EXTERNAL SCIENTIFIC REPORT

Specification of data collection on animal diseases to increase the preparedness of the AHAW panel to answer future mandates – CFP/EFSA/AHAW/2010/01[1][2]

C Bellet, M-F Humblet, M Swanenburg, J Dhé, S Vandeputte, A Thébault, F Gauchard, P Hendrikx, C De Vos, A De Koeijer, C Saegerman, M Sanaa

French Agency for Food, Environmental and Occupational Health and Safety, France
Unit of research in epidemiology and risk analysis, University of liege, Belgium
Central Veterinarian Institute (CVI), Wageningen, The Netherlands

ABSTRACT

Risk assessment is EFSA's core activity. The lack of quantitative data and standardized methodology to collect data for risk assessment may limit the use of formal and structured quantitative risk assessments and delay the production of accurate scientific opinions. The aim of this project was to develop a methodological framework to identify data needs in accordance with the risk questions and the relevant risk assessment methods. A review of AHAW panel opinions in the area of animal health, adopted between 2004 and 2010 (WP1) was conducted in order to identify and to categorize the most recurrent risk questions, the suitable risk assessment methods and data needs. Subsequently an inventory of the possible sources of required data and an assessment of their availability and accessibility were performed (WP2). Facts and metadata were distinguished from WP1 initial list of data needs and a metadata model for facts collection and documentation was then proposed (WP3). The outcomes and conclusions from the three first work packages was used to suggest a structured approach to collect data (facts and metadata) in regard to expected future risk questions. Based on four case studies related to *Echinococcus multilocularis*, *E. granulosus*, Porcine Reproductive and Respiratory Syndrome and Venezuelan Equine Encephalitis (WP4), we concluded that this approach could lead to more efficient preparatory data collection, and therefore enable more rapid response to new risk managers questions.

KEY WORDS

Risk assessment, risk questions, animal health, data, fact, metadata

[1] EFSA-Q-2010-00903

[2]Acknowledgement: This study was carried out within the framework of the EFSA art.36 project (CFP/EFSA/AHAW/2010/01). We warmly thank the ANSES, CVI and ULg consortium team. We are grateful to all the experts from Belgium, France and Netherlands, who helped evaluate the needed data related to WP4 case studies

Any enquiries related to this output should be addressed to AHAW@efsa.europa.eu

Suggested citation: C Bellet, M-F Humblet, M Swanenburg, J Dhé, S Vandeputte, A Thébault, F Gauchard, P Hendrikx, C De Vos, A De Koeijer, C Saegerman, M Sanaa. Specification of data collection on animal diseases to increase the preparedness of the AHAW panel to answer future mandates – CFP/EFSA/AHAW/2010/01. Supporting Publications 2012:EN-354. [215 pp.]. Available online: www.efsa.europa.eu/publications

SUMMARY

Risk assessment is one of the main tools used by EFSA AHAW Panel to address questions from risk managers. The lack of quantitative data and standardized methodology to collect data for risk assessment may delay the production of accurate scientific opinions and limit the use of formal and structured quantitative risk assessments. Therefore, the aim of this project was to develop a methodological framework stating how to specify the data needs in accordance with the type of risk questions and the relevant risk assessment methods. The project was organised in four interdependent work packages.

The first work package, "Typology of risk questions and identification of data needs", defined a comprehensive list of relevant questions in regard to Animal Health (AH) issues and reviewed the historical scientific opinions (2004-2010), in order to identify the most recurrent types of questions and the main difficulties encountered by experts to respond to it. We named them **risk questions,** acknowledging that several risk questions could be included in a given AHAW opinion's terms of reference (ToR). The following types of questions were identified: 1) Host/Pathogen characteristics, 2) Disease status, 3) Potential risk of spreading to susceptible population/Pathways of transmission & speed of the spread, 4) Risk of (re)introduction, 5) Risk of establishment, 6) Effectiveness of surveillance measures, 7) Diagnostic tools availability and efficiency, 8) Effectiveness of prevention tools, 9) Effectiveness of control measures, 10) Effectiveness of bio-security measures, 11) Treatment availability and efficiency, and 12) Vaccine availability and efficiency. The three risk questions, the more often asked, were by decreasing order: the risk of (re)introduction (21 opinions), the risk of potential spread to susceptible population, the pathways of transmission and speed of the spread (20 opinions), and the effectiveness of control measures (17 opinions). Scientific opinions were in general dealing with more than two types of questions. One of the main difficulties pointed out by AHAW panel to address quantitatively the risk questions was the data availability. We noticed an absence of a formal working procedure that helps achieving adequacy between used methodologies, risk questions and available data. Data gaps were in general recognized ; however, needs of collecting new data were not prioritised in regard to their added value on facilitating the answer to risk managers' questions. Data needs were classified in five categories with references to the previous risk questions typology: disease general information, descriptive epidemiological data, analytical epidemiological data, prevention and control data, and disease surveillance data. Within each category, three levels of subcategories were defined. This classification allowed the establishment of a comprehensive list of data needs.

The second work package, "Data availability", aimed at first assessing WP1 data availability, identifying the different sources and forms of the data, but also to evaluate data accessibility. This process was carried out focusing on four diseases (Venezuelan Equine Encephalitis (VEE), Porcine Respiratory and Reproductive Syndrom (PRRS), *Echinococcosis granulosus and E.multilocularis*) and four Member States (MS) (BE, FR, NL and SP). It was concluded that differences of data availability either depend on the status of the targeted disease, or on the category of the data itself (e.g. descriptive epidemiology vs. public health). As for examples, few data are available on VEE, an exotic disease, and on PRRS, a disease of limited interest for human health. Numerous types of data sources were identified. Nevertheless, for the vast majority of diseases, the disease-specific information was found in textbooks, papers, scientific literature, etc. Websites represented the majority of sources regarding MSs' data. The main forms of data sources were PDF and HTML files. Although raw tabulated data are more appropriate for risk assessment, these sources were not often available and sometimes difficult to access (e.g. restricted or paying access); this latter might represent a constraint to conduct a

risk assessment in a short timeframe. Few data were in a form directly usable, maybe because of the common difficulty to access to some relational databases which are often more restricted or not directly available through a classical search. The lack of availability was mainly observed for prevention and control, surveillance and public health. The accessibility of data was generally related to their availability. It is worse noting that sources gathered within WP2 are useful for future EFSA and/or national risk assessments, among others, to prepare future mandates. Because one data source could cover several data needs, a total of 471 different data sources were listed.

The third work package, "Data specification, validation and management", worked on **facts** and **metadata** modelling; both facts and metadata (see glossary) representing the **data** needed by experts to respond the different risk questions (WP1). The ultimate goal was to ensure the proper use and interpretation of available facts or collected facts, in relation to a specific AHAW Panel task. Based on WP1 terminology, tables were created to describe all the characteristics and attributes of the required facts and metadata. The metadata model made references to the approved Dublin Core Metadata Element Set (DCMES) (Dublin Core Metadata Initiative, 2010), as well as on existing EFSA and international data standards (WP3). In order to facilitate facts' specification, facts and groups of facts were defined, along with their metadata. Around 60 facts and 34 metadata were thus specified. The final model allowed linking a risk question to the required facts and metadata. The connection between all tables (data categories, type of questions, facts and metadata) enables different types of **queries**, starting from the type of questions, ending with the needed facts, or group of facts, their associated metadata and their possible sources (WP2). When data (either facts or metadata) were not available, proposals for surrogate data were done. The final 63 technical cards that list the metadata of each required fact could constitute a new way to systematize EFSA AH data collection queries to MSs (annex WP3).

The fourth work package, "Methodology framework", reviewed the risk assessment methodologies and tools that could be relevant for AHAW Panel tasks. Three main types of methods were distinguished: 1) **Statistical methods** that may be suitable for example for endemic diseases, where it is needed to assess the disease frequency (prevalence/incidence) and to assess possible growing or declining of the disease frequency. This method may also be useful to assess specific correlation or association between exposure and occurrence of diseases, or effects of intervention measures at the point of its application; 2) **Linear probabilistic risk assessment** that may be appropriate for example to assess the probability of an exotic agent entrance; and 3) **Mechanistic or dynamic models** that may be suitable for example to describe and assess the spread of a disease in a given population. The suggested general methodological framework was thoroughly tested in the three case studies selected in agreement with EFSA-AHAW unit. The three case studies questions were: 1) what is the added value of meat inspection with respect to disease prevention and control in the definite host (dog) and intermediate hosts (sheep, goat, cattle and swine) for *Echinococcus granulosus* 2) what are the surveillance measures needed on domestic and wild canids to demonstrate freedom for *Echinococcus multilocularis*; 3) what is the effectiveness of intervention measures in France at reducing the prevalence for Porcine Reproductive and Respiratory Syndrome (PRRS). Intervention measures to be considered included: vaccination, herd management, and biosecurity; and 4) what is the risk of EU introduction, taking into account risk reduction measures in place, for Venezuelan Equine Encephalitis. For each case study, data needs were discussed and linked to both WP1's types of questions and the selected method. The methodological framework started first with the question, secondly with the related list of data needs (WP1), for which both availability and accessibility were evaluated (WP2), and finally with the specification of the data, once the data sources were identified (WP3). Data specifications were used to extract data from available sources or to build a questionnaire template to collect the data, directly from MSs' authorities, scientific or professional organisations. With reference to the nature of the available data and the type of the risk question, the more suitable

methods (Statistical methods, linear probabilistic risk assessment or Mechanistic or dynamic models) were at the same time selected.

