed to be done on the output genome sequences. For SARS3, the overall quality control needed to be done on the input genome sequences. For SARS4, the specific quality control needed to be done on the S-gene sequence. For influenza, some samples of this EQA had lowly covered regions, or no influenza sequences at all. Participants needed to indicate for each sample whether the overall quality was OK or if there were coverage issue in some segments or overall.

Consensus sequence generation

The purpose of the SARS1/2 and INFL1/2 components (Table 1) was to assess the generation of full-length SARS-CoV-2 and influenza virus genomes from amplicons sequenced by the Nanopore platform (SARS1 and INFL1 components) and the Illumina platform (SARS2 and INFL2 components), respectively, by comparing them with a comparator sequence for each sample in the dataset. The methodology for generating the comparator sequences and the composition of the respective datasets are described per component in the sections below.

SARS1 and SARS2 components

Two consensus sequences were generated (strict and a less strict), corresponding to a high and a low coverage threshold. This was done using the ViroConstrictor pipeline v1.4.1 [12] with a modified version of TrueConsense specific for this EQA. The Nanopore reads were created with R9.4.1 and R10.4.1 flow cells and base calling was done with the High-Accuracy preset of Guppy [13]. The high and low coverage thresholds were set to 50x and 20x, respectively, and for Illumina reads to 20x and 10x. Ambiguity symbols representing two nucleotides were put when coverage was sufficiently high and the difference between the most frequent and the second most frequent nucleotide was at most 10% or 20%, respectively, for the strict and the less strict consensus genome. For influenza virus a dedicated pipeline was used, with coverage thresholds set to 10x for both types of reads. Ambiguity symbols were assigned when the difference between the most frequent and the second most frequent nucleotide was at most 10% for the strict and the less strict consensus genomes.

Low coverage regions were assigned Ns for both types of genomes and, subsequently, manually curated by a person doing this curation routinely for SARS-CoV-2 and influenza national surveillance, respectively. A comparator sequence was created from that, which allowed for each sequence position both the strict and the less strict version. As such, an acceptable range of variation in the consensus sequence was created to reflect the fact that there are no exact criteria for defining correctness of a sequence in routine application. In addition, the sequence regions upstream and downstream the first/last primer were set to allow both N and gaps, and positions where an ambiguity symbol (not N) was allowed, were also set to allow N.

The genome sequences returned by the participants were compared with the comparator sequence and the differences found were classified into eight types of differences, listed in Table 2. Among these eight types, an insertion, deletions and wrong nucleotides were considered critical differences since they are errors rather than reducing the information it's possible to extract. The SARS1 and SARS2 datasets consisted of 20 samples that were selected as described in Table 2A and Table 3A.

Table 2. Types of differences found when the consensus sequence was compared with the comparator sequence, ESIB-EQA, 2024

Type of difference	Description
Insertion	An insertion was introduced versus the comparator sequence. The entire insertion is counted as a single difference.
Deletion	A deletion was introduced versus the comparator sequence. The entire deletion is counted as a single difference.
Wrong nucleotide	A different nucleotide or ambiguity symbol (other than N) than the one allowed was provided, or, in case an ambiguity symbol (not N) was allowed, the provided nucleotide is not represented by it.
Nucleotide instead of ambiguity	An actual nucleotide was provided where only an ambiguity symbol (other than N) was allowed.
Ambiguity instead of nucleotide	An ambiguity symbol (other than N) was provided where only an actual nucleotide was allowed.
Nucleotide stretch instead of stretch of Ns	A stretch of actual nucleotides or ambiguity (other than N) symbols was provided for a particular region (including a region of length 1), where only Ns were allowed. The entire stretch is counted as a single difference.
Stretch of Ns instead of nucleotide stretch	A stretch of Ns was provided for a particular region (including a region of length 1), where only actual nucleotides or ambiguity (other than N) symbols were allowed. The entire stretch is counted as a single difference.
Nucleotide stretch outside amplicon range	A stretch of actual nucleotides or ambiguity (other than N) symbols was provided before the first amplicon or after the last amplicon, instead of Ns or gaps. The entire stretch is counted as a single difference.

The types of differences are sorted in decreasing order of overall impact on the outcome. Insertions, deletions and wrong nucleotides are considered critical differences.

INFL1 and INFL2 components

The purpose of these two components was to assess the quality of the generation of full-length influenza genomes from RNA sequencing or from PCR amplicons sequenced by the Oxford Nanopore platform (INFL1 component) and the Illumina platform (INFL2 component). For each sample in the 10 respective datasets (described in Annex 2, Supplementary texts 1 and 2), a consensus genome was generated from the simulated reads with the internal routine workflow for Illumina, and a combination of IRMA [14] and wf-flu (https://github.com/epi2me-labs/wf-flu) for Nanopore. Coverage threshold was set to 10x for both Nanopore and Illumina. In some cases of this exercise, using data with 10x coverage allowed for reconstruction of a larger proportion of segments, providing higher confidence for the clustering question. Ambiguity symbols representing two nucleotides were put when coverage was sufficiently high and the difference between the most frequent and the second most frequent nucleotide was at most 20%.

Consensus genomes were subsequently manually curated by an expert curator from the organising institutes, and a reference dataset was created from it, allowing potential ambiguities at some sequence positions. Similar to the SARS1 and SARS2 components, regions before and after the first and last primer were allowed to be set to either N or gaps. Ambiguous symbols were set to also allow both the ambiguity symbol and N.

Participants' provided genomes were compared with the reference datasets and the differences found were classified into the same categories as described in Table 2.

Clustering and classification

SARS components

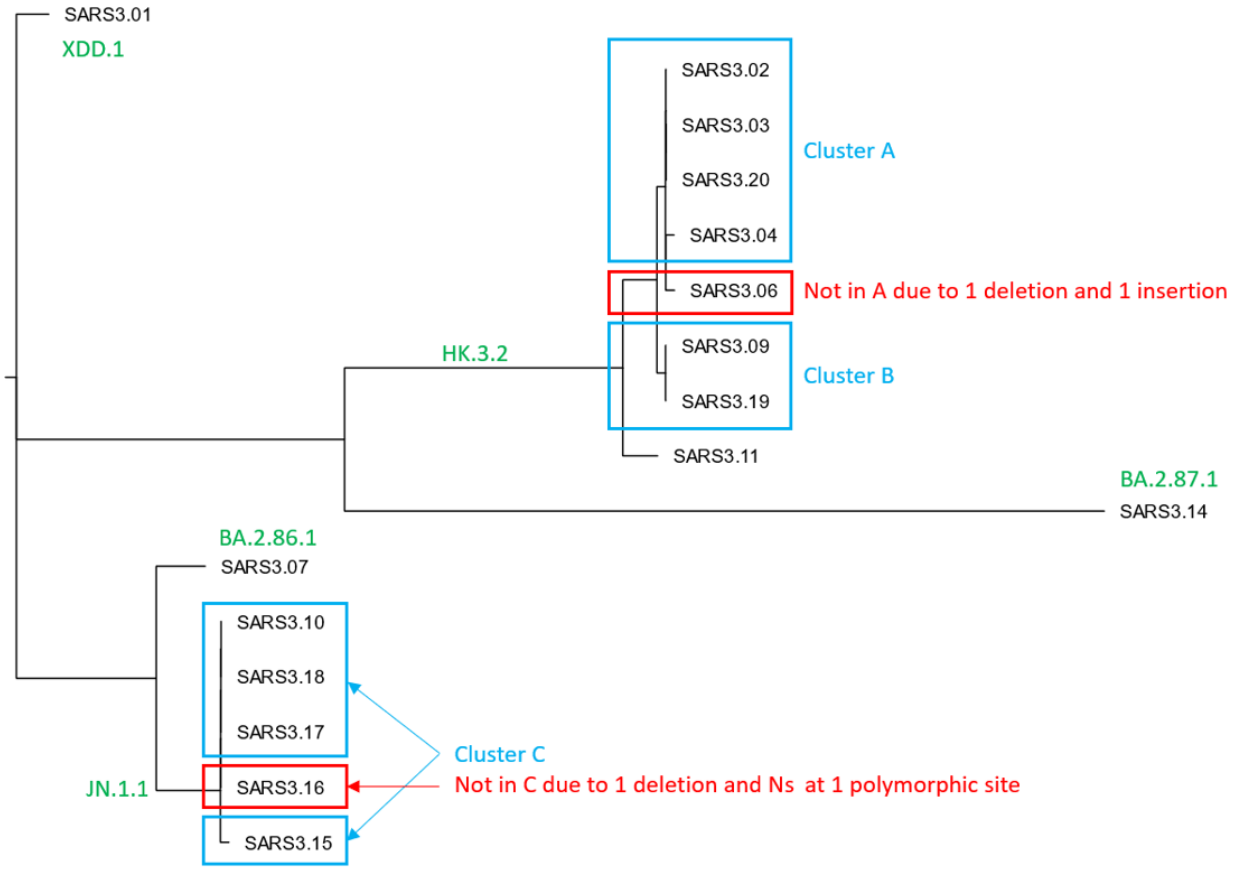
The purpose of the SARS3 component on clustering and classification was to assess the ability to identify clusters of samples that are phylogenetically so similar that they are likely part of a very short transmission chain – less than two weeks – and as such could be used as a basis for initiating contact tracing. In addition, the ability to perform classification at sub-species level was assessed. In the absence of an official recommendation for such investigation, comparator sequence clusters were defined using a cut-off of a single mutation (substitution or indel). Participants' provided cluster codes were mapped in all possible combinations to the comparator sequence

cluster codes, removing first singleton clusters. The combination that scored best (i.e. had the least number of cluster assignment mismatches) was used for further assessment.

