



---

**Deliverable: Translation of the source attribution methodological approach to the One-Health context for surveillance of Salmonella, Campylobacter, VTEC, and antimicrobial resistance: options, needs**

**D-JRPFBZ-1-WP5.3**

**Workpackage 5**

Responsible Partner: ISS

Contributing partners: DTU, ANSES, APHA, BfR, DAFM, INIAV, INSA, ISS, NIPH, NVI, PHE, PIWET, RIVM, SSI, SVA, Teagasc, UUNCOH, VRI, WBVR, VISAVET



## GENERAL INFORMATION

European Joint Programme full title	Promoting One Health in Europe through joint actions on foodborne zoonoses, antimicrobial resistance and emerging microbiological hazards
European Joint Programme acronym	One Health EJP
Funding	This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 773830.
Grant Agreement	Grant agreement n° 773830
Start Date	01/01/2018
Duration	60 Months

## DOCUMENT MANAGEMENT

JIP/JRP deliverable	D-JRPFZ-1-WP5.3
Project Acronym	DiSCoVeR
Author	Gaia Scavia (ISS) Tine Hald (DTU) Lapo Mughini Gras (RIVM) Sara Pires (DTU)
Other contributors	All projects partners have contributed to the results and discussions feeding into this documents.
Due month of the report	57
Actual submission month	60
Type <i>R: Document, report</i> <i>DEC: Websites, patent filings, videos, etc.;</i> <i>OTHER</i>	R Save date: 22/12/2022
Dissemination level <i>PU: Public (default)</i> <i>CO: confidential, only for members of the consortium (including the Commission Services)</i>	PU This is the default setting. If this project deliverable should be confidential, please add justification here (may be assessed by PMT): ..... .....
Dissemination	OHEJP WP 1 <input type="checkbox"/> OHEJP WP 2 <input type="checkbox"/> OHEJP WP 3 <input type="checkbox"/>



<b>Author's suggestion to inform the following possible interested parties.</b>	OHEJP WP 4 <input type="checkbox"/>	OHEJP WP 5 <input checked="" type="checkbox"/>	OHEJP WP 6 <input type="checkbox"/>
	OHEJP WP 7 <input type="checkbox"/>	Project Management Team <input checked="" type="checkbox"/>	
	Communication Team <input type="checkbox"/>	Scientific Steering Board <input checked="" type="checkbox"/>	
	National Stakeholders/Program Owners Committee <input type="checkbox"/>		
	EFSA <input checked="" type="checkbox"/>	ECDC <input checked="" type="checkbox"/>	
	Other international stakeholder(s): .....		
	Social Media: .....		
	<b>Other recipient(s):</b> .....		