The analysis of the four case studies showed how the general methodological approach could lead to more efficient preparatory data-collection, and therefore enable more rapid response to new risk managers' questions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

BACKGROUND AS PROVIDED BY EFSA

The categorisation of animal diseases, including those transmissible to humans, and priorisation of interventions may require recurrent assessment due to the varying nature of animal diseases, their epidemiology, and their impact on animal or human populations. Such an analysis must be subjected to regular updating and review based on:

- improvement of the relevant scientific knowledge;
- evolution of a situation (occurrence or introduction of diseases, changes in the epidemiological situation, available disease control options);
- and evolution of the economic context (imports and exports of animals and their products, development of particular animal production systems, trends...)

Criteria for listing of animal diseases have been developed by the World Animal Health Organization[1] but unexpected and potentially rapid epidemiological changes, such as the recent issue of Q-fever are difficult to predict. In the recent past, the AHAW panel has been asked to deliver opinions on a range of questions on animal diseases which importance was related to the impact attributed to them at a given time. To be referred that the terms of reference (ToR) specified in each of the previous AHAW mandates were often found to be similar. In particular, the AHAW panel was repeatedly asked to provide an actual epidemiological description of a disease, to evaluate the efficacy of currently available control measures, or to carry out an import risk assessment for diseases emerging in neighbouring countries of the EU.

The information and data needed to answer these questions were often similar. For example, systematic literature reviews for disease characterisation were carried out and quantitative and qualitative data on control measures were searched (e.g. diagnosis and monitoring tools, trade and movement measures, available vaccines or medical treatments, biosecurity measures). It is believed that a review of past opinions of the AHAW Panel will assist in developing a methodological framework to specify and prioritise the data needs.

Due to short deadlines or limited accessibility or availability of the data (either due to non-gathering or non-existence of data) at domestic or international levels, the preparatory work for the opinions tended to focus on the review or data collection process itself rather than providing answers as a result of the risk assessment. Such a situation may result in the opinion being more data-driven than demand-driven. A concise overview of the model types used in previous animal health related opinions was the focus of a recently published procurement report[2]. It was also explored the relationship between the terms of references and the models applied and an important outcome of this work was the classification of the ToR in terms of the tools requested that can be translated into data needs.

The EU animal health strategy establishes that animal health policy must be based on sound and reliable science and it is therefore essential to have sound knowledge on animal diseases and threats that are directly linked to them. Article 36 of the European parliament and Council Regulation (EC) No 178/2002 foresees the possibility to financially support a networking of organisations operating in the fields within the EFSA's mission. Activities include data collection, preparatory work for the development of the scientific opinions, and other scientific and technical support. The objective of this Article 36 call for proposals will be to assist the AHAW Panel in its level of preparedness to answer to future mandates by developing and maintaining a methodological framework to specify and prioritise

its data needs, and by developing a systematic data collection approach needed to provide rapid and accurate responses to risk managers.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The project will seek to achieve the following objectives:

- carry out a retrospective analysis of past mandates of the AHAW Panel to determine their most important, recurrent Terms of Reference,
- review the corresponding scientific assessment performed in order to specify and prioritise the type of data that are needed to answer future equivalent questions,
- assess the availability of the above mentioned data and specify where and how they can be accessed. Specifically, the project should distinguish between available data and gaps in knowledge. When data are inexistent, the project should establish what is necessary to assess the issue (e.g. collection of indirect related data, possibility for application of theoretical models...);
- define the methodology for collection, harmonisation, validation and compilation of data and at the same time, to maintain them updated; precise the format and minimum requirements in which these data need to be provided, special consideration should be given to their granularity;
- establish the methodological framework for the use of the data in a scientific assessment to address questions relevant to disease categorisation and priorisation of actions;
- participate to a workshop on data requirements for risk assessment in animal health

Thisgrant was awarded by EFSA to the consortium made between the French Food Safety Agency (ANSES-coordinator), with Moez Sanaa, Camille Bellet, Jennifer Dhé, Françoise Gauchard, Anne Thébault and Pascal Hendrikx, the University of Liege (BL), with Claude Saegerman, Marie-France Humblet and Sébastien Vandeputte, and the Central Veterinarian Institute (NL), with Aline de Koeijer, Manon Swanenburg and Clazien de Vos.

Beneficiary: ANSES/ULG/CVI

Grant title: Specification of data collection on animal diseases to increase the preparedness of the AHAW panel to answer future mandates

grant number: CFP/EFSA/AHAW/2010/01

INTRODUCTION AND OBJECTIVES

Different food safety agencies and authorities, such as EFSA (European Food Safety Agency), are requested to perform risk assessments in order to inform Animal Health (AH) and Public Health control policies. Specifically, AHAW (Animal Health and Animal Welfare) expert panel is asked to address assessment of risks related to animal diseases and welfare. The panel is invited to conduct qualitative and where possible quantitative risk assessments. Based on risk assessment outputs, the opinion recommends a range of practical options, which could be considered in order to decrease the incidence and the consequences of the considered animal disease. Risk assessment is generally produced with limited knowledge and data. The idea of risk assessment is to systemize and describe all the available knowledge and lack of it, relevant to a specific AH problem. It provides a science based, valid, and reproducible framework to address specific AH problems within a limited time and with currently available scientific data. The method for identifying, selecting, appraising and synthesising the input data for the risk model should be thoroughly considered and clearly documented.

When dealing with AH issues, the approach could differ from one assessment to another depending on the disease characteristics: pathogen type and its mode of transmission, and the presence or absence of the disease (endemic versus exotic/emerging) in the considered region. When the considered disease is exotic or emerging, the goal of scientific opinion can be the evaluation of the current preventive AH system and specially its capacity to prevent the introduction of the disease and assess its ability to prevent the disease establishment if it is introduced. The scientific opinion can aid to define and propose different sorts of management measures such as restrict animal importation and contact with wildlife, adapt quarantine food inspection system, require measures before importing animal or animal products, inform travellers etc. In addition to the protective (or safeguard) measures, the scientific opinion can address the exotic/emerging disease preparedness and evaluate the emergency management options including the development of surveillance, monitoring and early warning systems. On the other hand, when the disease is endemic, the goal of the scientific opinion will be to evaluate options and measures to be implemented to reduce or control the endemic level. The options and measures are intended for example to break the maintenance cycle for vector-borne diseases; reduce transmission between animals, and/or between animals and humans. The scientific opinion can help to show how to optimize AH programs, control the contamination of animal products, control food or water contamination, and when relevant, control the spread of the disease within wildlife and other possible reservoirs including environmental ones. As one could expect, there are several permutations or approaches for conducting AH scientific opinion.

Whatever the disease (e.g. exotic vs endemic) being considered and the population concerned (e.g. domestic animals vs wildlife), the scientific opinions consist of a series of answers to different relevant questions whose ultimate aim is to better inform risk managers and suggest scientific based measures to prevent or to control the disease.

MATERIALS AND METHODS

The overall objective of this project was hence to develop a methodological framework to specify data needs in relation to criteria for categorisation of animal diseases as defined by risk managers.

The project seeks to achieve the following objectives:

- establish a typology of risk questions and identify data needs (WP1);

- perform a retrospective analysis and critical review of Scientific Opinions adopted by the AHAW Panel in the area of AH (WP1);
- identify the data needs in relation to criteria for categorisation of animal diseases as defined by risk managers (WP1);
- identify sources of information and its accessibility (WP2);
- develop a methodology for data collection, including the definition of metadata standards for outcomes values to support facts validation and quality assessment (WP3);
- establish a methodological framework for use of the data in a scientific assessment to address questions relevant to animal diseases (WP4).

In order to reach these objectives, settled by EFSA, the project was organised in four interdependent work packages. Figure 1 gives a global overview of the structure of the report and its annexes.

