



**Data inventory with
assessment of availability,
quality and fitness for data-
driven surveillance JRP6 -
NOVA - FBZ1 - 1st Call**

NOVA D3.2

Responsible Partner: ANSES, SVA,



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-3.2. Data inventory with assessment of availability, quality and fitness for data-driven surveillance
Join Integrative/Research Project	JRP6 - NOVA - FBZ1 - 1 st Call
JIP/JRP Leader	Jenny FRÖSSLING (SVA)
Other contributors	Adeline Huneau, Fernanda Dórea, Petter Hopp, Gry Grøneng, Malin Jonsson, Viviane Hénaux, Carole Sala, Renaud Lailier, Madelaine Norström, Kyrre Kausrud
Due month of the deliverable	M10
Actual submission month	M
Type <i>R: Document, report</i> <i>DEC: Websites, patent filings, videos, etc.</i> <i>OTHER</i>	R: Report
Dissemination level <i>PU: Public</i> <i>CO: confidential, only for members of the consortium (including the Commission Services)</i>	PU



Data inventory with assessment of availability, quality and fitness for data-driven surveillance

Introduction

Syndromic surveillance (SyS) consists in the real-time collection, analysis, interpretation and diffusion of non-specific indicators in order to detect any changes in human or animal health as early as possible. SyS systems are expected to have a better timeliness and be more cost effective than traditional surveillance systems, since they make use of already existing data, collected for other purposes than surveillance. From a “One Health” perspective, no SyS system is currently operational in EU, including data from animal sector, environmental sector, food industry and human sector. According to a recent review from Wendt et al. (2016), existing systems collecting data from those different sectors, as media-source surveillance based on reports and notifications of diseases (MediSys, ProMed etc.), do not integrate them into a global analysis. NOVA’s WP3 was therefore designed to assess whether SyS in animal and food sectors could be a valuable add-on to the current or future surveillance of (Foodborne diseases) FBD in humans. The expected impacts are a better understanding of the interplay between animal and human health, an improvement of the prediction of human FBD outbreaks, and a help in the retrospective identification of the disease source.

The objective of the task 3.1 in the NOVA project was to identify potential data sources in Member states contributing to WP3 in order to develop specific surveillance components for FBD based on “secondary data” in tasks 3.2 and 3.3. Secondary data are already available information, sometimes collected for other purposes than health surveillance, that may be of interest for health surveillance.

Methods

The methodology for the inventory of data sources was defined in the kick-off meeting in March 2018. The terminology we use, evolved from SyS (relative to the surveillance of syndromes only) to “data driven” surveillance, i.e. surveillance based on the use of secondary data. These unspecific indicators are directly or indirectly linked to the health status of humans and animals (diagnosis codes, laboratory data, milk yield, mortality, drug sales...) in veterinary and public health sectors.

A table was then designed to collect data potentially useful and available for data-driven surveillance of FBD (Table 1). The table was filled in based on expertise of NOVA WP 3 members : ANSES, SVA, NVI, NIPH, RIVM.



Table 1: Meta-data collected on data sources for FBD and AMR surveillance for MS (Norway, France, Sweden, UK)

NB: modalities may be given as examples and do not reflect exactly all the modalities used in appendix

Variable	Modalities	Comment
Country	EU SW, NO, FR	
Data generator	Farmer Vet General practitioners Food Business Operator...	Who does the observation? Who reports the event?
System manager/data owner		In case of non-organized system, who is the owner of the data?
NOVA partner access	Yes Upon request to data owner No direct access	
Surveillance system implemented	Yes No – data are not analysed No – data are not centralized No – data are not registered on a database	To which extent are the data used for surveillance or ready for that use?
Species/sector	Feed Animal (all species) Cattle Pig Poultry Food (1 st processing to retail) Human	
Analytical unit / aggregation level	Sample, order etc. Individual animal/human/product Geographical aggregation (region, country...) Aggregation at operator level (farm, company etc.)	
Data type	Laboratory test Clinical report Production data Other	Production data include mortalities, condemnations, movement, etc.
Health indicator		Describe / give detail on the health related indicator
Data specificity	High Medium Low	low: health related event but very low specificity medium: first signal but not confirmed by a formal diagnostic high: lab confirmation, complete diagnostic
Frequency of data collection		At what frequency are the data collected for surveillance purpose?
Delay in data centralization	Time between data generation and transfer for centralization	Raw estimate (min-max or frequency of data transfer for centralization)