**Preamble - Rationale.....4**

**Setting the DISCOVER source attribution process .....5**

**Methods ..... 5**

**Data need, data availability at the EU level and critical evaluation of the DISCOVER datasets..... 6**

    Optimal data requirements..... 6

    Data quality ..... 7

    Data availability at the EU level..... 7

    The DISCOVER datasets..... 8

    Evaluation of the DISCOVER datasets ..... 9

    Optimisation of dataset for SA..... 12

**Skill ..... 12**

    The DiSCoVeR scientific community..... 12

**Conclusions .....13**

**References: .....14**



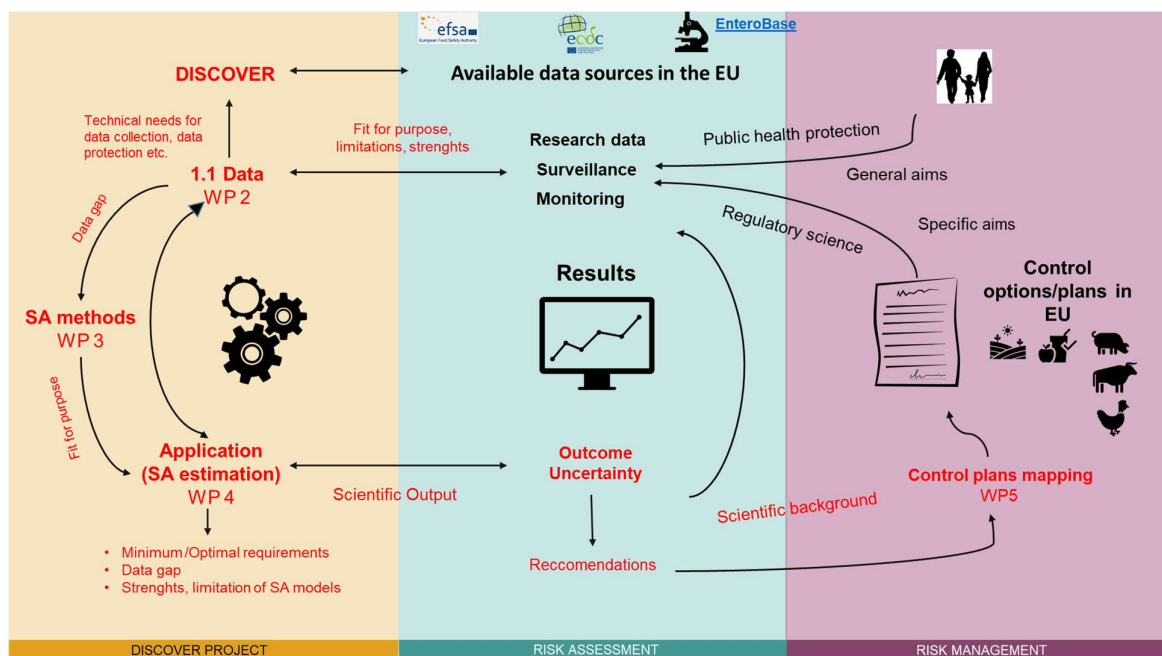
## Preamble - Rationale

The source attribution (SA) activities carried out by the DiSCoVeR project aimed at evaluating through different methodological approaches how much different animal reservoirs, food production chains and transmission routes contribute to the burden of disease in humans caused by *Salmonella*, *Campylobacter*, Shigatoxin-producing *E.coli* (STEC) and antimicrobial resistance (AMR) with particular focus on Extended-Spectrum Beta-lactamase *E. coli*.

SA has been defined as the partitioning of a human disease burden to specific animal sources (Pires et al. 2009). However, the application of SA studies to animals and to livestock may be sub-optimal for zoonotic foodborne pathogens with complex epidemiology that can also spread to humans by non-foodborne transmission routes such as direct animal contact or contaminated environmental sources (drinking water, wastewater, soil, top-soil improvers, etc.) or non-food producing animal reservoirs (wildlife, pets including horses, etc.) or by inter-human transmission. This is the case of the hazards targeted by the DiSCoVeR, which are considered top-priority zoonotic foodborne hazards in the EU, although it is well known that they are non-100% foodborne pathogens. They can in fact be transmitted to humans from multiple known and unknown reservoirs and transmission routes, which go beyond the foodborne route.

SA is important to guide intervention to control the spread of zoonotic pathogens from animal reservoirs and non-human sources to humans and support decision making.

In this document, the main components and activities of SA (figure 1) within the DiSCoVeR are summarised and critically evaluated to provide suggestions to optimise and ease the translation of the SA DiSCoVeR approach into the field, support capacity-building for SA and translate the results into actions in the field of the DiSCoVeR project.



**Figure 1. Main components and activities of the SA process within the DiSCoVeR (left panel) and the connections with the sectors of risk assessment (middle panel) and risk management (right panel).**



## Setting the DiSCoVeR source attribution process

The DiSCoVeR project applied multiple methodological approaches to maximise the exploitation of existing data and data collected during the project (DiSCoVeR partners' data, data obtained through partners' network, as well as public available data) and to encompass various levels of the epidemiological transmission pathways, i.e. the reservoir level and the exposure level. This allowed to comprehensively explore the fit-for purpose of the different methods in relation to the quality and the amount of data and conversely to highlight the current data gaps to be covered to optimise future SA exercises. SA is a data driven process, where the quality of data (e.g. information on the hazard) and metadata (information on the sample and the sampling context), the criteria for sampling and source selection and prioritization, are key elements.

### Core components of the DiSCoVeR SA are:

- Methods
- Data
- Skills

### Methods

SA analyses in the DiSCoVeR were carried out using various methodological approaches to obtain comparable results. The objective was to maximise the exploitation of available data, typically collected for many different purposes using a wide range of study designs, in order to better characterise and quantify sources and transmission routes through the comparison or combination of outcomes from different approaches. An exhaustive description of existing attribution methods and principles is available in recent article by Mughini-Gras et al. (2019). As part of DiSCoVeR, quantitative SA methods have been catalogued and analysed in the D-JRPFBZ-1WP3.2 (*Report on the critical and quantitative evaluation of existing and novel source attribution methods*).

Briefly, methods applied in DiSCoVeR for SA, by pathogen are summarised in table 1 and comprised the following approaches:

- **SA methods at the reservoir level:** these methods identify the primary source of human infection but do not fully clarify the transmission route(s) of zoonotic infection (e.g. foodborne, waterborne, direct contact etc.). Methods focusing on reservoirs are based on microbial typing/subtyping and aim to compare the distribution of hazard subtypes in the different animal reservoirs and environmental sources and in humans to quantify the relative contribution of the sources to human infection. Typing/subtyping-based methods for SA have evolved in parallel to the technical evolution and have moved towards large-scale application of more discriminatory characterisation methods, mainly due to reduction of costs. Phenotypic-based methods (i.e. using serotyping, phagotyping, ribotyping, AMR profiling etc.) have been progressively replaced by genotype-based methods (MLVA, MLST, virulotyping) and Whole-genome-sequence (WGS) based methods.
- Frequency matching SA methods based on phenotyping and also genotyping (e.g. MLVA) data has been widely used by for many years to guide control actions e.g. for *Salmonella*. Information on serotypes, phagetype and antimicrobial resistance etc. is/has been collected in case-based surveillance making this approach suitable for the application to 'large' amount of data collected in surveillance. This microbial subtyping approach requires the minimum



set of data, at current. SA based on frequency matching has been applied to *Salmonella* and STEC (table 1). For further details, see Deliverables D-JRPFZ-1-WP4.1/2/3/4.