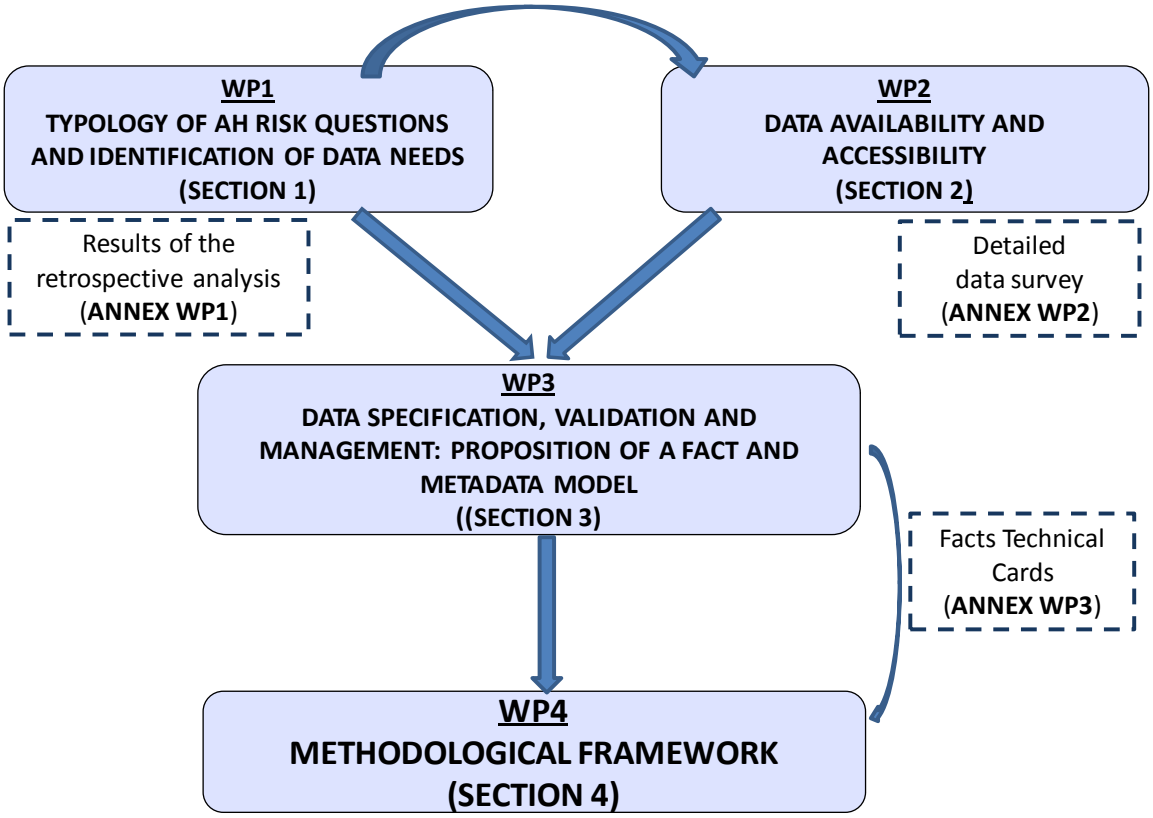


Figure 1: Presentation of the DATASPEC project Flowchart

1. WP1-Typology of risk questions and identification of data needs

1.1. Objectives of WP1

The aim of this first work package was to define a comprehensive list of relevant AH questions, based on the retrospective review of EFSA AHAW scientific opinions, and to identify the related AHAW experts' data needs.

1.2. Retrospective analysis of AHAW scientific opinions (2004 to 2010)

1.2.1. Selection of scientific opinions

We included all the scientific opinions dealing with AH issues and adopted by AHAW panel between 2004 and 2010. Animal Welfare scientific opinions were excluded, thus including a total of 38 opinions out of 75 AHAW collected. The complete list is included in appendix A.

1.2.2. Workflow of analysis

The review Workflow is displayed in appendix B. The scheme aims at showing the different steps undertaken during the assessment of the 38 reviewed opinions. Two steps of readings were done for each opinion:

Step 1: One single scientist (responsible scientist) reviewed all the terms of reference of the 38 opinions in order to identify the different types of risk questions asked to AHAW panel within each mandates. The goals were to create a first level of risk questions' typology and grouping (table 1) and to distribute the reports in between the eleven consortium members, regarding their fields of expertise.

Step 2: The members of the consortium analysed one to seven opinions, and were asked to extract the following information from each:

- the type(s) of risk question(s) present in the opinion,
- the methodology used to answer each identified risk question
- the list of data and sources used for each identified risk question
- the list of data gaps related to each identified risk question.

In order to standardize the review, a form (appendix C) was given to the different consortium members. One form was used by opinion. Meetings were organized between consortium members and the responsible scientist, avoiding any subjective misinterpretation of WP1's workflow and guaranteeing a common definition of methodological terms used by reviewers. All forms were finally collected and analyzed by this same responsible scientist, assuring the consistency of the review although the given limited timeframe.

1.2.3. Typology of risk questions

The review pointed out twelve main types of risk questions (figure 2 and table 1), which categorisation was done, contingent on A. Singer previous risk questions' classification (Singer, 2010).

1. Host/Pathogen characteristics;

2. Disease status;

3. Potential risk of spreading to susceptible population/Pathways of transmission & speed of the spread;

4. Risk of (re)introduction;

5. Risk of establishment;

6. Effectiveness of surveillance measures;

7. Diagnostic tools availability and efficiency;

8. Effectiveness of prevention tools;

9. Effectiveness of control measures;

10. Effectiveness of bio-security measures;

11. Treatment availability and efficiency;

12. Vaccine availability and efficiency;

The distribution of the risk questions per opinion vary from 1 to 8 with an average of 4.

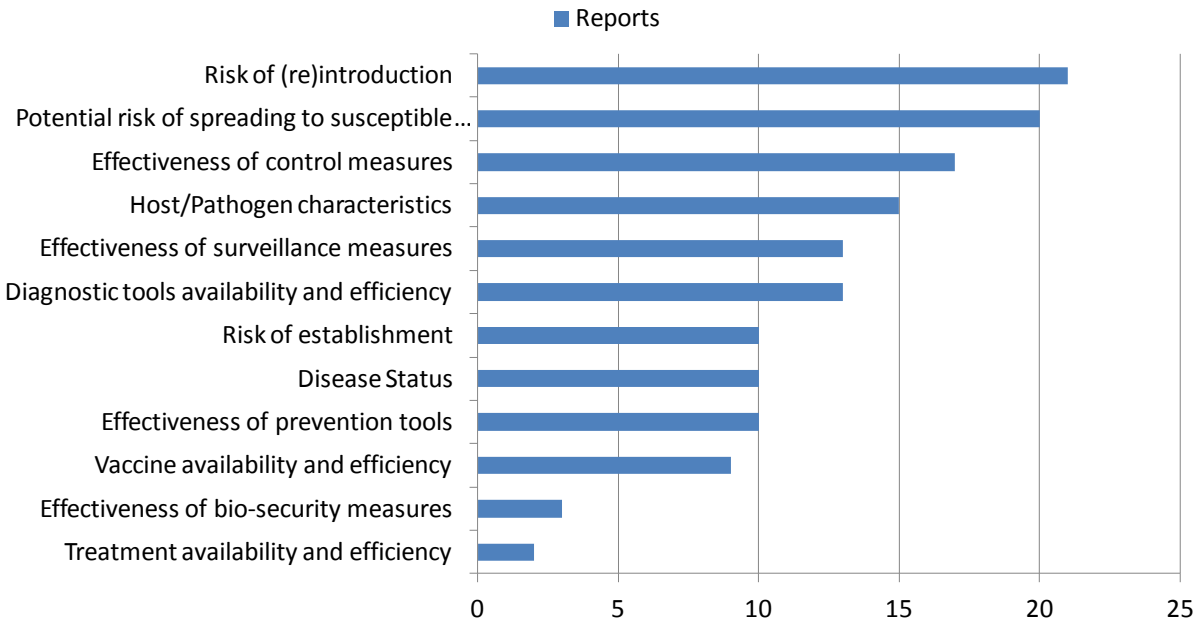


Figure 2: Distribution of AHAW risk questions' types among the 38 opinions

Table 1: Typology of the risk questions

Type of risk questions	Description	Examples of AHAW opinions
Disease Status	Status of a country or a zone with respect to an animal disease, according to the criteria listed in the Terrestrial Code (OIE)	Foot and Mouth Disease (EFSA-Q-2004-113)
Host/Pathogen characteristics	Host's epidemiological status (e.g. principal/accidental, reservoir, vector), host/pathogen's type (e.g. domestic/wildlife, human, bacteria/virus/parasite), habitat, resistance	Rift Valley Fever (EFSA-Q-2004-050), Besnoitiosis (EFSA-Q-2009-00879)
Potential risk of spreading to susceptible population/Pathways of transmission & speed of the spread	Disease potential to spread geographically and to cross natural barrier, route and transmission pathway(e.g. food born, contact, air, vector...), spread capacity of the disease (e.g. highly/barely contagious)	Dairy Products (EFSA-Q-2004-161), Porcine Reproductive and Respiratory Syndrome (EFSA-Q-2004-100)
Risk of (re)introduction	Release and exposure assessments	Foot and Mouth Disease (EFSA-Q-2004-113)
Risk of establishment	Consequence assessment	Classical Swine Fever in Wild Boar (EFSA-Q-2007-200)
Effectiveness of bio-security measures	Type and quality of bio-security protocols	BoHV-1 (EFSA-Q-2005-018)
Effectiveness of prevention tools	Description (e.g. zoning) and evaluation of the control practices at the borders for trade and movement…	Rabies (EFSA-Q-2006-014)
Effectiveness of control measures	Disease monitoring, description of the tests/control (e.g. at the border, slaughter…)/vaccination/treatment, description of the control experience and success in other countries (Member States, Developed & Developing countries), host traceability…	Bluetongue vectors and insecticides (EFSA-Q-2007-201)
Diagnostic tools availability and efficiency	Lists of the available and efficient tests/kits, Se, Sp, skills of the national laboratories…	Brucellosis diagnostic methods (EFSA-Q-2005-060)
Treatment availability and efficiency	Lists of the available and efficient tests, strategies of treatment, countries specificities, and residues of treatments…	Besnoitiosis (EFSA-Q-2009-00879)
Vaccine availability and efficiency	Type of vaccines (lived/attenuated/killed/sub unit), induced immune-response, composition of the vaccine, DIVA strategy, vaccine market, adverse and side-effects…	Vaccination again avian influenza of H5 and H7 subtypes (EFSA-Q-2006-156)

The top three risk question types were, by decreasing order (figure 2):

1. The risk of (re)introduction (21 opinions),

2. The risk of potential spread to susceptible population, the pathways of transmission and speed of the spread (20 opinions),

3. The effectiveness of control measures (17 opinions).

1.3. Risk assessment's steps

Most AH risk assessments are conducted according to the OIE guidelines. The Terrestrial Animal Health Code (OIE, 2010), which governs animal import risk assessment describes four steps: (i) Release assessment, (ii) Exposure assessment, (iii) Consequence assessment, and (iv) Risk estimation. Full risk assessment, including the four steps, was found only in ten AHAW opinions. Figure 3 shows the frequency of the different risk assessment's steps and hazard identification.

