Data sharing for European Surveillance

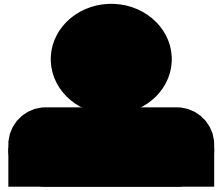
ECDC's mission

- **identify, assess and report** on current and emerging threats to human health from communicable diseases
- **provide recommendations** for response at the EU and national levels (and at the regional level, if necessary)

In order to meet these goals ECDC has to collect, collate, evaluate and disseminate relevant scientific, epidemiological and genomic information

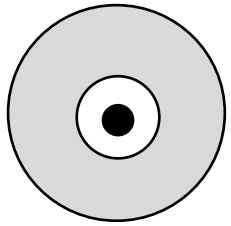
Source: EU-lex.europa.eu

Which data do we collect?



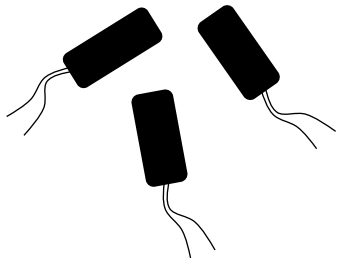
Age	Gender	Clinical presentation
39	M	Severe
26	F	Mild

Epidemiological data



```
Erythromycin: Resistant
Tetracycline: Susceptible
[…]
```

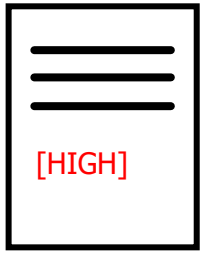
Phenotypic data



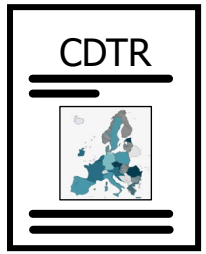
```
>E1101_kpneumoniae
ATGCAAGCTGACTG…
```

Molecular sequence data

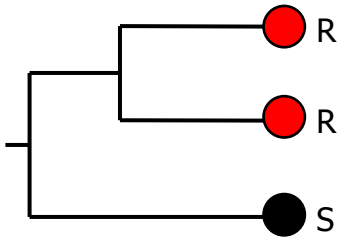
What do we do with the data?



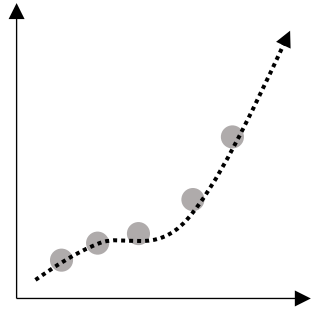
- Threat assessment Brief (TAB)
- Rapid Risk assessment (RRA)
- Rapid Outbreak Assessment (ROA)
- Joint Notification Summary (JNS)



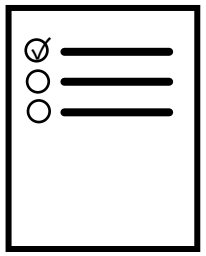
Weekly CDTR



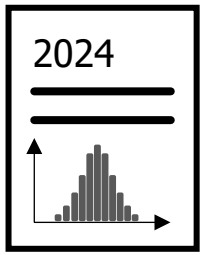
Genomic analyses and annotations



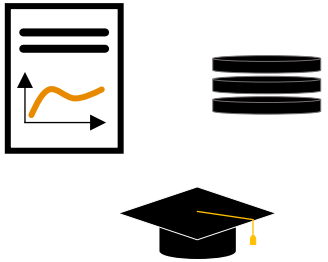
Epidemiological modelling



Recommendations and guidelines

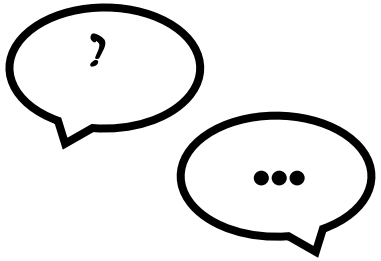


Annual epidemiological reports



Reports, publications, datasets, trainings, …

Why data sharing makes sense



Sharing data allows ECDC to **produce outputs with added value** to the countries. Main objective is to **assess risks**. All products are part of a **continuous dialogue** and exchange of information between public-health institutions and disease networks within Europe and beyond. This is required to face current and future public-health threats.

Long-term objective: strengthen Europe's defences against infectious diseases.

Why molecular data are so important?