- The application of SA methods making use of WGS data (machine-learning, population structure) allowed to reduce remarkably the uncertainty compared to previous studies since the use of sequence based characterisation methods have a high discriminatory power. In DiSCoVeR this approach was used for all hazards (Table 1).

Methods based on microbial subtyping approaches were used to perform the analysis at different geographical level, mainly nationally, regionally and/or more generally for the EU if possible and appropriate

- **At the exposure/near exposure level, methods** applied for SA in the DiSCoVeR were (Table 1):
  - Metanalysis of case-control studies from the scientific literature.
  - Methods based on aggregation of data from outbreak investigation, in collaboration with EFSA.
  - Exposure assessment based approaches which looks into exposure (e.g. animal contact) and hazard prevalences at the exposure level
  - Infectious disease modelling based on compartmental multidirectional dynamic models for AMR

A critical assessment of the existing quantitative model for SA and how to improve existing models including those that account for multidirectional transmission mode of agent between reservoir and humans have been carried out under WP3. Results are summarised in Deliverable D-JRPFZ-1-WP3.2

**Table 1: Summary of the methodological approaches used for the source attribution analysis, by pathogen (DiSCoVeR, 2022)**

	Methods						
	Microbial Subtyping approach			Meta-analysis of case control studies	Analysis of outbreak data	Comparative exposure assessment	Multi-directional dynamic model
	Frequency-Matching	Machine Learning	Population genetic				
<i>Salmonella</i>	yes	yes	yes	yes	yes	yes	
<i>Campylobacter</i>		yes	yes	yes		yes	
STEC	yes	yes	yes			yes	
AMR (ESBL E.coli)	yes	yes				yes	yes

## Data need, data availability at the EU level and critical evaluation of the DiSCoVeR datasets

### Optimal data requirements

Data needs for SA according to the different methodological approaches are shown in figure 2 and include information on infectious agents, the source, the sampling context, the production systems, human hosts including information on the clinical condition, demographic and animal population data (human population, animal population, food products, animal and products volumes,



companion animals population, wildlife numbers and distribution).

Method	Point of attribution				Pathogen data		Human data			Source data		
	Reservoir*	Exposure	Transmission route	Risk factor	Subtyping	Fitness	Subtyping	Sporadic or epidemic status	Exposure & risk factors**	Subtyping	Prevalence and exposure***	Need to consider all potential sources
Epidemiological (case-control study)	No	Yes	Yes	Yes	No	No	No	Sporadic	Yes	No	No	High
Epidemiological (outbreak investigation)	No	Yes	Yes	Yes	No	No	No <sup>†</sup>	Epidemic	Yes	No	No	High
Microbiological (frequency-matching models)	Yes	Yes	Yes	No	Yes	No <sup>†</sup>	Yes	Sporadic and epidemic	No <sup>†</sup>	Yes	Yes	Medium
Microbiological (population genetics models)	Yes	Yes	Yes	No	Yes	No	Yes	Sporadic or epidemic	No	Yes	No	High
Quantitative exposure/risk assessment	Yes	Yes	Yes	Yes	No	Yes	No	N/A	Yes	No	Yes	Medium
Expert elicitations	Yes	Yes	Yes	Yes	No	No	No	Sporadic or Epidemic	No	No	No	Low

N/A, not applicable. \*Or amplifying hosts. \*\*Including travel status. \*\*\*In this case, food consumption weights. <sup>†</sup>But can be used if available.

The use of different categories of data type depend on the models and approaches. DiSCoVeR got data from many different sources. Many were from surveillance and monitoring of the hazards in individual countries, survey and studies from published and un-published, national and international statistics including industry annual report, and expert opinions.

### Data quality

Good quality data underpinning SA analyses is key for good quality SA estimates. In particular data should be accurate and representative (i.e. they should ideally be obtained from, but at least be representative of the population under investigation). Data and isolates collected under surveillance and monitoring programs in human and non-human reservoirs would ideally represent the best option because they are usually associated with high quality attributes/metadata (i.e. detailed information on source/sample/risk factors) on the domain of provenance of the samples/ isolates and are collected within specific sampling framework. These characteristics are important to both optimise the quality of input data (i.e. making it possible to select the more appropriate isolates according to the SA method being applied) and support the inference process of SA outputs, which is also important for the evaluation of the uncertainty around SA results.

### Data availability at the EU level

Information on typing/subtyping of *Salmonella*, *Campylobacter*, Shigatoxin-producing *E. coli* and Extended-Spectrum Beta-lactamase *E. coli*, and the associated metadata (i.e. data describing the sampling context and the characteristic of the host/source/reservoir) were needed for DiSCoVeR SA.