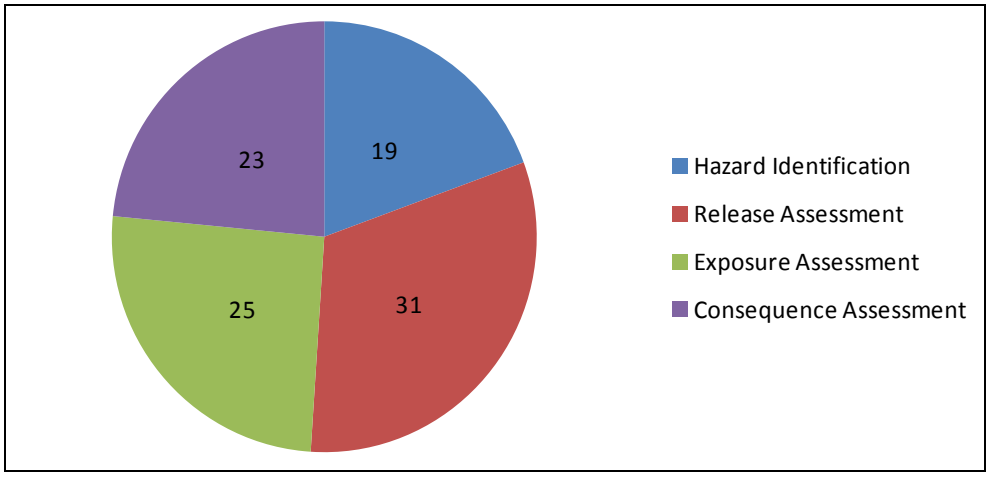


Figure 3: Risk assessment steps among the 38 opinions

Three general categories of risk assessment were identified: quantitative, qualitative and narrative risk assessment (see glossary). Figure 4 presents the proportion of the different risk assessment categories used in the 38 opinions.

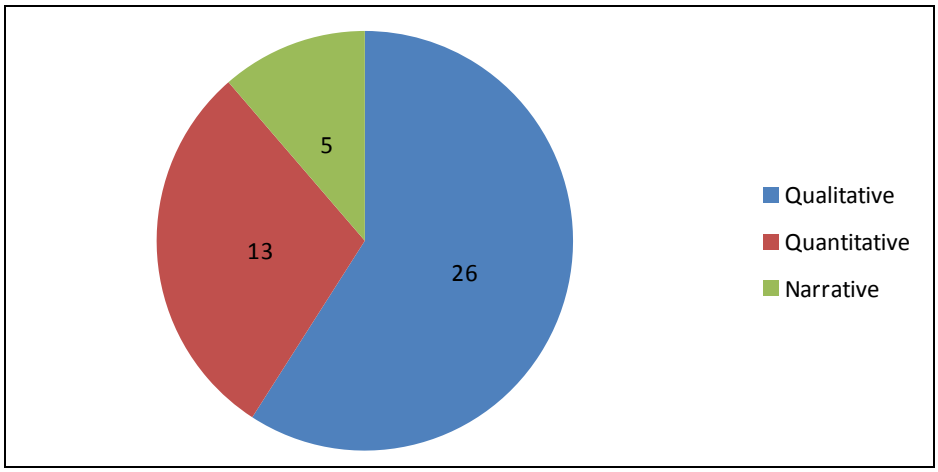


Figure 4: Proportion of risk assessment's type among the 38 opinions

In general, only one approach by report was used. Qualitative risk assessment was more often used than quantitative risk assessment. The use of qualitative and narrative risk assessment was occasionally justified by the wideness of the Terms of References (ToRs), but mainly by the lack of quantitative data.

1.4. Data gaps identified

Based on the data used and the expressed data needs observed in the reviewed AHAW opinions, examples of current data gaps of relevance (not prioritised), on the considered animal health issues, could be presented as follows:

- data related to the description of the disease and the mechanisms of hosts-pathogens interactions pathogenicity that allow a clear case definition;
- data from MSs to assess the level of harmonisation and standardisation, regarding for example characteristics of laboratory tests and vaccine protocols;
- accurate data allowing the comparison of infectious diseases' prevalence, provided by the different MSs;
- data describing the disease transmission modalities and pathways;
- accurate data describing animal populations and animal production systems (e.g. data on animal movements and trade and on housing and feeding systems).

It is important to distinguish between current data gaps and lack of data accessibility. Current data gaps may be due to the absence of evidence about a causal relationship between an exposure factor and a certain disease, an absence of knowledge about transmission mechanisms, or an absence of surveys assessing disease prevalence. Lack of accessibility may be associated to the difficulties to centralise available data in MSs. A general comment was frequently mentioned by experts and reviewers, regarding the difficulty to extract MSs' National databases, compiled by National Laboratories or Agencies.

Data gaps could be classified in four categories:

- informational (= *what*)
- temporal (= *when*)
- spatial (= *where*)
- and quality (= *how*)

1.5. Inventory of the data needs

1.5.1. Identification and categorization of the data used

Eight main data sources were identified when analysing the 38 AHAW opinions (figure 5). Detailed descriptions of the data used were organized in an Excel file (annex WP1).

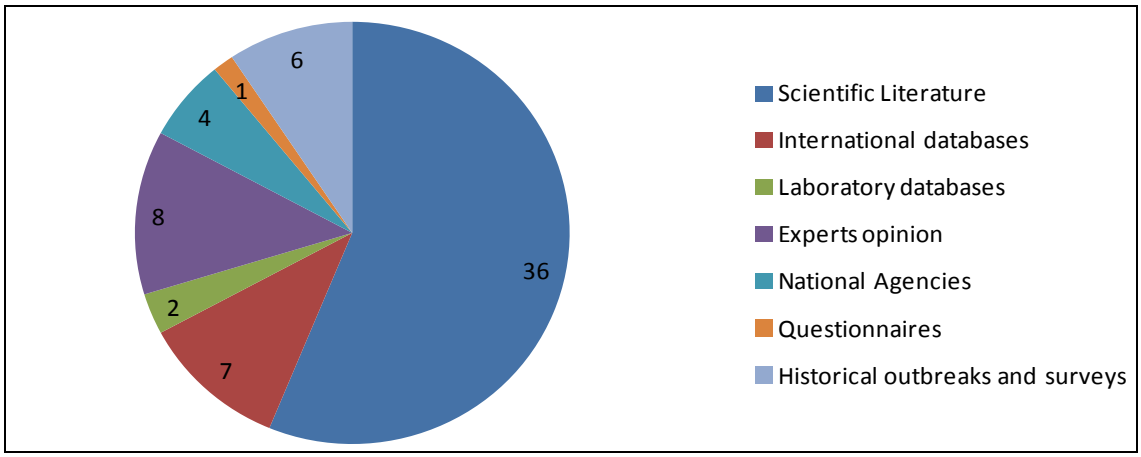


Figure 5: Types of the data sources used among the 38 reviewed AHAW opinions

The different types of sources used within AHAW opinions were:

- scientific literature, meaning published papers in peer reviewed journals;
- international databases, such as those from OIE, EUROSTAT, FAOSTAT, EFSA ADNS;
- laboratories databases, such as the one from Pirbright;
- national agencies, meaning the data or assessments coming from Member States' national institutions;
- expert opinions;
- questionnaires, as for example questionnaires sent to manufacturers;

- historical outbreaks and surveys, meaning data from disease surveillance programmes.

As shown in the figure 5, published scientific papers represented the main source of data used by AHAW panel experts in the 38 reviewed opinions (55%).

1.5.2. Categories of data needs

On combining the list of used data and the identified data gaps corresponding to the questions addressed in the 38 analysed opinions, a list of data needs that grouped in five categories was established (tables 2 to 6), data concerning:

- disease general information,
- descriptive epidemiology of the disease,
- analytical epidemiology of the disease,
- disease prevention and control,
- disease surveillance.

Within each category, subcategories of data were created in order to facilitate the development of the data specification's methodological framework (see section 3).