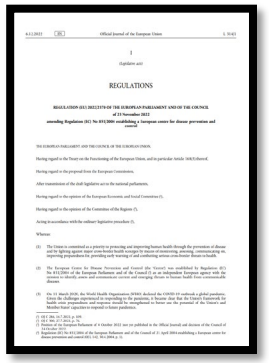
- We can infer some properties of the pathogen
  - Virulence and transmissibility
  - Resistance to antimicrobials
  - Effectiveness of vaccines

- Investigate the relationship between isolates/samples/cases
  - Allows for very precise case definitions
  - Outbreak delineation with a high degree of confidence
  - May identify additional cases that initially have no obvious epidemiological link to other outbreak cases

Nov 2022: new ECDC mandate and regulations

Regulation on serious cross-border threats to health

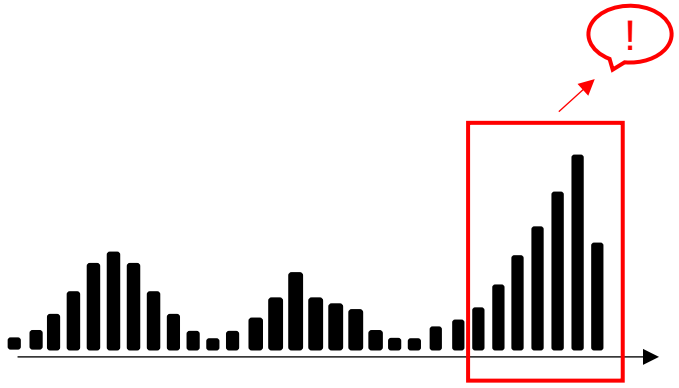
The national competent authorities [..] shall communicate the following information [..] to the participating authorities of the network for epidemiological surveillance:
[…]
molecular pathogen data, if required for detecting or investigating serious cross-border threats to health



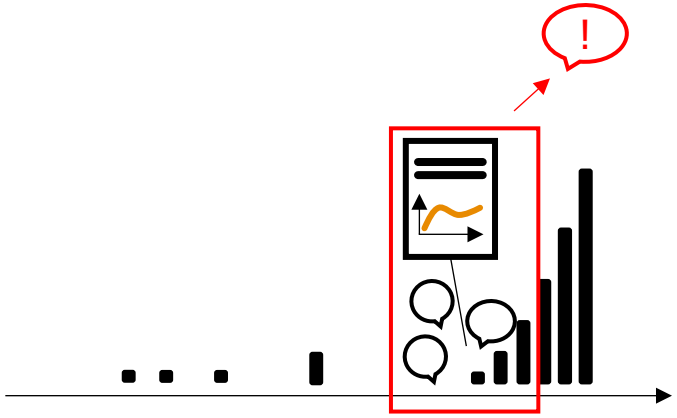
ECDC mandate

The Centre should broaden its collection and analysis of data in terms of epidemiological surveillance and related special health issues, progression of epidemic situations, unusual epidemic phenomena or new diseases of unknown origin, including in third countries, **molecular pathogen data** and health systems data.

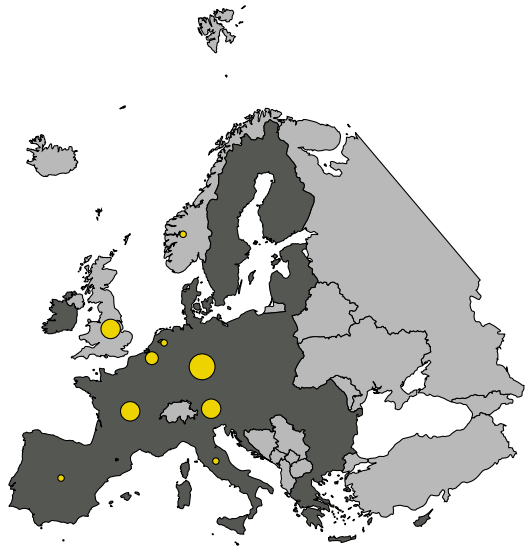
Different types of surveillance



Indicator-based surveillance
(e.g. sentinel surveillance)

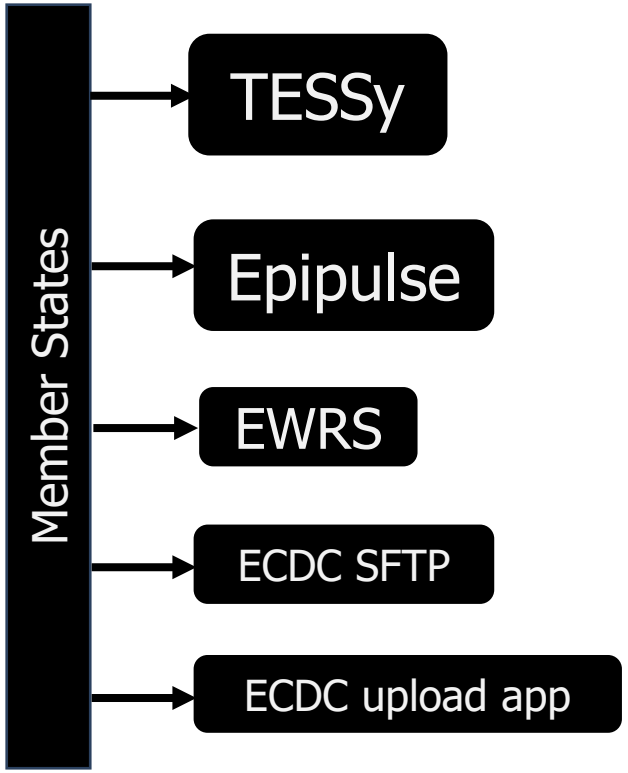


Event-based surveillance
(e.g. outbreaks)