At the EU-level, data on zoonotic agents typing/subtyping in humans and in animals, food and feed are collected from Member States within the case-based surveillance in humans (ECDC-Atlas, 2021 ) and the EFSA monitoring zoonotic agents (EFSA, 2022). Typing/subtyping Information focuses on species, serogroup/serotypes, phagetypes, AMR profiles and virulence genes types and subtypes (for STEC) despite descriptive metadata are generally highly detailed. However, the reporting of microbial typing/subtyping is not mandatory (except for data collected within Reg. (CE) 2160/2003 on Salmonella National Control Programs and Reg. (CE) 2073/2005 on *Salmonella* and STEC microbiological food-safety criteria).

WGS-based data collection, which for SA methods is the best option due to the high discriminatory power, is not well established within the surveillance and monitoring framework for foodborne pathogens and AMR at the EU level, although the main EU bodies have recently made important progresses. WGS-data are widely used at country level or EU level to support cluster detection and outbreak investigation under the framework of event-based surveillance (EpiPulse). EFSA and ECDC have recently developed a One-Health interoperable WGS system for *Salmonella*, *Listeria* and STEC which could be queried in order to discover matching profiles to mainly support cluster detection,



outbreak investigation and source tracking (EFSA, 2022b). Data are stored according to specific legal framework and the interoperability of the two systems is subjected to specific business rules.

### The DiSCoVeR datasets

The DiSCoVeR SA activities allowed us to use a wealth of data that were either originally produced for the DiSCoVeR needs or collected for other purposes by the DiSCoVeR partners (i.e. data from surveillance & monitoring, survey, official control, research, literature) or available in public repositories (e.g. EFSA dataset for foodborne outbreaks). This approach together with the production of secondary evidence studies (e.g. meta-analysis, analysis of foodborne outbreaks) matches well with international policies recommending re-use of data such as FAIR (Findability, Accessibility, Interoperability, and Reuse) (<https://www.go-fair.org/fair-principles/>). At the same time, the re-use of data was remarkably challenged by barriers<sup>1</sup> and/or complexity hampering accessibility to existing data in either the public or private domain or due to the lack of harmonisation.

Barriers for data sharing with specific focus on the use of WGS data in the context of foodborne outbreak detection and investigation were discussed with EFSA and ECDC with main focus on:

- **Technical:** e.g. lack of harmonization and standardization of methodological processes (e.g. pipelines for WGS analysis)
- **Practical:** lack of open platforms for sharing genomic sequences associated with metadata
- **Formal:** barriers associated with restrictive rules and policies for data protection and data usability that also go beyond the legitimate legal requirements (i.e. ownership of data and data usability subjected to specific restrictive terms of reference etc.). Overcoming these barriers would be greatly beneficial for future translation of SA into the field and call into action also actors not traditionally involved in the One-Health paradigm such as lawyers, bio-ethicists, food companies etc. to commit on how to do this in the future, as part of the societal engagement and responsibility in protecting public health

Inventory of available data for *Salmonella*, *Campylobacter*, STEC, and ESBL *E.coli* and metadata was done by the DiSCoVeR WP2. All the DiSCoVeR datasets were single structured stand-alone data collections, which received active contribution from the OHEJP partners. As such, datasets have been implemented for the purpose at the best of partners' ability in order to comply with the specific 'optimal' data requirements according to different SA methods implemented under the DiSCoVeR (DiSCoVeR D-JRPFZ-1-WP2.6-data description). The DiSCoVeR datasets include typing/subtyping information as well as descriptive metadata (description of the samples and sampling context, information on area of sampling and date of sampling etc.):

- Inventories of available data and dataset for each hazard targeted by the DiSCoVeR are described in deliverables D-JRPFZ-1-WP2.6 and are available in Zenodo ([link](#))
- Genomic sequences of isolates collected for SA within the DiSCoVeR are/will be made public available in a DiSCoVeR dedicated project in the European Nucleotide Archive (ENA)

---

<sup>1</sup> Several OH-EJP projects have been investigating barriers for data sharing on surveillance of zoonoses: Cohesive, NOVA, Orion. In addition, in 2020 a Workshop was funded by the Med-Vet-Net Association with EFSA, ECDC and other stakeholders focused on this topic in WGS data use foodborne zoonoses. <https://www.foodsafetynews.com/2020/10/experts-assess-use-of-genome-sequencing-in-multi-country-outbreaks/>





- Metadata associated with genomic sequences are/will be made public available in Zenodo. Reference key are/will be the accession number.
- The datasets will serve the DiSCoVeR aims and will not be updated in the future, which raises the question of the obsolescence of the datasets. This represents a major limitation and raises the question for the future sustainability of the DiSCoVeR experience for the translation of the SA into the One-Health 'surveillance & monitoring' in the EU.

### Evaluation of the DiSCoVeR datasets

SA is a data driven exercise and covering with accurate, detailed and representative data the different sources, is very important. Precision of the estimates depends on both the quality of data and the amount of information collected.

Compared to previous studies the DiSCoVeR aimed at increasing the data coverage in different sources also targeting non livestock reservoirs and non-food transmission routes. Data have been obtained, by collecting and collating data from DiSCoVeR partners where possible, also on non-livestock reservoirs (i.e. wildlife) and environment to fill out knowledge gap that the DiSCoVeR had identified.