For the ability to perform classification at sub-species level (i.e. lineage) the assignment of Pango lineages was assessed. Comparator sequence Pango lineages were assigned with Pangolin version 4.1.3 of 26 July 2023 (https://github.com/cov-lineages/pangolin) [15] and Pangolin data version 1.25.1 of 14 February 2024. Provided values were compared with comparator sequence values on the basis of identical matches.

The SARS3 dataset consisted of 20 samples that were selected as described in Table 4A. The annotated phylogenetic tree for the SARS3 component, limited to OK quality sequences, is shown in Figure 1. This tree was made using IQ_TREE COVID-19 release v2.2.0 [16].

Figure 1. Phylogenetic tree of SARS3 samples, excluding bad quality samples, ESIB-EQA, 2024



Samples or subtrees are annotated with their Pango lineage in green and in blue the cluster to which they belong, if any. Specific cases that may be included in a cluster but should not be are indicated in red along with the explanation.

INFL components

The ability to perform classification at sub-species level (i.e. subtyping), and especially to assign influenza genetic clades, was assessed as well. Clade assignments were based on Nextstrain [17] v.3.4.0, BV-BRC (https://www.bv-brc.org) and LABEL [18], and compared with provided values for identical matches. The classification of the 10 influenza datasets is given in Annex 2, Supplementary Text 2.

The INFL3 component was similar to that of SARS3 (i.e. assess the ability to identify clusters). For sub-species classification, the ability to perform classification at genetic clade level was assessed. Reference dataset clusters were defined using a cut-off of one to three mutations (substitution or indel) for each influenza segment. Description of the expected clustering results is given in Annex 2, Supplementary Text 3.

SARS-CoV-2 mutation detection

The purpose of this component, SARS4, was to assess the enumeration of amino acid substitutions present in a sample that had known phenotypic associations. The list of to be detected mutations for different Pango lineages was provided as part of the dataset for the component and was based on ECDC's list of variants of concern (including de-escalated ones) [19]. Additional mutations that were present in the sample were allowed to be reported as well, while any non-existing mutations or missed mutations compared with the list were considered a difference.

The S-gene mutations to be detected for different lineages are given in Table 3. The SARS4 dataset consisted of 10 samples that were selected as described in Table 5A.

Table 3. SARS-CoV-2 amino acid substitutions in the S-gene (S) that should be detected for the SARS4 component, by lineage, ESIB-EQA, 2024

Lineage	Amino acid substitutions
BA.2.86	S:I332V, S:D339H, S:R403K, S:V445H, S:G446S, S:N450D, S:L452W, S:N481K, S:483-, S:E484K, S:F486P
BA.2.87.1	S:G75D, S:S98F, S:V126A, S:W152L, S:R190S, S:K417T, S:K444N, S:V445G, S:L452M, S:N481K, S:V642G, S:K679R, S:S691P, S:T791I, S:Y796H, S:D936G
XBB.1.5-like	S:N460K, S:S486P, S:F490S
XBB.1.5-like + F456L (e.g. EG.5, FL.1.5.1, XBB.1.16.6, and FE.1)	S:F456L, S:N460K, S:S486P, S:F490S
XBB.1.5-like + L455F + F456L	S:L455F, S:F456L, S:N460K, S:S486P, S:F490S

Influenza mutation detection and prediction of susceptibility to selected antivirals

For influenza components, the aim was to list the interesting sample features potentially relevant to public health, such as mutations known to increase virulence or transmission in specific hosts or affecting sensitivity to antivirals. Description of the expected mutations (minor and major) is given in Annex 2, Supplementary text 4.

The purpose of the INFL4 component is to assess the detection of a substitution associated with potential changes in susceptibility to antivirals (adamantanes, oseltamivir, zanamivir, and baloxavir), and a specific question of treatment was added for one sample of some scenario (INFL1/2.02, INFL1/2.03, and INFL1/2.07). Some of the changes were not designed to be present at the consensus level of the reconstructed genomes, but within the intra sample diversity. The list of mutations to be detected for different clades was based on the World Health Organization (WHO) lists of amino acid substitutions associated with reduced inhibition [20]. The list of expected results for susceptibility to antivirals is given in Annex 2, Supplementary text 5.

3 Results

Participation

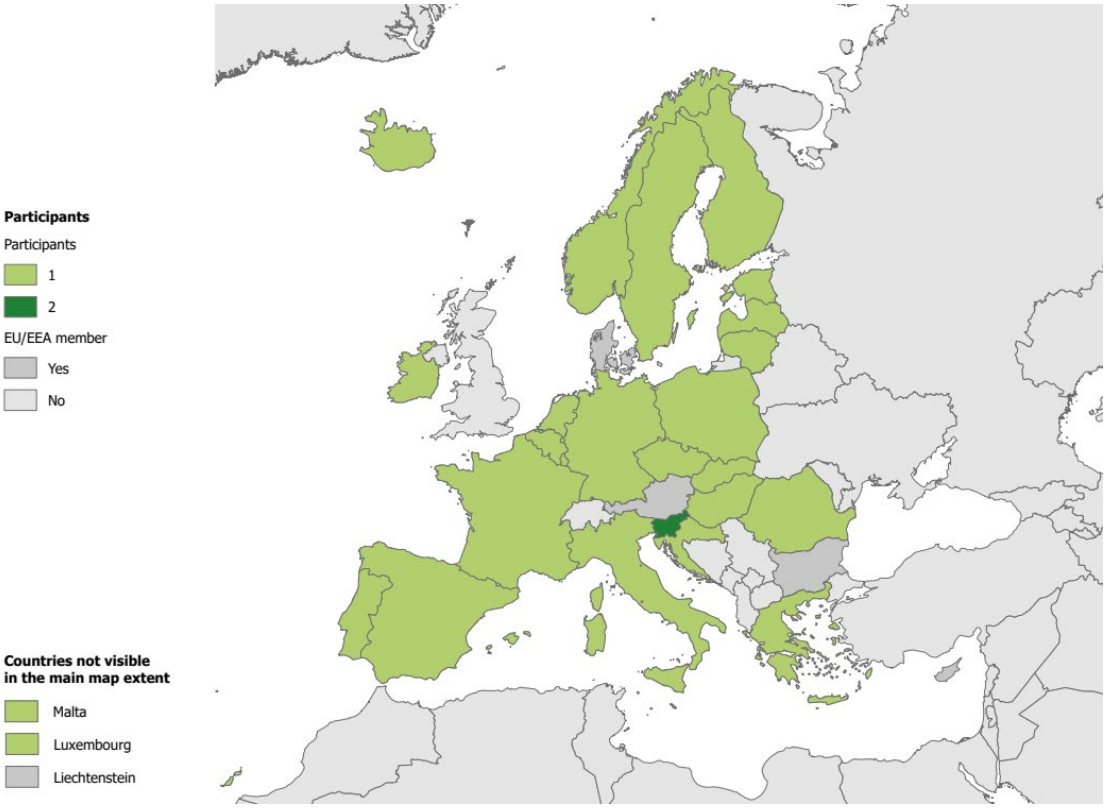
Of 30 EU/EEA countries, 25 participated in this EQA. Western Balkan countries and Türkiye did not participate (Figure 2). One country had two participating laboratories, while all others had one. One laboratory withdrew from the EQA altogether because of unexpected resource constraints and is not included in the statistics. A total of 11 laboratories participated in both SARS1 and SARS2 (i.e. performing the same task from different sequencing platforms), whereas 12 participated in only one of these two components (Table 4). For INFL1/3/4 and INFL2/3/4, these numbers were, respectively, 7 and 14 laboratories. A total of 18 laboratories also participated in the first iteration of the EQA. The list of participating laboratories is provided in Table 1A.