Strategy-oriented surveillance
(e.g. structured surveys)

How do we collect the data?



Designed for

Collecting, analysing and disseminating surveillance data.

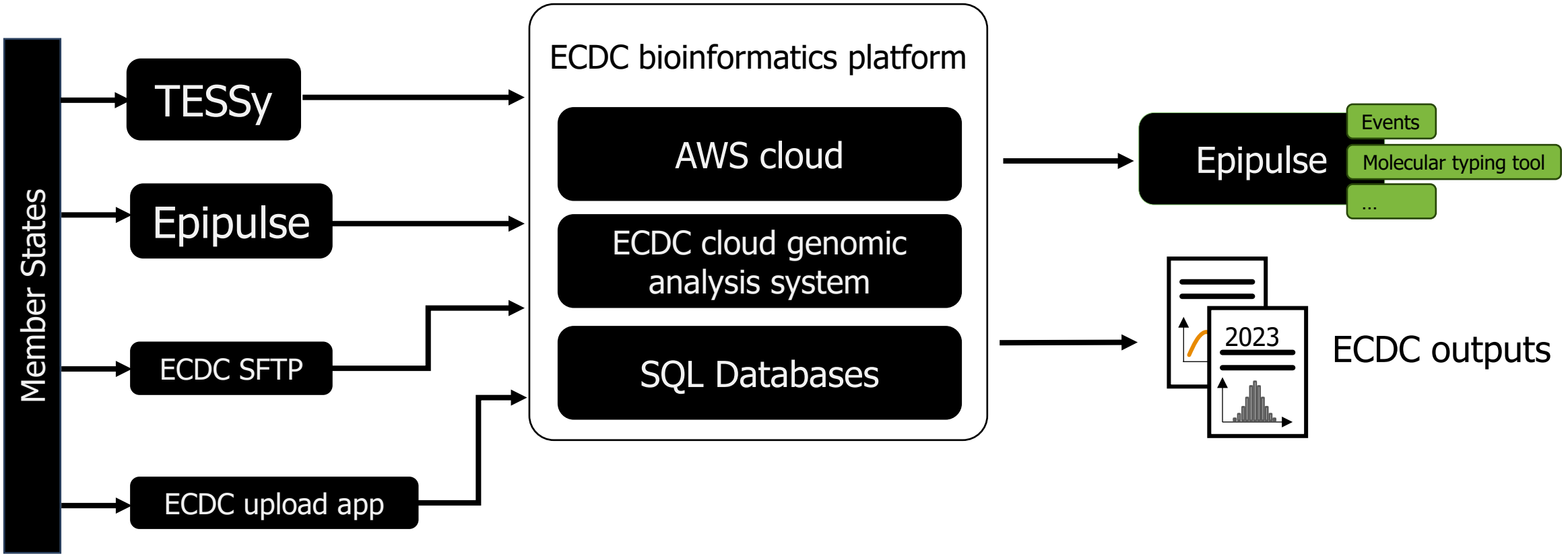
Collecting, analysing, sharing, and discussing infectious disease data for threat detection, monitoring, risk assessment and outbreak response. Data are shared on a voluntary basis.

Formal notification of outbreaks, clusters and contact tracing data. Data are used for risk management.

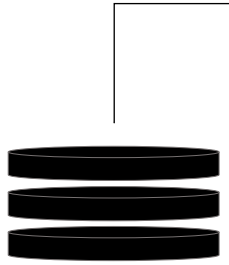
Genomic data upload (ad-hoc reporting, GDPR-compliant space).

Genomic and epidemiological data upload.

Current ECDC ecosystem



TESSy (The European Surveillance System)



- Main data source for indicator-based surveillance
- Enables data standardization and harmonization
- Allows the reporting of case-based or aggregated data
- Policies in place to regulate data access/visibility
- The format and type of data of all TESSy entries are defined by a “record type”.