Figure 3 describes the final size of the DiSCoVeR data inventory for *Salmonella*, *Campylobacter* and STEC in years 2015-2019 (for STEC 2010-2020) and coverage in non-human sources, by sector and type of source. The distribution was rather unbalanced with the vast amount of data available for few animal species, namely the species whose epidemiological role of reservoir is well known for the given pathogen:

- *S. Enteritidis* in Layers and broilers
- *S. Typhimurium* (including monophasic variant) in pigs
- *S. Infantis* in broilers
- *S. Derby* in pigs
- *Campylobacter* in broilers
- STEC in cattle

Altogether, the hazard/source combinations listed above accounted for more than half of total data (54.7%) collected in years 2015-2019. Data from livestock species (pigs, cattle, sheep and goats, *Gallus gallus*, turkeys, ducks and goose) accounted for the 84.2 % of all data. Pets and cats were the sources most frequently reported among non-food producing animals (6.2%) followed by wild birds (4.2%), horses (1.7%), reptiles (1.3%) and other wild animals (0.5%). However, data from non-livestock sources were contributed with by only a few countries. Environmental samples were scarcely represented for all hazards except for *Campylobacter* (17.2%) with only two countries contributing to this source and hazard. Distribution of data on positive samples for *Salmonella*, *Campylobacter* and STEC collected by EFSA in the same period (years 2015-2019<sup>2</sup>) is shown in figure 4 for comparison with the DiSCoVeR datasets and get information on the representativeness of the DiSCoVeR data collection against monitoring data of the same pathogens in animals, in the EU.

---

<sup>2</sup> Data are available in the EFSA on-line resources and in the EFSA Knowledge Junction in Zenodo: 2015 (<https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2016.4634>); 2016 (<https://doi.org/10.5281/zenodo.1044742>); 2017 (<https://doi.org/10.5281/zenodo.1475841>); 2018 (<https://doi.org/10.5281/zenodo.3527706>); 2019 (<https://doi.org/10.5281/zenodo.4298993>).



**Figure 3 - Distribution of positive samples by zoonotic hazard and source in year 2015-2019 (data source DISCoVeR inventory). For STEC the sampling period is 2010-2020.**

5 years sum of positive samples	S. Enteritidis		S. Typhimurium+ Monoph		S. Infantis		S. Newport		S. Derby		Campylobacter		STEC	
	N positive	N country	N positive	N country	N positive	N country	N positive	N country	N positive	N country	N positive	N country	N positive	N country
Pigs	2	2	494	9	12	3	20	2	92	3	146	4	119	5
Cattle *	6	2	57	6	3	3	15	1	6	1	221	3	773	10
Sheep/goats	1	1	19	2	1	1	1	1	0	0	107	1	209	7
Chicken (broilers)	248	6	77	5	86	5	65	2	15	2	755	6		
Layers	369	6	70	5	9	4	15	2	7	2	55	1		
Gallus gallus (other)														
Turkeys *	10	3	62	4			36	2	30	2	35	2		
Ducks/goose *	34	2	35	3	13	1	7	1	1	1				
Pets (dogs/cats)	15	2	94	4	17	2	21	2	7	1	161	3	5	5
Pets (reptiles)	20	2	7	2	17	2	21	2						
Wild animals *	12	4	7	3	3	1			1	1	2	1		
Wild birds	18	3	9	2	1	1			0	0	190	3		
Horses *	13	1	49	2	2	1	21	1	0	0				
Other animals *													98	7
Zoo animals														
Shellfish														
Environment (Water)	9	3	4	2	0	0	1	1	6	1	287	2	31	4
<b>Total</b>	<b>748</b>	<b>9</b>	<b>980</b>	<b>9</b>	<b>164</b>	<b>9</b>	<b>222</b>	<b>9</b>	<b>159</b>	<b>9</b>	<b>1672</b>	<b>6</b>	<b>1204</b>	<b>10</b>

For STEC the sampling period refers to years 2010-2020.





As expected, data on environmental sources were absent from the EFSA dataset but present in the DiSCoVeR. Main differences between the DiSCoVeR dataset and EFSA data summary statistics = were a much higher proportion of samples from non-food producing animals (Dogs/Cats, Reptiles, Wild animals, Wild birds, Other animals, Zoo animals) in the DiSCoVeR dataset for all pathogens with few exception. Proportion of *S. Typhimurium* including monophasic and *S. Derby* in pigs were higher in the DiSCoVeR and the same was for STEC in cattle. *Campylobacter* and *S. Infantis* in broiler chicken were reported in lower proportion in the DiSCoVeR than in the EFSA dataset, and the same was observed for *S. Newport* in turkey.

Comparison on data on *Salmonella*, *Campylobacter*, and STEC in human cases reported to DiSCoVeR and ECDC as well as coverage of non-human sources by agent and by countries are reported in the Appendix.