Table 4. Participation and completion of individual components by all laboratories that participated in at least one component, ESIB-EQA, 2024

Status	SARS1	SARS2	SARS3	SARS4	INFL1/3/4	INFL2/3/4
Completed	14	20	23	23	13	17
Participated but not completed	0	0	0	0	2	2
Not participating	12	6	3	3	11	7
Completed for both iterations of the EQA	9	14	17	17	8[a]	13[a]

[a] *Based on INFL1 and INFL2 component of the first iteration.*

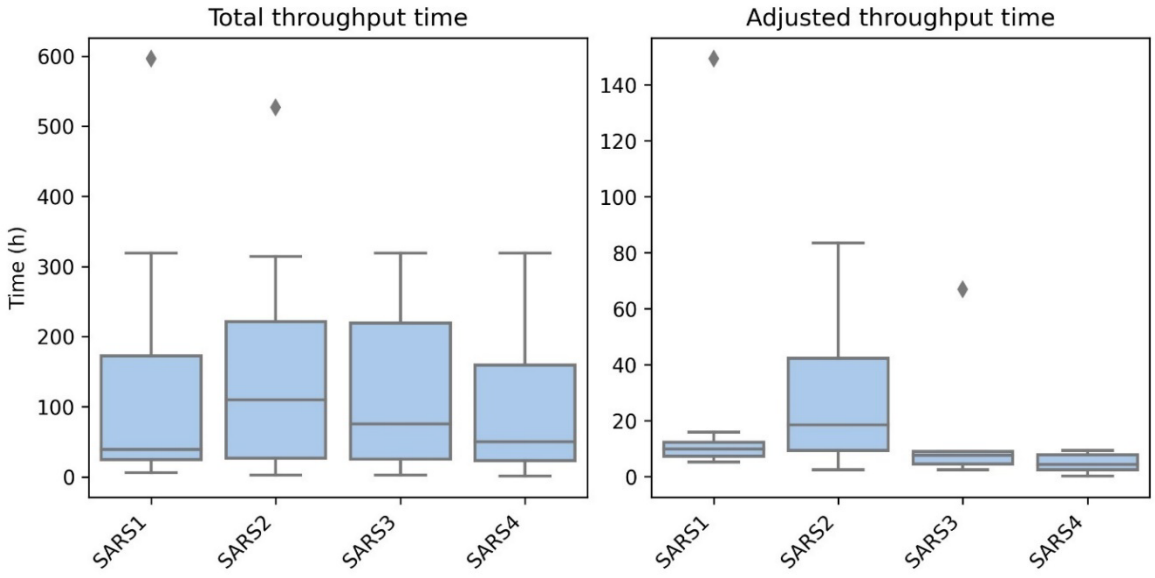
Figure 2. Number of participating laboratories in the EU/EEA, the Western Balkans and Türkiye, ESIB-EQA, 2024



Throughput time

The total and adjusted throughput times for the SARS-CoV-2 components are shown in Figure 3. For the two consensus sequence generation components, the interquartile range (IQR) of the adjusted throughput time was 7.4 to 42.2 hours or one to five working days. The median ratio between Nanopore reads (SARS1) and Illumina reads (SARS2) adjusted throughput time was 0.52 (IQR: 0.39–1.19), indicating that the platform technology had some impact on throughput time.

Figure 3. Total and adjusted throughput time for the different components, ESIB-EQA, 2024



Distributions are represented as box plots with whiskers at maximum 1.5x the interquartile range.

Quality control

Most of the samples were classified correctly according to their quality. One laboratory had 100% correct classification for SARS1. Combining SARS1 and SARS2 results, 13/34 (38%), 8/34 (24%) and 8/34 (24%) laboratories incorrectly classified one, two or three sequences, respectively. The remaining 4/34 (12%) incorrectly classified between four and six sequences. The number of samples for which the quality assessment was identical or different between comparator sequence and provided value is given in Table 5.

For the BAD/OK combination (i.e. where the sample was classified as bad quality in the comparator sequence but not by the laboratory), 12/16 (SARS1) and 15/23 (SARS2) were reported for samples where a co-infection with distinct variants was observed (Annex 2), most likely because the variants were not detected by the laboratory or the samples were not considered as bad quality. Another 4/16 (SARS1) and 8/23 (SARS2) were most likely due to having too high a threshold on the maximum number of missing nucleotides (Ns).

For the OK/BAD combination, (i.e. where the sample was classified as good quality in the comparator sequence but not by the laboratory), 8/28 (3/5 SARS1 and 5/23 SARS2) were reported for samples where the generation of the participant sequence introduced a substantial amount of quality issues that were not in the comparator sequence. Another 12/28 were reported for samples where the presence of a substantial amount of 'private' mutations (i.e. mutations that do not occur in a recent reference common ancestor) most likely led to a classification as bad quality. The likely cause for the remainder of the discrepancies was unclear.

For INFL1 and INFL2, the differences in terms of quality assessment were generally due to distinct criteria on minimal segment coverage or read quality. For INFL1/2, two samples had low-quality segments (PA and NS for sample 1, HA and NS for sample 6) with coverage dropping below 10x. More generally, several samples were considered by laboratories as bad quality because of low coverage on segments other than HA, PA, or NS. But overall, a large majority of the laboratories correctly classified the samples according to their quality, including one laboratory with 100% correct classification for a specific component.

Overall, among all samples analysed by all the participating laboratories (except sample 4, which was not an influenza virus sample, see Annex 2), 166/270 (62%) samples were classified as good quality (same as the comparator sequence), and 8/270 (3%) were classified as low-quality sequences (same as the comparator sequence). However, 35/270 (13%) samples were classified as low quality even though they were good quality in the comparator sequence. No laboratory reported good-quality segments for low-quality samples of the comparator sequence. Moreover, among the low-quality samples, the inconsistencies between the comparator sequences and the quality reported by the laboratories only involved the set of low-quality segments found. In total, 49/270 (18%) samples did not have the exact set of low-quality segments compared with the comparator sequence. Finally, no answer at all was given for 12 samples.

Table 5. Number of samples with correct and incorrect classification, ESIB-EQA, 2024

Component	Correct classification		Incorrect classification	
	OK/OK	BAD/BAD	OK/BAD	BAD/OK
SARS1	219	40	5	16
SARS2	297	57	23	23
SARS3	367	83	1	9
SARS4	150	37	11	32
INFL1	75	6	10	0
INFL2	91	2	25	0

Number of samples for which the quality assessment was identical (OK/OK or BAD/BAD) between comparator sequence/laboratory provided value, and number of samples where the assessment was different (OK/BAD or BAD/OK). The latter two should ideally have zero samples.

Consensus sequence generation

SARS1 and SARS2 components

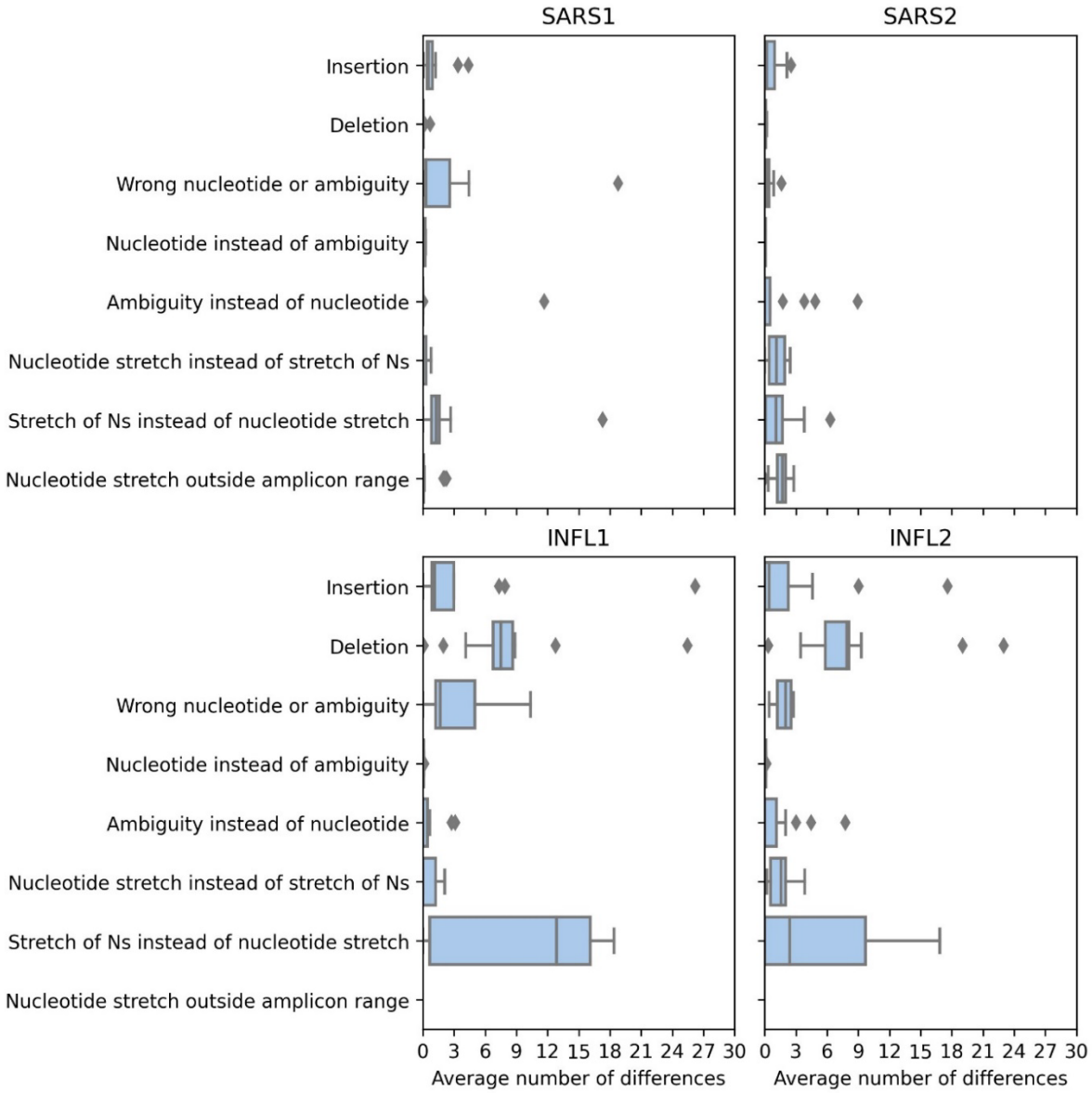
The average number of sequence differences per good-quality, non-educational sample, stratified by the type of difference, is shown in Figure 4. See Table 2 for a description of the types of differences that were distinguished between. Among these, the critical differences of insertions, deletions and wrong nucleotides are summarised by type in Table 6, and subdivided by potential cause. It should be noted that these differences, including insertions and deletions (indels), are with respect to the comparator sequence (GS) rather than the Wuhan-1 global reference (GR). An insertion in a provided sequence compared with the comparator sequence can therefore be at the same locus as where the GR has an insertion compared with the comparator sequence, or at a different locus. The same applies to deletions compared with the comparator sequence.