More info at: https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy

An example of record type: SALMISO

Record type	Variable	Description	Type	…
SALMISO	RecordId	Unique identifier for each isolate within the data source / lab system.	TEXT	…
SALMISO	Age	Age of patient in years as received or at the date of sampling.	NUM	…
SALMISO	DateUsedForStatistics	The most epidemiologically relevant date for the isolate. Equal to the date of sampling if available. […]	DATE	…
SALMISO	Serotype	Salmonella serotype predicted from molecular methods.	CV	…
SALMISO	Specimen	The relevant specimen type used for diagnosis of the case	CV	…
…	…	…	…	…

- There are 172 different record types in TESSy (MetaDataSet 53, last update: 2023-10-17)
- Some variables are mandatory, while others are optional

SpecimenSALM (CVs):

- BLOOD = Blood
- CSF = Cerebrospinal fluid
- FAECES = Faeces
- NA = Not applicable
- OTHER = Other
- …

SerotypeSalm coded values (CVs):

- AACHEN
- AARHUS
- …

Reporting a severe case of salmonellosis on TESSy

Step 1: choose the right record type. We will use SALMISO

- see MetaDataSet 53 for further information.

↓

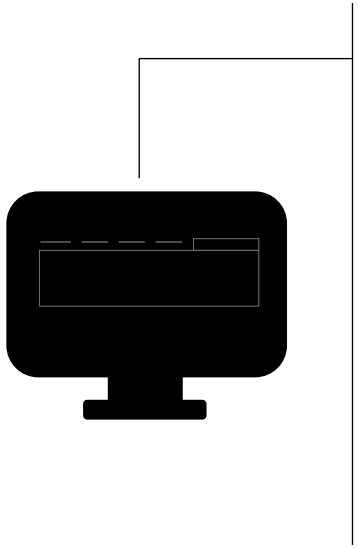
Step 2: collate the information (*e.g.* spreadsheet) and save in csv format

	A	B	C	D	E	F	G	H	I	J	K
1	RecordId	Subject	RecordType	Status	ReportingCountry	DateUsedForStatistics	Age	DateOfReceiptReferenceLab	Serotype	Specimen	DataSource
2	455e83	SALMISO	SALMISO	NEW	SE	2024-03-06	40	2024-03-08	HAVANA	UNK	SE-SALMISO

↓

Step 3: upload to TESSy (nominated users). Notes: (1) You will go through data validation; (2) Your submission might need to be approved.

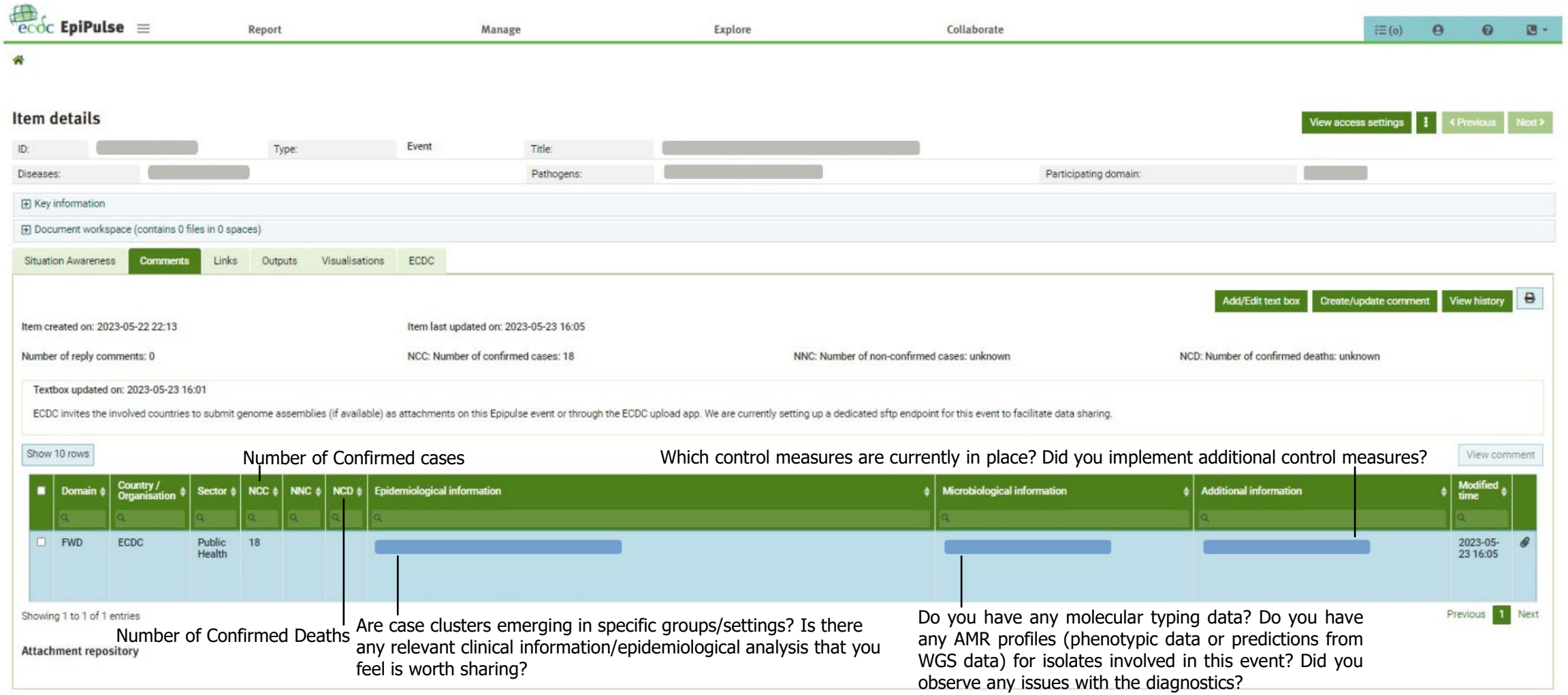
Epipulse (*the European surveillance portal for infectious diseases*)