Population coverage	Exhaustive Volunteers	Raw estimation of the coverage of the targeted population: is it a mandatory surveillance? Is it a surveillance system based on a sub-sample of the population?
Relevance in NOVA	No relevant to highly relevant	
FBZ note		Is it of interest for FBD surveillance?
Gastrointestinal syndrome note		Is it of interest for gastrointestinal syndrome surveillance?
AMR note		Is it of interest for AMR surveillance?

Results and discussion

Over 49 data sources identified, 28 were dedicated to animal health surveillance only, 13 to public health, one to feed and one to environment (Table 2; Appendix). The six remaining data sources covered at least two sectors, generally human and veterinary health. A wide diversity of situations was seen from operational surveillance systems with daily alert systems, mostly in public health sector, to isolated data basis with no centralisation of data at national level.

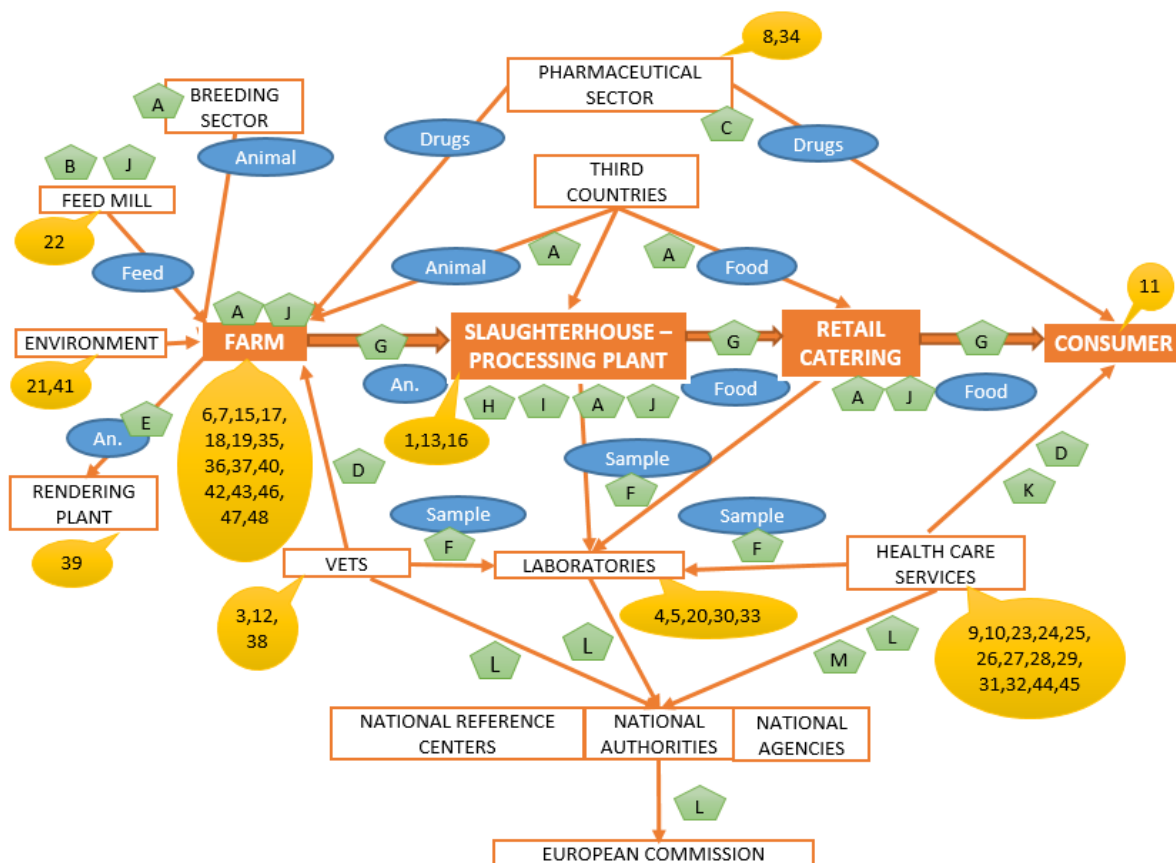
The identified data sources were placed on the food chain map (cf. Deliverable 3.1) to visualize the coverage of the food chain with the available sources. No data source specifically dedicated to food quality was identified; this observation may be due to a lack of expertise of the WP3 contributors or to a lack of information about private initiatives in food industry. Only the National and European control and surveillance plans provide data on food sanitary quality but not on a regular basis (annual reports).