Table 6. Number and type of critical sequence differences versus the comparator sequence for SARS1 and SARS2, ESIB-EQA, 2024

Type and subtype of sequence difference versus comparator sequence	SARS1		SARS2	
Insertion	**N=227**		**N=154**	
	N	**%**	**N**	**%**
Same locus as GR insertion vs GS	165	73	136	88
Extra nucleotide in homopolymeric stretch	17	7	2	1
No distinguishable pattern	0	0	8	5
Insertion in primer region	45	20	8	5
Deletion[a]	**N=16**		**N=20**	
	N	**%**	**N**	**%**
Missing nucleotide in homopolymeric stretch	3	19	0	0
Same locus as GR deletion vs GS	3	19	16	80
Missing deletions in or near primer regions	10	63	0	0
No distinguishable pattern	0	0	4	20
Wrong nucleotide[b]	**N=523**		**N=116**	
	N		**%**	
In or near primer regions	200	38	41	35
Same nucleotide as GR rather than GS	247	47	48	41
No distinguishable pattern	76	15	27	23

GS: comparator sequence; GR: global reference (Wuhan-Hu-1).
[a] For SARS1, a single laboratory accounts for 11/16 (69%) of the cases.
[b] For SARS1, a single laboratory accounts for 301/523 (58%) of the cases.
Bad quality sequences are excluded and educational samples are included. The main types are those described in Table 2 Subtypes are given when relevant, or when the cause of the difference is evident. Values for each subtype are expressed as the number and proportion (%) of the differences of that subtype within that type. The totals per type are shown in the header row for that type.

Figure 4. Average number of sequence differences per sample for all good-quality, non-educational samples as compared with the comparator sequence, ESIB-EQA, 2024



Results are stratified by the type of difference. Distributions are represented as box plots with whiskers at maximum 1.5x the interquartile range.

INFL1 and INFL2 components

Out of 19 laboratories that participated in INFL1, 18 completed the component. For INFL2, 22/24 laboratories completed the component.

The main sources of differences between consensus sequences and comparator sequences were the following:

- Lab consensuses were truncated, i.e. starting at the initial start codon (ATG) instead of the beginning of the segment (due to usage of references consisting only of the coding sequence part). In that case, nucleotides before ATG could not be called and were counted as deletions compared with the comparator sequence.
- In other cases, primer matching sequences were not masked (in 5' and 3'), and thus counted as nucleotides instead of Ns when compared with comparator sequences.
- In the manual curation phase, ambiguity in nucleotide calling was not resolved the same way in comparator sequences and in the consensus. This case ended up with wrong nucleotides, ambiguity instead of nucleotide, or nucleotide instead of ambiguity.
- Diversity of reference sequences and pipelines used for influenza components may explain the number of insertion-deletions and wrong nucleotides called, especially between H5N6 and H5N1 consensus reconstruction (Supplementary Text 2).

Clustering and classification

For the SARS3 component, 23/26 laboratories completed the component. A total of 314/368 (85.3%) good-quality sequences were correctly assigned, either to the correct cluster or as not belonging to any cluster. The 54 different assignments were inspected, and the likely causes determined (Table 7). Around two thirds of the laboratories likely did not take into account indels when performing the clustering (i.e. only considered substitutions, leading to around half of the wrong assignments). Two laboratories used a cluster cut-off of three or higher, which is clearly too high, and this accounted for another 22% of the wrong assignments. With respect to the assignment of the Pango lineage, 12 (52.1%) laboratories likely used an outdated version of the software, resulting in an ancestral, (i.e. less high resolution) lineage being assigned. The remainder of the Pango lineage assignments were identical.

Table 7. Likely causes of different cluster assignments for the SARS3 component, ESIB-EQA, 2024

Likely cause	Different assignments		Laboratories	
	No.	%	No.	%
Too large cluster cut-off	12	22.2	2	8.7
Cluster cut-off of 2 instead of 1	6	11.1	3	13.0
Cluster cut-off of 0 instead of 1	0	0	0	0
Not considering indels	26	48.1	15	65.2
Incorrect interpretation of tree	0	0	0	0
Unknown	10	18.5	3	13.0

For the INFL1/2 component, 23/26 laboratories completed the component A total of 222/300 (74%) sequences were correctly assigned. Discrepancies between comparator sequence and laboratories can be explained by:

- A different definition of the cut-off between the comparator sequence and what laboratories consider as a cluster;
- different ways of counting the number of mutations;
- poor quality consensus construction that did not allow a complete comparison of sequences; and
- the comparison being based only on the clades or subclades instead of analysing the sequences and counting the number of mutations.

SARS-CoV-2 mutation detection

Of 26 participants, 23 completed the SARS4 component. Only 2/161 (1.2%) incorrectly detected substitutions in the S-gene, each by a different laboratory. The likely cause for these incorrect results is not clear.

Influenza virus mutations of interest

For the INFL1/2 component, 21 different mutations of interest potentially relevant for public health (Supplementary Text 4) (increase in virulence or in transmission in specific hosts) were selected to be detected in the major variant. Across the 23 laboratories that completed the component, a total of 496/1 560 (31.8%) mutations of interest were found. All laboratories found at least one of the mutations of interest, but none found all the mutations. The scores across all samples and components ranged from 5.8% to 57.7%. Most laboratories reported additional mutations that were absent from the reference list. Classifying some of these additional mutations as 'of interest' is open to discussion, while others are found in most sequences of the clade and cannot be defined as of interest.

In addition to the mutations of interest in major variants (found in >50% of the reads), four samples harboured minor variants. Across the 23 laboratories that completed the component, a total of 17/120 (14.2%) mutations were found. Most of the laboratories (14/23; 60.9%) did not report any of the minor variants.

Influenza virus susceptibility to antivirals prediction

For the INFL1/2 component, resistance mutations to antivirals were expected among the mutations of interest (Supplementary Text 5), either among major or minor variants.

For oseltamivir, across all scenarios, resistance was not correctly reported for 55/90 (61.1%) samples. Most of these results are explained by sample 2, where the mutation was present in a minor variant and, as seen in the previous section, minor variants were usually not reported. Additionally, some laboratories have not reported the mutation R292K, which implies a resistance to both zanamivir and oseltamivir.

For zanamivir, resistance was correctly reported for 49/90 (45.6%) of the samples compared with the comparator. For the incomplete reports, the reasons were the same as for oseltamivir – mainly due to the minor variants. Additionally, some laboratories reported a resistance to baloxavir (NA:I97V+I297V) as a resistance to zanamivir, although it is defined as neuraminidase inhibition according to WHO guidelines [21].

For baloxavir, resistance was correctly reported for 75/90 (83.3%) of the samples. Most of the discrepancies between answers and the comparator dataset were caused by the reporting of unexpected resistance mutations (e.g. PA:I38M), which has been described in seasonal influenza but not the given subtype (e.g. A(H7N9)), according to WHO guidelines [22].

For adamantanes, most of the laboratories correctly reported the expected resistance, with 79/90 (87.8%) of samples correctly assigned. The incorrect resistance reports were due to missing the mutation M2:S31N in the provided consensus or missing its effect on adamantanes.

4 Discussion

Throughput time

Throughput time was only measured for the SARS components in this second iteration of the EQA. For consensus sequence generation, Illumina short reads took twice as long as Nanopore reads to process based on the median time. This contrasts with the first iteration of the EQA, where both took around the same time. It is therefore possible that this difference is due to experimental error. In either case, around one to two working days were needed by the second quartile of the laboratories (i.e. not the fastest 25%, but the subsequent 25% of laboratories) to complete one of these components, which is the same as for the first iteration of this EQA. This throughput time is fast enough to allow for timely data sharing, e.g. to GISAID [23]. For the SARS3/4 and INFL3/4 components, it is less clear what may have driven the throughput time.

Quality control

For the SARS-CoV-2 consensus sequences, the majority of sequences incorrectly classified as OK quality were likely due to not properly considering the possibility of samples corresponding to co-infections with distinct variants. This was also an issue in the first iteration of this EQA. Such co-infections should lead to ambiguity symbols in the consensus sequence, and should therefore be considered for the quality assessment. At the same time, in routine lab work co-infections are rare events and there was an unusually high proportion of them in this dataset, so the overall impact of this issue is likely not very high. Among the sequences incorrectly classified as BAD quality, the main cause that could be identified was considering the presence of a substantial amount of 'private' mutations (i.e. mutations that do not occur in a recent reference common ancestor) an indicator for bad quality. However, in the absence of ambiguities, which would indicate a co-infection or otherwise mixed sample, such samples need not necessarily be considered of bad overall quality since the 'private' mutations may simply be correct. At the same time, if no ambiguities are assigned (i.e. only nucleotides A, C, T or G or N are assigned), this distinction cannot be made in the first place.

