Accreditation and inter-EURL working group guidelines

Valeria Michelacci

September 2nd 2026

INTER EURLs WORKING GROUP ON NGS

AIM

to promote the use of WGS across the EURLs' networks, build WGS capacity within the EU and ensure liaison with the work of the EURLs and the work of EFSA and ECDC on the WGS mandate sent by the Commission

Active since November 2017

Meeting twice a year
(**18** meetings done)

- Publication of supporting documents for WGS
- Joint activities: webinars, trainings for NRLs

Direct link!

https://zenodo.org/communities/eurls-biorisks_wg_on_ngs

Supporting documents available

in continuous update

«Guidelines for the validation of Whole Genome Sequencing for accreditation»

- Overview of conducted and planned PTs – curated by EURL Antimicrobial Resistance
- Reference Whole Genome Sequencing collection – curated by EURL Salmonella
- Guidance document for WGS-laboratory procedures – curated by EURL Parasites
- Bioinformatics tools for basic analysis of Next Generation Sequencing data – curated by EURL VTEC
- Guidance document for Whole Genome Sequencing - cluster analysis – curated by EURL Campylobacter
- Guidance document for NGS-Benchmarking – curated by EURL Listeria monocytogenes
- Inventory of training supports – curated by EURL Coagulase Positive Staphylococci
- Survey on the use of NGS across the NRLs networks – curated by EURL VTEC
- Supporting document for preparing high quality DNA for Whole Genome Sequencing – curated by EURL VTEC
- Guidelines for Dry-Lab Quality Control Metrics and Acceptance Thresholds in Short-Read WGS of zoonotic and indicator bacteria - curated by EURL Listeria monocytogenes
- Guidelines for amplicon sequencing of the capsid region of noroviruses - curated by EURL Foodborne Viruses
- Guidelines for the validation of Whole Genome Sequencing for accreditation - curated by EURL VTEC

Context

COMMISSION IMPLEMENTING REGULATION (EU) 2025/179

Whole Genome Sequencing (WGS) shall be performed on isolates of food-borne pathogens associated with food-borne outbreaks. For this, **National Reference Laboratories and/or Official Laboratories shall be accredited for this method according to EN ISO/IEC 17025:2017**

Species concerned:
Campylobacter coli and *C. jejuni*, *Shiga toxin-producing E. coli*, *Listeria monocytogenes* and *Salmonella enterica*

From FAQ 13: "*…A dedicated working group of EU Reference Laboratories is preparing guidelines on what is expected by official laboratories to be accredited for WGS.*"

Guidelines for the validation of Whole Genome Sequencing for accreditation

Version 01 – 30/03/2026

European Union Reference Laboratory for *Escherichia coli*,

Istituto Superiore di Sanità, Rome, Italy

(updated on 27/08/2026)

The document produced by the Inter EURLs WG for NGS does not replace the procedures of the Quality Assurance Systems in use in the accredited laboratories

Species concerned:
Campylobacter coli and *C. jejuni*, *Escherichia coli (focusing on Shiga toxin-producing E. coli)*, *Listeria monocytogenes* and *Salmonella enterica*

Scope of WGS data production

EU Reg 2025/179

- Data are produced and sent (raw) to EFSA — Accreditation of raw data production
- Data are submitted to EFSA after analysis to a certain extent (e.g. assemblies or cgMLST strings) — Partial accreditation of bioinformatic workflow

- Data are analysed for national purposes — Complete accreditation of bioinformatic workflow

Structure

- Producing WGS data
- Acceptance criteria for producing WGS data
- Bioinformatics analyses of WGS data
- Acceptance criteria for bioinformatic pipeline validation

Producing WGS data

References:

- Bacterial Manipulation: EN ISO 23418:2022
- DNA Isolation: EN ISO 23418:2022 criteria and controls in Table 1
- Library preparation: Primary EN ISO 23418:2022 criteria and controls in Table 1; Secondary ISO 20397-1:2021
- Sequencing process: Primary EN ISO 23418:2022; Secondary ISO 20397-1:2021
- Data quality metrics: ISO 20397-2:2021

Parameters:

• Repeatability (accuracy/precision) = (number of within-run replicates in agreement)/(total # within run replicates) x 100%

• Reproducibility (accuracy/precision) = (number of between-run replicates in agreement)/(total # between-run replicates) x 100%

• Agreement with other methods (accuracy/trueness) = [Number of correct results (positive and negative)/total number of samples] x 100%.

Table1

Step	Validation stage	Repeatability	Reproducibility	Agreement with other methods
1	Pure culture	Include cultures prepared on same day by operator 1. Use strain collection A[1].	Include cultures prepared on different days[3] by operator 2[4]. Use strain collection A[1].	Include strains with different characteristics. Use strain collection A+B[2].
2	DNA extraction	Include DNA extractions produced in triplicate from same culture (step 1) prepared on same day by operator 1, using the same batches of reagents.	Include DNA extractions prepared from cultures (step 1) on different days[3] by operator 2[4], using different batches of reagents.	Include DNA extractions prepared from cultures (step 1) from strains with different characteristics.
3	DNA sequencing	Include libraries prepared from DNA extractions (step 2) by operator 1 on the same day and (if possible) in the same run.	Include libraries prepared from DNA extractions (step 2) by operator 2[4] on different days[3] in different runs[3] of the same instrument, and on different instruments if relevant.	Include libraries prepared from DNA extractions (step 2) from strains with different characteristics.

[1]Strain collection A: A subset of strains representative of the scope of the validation (less than 10 strains, considering the maximum sequencing capacity in one run).

[2]Strain collection B: A subset of strains representative of different characteristics of the target organisms of the scope of the validation.