### **Optimisation of dataset for SA**

In summary, both DiSCoVeR and EFSA datasets are characterised by unbalanced distribution of positive samples among the different sources. This is probably due to a most intensive sampling activities carried out in specific sources according with the targeted zoonotic hazards. It also reflects a higher probability to obtain isolates for further analysis in those sources (e.g. *S. Enteritidis* in layers). Environmental samples (water) are only available in DiSCoVeR dataset in particular for *Campylobacter*.

Datasets characterised by unbalanced distribution with few oversampled sources (e.g. *Campylobacter* in poultry) and others missing information or with few data available may be a problem, because they can lead to biased attribution estimates.

For this reason efforts should be made to balance the availability of data with active and supplementary sampling in specific sources, to improve the quality of SA estimated based on typing/subtyping method.

A sustainable way to optimise the availability of data could be to promote sampling through synergies with other already existing sampling programs (e.g. pigs or turkeys monitoring survey for AMR could be used to get samples to be tested for *Campylobacter*). Synergies with regular sampling from baseline surveys would allow to get comparable data from all countries.

Optimisation of SA means finding a compromise between the need to enlarge the dataset and the coverage of the different sources and the quality of data being introduced in the model. Increase of data to enlarge source coverage may result in increased uncertainty, if data are not as accurate and representative as needed (e.g. data collected outside sampling survey or surveillance and monitoring). For this reason the quality of data especially for reservoir and sources less frequently populated in terms of data should be carefully evaluated in advance. Optimal data requirements for the DiSCoVeR has been specified in deliverable D-JRPFZ-1-WP2.6.

## **Skill**

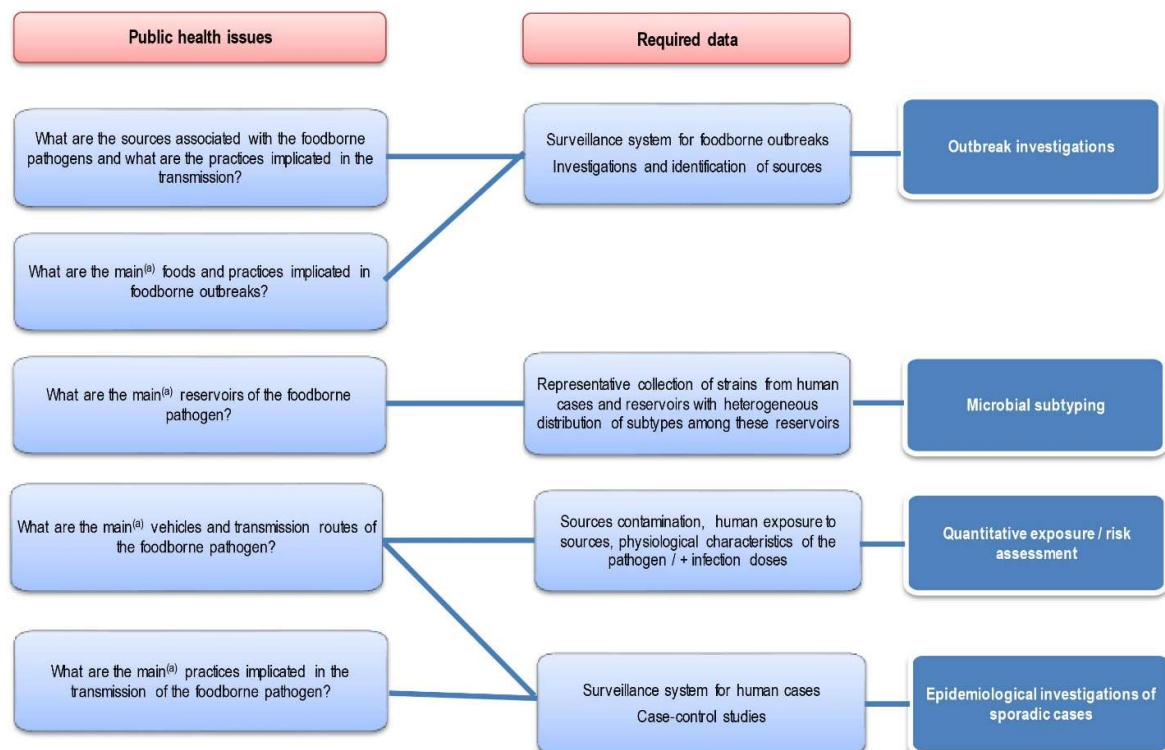
### **The DiSCoVeR scientific community**

Under the umbrella of the DiSCoVeR project, expertise and skills from multiple sectors and disciplines have been brought together into a small scientific One-Health community. Collaboration, integration, networking as well as exchange of knowledge and skills among partners' experts (19 partners from 13 countries) allowed to comprehensively approach 'SA science' including its main components (methods, data, knowledge, operational framework). This was an added value of the DiSCoVeR and an important achievement for future translation of the capacity building into the EU context.