- ECDC platform for event-based surveillance
- Also allows for continuous surveillance using WGS data (Molecular Typing tool)
- allows interactions between public-health institutes and ECDC
- designated officials from Member States can notify outbreaks / clusters
- domains span all the infectious diseases covered by ECDC

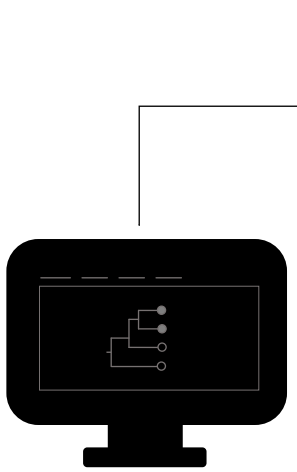
Link: https://epipulse.ecdc.europa.eu/

Epipulse events



Path on Epipulse: ☰ > [Explore tab] Events, Forum & News > {Item ID}

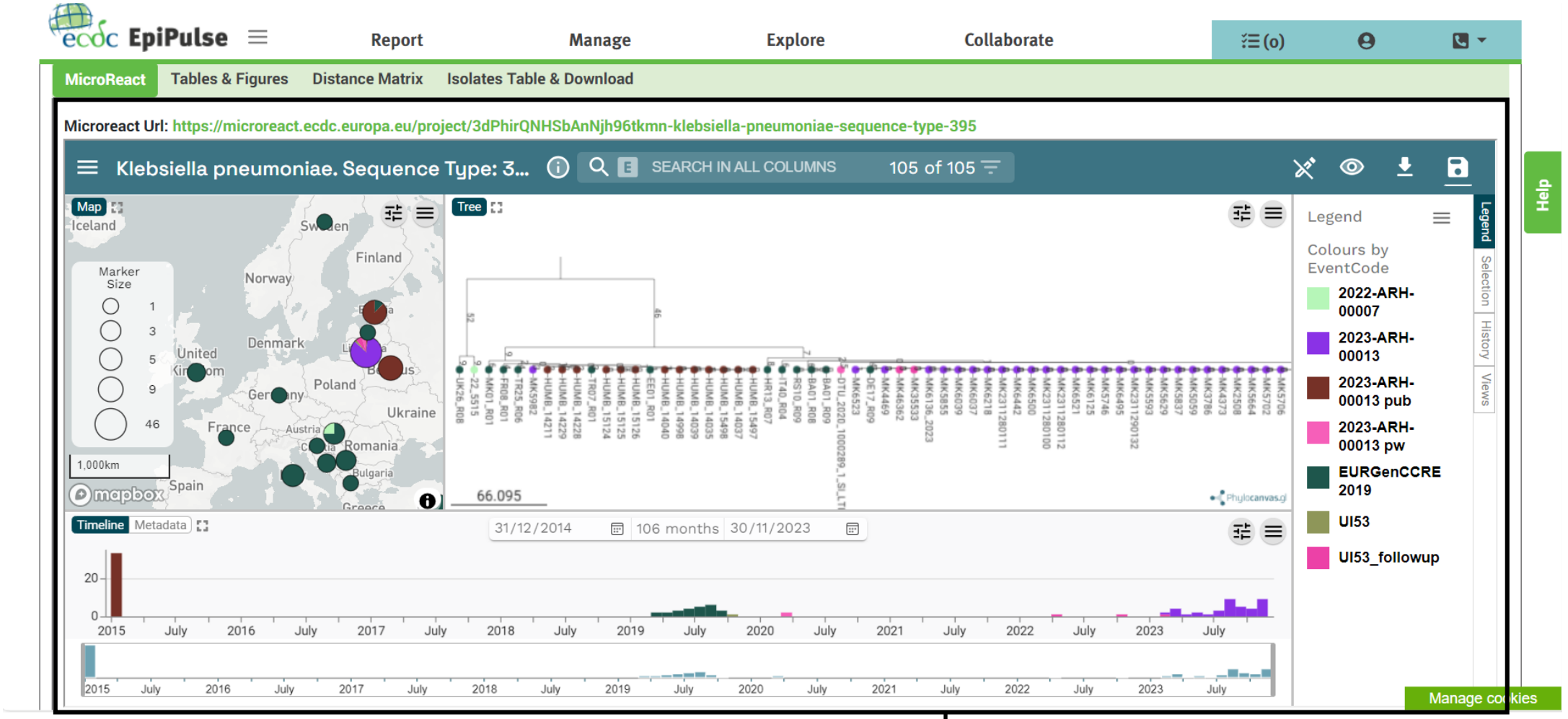
Epipulse: the Molecular typing tool



- The Molecular Typing tool is the Epipulse tool to visualize and analyse genetic clusters
- When genomic sequences are shared with ECDC, they are automatically analysed to:
  - find if they are genetically related to other known sequences (*i.e.* are part of a cluster) → can generate signals. Cut-offs to define clusters depend on the typing scheme used.
  - generate antimicrobial resistance profile predictions
- It is possible to associate clusters identified by the Molecular Typing tool to Epipulse events
- You can perform queries using your isolates ID and even set your own AD thresholds. You can export data (e.g. AMR profiles, typing data).

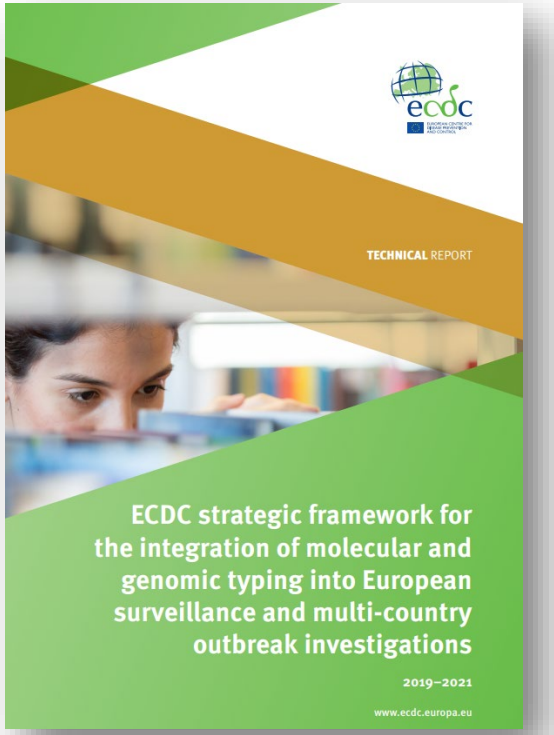
Path on Epipulse: ☰ > [Explore tab] Molecular typing tool

Explore molecular data with the Molecular typing tool



Microreact