Table 2: Disease general information data needs

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
DISEASE GENERAL INFORMATION	**Disease information**	Etiologic agent	Viral	General categories, taxonomic classification (family, genus, species), common name of the disease or condition, pathogen strain or serotype
			Bacterial	Classification of the causative agent
			Fungal	Classification of the causative agent
			Parasite	Classification of the causative agent
			Prionic	Classification of the causative agent
		Host(s)	Principal	List of susceptible species (man included if zoonotic/common disease)
			Intermediate	List of susceptible species (man included if zoonotic/common disease)
			Accidental	List of susceptible species (man included if zoonotic/common disease)
		Pathogenicity	Source	Sources of the causative agent(s)
			Infection	Incubation period, latency period, pathogenicity, tissue pathogen loads, clinical signs (live animals), gross lesions, duration of clinical signs
			Transmission	Routes of transmission (direct/fomites), dose-response, excretion
		Diagnosis	Clinical diagnosis	Pathognomonic signs, differential diagnosis of the disease
			Lab diagnosis	Sample type, test used (procedure, sensitivity, specificity, threshold...)
		Immune response		Type of immune response, maternal derived antibody protection, level of immune response, duration of immune response (Ab+), serotype/strain cross-protection
		Genetics	Pathogen	Evolutive character of the pathogen (mutations, genetic re-assortments)
			Host	Genetic resistance against the pathogen
Category	**Sub-category level 1**	**Sub-category level 2**	**Sub-category level 3**	**DATA NEEDED**
DISEASE GENERAL INFORMATION		Resistance of the pathogen		Pathogen resistance to chemical (including treatment) and physical agents
	Vector information	Vector species		List of vector species

	Vector competence		Factual evidence of local vector(s) competence, list of potential local vector species
	Habitat / Environment		Vector type of habitat, optimal range of T°, relative humidity, wind speed, altitude range
	Distribution area		Distribution area (continent/country/region), density of vectors
	Vector activity		Type (indoor/outdoor), period (day/night), duration and seasonality
	Cycle		Duration, lifestages, overwintering
	Pathogen transmission		Infectious load, type of transmission (trans-stadial, trans-ovarial), maximal distance of potential spread
	Resistance to insecticides		Reported resistance(s) to insecticides, list of insecticides for which resistances have been reported
	Climatic information	Climate Classification	Köppen-Geiger climate classification
		Climatic data	Average/minimal/maximal temperatures, average relative humidity, rainfalls, sunlight, wind speed and prevailing direction
		Drought	Palmer Drought Severity Index
	Altitude		Altitude

Table 3: Descriptive epidemiology data needs

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
DESCRIPTIVE EPIDEMIOLOGY	**Morbidity**			Prevalence, Incidence
	Mortality			Mortality rate
	Case-fatality			Case-fatality rate
	Spatio-temporal distribution			Annual prevalence, annual incidence, form of the disease (epizootic, endemic, etc.)...
	Demography of hosts			Populations and subpopulations, herds/flocks/groups of animals

Table 4: Analytical epidemiology data needs

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
ANALYTICAL EPIDEMIOLOGY	**Disease factors**	Reservoir		List of reservoir species
		Optimal climatic and environmental conditions for pathogen survival		Optimal relative humidity, optimal soil pH, maximal radiant flux, duration under optimal conditions
		Ecology		Land cover, length of pasture-edges, symbiotic factors
		Industry and management		Industry and management factors affecting disease transmission and spread: confinement operation, biosecurity practices, industry awareness…
	Factors of disease introduction	Importations/entries	Animals	N importations of livestock species and horses , N entries of domestic and exotic pets, vaccination status of arriving animals, illegal trade of live animals, migratory birds
			By-products	N importations of livestock and horse by-products by categories (1/2/3), illegal trade of by-products by categories (1/2/3)
			Vectors	Capacity of vector spreading by the air, transports…
			Biological products	N importations of lived and attenuated vaccines
			Humans	Human entries (arrivals) from areas at risk
			Plants	List of the risky plants, N importations of plants by type of plants

		Fomites	List of the fomites, N importations of high-risk fomites
Factors of disease spreading and establishment	Exportations/exits	Animals	N exportations of livestock species and horses , N exits of domestic and exotic pets
		By-products	N exportations of animal by-products by categories (1/2/3)
		Vectors	Capacity of vector spreading by the air transports…
		Biological products	N exportations of semen, embryos, lived and attenuated vaccines
		Humans	Human exits
		Plants	N exportations of plants
		Fomites	N exportations of high-risk fomites
	Basic reproductive rate		Basic reproductive rate (R0)
	Reproduction programme		Type of reproduction programme (AI vs. natural)
	Production system	Industries	Cattle production systems, N animals in meat and dairy industries, N livestock herds/flocks in meat and dairy industries
			Goat production systems, N animals in meat and dairy industries, N flocks in meat and dairy industries
			Sheep production systems, N animals in meat and dairy industries, N flocks in meat and dairy industries
			Swine production systems, N animals in meat industries, N herds in meat industry
		Farming systems	Cattle: N animals and herds/flocks under intensive vs. extensive farming systems
			Goats: N animals and herds/flocks under intensive vs. extensive farming systems
			Sheep: N animals and herds/flocks under intensive vs. extensive farming systems
			Swine: N animals and herds/flocks under intensive vs. extensive farming systems
	Local establishment		Risk of establishment in local vector(s), domestic/wild host(s) and environment

Table 5: Prevention and control data needs

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
PREVENTION - CONTROL	**Veterinary Services (VS)**	Identification		Name of the veterinary service
		Veterinary Statutory Body		Level of authority, capacity to implement its functions and objectives in conformity with OIE standards
		Evaluation	Resources	Professional profile of VS staff, skills of veterinarians and para-professionals, N veterinarians involved in epidemiosurveillance, N vets employed at border inspection points, access to physical resources, ability to access financial resources for continued operations and in emergency situations
			Organisation	Stability of structures and sustainability of policies, centralization vs. decentralization, internal/external coordination, mechanisms for consultation with stakeholders
			Education	Existence of continued education and awareness campaigns for professionals involved in animal health, public targeted, frequency
		Legislation		Authority/capability to participate in the preparation of national legislation and regulations, programmes/activities to ensure stakeholder compliant with relevant legislation and regulations, authority/capability to be active in the international harmonisation of regulations and sanitary measures and to ensure national legislation/regulations take into account international standards, authority/capability to certify animal, animal products, services and processes in accordance with national legislation/regulations and international standards, authority/capability to negotiate, implement and maintain equivalence and other types of sanitary agreements with trading partners
		Compartmentalisation		Authority/capability to establish and maintain disease-free zones/compartments in accordance with OIE criteria
	Disease inspection practices	Quarantine		Number of inspection points, N animals quarantined, average duration
		Mandatory slaughters		N mandatory slaughters
		Animal movement and traceability		Identification of live animals and by-products, follow-up of movements (live animals and by-products)

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
PREVENTION - CONTROL	**Biosecurity practices**	Biosecurity protocol, control and certification	In the production systems	Biosecurity protocol, control and certification in farms/inspection posts (livestock herds, insemination centres, pet shops, horse farms, riding schools, border posts) and for transports
			During animal transport	Biosecurity protocol, control and certification for transport of live animals and by-products, for embryo transfer, at animal groupings (markets, shows, competitions), at slaughterhouses
		Manufacturing process of the feed	Feed composition	Type of animal feed (pasture, concentrates,etc.), nature of animal feed (animal vs. non-animal), nutritional characteristics
			Certification	Protocol of good manufacturing practices (GMP), control of GMP, existence of ISO standards, control of ISO standards
			Treatment	treatment of feed, chemical treatment of water
			Control	Existence of feed and water control, type of samples, sample size, frequency
		Distribution process of the feed	Origin	Country of origin, year of manufacturing, origin (farm vs. manufacture), animal vs. non-animal, Quantity of feed distributed
			Network	N intermediates between supplier and user, list of sale points, control of sale points, frequency of controls, type of water supply
		Waste management		Type of waste management for feed, water, contaminated materials and instruments (veterinary)
		Supplementary inputs		Existence of supplementation with nutritional additives, list of nutritional additives used in routine
		Systematic treatment		Existing systematic treatment distributed with feed, list of specific treatments used as feed supplements (antibiotics, etc.)
	Reservoir(s) and vector(s) control	Reservoir(s)		Existence of a control of reservoir(s), methods, place and effectiveness
		Vector(s)		Existence of a control of vector(s), methods, place and effectiveness
	Treatment	Treatment strategy		Authorized vs. prohibited, category of molecule(s), duration, posology, withholding periods in milk and meat
	Vaccination	Vaccine	General information	Type of vaccine, serotype/strain, DIVA

			Protocol	Route of administration, N doses for a regular protocol, N doses primo-vaccination, interval between 2 doses of primo-vaccination, interval between primo-vaccination and booster, frequency of vaccination
			Efficacy	Duration of vaccinal protection, existence of cross-protection, list of pathogens for which cross-protection is reported, potential risk of pathogen excretion
			Safety	List of reported side-effects and adverse-effects, subpopulations exempted from vaccination, vaccine safety surveillance system
		Programme	Mandatory	Status of vaccination (mandatory/facultative/prohibited)
			Strategy	Doses of vaccine used at the national level, N animals vaccinated at the national level, frequency of vaccination, type of vaccination programme (national, regional)
			Distribution	Delay between exit from pharmaceutical industry and use in the field, vaccinator, actors of the distribution network
			Conservation	Duration of vaccine conservation, temperature conditions, percentage of doses really used