For influenza sequences (INFL1 and INFL2), the differences in terms of quality assessment were generally due to distinct criteria on minimal segment coverage or read quality, or not accounting for all segments.

Consensus sequence generation

A total of eight types of differences were distinguished as part of the assessment (Table 2). The clearly wrong – and therefore critical – types of differences were insertions, deletions or a wrong nucleotide compared with the comparator sequence. The other types have less of an impact, as they are not necessarily wrong or merely reduce the available information, and are therefore considered non-critical differences: nucleotide instead of ambiguity, ambiguity instead of nucleotide, nucleotide stretch instead of stretch of Ns, stretch of Ns instead of nucleotide stretch, nucleotide stretch outside amplicon range.

SARS1 and SARS2 components

In most cases, insertions compared with the comparator sequence were likely caused by over-reliance on the global reference sequence in reference-sequence based pipelines, irrespective of the sequencing platform (i.e. Nanopore for SARS1 or Illumina for SARS2). On the other hand, insertions or deletions in a homopolymeric stretch, a type of error specific to Nanopore reads, were not observed very frequently. Overall, deletions were also observed much less frequently than insertions. The large majority of wrong nucleotides occurred either in or near primer regions, or had the global reference nucleotide assigned rather than that of the actual reads. Both these types of errors occurred with around the same frequency and irrespective of the sequencing platform.

The problem of 'filling in' the global reference is one of the main issues that should clearly be improved in consensus sequence generation for SARS-CoV-2. In addition, the absence of or incomplete trimming of primers from reads is the likely root cause for most of the wrong nucleotides in primer regions. As a result, when the primer has a difference compared with the actual sequence, the reference (primer) nucleotide is called, or an ambiguous nucleotide is called. This is a second main issue that should be improved about consensus generation for SARS-CoV-2.

For the non-critical differences, the pattern is also quite similar between Nanopore and Illumina reads. The most frequently occurring one is one or more nucleotide stretches outside the amplicon range, instead of putting Ns or gaps. Here as well, it is possible that when a global reference is used as part of the algorithm for consensus sequence generation, the global reference is filled in outside the amplicon range. Alternatively, it is also possible that the presence of reads from the original material (i.e. not from amplicons) leads to still being able to resolve (parts of) the sequence outside the amplicon range. The 'filling in' of the global reference nucleotides outside the reference range should in any case be avoided, since the actual sequence may have differences, and can likely be

addressed in the pipelines used by the laboratories at the same time as the analogous issue with the critical differences. The second most frequent non-critical difference was putting a stretch of Ns instead of nucleotides.

In cases where this is related to applying too high a coverage threshold, correcting this is likely only a question of parameterisation, by lowering the minimum coverage threshold to e.g. 20–50x and 10–20x for Nanopore and Illumina reads, respectively, as described in the study design. The opposite issue, where a nucleotide stretch was put instead of a stretch of Ns, occurred substantially less frequently. Finally, putting an ambiguity rather than a nucleotide while there is sufficient evidence for only one particular nucleotide occurred more frequently for Illumina reads; however, no likely cause was found.

INFL1 and INFL2 components

One of the main issues found in these components relate to missing nucleotides at the beginning and at the end of segments which, in most cases, corresponded to flanking and non-coding stretches in 5' or 3' regions; some laboratories produced segments beginning at the first start codon and ending at the last stop codon of the related genes. This is mainly due to the usage of software or pipelines that focus on the coding part of segments instead of the whole segment. In addition, some laboratories did not mask the primers at the beginning and at the end of the segments. Both particularities explain the numbers of deletions and nucleotides instead of Ns.

Another main issue appeared with some avian influenza virus sequences that were not analysed thoroughly by some laboratories. For example, some laboratories did not reconstruct the NA segment of H5N6 samples because a H5N1 reference was used. This issue reveals that:

- Some currently used software do not propose a diverse enough reference dataset and
- A comprehensive influenza reference sequence dataset that laboratories could readily use would be desirable.

One last point is related to defective genomes that are particularly difficult to analyse and require time-consuming manual curation when detected. These samples are the source of differences between comparator sequences and laboratory consensus.

Clustering and classification

For SARS-CoV-2 clustering, most of the discrepancies arose from not considering the presence of indels in the clustering. This is indeed an issue that would need to be improved, since these evolutionary events are at least equally important to consider, in particular when deciding on public health actions. In contrast to the first iteration of the EQA, only a few of the discrepancies arose from the use of a cluster cut-off that was one higher than the one used for the comparator sequence. For the SARS-CoV-2 Pango lineage classification, the main issue was, as in the first iteration, that several laboratories did not use sufficiently up-to-date software or rather the continuously updated lineages definitions (pangolin-data), resulting in less precise lineage assignments. This issue should also be addressed.

For influenza virus, most of the discrepancies came from the poor quality of some consensus sequences that were reconstructed by the laboratories. In such cases, it was impossible to correctly compare the samples and extract the right number of mutations. Another major issue was the strategy used: some laboratories based their decision on whether the samples shared the same clade or subclade without checking the mutations. In the future, these issues could be addressed by more specific guidelines on how to perform the exercise, with a clearer definition of the cut-off and analysis strategy.

Influenza virus mutations of interest

One major issue was the definition of the (non-resistance) mutations of interest. FluSurver at GISAID [23] was generally used to generate a list of mutations described as increasing resistance or inducing a shift in host specificity. However, those mutations were not further examined to determine their relevance, e.g. by verifying that the mutation is rare in the clade or by checking if they have been documented in the literature for the specific clade. In parallel, host adaptation and other mutations of interest in avian influenza viruses remain debated. While some mutations are well described in the literature for some subtypes, others are still under verification. In this EQA, we proposed a subset of mutations of interest (see Supplementary Text 4) that triggered discussions with some laboratories, underscoring the need to further refine the definitions. In the future, we may focus on subsets of specific mutations that are well described in the literature, and give a list of relevant articles.

Influenza virus susceptibility to antivirals prediction

This component did not reveal any major issues. The main challenge was detection of minor variants, when these were the ones carrying the resistance mutations. Additionally, a common mistake was to report PA:I38M to justify a resistance to baloxavir for an H7N9 sample, while it has been described only for seasonal influenza.

SARS-CoV-2 mutation detection

Similar to the first iteration of this EQA, this component did not reveal any particular issues to be addressed.

Feedback from laboratories

A total of five laboratories provided feedback through the questionnaire. In all cases, the time period to complete all the components was deemed sufficiently long. With respect to how satisfied laboratories were with the practical organisation of the EQA in general, regardless of the content of the different components, four answered 'Very satisfied', and one 'Somewhat satisfied'. With respect to the usefulness of the individual EQA report, three answered 'Very satisfied' and one 'Somewhat satisfied'. One laboratory commented that reporting results through Excel sheets could be done differently, but was also ok if it were kept as such. Compared with the first iteration of the EQA, and based on comments provided then, several improvements were made for the second iteration. These included the possibility to report sequences in Fasta format, fully self-contained instructions per pathogen, and adjusting the naming convention of sequences in Fasta files provided to the laboratory (underscores instead of dots to avoid problems with some pipelines).

5 Conclusions

For this EQA, 25 of 30 EU/EEA countries participated. None of the six Western Balkans countries participated, nor did Türkiye. Overall, this level of EU/EEA participation is good for the second iteration of this EQA; however, for the Western Balkans countries and Türkiye it would be ideal to obtain participation in future iterations. There was a slightly higher participation rate for SARS-CoV-2 than for influenza virus. About two thirds of the laboratories also participated in the first iteration of the EQA.

Similar to the first iteration, for the consensus sequence generation, about half of the laboratories signed up for both Nanopore and Illumina reads based components and the other half only for one, indicating that both technologies are often available in these laboratories. No substantial difference in throughput time for both technologies could be observed for the second quartile of laboratories, indicating a similar workload for the consensus generation process. A number of recurring issues in the generation of the consensus sequence were observed and comparatively more issues were observed for influenza virus than for SARS-CoV-2. This may mainly be due to:

- usage of references with only coding sequences;
- usage of the wrong reference sequence (per segment);
- non-masking of primer sequences; and
- difficulty to analyse (either automatically or manually) data corresponding to defective genomes.

Altogether, consensus sequence generation should be improved for a number of laboratories and in general for influenza virus. The most important of these improvements are described further in the Recommendations section.

Similar to what was observed in the first iteration, for clustering, the main issue was not considering indels when determining phylogenetic relatedness. Apart from that, for SARS-CoV-2, a few laboratories used a cluster cut-off of two instead of one, which was used for the assessment. Contrary to the first iteration, no laboratories used a cut-off of zero. However, in the absence of guidance on which cut-off or cut-off range to use for which purpose, such differences cannot be considered as critical issues. For classification, in some cases the data definitions used were outdated for SARS-CoV-2, which should be avoided. For influenza virus, main classification and clustering issues were due to wrongly reconstructed consensus sequences.