[3]On any day, except when operator 1 is doing step 1, 2 or 3, respectively for each step.

[4]The number of operators depends on routines of the laboratory.

Example of interpretation of validation of workflow stages as illustrated in Table 4 of ISO 23418:2022

The number of strains used for testing repeatability and reproducibility can be smaller than that of testing agreement with other methods.

When testing agreement with other methods it is important to cover a representative subset of relevant target characteristics.

The EURLs can provide strains for the validation, upon request by contacting the EURL for the specific pathogen

Acceptance criteria for producing WGS data

The following QC parameters can be determined (examples of acceptance criteria are available in Table 1 of Deliverable 10 or Table A.2 of EN ISO 23418:2022) to evaluate **repeatability and reproducibility**:

- minimum read depth
- base quality score
- genome size
- GC content
- level of contamination
- percentage of missing cgMLST loci
- identification of relevant targets

QC thresholds for each target species detailed in Table 1 of the Inter EURL WG document **"Guidelines for Dry-Lab Quality Control Metrics and Acceptance Thresholds in Short-Read WGS of zoonotic and indicator bacteria"**

The **accuracy** of the method compared to other methods: how well the sequences generated match relevant target characteristics (e.g. species, ST, virulence genes…) and group or exclude strains based on genetic relatedness, compared to expected results (obtained with a different method or available as reference data).

Repeatability and reproducibility = 100%
Accuracy (agreement with other methods) ≥ 90-95%

“Guidelines for Dry-Lab Quality Control Metrics and Acceptance Thresholds in Short-Read WGS of zoonotic and indicator bacteria”

e.g. Table 1 for *L. monocytogenes*

5.5 Listeria monocytogenes

QC steps	*Listeria monocytogenes*
Species confirmation	**Mash**
Reads QC: **Fastqc or Fastp**	**Fastp**
Reads QC: Minimum percent of High Quality of read (%)	≥ 80%
Reads QC: **tool_contamination_threshold**	**Kraken2 ≤ 2%** **Confindr ≤ 2%**
Q30 threshold (%)	_
Minimum depth (X)	**50X**
Assembly QC: N50 threshold (bp)	Determined but not critical
Assembly QC: N. contigs threshold	Determined but not critical
Assembly QC: **Min genome size (Mbp)**	**2.8**
Assembly QC: **Max genome size (Mbp)**	**3.2**
Tool for cgMLST calling	ChewBBACA / Ridom® SeqSphere
Schema used for cgMLST	ChewieNS (Moura) and Ridom® SeqSphere (Ruppitsch)
Accepted % missing cgMLST loci	**5%**
% GC content	

Bioinformatics analyses of WGS data

Reference: EN ISO 23418:2022 (Table 4)

Parameters:

- Repeatability (accuracy/precision): identical results should be obtained from the same data set by repeating the analysis on the **same infrastructure, using the same version of the software and the same parameters**.
- Reproducibility (accuracy/precision): consistent results should be obtained by analysing the same dataset in a **different computational environment** (e.g. different computer or computing cluster or supercomputing node), **using the same version of the software and the same parameters**.
- Agreement with other methods (accuracy/trueness): comparable results should be obtained by analysing the same dataset with **a different pipeline developed for the same application**

Sensitivity: proportion of true positive results obtained from samples known to contain the target analyte
Specificity: proportion of true negative results obtained from samples known not to contain the target analyte

According to EN ISO 16140-7 (on validation of identification methods) the number of strains to be tested at species level is **(at least) 25 strains per species,** from the most important/most frequently found serogroups/serovars/types per species.

Reference Data shall be used for validating the pipelines.
They may be provided by the EURLs upon request.

Acceptance criteria for bioinformatic pipeline validation

- Repeatability (accuracy/precision) = 100%
- Reproducibility (accuracy/precision) = 100%
- Agreement with other methods (accuracy/trueness) ≥ 90-95%
- Sensitivity and specificity reported in annex, different for the different species

e.g. Shiga toxin-producing *E. coli*

Determination	Sensitivity	Specificity
stx1, stx2, eae genes	100%	100%
Additional virulence genes	>95%	100%
O antigen*	>95%	100%
H antigen*	>95%	>95%
Sequence type (ST, result of MLST)*	>95%	>95%
stx subtype*	>70%	>70%

* For the O antigen (serogroup), the H antigen, the sequence type (ST), and the stx subtype, sensitivity can be calculated globally, whereas specificity can be determined for each individual characteristic (e.g. individual serogroups or H antigens)

Cluster analysis

The document does not comprise the accreditation of the cluster analysis

as it is meant to facilitate the accreditation of WGS workflow to respond to the Commission Implementing Regulation (EU) 2025/179

produce data and send to EFSA

Nevertheless, the laboratory may want to validate this process:

- Refer to ISO 23418:2022 Table 4
- Use data with known evolutionary relationships:
  - 5 genomes related + 5 unrelated
  - OR, where not available, by comparing the results obtained with other pipelines with the same scope (e.g. EFSA WGS pipeline, available for download and setup at this link)

Metadata

Metadata are integral to the process of outbreak investigation as are the WGS data.

From FAQ 6 appended to EU IR 2025/179:

"A genetic match of WGS profiles between a human isolate and a food isolate alone does not confirm the food as the outbreak source"

The clusters identified shall be further investigated, by assessing the sequences metadata

Refer to the EN ISO 23418:2022 for standardization of the ontologies

Thank you for your attention!

Questions?

Questions and Answers on Deliverable 12:

Comments on the Inter-EURL guidance for the validation of Whole Genome Sequencing for accreditation

Version 01 - 30/06/2026

European Union Reference Laboratory for *Escherichia coli*,

Istituto Superiore di Sanità, Rome, Italy