Table 6: Surveillance data needs

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	DATA NEEDED
SURVEILLANCE	**Surveillance Networks**	Organisation		Stakeholders (policymakers, information users, beneficiaries of the surveillance information, data providers) and responsible parties (responsible of surveillance system design, implementation and leadership, data application design, development and implementation, data application support and maintenance, data collection, laboratory testing, detection of the cases, confirmation of the cases, reports of the cases, training of data collectors, data analysis and interpretation, result dissemination and reporting, action based, action based on surveillance findings, review of surveillance system effectiveness): role, title, group name or agency, N officers involved, N levels

	Characteristics	Field of surveillance networks, type of surveillance, situation of the disease(s) surveyed, population(s) surveyed, mode of data collection, dependence towards fighting actors
	Evaluation	Existing evaluation, type of evaluation (internal/external), list of performance indicators, frequency
Objectives and decision criteria	Objectives	Description of the purpose and rationale for surveillance: estimate the magnitude and baseline status of a problem, determine the geographic and demographic extent of an outbreak, predict possible spread and provide data for disease regionalization, describe the natural history of a pathogen or disease, detect unusual clusters of disease, providing for early detection, generate hypothesis and stimulate research, define or assess the health status of a population, providing the foundation for market confidence, detect changes in health practices, risk factors or exposure, facilitate planning of national control or eradication programs and strategies, evaluate control measures and intervention efforts, identify factors associated with a disease agent that may be used in conducting surveillance elsewhere and in modelling pathogen spread and determine times of year when most case are observed
	Decision criteria	List of criteria for surveillance (criteria prioritization, such as impacts on trade and productivity, animal welfare concerns, feasibility of control, cost of surveillance and public health implications), list of decision criteria for active vs. passive surveillance
Study design	Type of the surveillance	Type of the surveillance (active vs. passive), proportion of active vs. passive surveillance, surveillance at markets, border inspection points, animal groupings, wildlife, clinical vs. syndromic surveillance
	Expected outcomes	Information resulting from the surveillance effort, which is then used for decision-making, policy development and action
Population description and characteristics	Sampling Units	Simple units (individuals) or aggregated units (herds or flocks), geographical or spatial measure included, time constraints, if present, are included
	Target population	Population about which statistical inference will be made (general population at risk), should be identified and clearly defined or estimated, if different from the study population, the rationale for inference should be provided, size of target population

		Study population	Population from which the sample is to be drawn, size of study population, sample frame (list of units to be sampled)
		Targeted population	Population defined by specific disease variables inherent to the disease in question
		Administrative units	Which units are included in the surveillance system (states, province, sample grid reference…)
		Size of sample	Number of reporting unit, should include geographic area serviced per unit sampled, number of eligible units served by reporting unit (per unit of geographic area being serviced)
		Animal and group type	Species, breed and type (if applicable) of animals should be evident; include breeds and crosses, define the animal by appropriate production phase concept, age categories, including all appropriate categories pertinent to the surveillance objectives
	Case definition	Clinical description and case definition	List of criteria for positive case, negative case, and others as applicable
		Epidemiological criteria and restrictions	Criteria that: may restrict case definition to individual animals, herds, flocks or premises that possess specific epidemiological characteristics, may relate to the geographic location of an animal, farm or premises; a particular point in time or season of the year; a particular behaviour associated with disease transmission or risk factor, may compartmentalized the surveillance within a segment of a vertically integrated industry, age group or commodity type, may include variables related to habitat, environmental conditions, seasonality, climate,...
		Laboratory criteria	Description of the tests' specificity, sensitivity, identification of the limitations of the tests used for the disease confirmation, type of the diagnostic test and cut-off point or dilution used to define categories of cases, particular additions specific to the testing
		Case classification	Definition of the so called suspect, probable and confirmed case categories, levels of the classification certainty
	Outbreaks	Definition	Decision criteria of an outbreak notification
		Notification	N outbreaks reported and N animals affected per outbreak

	Sampling methods		Description of the field and laboratory data collection techniques: simple random, systematic, cluster, stratified or complex sampling, probability sampling, methods for randomization and stratification, level of detection, statistical level of confidence, diagnostic sensitivity of the sampling, predictive value, time intervals and frequency of data collection, geographic extent of the study area under surveillance, methods of data collection and handling (how raw data are gathered from the field, sample handling protocol, cold chain measures, sample degradation factors), sources of potential bias, trigger for data collection, transmission of collected data (web-based data, e-mail, fax, software...)
	Laboratory	Diagnostic test	Date of sampling collection and lab test, test(s) used, sensitivity and specificity, chosen threshold of detection
		Accreditation	Existence of accreditation programs, list of standards for accreditation programs, existence and frequency of proficiency testing
	Early warning system		Existence of an early warning system, type, list of indicators used, website
	Risk and exposure factors		Population risk factors that may influence the outcomes of the study, confounders should be included
	Communication		Awareness Campaigns, Communication and feedbacks

2. WP2 - Data availability, accessibility and form

2.1. Objectives of WP2

In the previous section we proposed an exhaustive list of data required by experts to answer AH risk questions. In WP2, we intended to assess the availability of these data and identify the manner they could be accessed on a sustainable way for further analysis. Therefore, we performed a systematic and standardised assessment of existing data resources, at a national and international level.

The main objective of WP2 was to make available a direct link between the identified AHAW data needs and the database compiling the data resources.

Because of the limited timeframe, it was decided to focus on four MSs (Belgium, France, Netherlands and Spain) and four diseases consistent with EU Commission and EFSA requirements. The diseases, selected as case studies to apply the proposed methodological framework, were:

- an exotic and vector borne disease, the Venezuelan Equine Encephalitis (VEE).
- an endemic disease, the Porcine Reproductive and Respiratory Syndrome (PRRS).
- a zoonotic parasitic disease, involving wildlife and pets in its cycle, the Echinococcosis, with *Echinococcus granulosus* (EG) and *Echinococcus multilocularis* (EM).

2.2. Materials and methods

Different methodologies were used to provide a general overview of data availability. A direct and an indirect survey were conducted to identify existing data resources. The direct survey was based on a comparison of MSs' data collection systems, a workshop organised in Paris and a questionnaire sent to governmental contact points of each elicited MS, whereas the indirect survey was based on a more global web-based search. Figure 6 summarizes the process applied in WP2.

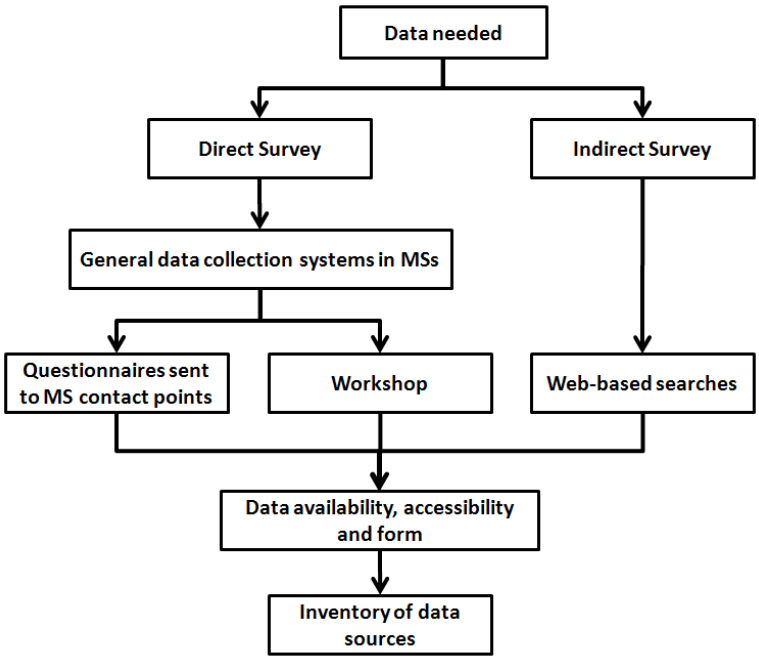


Figure 6: Review of the resources of WP1 required data: WP2 flow chart

2.2.1. Assessment of data availability: direct survey

Animal Health data accessibility, and therefore availability, is first met when authorized animal health professionals are provided with the means to find the data of their interest (Kirch, 2008). It often relies on the knowledge of the system data collection organisation, what had indeed to be searched for, before the direct survey started.

Because Spain finally declined its participation to the direct survey, the general description of the data collection systems was only done for BE, FR and NL.