The SARS-CoV-2 mutation detection component did not reveal significant issues. It should be noted that this component is intended to be converted into prediction of antiviral resistance, as antivirals and corresponding resistance data become available. The zoonotic influenza virus mutation analysis was more difficult than expected, especially because of the lack of widely used, standard mutation of interest datasets. Clearer instructions may be needed to help laboratories find the good mutations.

6 Recommendations

With the first iteration of this EQA, a baseline was established for the quality that the processes of consensus sequence generation, clustering, classification, mutation detection and antiviral resistance prediction can be expected to have. Subsequent iterations can also assess the evolution of the quality of these processes, and apply more stringent quality criteria where necessary.

All the recommendations derived from the previous sections have been compiled in Table 8. In order to improve quality, laboratories that have quality issues and that are at present not well-placed to address them, should be given the opportunity to participate in relevant training. This training can cover consensus sequence generation, e.g. understanding the differences between and limitations of existing tools, appropriate parameterisation of tools, local installation of tools, quality control and manual curation. For clustering, the usage and limitations of tools, as well as guidance on appropriate cluster cut-offs, may also be helpful. Over time, results from this EQA can then also be used to monitor the impact of training and support.

Table 8. Recommendations based on the second ESIB-EQA, 2024 (continued on page 20)

ID no.	Target audience	Process	Description
1[a]	Individual laboratories	Consensus sequence generation	The introduction of indels in the consensus sequence (that are not present in the raw data) due to alignment to a reference sequence should be avoided. This can be done through e.g. manual curation, less stringent reads mapping with respect to gaps, and/or additional steps of aligning reads to the generated (intermediate) consensus sequence.
2[a]	Individual laboratories	Consensus sequence generation	The introduction of wrong nucleotides or wrong non-N ambiguities due to absence of or incomplete removal of primers should be avoided. The primer removal step should be well validated.
3[a]	Individual laboratories	Consensus sequence generation	Clearly ambiguous positions should be assigned a non-N ambiguity symbol so that subsequent analyses, including quality controls, can take this information into account. When e.g. the two most frequent nucleotides at a certain position have less than 10% difference in frequency, the corresponding ambiguity symbol should be put instead of that of the most occurring nucleotide.
4[a]	Individual laboratories	Consensus sequence generation	The minimum coverage threshold that is normally applied should be neither too high nor too low, to avoid stretches of Ns where actual nucleotides can be put and vice versa. For Nanopore reads, a threshold in the range of e.g. 20–50x could be applied, whereas for Illumina reads a threshold in the range of e.g. 10–20x could be applied. It should be kept in mind that these figures are also dependent on the exact protocol applied and that as technology evolves, increased accuracy may lower these thresholds.
5[a]	Individual laboratories	Consensus sequence generation	The completion of the genome in 5' of the first amplicon and in 3' of the last amplicon with the corresponding part of a reference sequence need to be avoided, since this wrongly assumes that no mutations may be present in these regions. Instead, an equivalent number of Ns can be put or no nucleotides at all.
6	Individual laboratories	Consensus sequence generation – influenza	Up-to-date and complete reference influenza genomes should be used. Each reference segment may be chosen among a large database of possible sequences, and should include sequences outside the coding sequence, as well as properly named segments.
9	Individual laboratories	Consensus sequence generation – influenza	Samples indicative of defective viral genomes, which may lead to many wrong nucleotides being reported, should be detected and manual curation performed on them.

ID no.	Target audience	Process	Description
10[a]	Individual laboratories	Sequence quality control	The presence of non-N ambiguity symbols in the sequence should be part of quality control since this can be indicative of co-infection with distinct variants or contamination with material from other samples. Such sequences should not be considered to be of ok quality, and only used in specific analyses such as on co-infections.
11	Individual laboratories	Sequence quality control – SARS-CoV-2	The presence of a substantial amount of 'private' mutations (i.e. mutations that do not occur in a recent reference common ancestor) is not necessarily an indicator of bad quality. In the absence of ambiguities, which would indicate a co-infection or otherwise mixed sample, such samples shouldn't necessarily be considered of bad overall quality since the 'private' mutations may simply be correct. If no ambiguities are assigned (i.e. contrary to Recommendation 3), this distinction cannot be made in the first place
12[a]	Individual laboratories	Clustering	When determining the phylogenetic distance between sequences, indels should also be considered, rather than only substitutions.
13[a]	Individual laboratories	Classification	Sufficiently recent versions should be used for classification algorithms such as Pangolin for SARS-CoV-2. This also applies to any (reference) dataset that is used. If these algorithms are run as part of a broader pipeline, the versions should be checked as well.
14[a]	ECDC	Clustering	Guidance in terms of which cut-offs should be used for clustering and for which purpose could be helpful.
15	Individual laboratories	Mutation analysis – influenza	Mutations of interest must be listed with caution when using FluSurver, especially for avian influenza. As long as no exhaustive reference list of mutations exists, each mutation must be checked manually, in the literature.

[a] *Recommendation was also made in the report for the first iteration of this EQA.*

References

1. Grubaugh ND, Hodcroft EB, Fauver JR, Phelan AL, Cevik M. Public health actions to control new SARS-CoV-2 variants. Cell. 2021 Mar 4;184(5):1127-32. Available at: https://www.ncbi.nlm.nih.gov/pubmed/33581746
2. Abdool Karim SS, de Oliveira T. New SARS-CoV-2 Variants - Clinical, Public Health, and Vaccine Implications. N Engl J Med. 2021 May 13;384(19):1866-8. Available at: https://www.ncbi.nlm.nih.gov/pubmed/33761203
3. World Health Organization (WHO). Factsheet - Influenza (Seasonal) 2023. Geneva: World Health Organization; 2023. Available at: https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal)
4. European Centre for Disease Prevention and Control (ECDC). Public health impact of SARS-CoV-2 variants of concern: scoping review protocol. Stockholm: ECDC; 2021. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/Public-health-impact-of-SARS-CoV-2-variants-of-concern-scoping-review-protocol.pdf
5. XEC Initial Risk Evaluation 09 December 2024. Geneva: World Health Organization; 2024. Available at: https://www.who.int/docs/default-source/coronaviruse/09122024_xec_ire.pdf?sfvrsn=13695ab6_2
6. European Centre for Disease Prevention and Control (ECDC). Surveillance and targeted testing for the early detection of zoonotic influenza in humans during the winter period in the EU/EEA. Stockholm: ECDC; 2024. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/surveillance-testing-early-detection-zoonotic-flu.pdf
7. World Health Organization (WHO). Interim guidance – Maintaining surveillance of influenza and monitoring SARS-CoV-2 – adapting Global Influenza surveillance and Response System (GISRS) and sentinel systems during the COVID-19 pandemic. Geneva: World Health Organization; 2020. Available at: https://iris.who.int/bitstream/handle/10665/336689/WHO-2019-nCoV-Adapting_GISRS-2020.1-eng.pdf?sequence=1
8. European Centre for Disease Prevention and Control (ECDC). Website – What we do: Surveillance. Stockholm: ECDC; 2023. Available at: https://www.ecdc.europa.eu/en/about-ecdc/what-we-do#Surveillance
9. European Centre for Disease Prevention and Control (ECDC). Website – European Reference Laboratory Network for Human Influenza (ERLI-Net). Stockholm: ECDC; 2017. Available at: https://www.ecdc.europa.eu/en/about-us/networks/disease-and-laboratory-networks/erlinet-about
10. European Centre for Disease Prevention and Control (ECDC). Website – European COVID-19 reference laboratory network (ECOVID-LabNet). Stockholm: ECDC; 2022. Available at: https://www.ecdc.europa.eu/en/about-ecdc/what-we-do/partners-and-networks/disease-and-laboratory-networks/european-covid-19
11. European Centre for Disease Prevention and Control (ECDC). European SARS-CoV-2 and influenza Bioinformatics External Quality Assessment (ESIB-EQA), 2023. Stockholm: ECDC; 2024. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/SARS-CoV-2-influenza-bioinformatics-external-quality-assessment-2023.pdf
12. Zwagemaker F HK, Schmitz D, Kroneman A, The RIVM-IDS Bioinformatics team. ViroConstrictor: Zenodo; 2024. Available at: https://github.com/RIVM-bioinformatics/ViroConstrictor/tree/v1.4.6
13. Oxford Nanopore Technologies. Guppy basecalling software. Available at: https://nanoporetech.com/document/Guppy-protocol
14. Shepard SS, Meno S, Bahl J, Wilson MM, Barnes J, Neuhaus E. Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. BMC Genomics. 2016 Sep 5;17(1):708.
15. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature Microbiology. 2020 2020/11/01;5(11):1403-7. Available at: https://doi.org/10.1038/s41564-020-0770-5
16. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution. 2020;37(5):1530-4. Available at: https://doi.org/10.1093/molbev/msaa015
17. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34(23):4121-3. Available at: https://doi.org/10.1093/bioinformatics/bty407
18. Shepard SS, Davis CT, Bahl J, Rivailler P, York IA, Donis RO. LABEL: fast and accurate lineage assignment with assessment of H5N1 and H9N2 influenza A hemagglutinins. PLoS One. 2014;9(1):e86921.
19. European Centre for Disease Prevention and Control (ECDC). Website – SARS-CoV-2 variants of concern as of 20 December 2024. Stockholm: ECDC; 2024. Available at: https://www.ecdc.europa.eu/en/covid-19/variants-concern