ECDC strategy for integrated genomic typing



Support to multi-country outbreak investigations through sequence-based typing: *Campylobacter* spp., *Clostridium difficile*, hepatitis A virus, *Legionella* spp., *Listeria monocytogenes*, multidrug-resistant *Mycobacterium tuberculosis* (MDR TB), *Neisseria meningitidis*, *Salmonella enterica*, Shiga-toxin producing *E. coli,* West Nile virus and emerging multi- or extensively drug-resistant (MDR or XDR) bacteria, new pathogens or new modes of transmission of healthcare-associated or community pathogens.

EU-wide sequence-based continuous surveillance: influenza virus, *Listeria monocytogenes*, MDR TB, *Neisseria meningitidis*, *Salmonella enterica* and Shiga-toxin producing *E. coli*.

Sentinel surveillance or surveys: antibiotic-resistant *Neisseria gonorrhoeae, Bordetella pertussis*, carbapenem- or colistin-resistant Enterobacteriaceae, carbapenem-resistant *Acinetobacter baumannii*, HIV-transmitted drug resistance, and *Streptococcus pneumoniae*.

Source: ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigation

EU/EEA molecular surveillance roadmap: requirements

- CLEAR SURVEILLANCE OBJECTIVES, INDICATORS and OUTPUTS.
  - Developed by ECDC in close collaboration with countries.

- PROTOCOLS in place for data collection (e.g. sample preparation, sampling, QC).
  - Relevant to ensure data comparability. Developed by ECDC in close collaboration with countries.