Table 2: Characteristics of identified data sources (N=48)

Sector	Animal :28 Public health: 13 Feed: 1 Environment: 1 Animal and public health: 1 Feed and food: 1 Animal, food and public health: 1 Food, Public health: 2 Animal, food, feed, environment :1
Types of data	Clinical: 11 Laboratory diagnostic - analysis: 8 Clinical + Laboratory diagnostic - analysis: 7 Production: 12 Sale-exchange: 6 Other: 5 (web queries, meteorological data, etc.)
Data specificity	High: 13 (confirmed cases) Medium: 17 (clinical note, suspicion report) Low: 19 (productivity data, environmental data)
Frequency of collection	Daily: 30 (automatic registration and transmission) Weekly: 3 Monthly: 11 Yearly: 3



	Not known: 2
Delay in transmission	No delay (real-time): 15 About one week: 6 About one month: 6 More than one month: 5 Not evaluated or not known: 17
On-going surveillance system	Yes: 14 Pilot phase: 2 No: 33
Interest for FBD surveillance	High: 10 Medium: 10 Little: 13 Not relevant: 5 To be assessed: 11
Interest for AMR surveillance	High: 8 Medium: 5 Little: 2 Not relevant: 34
Availability for NOVA	Yes: 14 Upon request to data owner: 24 No: 11