20. World Health Organization (WHO). Website – Laboratory methodologies for testing the antiviral susceptibility of influenza viruses. Geneva: World Health Organization; 2024. Available at: https://www.who.int/teams/global-influenza-programme/laboratory-network/quality-assurance/antiviral-susceptibility-influenza

21. World Health Organization (WHO). Laboratory methodologies for testing the antiviral susceptibility of influenza viruses: Neuraminidase inhibitor (NAI). Geneva: WHO; 2024. Available at: https://cdn.who.int/media/docs/default-source/influenza/laboratory---network/quality-assurance/human-nai-marker-table_for-publication_final_20240918.pdf?sfvrsn=c6d153ec_3

22. World Health Organization (WHO). Laboratory methodologies for testing the antiviral susceptibility of influenza viruses: Polymerase acidic (PA) inhibitor, Baloxavir. Geneva: WHO; 2024. Available at: https://cdn.who.int/media/docs/default-source/influenza/laboratory---network/quality-assurance/antiviral-susceptibility-influenza/pa-marker-who-table_07-08-2024_updated_final-version.pdf?sfvrsn=5307d6fe_2

23. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. Euro Surveill. 2017 Mar 30;22(13):30494. Available at: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2017.22.13.30494

Annex 1. Participating laboratories

Table 1A. Laboratories participating in ESIB-EQA, by country and name, 2024

Country	Name
Belgium	National Influenza Center in Belgium, Sciensano
Czech Republic	National Institute of Public Health, National Reference Laboratory for Influenza and Noninfluenza Respiratory Viral Diseases
Germany	Robert Koch Institute
Denmark	Statens Serum Institut
Estonia[a]	Health Board
Greece[a]	National Influenza Centre of Northern Greece
Spain[a]	Institute of Health Carlos III
Finland[a]	The National Institute for Health and Welfare
France[a]	Plateforme genEPII, CNR des virus des infections respiratoires (dont la Grippe), Hospices Civils de Lyon
Croatia[a]	Croatian Institute of Public Health
Hungary	National Center for Public Health and Pharmacy
Ireland[a]	National Virus Reference Laboratory
Iceland[a]	Landspitali University Hospital
Italy	Istituto Superiore di Sanità
Lithuania[a]	National public health surveillance laboratory
Luxembourg[a]	Laboratoire Nationale de la santé
Latvia[a]	Riga East University Hospital, National Microbiology Reference Laboratory
Malta[a]	Mater Dei Hospital
Netherlands	National Institute for Public Health and the Environment
Norway[a]	Norwegian Institute of Public Health
Poland[a]	National Institute of Public Health NIH - NRI
Portugal[a]	Instituto Nacional de Saúde Doutor Ricardo Jorge
Romania[a]	Institutul National de Sanatate Publica
Sweden	Public Health Agency of Sweden
Slovenia[a]	National Laboratory of Health, Environment and Food
Slovenia[a]	University of Ljubljana, Faculty of Medicine, Institute of Immunology and Microbiology
Slovakia	Public Health Authority of the Slovak Republic

[a] Also participated in the first iteration of the EQA.

Annex 2. Composition of the datasets

Table 2A. Composition of the SARS1 dataset, ESIB-EQA, 2024

Samples	Description
1, 2, 3, 7, 14, 15, 16, 25, 27	No specific challenges. Consists of 7 Omicron, 1 Alpha and 1 Delta variant.
6, 10, 13, 17, 18, 28	Challenges in the S gene.
29	Challenges in the N gene.
19, 24, 30, 31	Challenges in the ORF1ab region.
8, 9, 20	Challenges in the ORF7 gene.
21, 32	Challenges in the ORF8 gene.
4, 5	Co-infected samples, of which 1 Alpha/Delta and 2 Omicron/Omicron. These are considered as bad overall quality, since they are not usable for routine analyses that do not specifically consider co-infection.
12, 26	Clearly bad overall quality.
11	Omicron/Omicron recombination.
22, 23	Hepatitis E virus, i.e. different species, samples. These are educational samples intended to indicate the importance of detecting potential miscommunications between wetlab and drylab/bioinformatics processes, i.e. in this case a sequencing run with several species where only one species was expected. Some SARS-CoV-2 reads are still present in each sample, likely due to the adapter ligation reaction continuing a little bit after pooling of the samples, complicating the detection of the issue.
13	V3 rather than V4.1 primers used. This is an educational sample intended to indicate the importance of detecting potential miscommunications between wetlab and drylab/bioinformatics processes, i.e. in this case a change in the primers used.

Table 3A. Composition of the SARS2 dataset, ESIB-EQA, 2024

Samples	Description
1, 2, 4, 8, 10, 11, 12, 19, 28, 29	No specific challenges. Consists of 8 Omicron, 1 Delta and 1 Alpha variant.
5, 21, 31	Challenges in the S gene.
24	Challenges in the N gene.
6, 14, 16	Challenges in the ORF1ab region.
3, 15, 25	Challenges in the ORF7 gene.
26	Challenges in the ORF8 gene.
9, 17, 32	Co-infected samples, all Omicron/Omicron. These are considered as bad overall quality, since they are not usable for routine analyses that do not specifically consider co-infection.
20, 27	Recombination, of which one Delta/Omicron and one Omicron/Omicron.
18, 30	Clearly bad overall quality.
22, 23	Influenza virus, i.e. different species, samples. These are educational samples intended to indicate the importance of detecting potential miscommunications between wetlab and drylab/bioinformatics processes, i.e. in this case a sequencing run with several species where only one species was expected.
7, 13	V4 rather than V3 primers used. These are educational samples intended to indicate the importance of detecting potential miscommunications between wetlab and drylab/bioinformatics processes, i.e. in this case a change in the primers used.

Table 4A. Composition of the SARS3 dataset, ESIB-EQA, 2024

Samples	Pangolin lineage	Description
1, 10, 15, 16	BA.2.30	Cluster A, zero SNPs and indels difference.
12	BA.2.30	Cluster A, one SNP difference from the rest of the cluster.
4, 5, 11, 18	BA.2.30	Cluster B, zero SNPs and indels difference.
9	BA.2.30	Not in a cluster; two SNPs difference from both cluster A and B.
3, 8, 17	BA.2.12	Cluster C, zero SNPs and indels difference.
19	BA.2.12	Cluster C, one SNP difference from the rest of the cluster.
2	BA.2.12	Not in a cluster, zero SNPs, one deletion and one insertion from cluster C.
6,	BQ.1	Not in a cluster.
13	XBB	Not in a cluster.
7	Unassignable	Clearly bad overall quality.
14	BA.5	Clearly bad overall quality.
20	BA.2.30	Clearly bad overall quality.

Table 5A. Composition of the SARS4 dataset, ESIB-EQA, 2024

Samples	S gene substitutions	Description
1	S:L452M, S:N440K, S:K444R, S:N460K	BA.2.3.20
2	S:W152R, S:G257S, S:I210V, S:F157L, S:N460K, S:N440K, S:G446S	BA.2.75 with good overall quality but with a portion of the S gene not resolved. Thus, it is considered as bad quality since the presence of specific mutations is important for this analysis and they may be masked by ambiguities. This sample was added for educational purposes, to indicate that care should be taken as to which results can be considered reliable.
3	S:F486V, S:R493Q	Clearly bad overall quality sample with substantial parts of the S gene not resolved and thus also bad quality.
4	S:K444T, S:F486V, S:L452R, S:N440K, S:R493Q	BA.5. Contains a reversion, S:R493Q, that should be detected.
5	S:F486V, S:L452R, S:N460K, S:N440K, S:R346T	BF.7
6	S:W152R, S:F490S, S:R346T, S:G257S, S:I210V, S:F157L, S:N460K, S:K356T, S:N440K, S:G446S	BN.1
7	S:K444T, S:F486V, S:L452R, S:N460K, S:N440K	BQ.1
8	S:W152R, S:K444T, S:L452R, S:R346T, S:G257S, S:I210V, S:F157L, S:N440K, S:G446S	CH.1.1
9	S:F490S, S:R346T, S:N460K, S:N440K, S:G446S	XBB
10	S:F490S, S:R346T, S:F486P, S:N460K, S:N440K, S:G446S	XBB.1.5

The S gene substitutions column only contains those substitutions that were included in the list of to be detected mutations.