## Conclusions

- Although past experience were reported of SA sequence-based studies, the DiSCoVeR project allowed for the first time to conduct large-scale SA exercise at the EU-level thanks to the availability of large high-quality (fit for purpose) data also including non-traditional tested sources (pets, wildlife, environment). This is an important achievement of the DiSCoVeR and OHEJP consortium in terms of scientific evidence outcomes and uncertainty that have been made available to end-users, i.e. primarily public health authorities, and other stakeholders including farmers, industry etc. to be taken into account for future development of policy of intervention to protect public, animal and environmental health and support sustainability of primary production.
- The release of the DiSCoVeR comprehensive large One-Health datasets in the public domain has been an important initiative for ease the translation of the SA into the field and represents an important achievement of the project, which, however, raises in the near future the question of the obsolescence and future sustainability of these important resources, as resources for updating the datasets are not available.
- At the same time, the DiSCoVeR highlighted the importance of building public and accessible repositories of WGS data and metadata not only to support investigation of outbreaks. It also provided a proof of concept for the long-term use of WGS-based data from and the structural integration of WGS data into monitoring and surveillance activities pursuing One-Health objectives. This would enable performing SA studies on a regular basis in order to support the monitoring of the effect of intervention programs and the development of new policies to protect public and animal health against priority foodborne zoonoses and AMR.
- Optimisation of SA approach is a complex task and selection of best appropriate approach can be guided by the aims, the quality of available data, the skills and the public health questions. Figure 5 provides a useful guidance for this purpose.





**Figure 5 Preferential choice of source attribution methods based on public health issues. (a) Ranking and/or quantifying the relative importance (from Mughini-Gras et. al 2019)**

- SA methods based on genomic data represent nowadays the gold standard for SA, since they have the highest discriminatory power. However, the choice of the best methodological approach for SA should take into account that genomic-based data requires high quality data and skills to be reliably implemented.
- On the other hand, frequency-matching SA models making use of traditional pheno/genotypic characterisation data for *Salmonella* and STEC continue to represent an interesting option at country level and EU level, since these data can be regularly collected within surveillance and monitoring of foodborne zoonoses in human and non-human sources.
- Training and dissemination to improve skills and capacity-building in SA is definitely a crucial resource for SA translation into the field.

### References:

- Mughini-Gras L, Kooh P, Fravallo P, et al. Critical Orientation in the Jungle of Currently Available Methods and Types of Data for Source Attribution of Foodborne Diseases. *Front Microbiol.* 2019 Nov 12;10:2578. doi: 10.3389/fmicb.2019.02578. PMID: 31798549; PMCID: PMC6861836.ECDC
- (European Centre for Disease Prevention and Control), Surveillance Atlas of Infectious Diseases FWD (2021 data) available at <https://atlas.ecdc.europa.eu/public/index.aspx> (last access 21/12/2022)
- EFSA (a) (European Food Safety Authority), Amore G, Boelaert F, Papanikolaou A, Rizzi V and Stoicescu A-V, 2022. Manual for reporting on zoonoses and zoonotic agents, within the framework of Directive 2003/99/EC, and on some other pathogenic microbiological agents for information derived from the year 2021. EFSA supporting publication 2022:EN-7130. 78 pp. doi:10.2903/sp.efsa.2022.EN-7130
- EFSA (b) (European Food Safety Authority), Costa G, Di Piazza G, Koevoets P, Iacono G, Liebana E, Pasinato L, Rizzi V and Rossi M, 2022. Guidelines for reporting Whole Genome Sequencing-based typing data through the EFSA One Health WGS System. EFSA supporting publication 2022:EN-7413. 29 pp. doi:10.2903/sp.efsa.2022.EN-7413



## Appendix

Comparison on data on *Salmonella*, *Campylobacter*, and STEC for human cases reported to DiSCoVeR and ECDC (data Source ECDC Atlas) and data coverage for the same agents in different animal species and other non-human sources within the DiSCoVeR and the EFSA datasets, in years 2015-2019

S. Enteritidis																	
Country	ECDC	D*	EFSA														
	Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water (various types)	Environment
Austria	4020																
Belgium	2534																
Bulgaria																	
Croatia	1621																
Cyprus	27																
Czechia	52413																
Denmark	1308																
Estonia	558																
Finland	690																
France	9999																
Germany	21586																
Greece	646																
Hungary	13179																
Iceland	77																
Ireland	428																
Italy	2139																
Latvia	1235																
Lithuania	3427																
Luxembourg	126																
Malta	130																
Netherlands	1505	86															
Norway	1601																
Poland	21055																
Portugal	509	15															
Romania	688																
Slovakia	23692																
Slovenia	601																
Spain	4895																
Sweden	1176																
United Kingdom	14454																

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.