2.2.1.1. General description of animal health data collection systems in three MSs

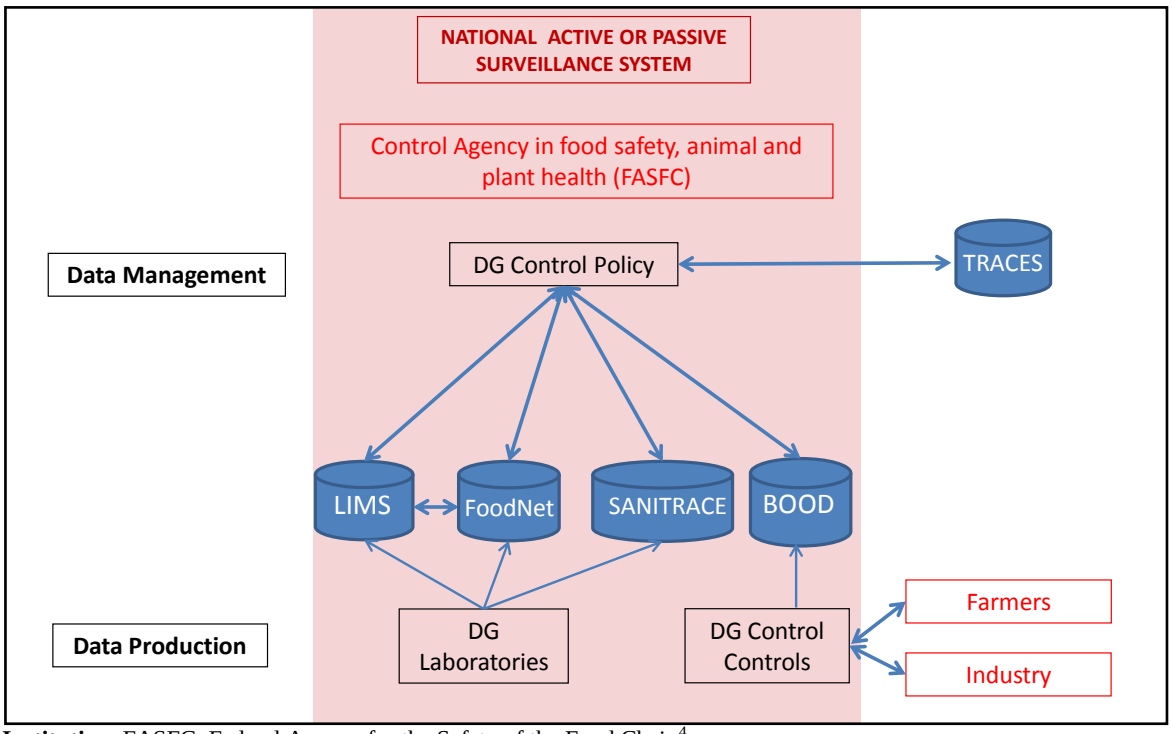
- **Belgian animal health data collection system**:

The Control Agency in food safety, animal and plant health FASFC uses multiple consequent databases which are continuously updated, interconnected and well centralized (figure 7). FASFC is responsible of the food chain data management. Beside FASFC, the Belgian Scientific Institute of Public Health[3] is mainly in charge of public health aspects including zoonotic infections in humans.

[3] http://iph.fgov.be/index.asp?Lang=FR

An advantage of the Belgian animal health data collection system resides in the fact that it is integrated with other databases covering the whole food chain. In general no additional private data are needed to perform scientific assessment studies and there is no limitation for the use of databases of the FASFC for purposes which are in accordance with the competences of the FASFC (e.g. FASFC Scientific Committee, 2007). When consulting databases, it is frequently observed that the results are incomplete or even incorrect; the main reasons are the lack of complete metadata, missing or non-informative data, inconsistent format (text format intermingled with numerical values, date format…), free fields, fields used for other information, insufficient preciseness, incorrect linking between different tables and inappropriate interpretation of the data.

Most FASFC databases are consultable via the Business Objects program after creation of appropriate universes. Extracting correct data out of the complex databases is not always simple. Indeed special attention has always to be paid to the formulation of the query in order to extract complete and informative data. The verification of the obtained information has to be done to be sure that correct linking or correct conditions have been used in the query and to draw correct conclusions afterwards (FASFC Scientific Committee, 2007).



Institution: FASFC: Federal Agency for the Safety of the Food Chain[4]

: Databases (DB): BOOD: Operator Database containing information about companies or persons in the food chain and on the localization of the facilities; **FoodNet**: database of the control activities of the FASFC; **SANITRACE** (fusion of Sanitel -live animals- and Beltrace –Slaughtering): Database for live animals (cattle, swine, poultry, ovine, caprine and cervids sectors) and slaughtering; LIMS: Laboratory Information Management System; TRACES: TRAde Control and Expert System.

Figure 7: The Belgian animal health data collection system

[4] http://www.favv.be/home-fr/

- **French animal health data collection system**:

The French animal health data collection system uses multiple decentralized databases (figure 8). Several actors are involved in the management of data.

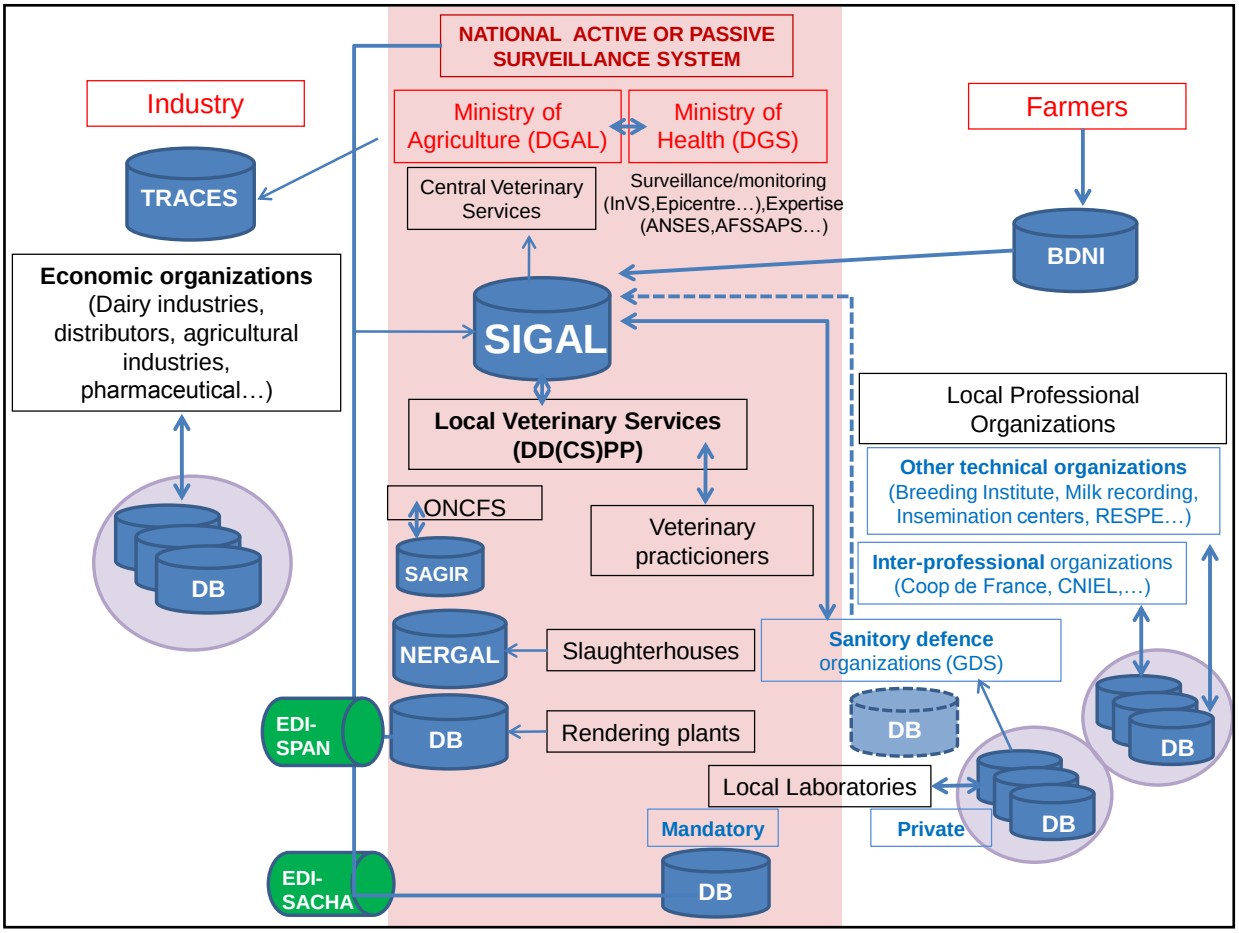
Since the beginning of the years 2000, the national veterinary services implemented a national database (known as SIGAL for *'Système d'Information de la DGAL'*) to integrate all data about farms (all species considered) and regulatory health monitoring activities linked to them. This database is monitored by the provincial veterinary services in charge of regulatory diseases surveillance and control in their zone. Data come from the veterinary services themselves (inspection activities), the veterinary practitioners (acting with an official sanitary mandate) and provincial veterinary diagnostic laboratories. Some of the activities are, in some provinces, delegated to professionals such as "Sanitary defence groupings" who have then access to some parts of SIGAL in order to contribute to monitor the database.

The advantage of SIGAL is its national coverage allowing a standardized collection of data. Integration of laboratory data or rendering plants data is done electronically which helps real time monitoring of diseases. SIGAL has a very good and operational link with other national databases such as the National cattle register. The weaknesses of SIGAL are linked with its size: it is difficult to make changes in the structure of the database which has an impact on its adaptability to new health problems and extraction facilities are at this stage not enough developed to have an easy to use set of data for analysis and interpretation both at provincial and central level. The system is very well developed for bovines but needs to be more developed for other species such as porcine or small ruminants.