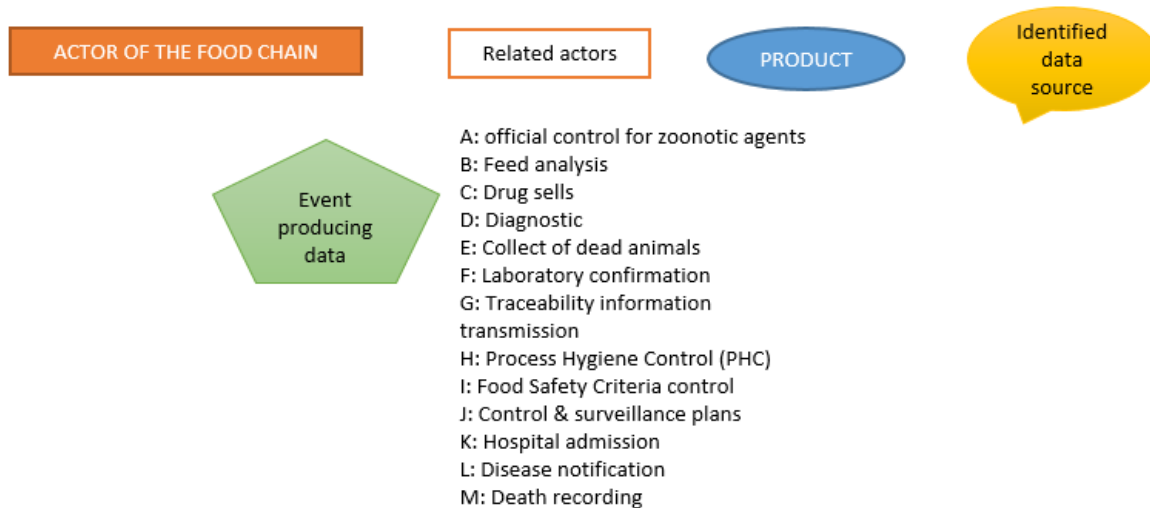


Figure: Scheme of the food chain and of the identified data sources (legends for data sources are in appendix)

The identified data sources are planned to be used in task 3.2. for the development of univariate surveillance modules for the surveillance of FBD. The main issue is to assess whether the identified data sources are fit for that purpose. Three main qualities are expected for using data with a surveillance purpose: completeness, timeliness and accuracy.

Completeness covers two different dimensions. Firstly, it refers to the number of missing-values in the dataset used. As the selected data sources have been already checked and completed by the actors in charge of their collection, this aspect is not an issue for their use in NOVA. Secondly, completeness also refers to the number of events that are registered over the total number occurring in the targeted population. This quality is far more difficult to assess. This may be approached by evaluating the coverage of the targeted population but this indicator is seldom available and rarely measured for the identified data sources.

A second quality expected from the data source is timeliness i.e. the length between the event occurrence and its availability in the dataset. This is of high importance in the objective of developing early detection systems. The timeliness of the data sources under study is highly variable from less than a day up to one year. There are 21 data sources that have a delay of one week or less and five data sources that have a delay of one month. These data sources have a sufficient timeliness for syndromic surveillance, while the data sources with more than one-month delay may be inadequate for syndromic surveillance. The delay in the reporting is often higher for systems that are not based on an automatic collection and transmission of the data.

The third quality to assess is the accuracy of the collected data: to what extent are they conform to the real situation? There are 20 of the data sources that were considered to have high or middle relevance for FBD surveillance while eleven additional data sources were under evaluation. Considering AMR surveillance, there were 13 data sources that were categorised as having high or middle relevance.



The further step is to select different sets (at least two) of data sources that are relevant for addressing FBD and AMR. Each data source of the sets will be assessed for its fitness for task 3.2.

D.G. Arts, N-F de Keizer and G-J Sheffer, 2002. *Journal of American Medicine Information Association*, 9: 600-611.

A. Wendt, L. Kreienbrock and A. Campe, 2016. *Zoonosis and Public Health*, 63: 503-514.



Appendix: Main characteristics of the identified data sources. The table presents the main meta-data reported from the partners in the project and was extracted from the information reported as presented in Table 1.