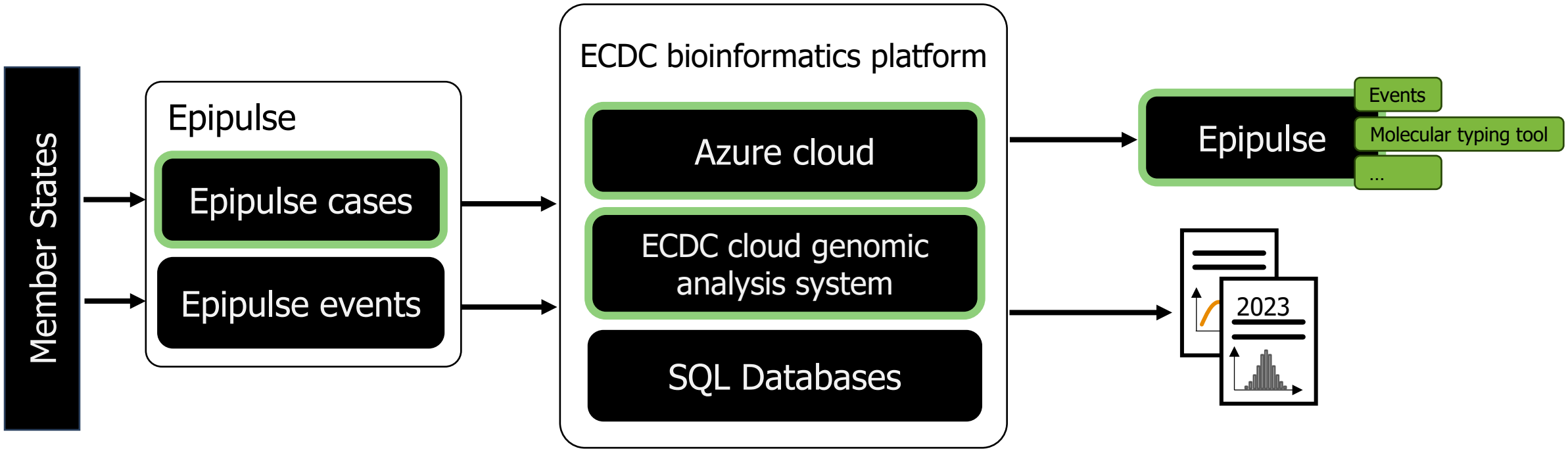
EU/EEA molecular surveillance roadmap: requirements

- INFRASTRUCTURE in place from data collection to output generation.
  - Integration of public data sources. Data upload should be perceived as seamless or as a low burden.

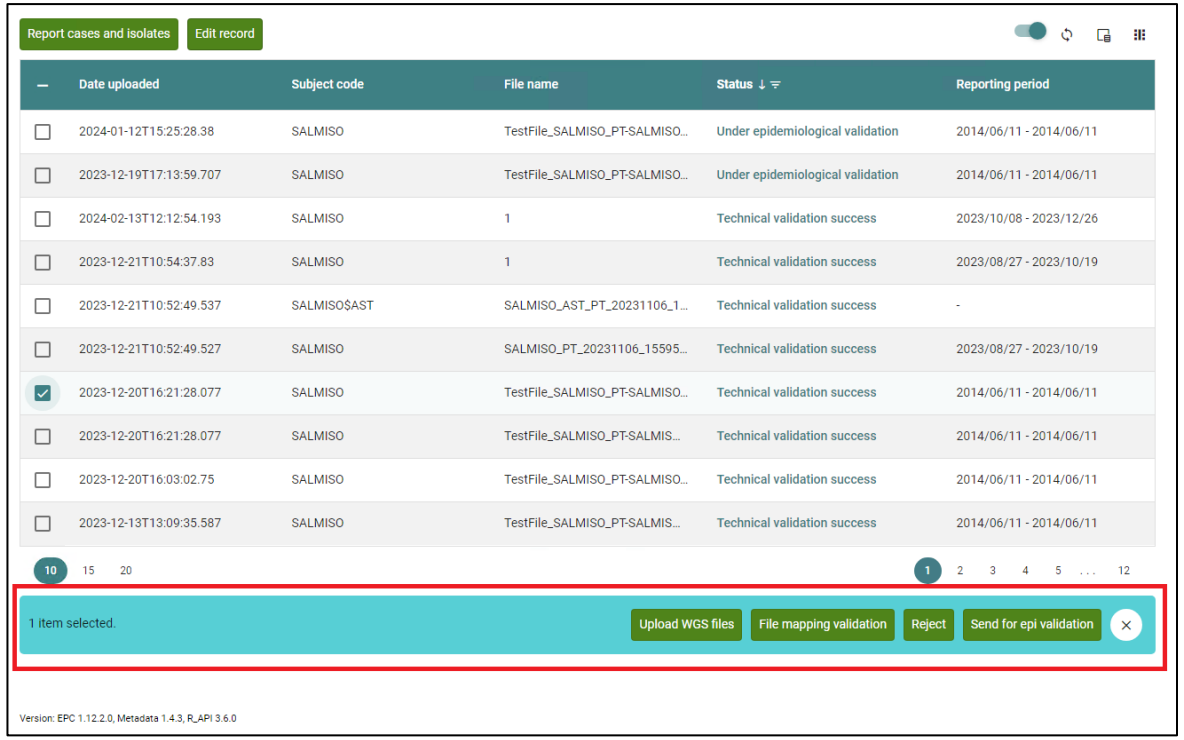
- DATA FLOW. Data are flowing to a rate and volume which is sufficient to fulfil the surveillance objectives and that allows take quick public-health actions.