Table 6A. Composition of the major influenza variants dataset, ESIB-EQA, 2024

Sample	PB2	PB1	PA	HA	NP	NA	MP	NS
INFL1/2.01	–	–	–	T110N,N224K Q226L,T318I	–	–	–	–
INFL1/2.02	E627K	–	–	T110N,N224K Q226L,T318I	–	–	–	–
INFL1/2.03	T271A,I292V E627K,D701N K526R,A588V	H99Y	–	–	–	R298K	S31N	–
INFL1/2.05	141-	–	–	–	–	–	S31N	–
INFL1/2.06	141-	–	–	–	–	–	S31N	–
INFL1/2.07	E627K,D701N	H99Y,I368V	–	H110Y,N158D G228S T160A,Q226L	–	I97V+I294V	–	–
INFL1/2.08	E627K,D701N	H99Y,I368V	–	H110Y,N158D G228S T160A,Q226L	–	–	–	–
INFL1/2.09	E627K	H99Y,I368V	–	H110Y,N158D T160A	–	–	–	–
INFL1/2.10	E627K	H99Y,I368V	–	N158D,T160A	–	–	–	–

Table 7A. Composition of the minor influenza variants dataset, ESIB-EQA, 2024

Sample	PB2	PB1	PA	HA	NP	NA	MP	NS
INFL1/2.01	E627K (30%)	–	–	–	–	–	–	–
INFL1/2.02	–	–	–	–	–	R293K (30%)	–	–
INFL1/2.03	–	–	I38M (45%)	–	–	–	–	–
INFL1/2.04	–	–	–	–	–	–	–	–
INFL1/2.05	–	–	–	–	–	–	–	–
INFL1/2.06	–	–	–	–	–	–	–	–
INFL1/2.07	–	–	–	–	–	–	–	–
INFL1/2.08	–	–	–	–	–	I97V,I294V (10%)	–	–
INFL1/2.09	–	–	–	–	–	–	–	–
INFL1/2.10	–	–	–	–	–	–	–	–

Please note that numbering is relative to our reference consensus sequences.

Annex 3. Description of the influenza scenarios

The two zoonotic influenza components (INFL1/3/4 and INFL2/3/4) in this bioinformatics EQA each consist of 10 samples (sequencing data simulated from artificial viral genome sequences) divided into four virtual scenarios. The four scenarios are summarised below:

In scenario 1, a father and a daughter are hospitalised following a car accident, and both test positive for influenza A. The two samples are:

- **INFL1/2.01** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the father.
- **INFL1/2.02** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the daughter.

In scenario 2, a woman (exotic bird collector) is hospitalised with severe respiratory distress. Molecular testing of a bronchoalveolar lavage sample from the woman was positive for influenza A. One of her neighbours also presented with a respiratory disease. The two samples are:

- **INFL1/2.03** corresponds to raw sequencing reads obtained from a bronchoalveolar lavage from the woman.
- **INFL1/2.04** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the neighbor.

In scenario 3, a woman (manager of a livestock farm) is hospitalised for severe influenza-like-illness. One of her employees is recovering from a mild respiratory disease that appeared after a trip to an international agricultural fair and has also been sampled. Molecular testing of the nasopharyngeal swab samples was positive for influenza A for both samples. The two samples are:

- **INFL1/2.05** corresponds to raw sequencing reads obtained from a nasopharyngeal swab of the manager.
- **INFL1/2.06** corresponds to raw sequencing reads obtained from a nasopharyngeal swab of the employee.

In scenario 4, cases of respiratory illness were reported in a countryside area where carcasses of wild birds have been noticed. The domestic cats from two neighbouring cereal farms fell severely ill. The owner of farm A (immunosuppressed following a kidney graft), and his son fell severely ill with respiratory symptoms. The manager of farm B presented with a mild respiratory disease. The local veterinarian (who had examined the cats) also developed a respiratory illness. All samples collected tested positive for influenza A. The four samples are:

- **INFL1/2.07** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the owner of farm A.
- **INFL1/2.08** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the son of the owner of farm A.
- **INFL1/2.09** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the manager of farm B.
- **INFL1/2.10** corresponds to raw sequencing reads obtained from a nasopharyngeal swab from the veterinarian.

The objectives for each scenario are as follows:

- Identify the virus reads from the dataset (Nanopore for INFL1 or Illumina for INFL2);
- If possible, reconstruct partial or full genomes of the virus(es) for characterisation, and comment on data quality;
- Identify the species and clade or lineage if possible (indicating the classification scheme);
- Report specific features that may be relevant to public health (analysis of mutations known to increase virulence or transmission in specific hosts, or affecting sensitivity to antivirals, or markers of adaption to mammals);
- Assess a potential link between the cases based on genomic information;
- For scenarios 1 (INFL1/2.02),2 (INFL1/2.03) and 4 (INFL1/2.07): Give advice on the possible treatment option(s): Antivirals to be considered are adamantanes, oseltamivir, zanamivir, and baloxavir.

Classification of influenza datasests

INFL1/2.01: Influenza A/H5N6, clade 2.3.4.4.h

INFL1/2.02: Influenza A/H5N6, clade 2.3.4.4.h

INFL1/2.03: Influenza A/H7N9 (HA, NA, PA, PB1, PB2) and A/H1N1pdm09 (MP, NP, NS)

INFL1/2.04: Rhinovirus

INFL1/2.05: Influenza A/H3N2s (HA), H3N2 (NA), H1N1pdm (PB2, PA, NP, MP), H1N1 eurav (PB1, NS), clade HA: 1990.4.a

INFL1/2.06: Influenza A/H3N2s (HA), H3N2 (NA), H1N1pdm (PB2, PA, NP, MP), H1N1 eurav (PB1, NS), clade HA: 1990.4.a

INFL1/2.07: Influenza A/H5N1, clade 2.3.4.4.b

INFL1/2.08: Influenza A/H5N1, clade 2.3.4.4.b

INFL1/2.09: Influenza A/H5N1, clade 2.3.4.4.b

INFL1/2.10: Influenza A/H5N1, clade 2.3.2.1.a

Clustering of influenza datasets

Scenario 1: Taking into account segments (or portions of segments) having good overall quality and sufficient coverage, it is likely that INFL1/2.01 and INFL1/2.02 are related. All segments seem to cluster together: they are all identical, except PB2, which has one mutation specifically in INFL1/2.02 (this mutation can be found in minor variants of INFL1/2.01).

Scenario 2: It is not possible to reconstruct any influenza consensus for INFL1/2.04, but we found Rhinovirus sequences, which makes the hypothesis of two separate infections more likely.

Scenario 3: Considering that segments (or portions of segments) have good overall quality and coverage, it is likely that INFL1/2.05 and INFL1/2.06 are related. All segments are strictly identical.

Scenario 4: INFL1/2.07 and INFL1/2.08 are highly similar (all segments are strictly identical, except NA with two mutations). It is likely that these two samples are related. INFL1/2.09 is similar to INFL1/2.07 and INFL1/2.08, but only for segments HA and NA (between four and seven mutations, compared to 65–270 mutations in the other segments). While these cases may be linked, a potential reassortment event may indicate a complex history that does not involve a direct link between these cases. It seems that INFL1/2.10 is not related to any other sample (50–300 mutations).

Description of the mutations of interest:

- PB2 T271A: Increased polymerase activity and replication in mammalian cells.
- PB2 I292V: Increased polymerase activity and replication in mammalian cells.
- PB2 K526R: Increased polymerase activity and replication in mammalian cells.
- PB2 A588V: Increased polymerase activity and replication in mammalian cells.
- PB2 E627K: Increased polymerase activity and replication in mammalian cells; increased virulence in mice.
- PB2 D701N: Virulence/host specificity shift.
- PB2 141-: This mutation has been artificially created for this EQA.
- PB1 H99Y: Host specificity shift.
- PB1 I368V: Host specificity shift, increased transmission among ferrets for A(H5N1) viruses.
- PA I38M: Potential reduced inhibition of baloxavir.
- HA T110N: Host specificity shift.
- HA N224K: Host specificity shift.
- HA Q226L: Host specificity shift/antigenic drift.
- HA T318I: Increased risk of transmission among ferrets; host specificity shift; and antigenic drift.
- HA H110Y: Host specificity shift.
- HA N158D: Host specificity shift.
- HA T160A: Increased virus binding to α2–6; increased transmission in guinea pigs.
- HA G228S: Host specificity shift.
- NA R292K: Potential reduced inhibition of oseltamivir, zanamivir, peramivir, and laninamivir.
- NA I97V+I294V: Potential reduced inhibition of oseltamivir.
- MP S31N: Potential reduced inhibition of adamantanes.

Susceptibility to antivirals

INFL1/2.02:

- R293K (R292K N2 numbering);
- Oseltamivir: HRI (15068) ;
- Zanamivir: RI/HRI (46) ;
- Baloxavir: Potentially susceptible;
- Adamantanes: Potentially susceptible.

INFL1/2.03:

- NA(H7N9):R289K (R292K N2 numbering); PA(H7N9):I38M; MP(H1N1pdm):S31N;
- Oseltamivir: HRI (>4 600);
- Zanamivir: RI (11–67);
- Baloxavir: Potentially susceptible; PA I38M has been described in seasonal influenza but not in H7N9. While it should not be advised to use baloxavir in the presence of such mutation, this solution remains the best one, but likely as a last resort;
- Adamantanes: HRI.

INFL1/2.07:

- NA: I97V + I294V (I117V + I314V N2 numbering);
- Oseltamivir: RI (16);
- Zanamivir: NI (1);

- Baloxavir: Potentially susceptible;
- Adamantanes: Potentially susceptible.

European Centre for Disease Prevention and Control (ECDC)

Gustav III:s Boulevard 40
16973 Solna, Sweden

Tel. +46 858 60 10 00
ECDC.info@ecdc.europa.eu

www.ecdc.europa.eu

Publications Office
of the European Union