Index	Database name	Country	NOVA access	SyS?	Target	type recod	Data specificity	Frequency of collection	FBZ relevance	AMR notes
1	EDI-SPAN	FR	YES	on test	Farmed animal	clin	low	Daily	to be tested	No
2	RESAPATH	FR	YES	NO	Farmed animal	lab-clin	high	monthly to yearly	to be tested	High
3	PJS	NO	YES	NO	Farmed animal	lab-clin	medium	Daily	little	High
4	SVALA	SW	YES	YES	Farmed animal	lab	high	Daily	high	High
5	Salmonella	FR	YES	NO	Animal, food, feed, envt	lab	high	Daily	high	No
6	BDNI	FR	YES	on test	Cattle	sale-exchange	low	weekly	little	No
7	CDB	SW	YES	YES	Cattle	other	low	Daily	little	No
8	ANMV tax found	FR	YES	NO	Cattle, swine, sheep	sale-exchange	low	yearly	not relevant	Little
9	Sykdomspulsen (NorSySS)	NO	YES	YES	Human	clin	medium	Daily	to be tested	No
10	VESUV	NO	YES	NO	Human	clin	medium	Daily	to be tested	No
11	Google analytics	NO	YES	NO	Human	other	medium	Daily	to be tested	Little
12	djursjukdata	SW	YES	YES	Farmed animal	clin	medium	Daily	medium	High
13	Delivery of carcasses registry	NO	YES	NO	Farmed animal	prod	low	monthly	little	No
14	VETREG	NO	YES	NO	Animal	sale-exchange	low	Daily	medium	Medium
15	Husdyrregisteret, storfe	NO	upon request	NO	Cattle	prod	low	Daily	little	No
16	Kassasjonsregister	NO	upon request	NO	Cattle, swine, sheep	prod	medium	Daily	little	No
17	SNIG	FR	upon request	NO	Cattle, swine, sheep	prod	low	Daily	little	No
18	RESYTAL	FR	upon request	NO	Cattle, swine, sheep	clin	high	monthly	medium	No
19	Kukontrollen	NO	upon request	NO	Cattle, swine, sheep	prod	high	Daily	little	Medium
20	INFOLABO	FR	upon request	NO	Cattle, goat	prod	medium	not known	to be tested	No
21	Meteorologicaol data	NO	upon request	NO	Environment	other	low	daily-hourly	little	No
22	OQUALIM	FR	upon request	NO	Feed	lab	high	monthly	to be tested	No
23	VESSy/NESSS	Eng	upon request	YES	Human	lab-clin	high	Daily	high	No
24	SGSS	Eng& Wales	upon request	YES	Human	lab	high	Daily	high	High
25	SOS medecin	FR	upon request	YES	Human	clin	medium	Daily	high	No
26	Sentinelles	FR	upon request	YES	Human	clin	medium	weekly	high	No
27	OSCOUR	FR	upon request	YES	Human	clin	medium	weekly	medium	No
28	Mortality	FR	upon request	YES	Human	clin	medium	Daily	medium	No
29	MDO (wintiac)	FR	upon request	NO	Human	lab-clin	high	Daily	high	No
30	MSIS	NO	upon request	NO	Human	lab	high	Daily	to be tested	Medium
31	Listeria	Eng	upon request	YES	Human (some food)	lab-clin	high	Daily	high	no



32	eFOSS	Eng	upon request	YES	Human (some food)	lab-clin	medium	not known	high	No
33	GDW	Eng& Wales	upon request	YES	Human (food, animal)	lab	high	Daily	high	High
34	NorPD	NO	upon request	NO	Human and animals	sale-exchange	low	monthly	no	High
35	REPAMO	FR	upon request	YES	Oyster; mussel	lab-clin	medium	monthly	little	No
36	BDPORC	FR	upon request	NO	Pigs	sale-exchange	low	Daily	no	No
37	BD Avicole	FR	upon request	NO	Poultry	sale-exchange	low	Daily	no	No
38	RNOEA	FR	upon request	NO	Poultry	clin	medium	monthly	to be tested	No
39	SI2A	FR	upon request	NO	Animal	prod	medium	Daily	medium	No
40	RENACEB/RENALAP	FR	upon request	NO	Rabbit	prod	low	yearly	medium	No
41	rodent population	NO	upon request	NO	Rodents	other	low	yearly	no	No
42	OVINFO	FR	upon request	NO	sheep, goat	sale-exchange	low	Daily	medium	no
43	Kokontollen	SW	No	NO	Cattle, swine, sheep	prod	low	monthly	little	Medium
44	Hospital labs	NO	No	NO	Human	Lab	high	Daily	to be tested	High
45	KPR	NO	No	NO	Human	clin-other	medium	Daily	to be tested	no
46	GTE	FR	No	NO	Pigs	prod	low	not known	medium	no
47	WinPig	SW	No	NO	Pigs	prod	low	monthly	little	Medium
48	Poultry prod. registers	NO	No	NO	poultry	prod	low	monthly	little	no
49	RASFF	Europe	upon request	NO	Feed, Food	Lab	medium	Daily	medium	Medium