Country	S.Typhimurium													S. Typhimurium, monophasic																						
	ECDC	EFSA												EFSA																						
Country	Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water	Environment	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water	Environment			
Austria	1157																																			
Belgium	6518																																			
Bulgaria																																				
Croatia	333																																			
Cyprus	65																																			
Czechia	3714																																			
Denmark	1399	73																68																		
Estonia	299																																			
Finland	319																																			
France	16646																																			
Germany	17873																																			
Greece	276																																			
Hungary	2749																																			
Iceland	75																																			
Ireland	471																																			
Italy	7379																																			
Latvia	452																																			
Lithuania	563																																			
Luxembourg	237																																			
Malta	80																																			
Netherlands	1757	1																5																		
Norway	618																																			
Poland	836																																			
Portugal	572	24																25																		
Romania	233																																			
Slovakia	1164																																			
Slovenia	321																																			
Spain	6585																																			
Sweden	1364																																			
United Kingdom*	9738																																			

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.





S. Infantis																	
	ECDC	EFSA															
		Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water
Country		Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water
Austria	375																
Belgium	288																
Bulgaria																	
Croatia	59																
Cyprus	9																
Czechia	525																
Denmark	96																
Estonia	150																
Finland	54																
France	1240																
Germany	1421																
Greece	15																
Hungary	1224																
Iceland	1																
Ireland	37																
Italy	452																
Latvia	62																
Lithuania	115																
Luxembourg	17																
Malta	56																
Netherlands	161	25															
Norway	57																
Poland	345																
Portugal	5	14															
Romania	20																
Slovakia	524																
Slovenia	55																
Spain	334																
Sweden	50																
United Kingdom	1176																

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.



S. Newport																	
	ECDC	EFSA															
Country	Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water	Environment
Austria	69																
Belgium	119																
Bulgaria																	
Croatia	6																
Cyprus	8																
<b>Czechia</b>	139																
<b>Denmark</b>	139																
Estonia	1																
Finland	124																
<b>France</b>	962																
Germany	371																
Greece	12																
Hungary	67																
Iceland	7																
<b>Ireland</b>	55																
Italy	84																
Latvia	1																
Lithuania	6																
Luxembourg	13																
Malta	1																
<b>Netherlands</b>	81	10															
Norway	138																
<b>Poland</b>	34																
<b>Portugal</b>	27	27															
Romania	2																
Slovakia	46																
Slovenia	8																
<b>Spain</b>	218																
<b>Sweden</b>	128																
<b>United Kingdom</b>	1523																

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.



S. Derby																	
	ECDC	EFSA															
	Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	Layers	Turkeys	Ducks/goose	Pets (dogs/cats)	Pets (reptiles)	Wild animals	Horses	Shellfish	Vegetable	Water	Environment
Country																	
Austria	28																
Belgium	217																
Bulgaria																	
Croatia	28																
Cyprus	2																
Czechia	87																
Denmark	70																
Estonia	14																
Finland	7																
France	727																
Germany	699																
Greece	3																
Hungary	263																
Iceland																	
Ireland	5																
Italy	536																
Latvia	24																
Lithuania	92																
Luxembourg	13																
Malta																	
Netherlands	62	10															
Norway	17																
Poland	54																
Portugal	1	9															
Romania	14																
Slovakia	71																
Slovenia	9																
Spain	142																
Sweden	15																
United Kingdom	162																

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.



Camylobacter jejuni																
	ECDC	D	EFSA													
	Humans	Humans	Pigs	Cattle	Sheep/goats	Chicken (broilers)	other Gallus gallus	Turkeys	Ducks/goose	Pets (dogs/cats)	Wild birds	Horses	Horses	Shellfish	Vegetable	Water
Country																
Austria	40522															
Belgium	48886															
Bulgaria	1171															
Cyprus	135															
Czechia	132676															
Germany	388148															
Denmark	26997	1230														
Estonia	1924															
Greece	1741															
Spain	82588															
Finland	25069															
France	42474															
Croatia	9344															
Hungary	42683															
Ireland	15982															
Iceland	742															
Italy	7538															
Lithuania	6225															
Luxembourg	3010															
Latvia	547															
Malta	1508															
Netherlands	19106	262														
Norway	18762															
Poland	4148	7														
Portugal	3513	299														
Romania	2973															
Sweden	49069															
Slovenia	7579															
Slovakia	42467															
United Kingdom	305929															

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2015-2019. For STEC, the reference period was 2010-2020. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.



Country	STEC							
	ECDC	D	EFSA					
	Humans	Humans	Pigs	Cattle	Sheep/goats	Pets (dogs/cats)	Other/zoo animals	Environment
Austria	2010							
Belgium	1160							
Bulgaria	2							
Cyprus	0							
Czechia	245							
Germany	22454							
Denmark	3245							
Estonia	65							
Greece	20							
Spain	858							
Finland	1273							
France	2651							
Croatia	60							
Hungary	136							
Ireland	6648							
Iceland	52							
Italy	675	364						
Lithuania	17							
Luxembourg	71							
Latvia	59							
Malta	165							
Netherlands	7660	2937						
Norway	2605	557						
Poland	53							
Portugal	9	6						
Romania	123							
Sweden	6138							
Slovenia	295							
Slovakia	55							
United Kingdom	13551	468						

Colored boxes in background represent data present in the ECDC/EFSA dataset for 2010-2020. Red boxes represent combination of source and country for which data are available in the DiSCoVeR. DiSCoVeR data from the United Kingdom only England and Wales.