- LEGAL ASPECTS. Must follow the regulations on processing of personal data.

INFRASTRUCTURE. Future ECDC ecosystem



INFRASTRUCTURE. Updates on Epipulse cases.



- Unified tool to report epidemiological and genomic data.
- New User Interface (UI) → objective: improve the user experience and lower the burden associated to data reporting.
- API to automate the data upload.
- Will enter in production in Q3 2024.
- The "subjects" that will be supported at launch will be the ones from the VPD disease group [epidemiological data] and some isolate subject codes [genomic and molecular typing data] including AMRISO. Implementation of all other diseases expected by Q3 2025.

INFRASTRUCTURE. Updates on the Molecular typing tool.

- The user interface will be revamped (objective: improved user experience) by the end of 2024.
- *A. baumannii* will be added to the list of available pathogens on the Molecular Typing tool by the end of this year.
- Klebsiella: Kleborate implemented by Q2 2024.
- [long-term] Expand the range of analyses for AMR pathogens making use of long-reads data, for instance to characterize mobile genetic elements.

LEGAL ASPECTS. Provisions on protection of personal data.

1. ECDC shall process personal data in accordance with Regulation (EU) 2018/1725 [which specifically addresses the processing of personal data by Union institutions, bodies, offices, and agencies] […].

2. Member States shall process personal data made available to or by ECDC in accordance with Regulation (EU) 2016/679 [GDPR] and with any national law applicable.

3. ECDC and the Member States are acting as joint controllers for any analysis on the data that is carried out jointly by ECDC and the Member States in the framework of a surveillance network and under the Protocol. The Member States shall inform data subjects about such processing operations (including through a data protection notice) and shall act as contact points for the data subjects that want to exercise the rights conferred upon them by data protection legislation.

LEGAL ASPECTS. Provisions on protection of personal data.

4. Each Member State is controller for any processing prior to the transfer of the data by the respective Member State to ECDC, including pseudonymisation. Each Member State is controller for any processing operation undertaken by the Member State for own purposes on data transferred from ECDC, including storage by the Member State and analysis of the data.

5. ECDC is controller for the processing operations related to storage by ECDC of the data provided by the Member States, as well as for any further processing that is undertaken by ECDC to fulfil its mandate.

6. In case of personal data breach, the entity where the incident occurs shall immediately (and in any case not later than 48 hours after the discovery of the breach) inform the Member States whose personal data is affected. The Member States shall then inform the data subjects, where this is required by law taking into account the risks related to the breach.

LEGAL ASPECTS. Provisions on protection of personal data.

7. ECDC shall refrain from activities aimed at re-identifying data subjects whose data have been subject to processing for the purposes described in the Protocol.
8. The *ECDC's policy on data submission, access, and use of data within TESSy* includes provisions on access by third parties to data stored in ECDC's surveillance platform. Such document applies for the data covered by the Protocol.
9. ECDC and the Member States shall designate contact points to discuss any data protection issue that might arise in the implementation of the Protocol, and to coordinate to ensure that data subjects can exercise their rights.
10. The courts of Stockholm shall have jurisdiction controversies arising between ECDC and the Member States on the interpretation and application of this Annex.

Conclusions

- Sharing molecular data is key to respond to public health threats quickly and appropriately.
  - The regulations on serious cross-border threats to health and the new ECDC mandate require Member states to report molecular data for cross-border events. A deeper integration of molecular and epidemiological data provide the best basis for robust data analyses and interpretations.
  - ECDC is making efforts to simplify and reduce the work burden associated with data reporting.
- In order successfully implement the molecular surveillance roadmap a continuous exchange with you is needed.
  - Which aspects should we improve? Which data would you like to see? How can we establish a data flow that is sufficient to make more rapid and impactful decisions?
- We have established a set of provisions to ensure that we follow the regulation on the protection of personal data.
  - This is a still ongoing process we will make sure to inform you on the progress we make and the challenges we face.

Thank you for your attention