As SIGAL is mainly monitoring regulated diseases, data on other diseases, when monitored, are to found in professional private organizations such as GDS France (for infectious bovine rhino-tracheitis for example) or Coop de France (for some porcine or poultry diseases). No links are established today between these systems.

The national veterinary services, ANSES, veterinary practitioners, veterinary diagnostic laboratories and the two main professional farmers organizations have decided to create a national Platform on epidemiological surveillance in order to group their efforts for the implementation of surveillance activities, have a better coordination between the surveillance networks and information systems. This Platform will be launched in September 2011[5].

[5] http://www.anses.fr/cgi-bin/countdocs.cgi?Documents/ANSES-Ft-PlateformeEpid.pdf



Institutions: DGAL: General Directorate for Food of the French ministry of Agriculture, Food, Fisheries and Rural Affairs; DGS: General Directorate of Health; DD(CS)PP: Departmental Direction of Social Cohesion and Population Protection; ONCFS: National Office for Wildlife and Hunting; RESPE: Equine disease epidemiological surveillance Network; **Coop de France**: French Agricultural Cooperative; CNIEL: National Inter-professional Centre of Dairy Economy.

: **Databases (DB)**: BDNI: National cattle base register; SIGAL: DGAL information system; TRACES: TRAde Control and Expert System; SAGIR: Wildlife epidemiological surveillance system; NERGAL: Slaughterhouse condemnation database; EDI-Sacha: Harmonized laboratory electronic data transfer[6]. EDI-Span: Harmonized rendering plants electronic data transfer

: Uncompleted transmission of data (not notifiable diseases). : Private databases difficult to access. : In construction.

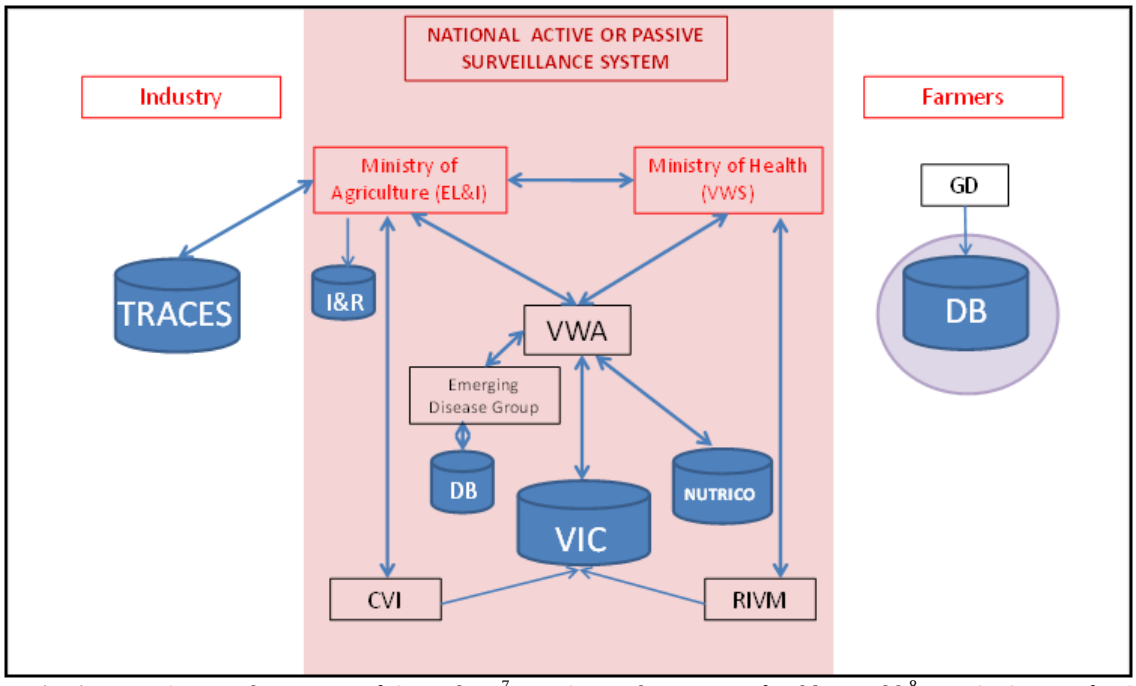
Figure 8: The French animal health data collection system

- **Dutch animal health data collection system**:

The Dutch animal health data collection system uses multiple interconnected databases (figure 9). Several actors are involved in the management of data.

[6] http://www.edi-sacha.eu/webpages.aspx?rID=325

The Dutch Ministry of Agriculture is responsible for the monitoring, surveillance and control of notifiable diseases in animals. The Dutch Ministry of Public Health is responsible of the public health, including zoonotic infections in humans. Different institutes collaborate with these two ministries: the Agency for Safety of Food and Consumer Products (VWA) works with both ministries and is responsible of notifiable diseases and zoonotic infections; the data used and maintained by the VWA, in CIV database, come from the Central Veterinary Institute (CVI) and the Institute for Public Health and Environmental Hygiene (RIVM). The RIVM mostly works for the Dutch Ministry of Public Health (eHealtHealth VWS): it diagnoses and monitors notifiable diseases in humans and is also in charge of vector surveillance. The CVI mostly works for the Dutch Ministry of Agriculture (EL&I) and performs diagnostics of notifiable diseases in livestock, pets and wildlife including cultured fish (the diagnostics are performed on routine, specific purchases, import/exports and so on). The Animal Health Services (GD) is a private business that works for farmers. It has its own databases that do not have a free access for public institutions, although access is often possible.



Institutions: EL&I: Dutch Ministry of Agriculture[7]; VWS: Dutch ministry of Public Health[8]; VWA: Agency for Safety of Food and Consumer Products[9]; RIVM: Institute for Public health and Environmental Hygiene[10]; CVI: Central Veterinary Institute[11]; GD: Animal Health Services.

: Databases (DB): TRACES: TRAde Control and Expert System; I&R: Database for cattle, pigs and sheep; NUTRICO: database on food safety including non-infectious aspects; VIC: Database on notifiable diseases.

: Private databases difficult to access

Figure 9: The Dutch animal health data collection system

[7] http://english.minlnv.nl/portal/page?_pageid=116,1640354&_dad=portal&_schema=PORTAL
[8] http://english.minvws.nl/en/
[9] http://www.vwa.nl/
[10] http://www.rivm.nl/en/
[11] http://www.cvi.wur.nl/uk

2.2.1.2. Experts workshop representing participating Member States

A workshop was held in Paris in June 2011, gathering experts of the four selected diseases from the participating MSs. Its main objectives were to evaluate the relevancy of the data to be incorporated in the MSs' questionnaires and to identify the possible risk questions to be targeted within the frameworks of WP4.

2.2.1.3. Questionnaires

Four questionnaires (one per case study) were elaborated and sent to MSs focal points. Starting from the generic exhaustive list of data needed, a switch on/off system was applied to focus on the pertinent questions to include in the questionnaires in relation with the specific case study. Data belonging to the same sub-category were sometimes grouped in a unique question when related to the same information, in order to shorten at the maximum the length of the questionnaire *(*between 48 and 55 questions*)*. A follow-up was conducted with the focal points by one person (standardisation), to ensure the correct understanding of all questionnaires items and to assess the potential difficulties encountered. For each question, the existence of data (availability) and their access were investigated. When data were available and easily accessible, it was asked to focal points to specify the data resource.

2.2.2. Assessment of data availability and accessibility: indirect survey

To be complete in assessing the availability and accessibility of data, web searches were performed. These web searches focused on data that were previously identified as specific to each case study (baselines of the elaboration of questionnaires) in order to create a list of web sites and data accessible via the web. In order to ensure the standardization of the process (compilation of data resources), a 10 minutes-web search was fixed per data set. Two types of searches were performed:

1) Based on specific keywords related to data needed, identified as specific to each case study (conventional search engines such as Google, Yahoo, etc. and scientific engines such as PubMed, Web of knowledge, Web of Science) ; the list of search engines and terms used was detailed in the annex WP2.

2) Search starting from international and national (participating MSs) websites: international (FAOSTAT, EUROSTAT, OIE, etc) and national (Ministries of Agriculture, National Institutes of Statistics, etc).

A distinction was made between:

1) Data in relation with general information on the disease: most data linked to the natural history of the disease and that are worth in all contexts and for all four MSs (e.g. incubation period, duration of clinical signs, etc.)

2) National/local data: data related to the national/regional context(s) of the MS (e.g. existence of a surveillance system for a specific disease)

All resources generated were subjected to relevance criteria before being included in the review: accuracy (resources including the specific terms of search), recency (from the year 2000 and after) and reliability (originating from recognized (inter)national organisations and other scientific resources).

2.2.3. Surrogate data

Surrogate data are data used to indirectly estimate other non available data. Surrogate data could be an alternative to the lack of available data in certain contexts and types of risk question. An ideal surrogate data should be available, easily accessible and easily monitored. Therefore we tried to propose and discuss several data, potentially useful as surrogate.

2.3. Results

2.3.1. Direct surveys

2.3.1.1. Workshop

The risk questions for WP4 case studies framework, as suggested by the workshop experts, were:

- The factors able to reduce the risk of introduction and spread/establishment of VEE in Europe.
- What is the effectiveness of intervention measures at reducing the prevalence of PRRS in France?
- The risk factors of the EM spill over from animals to the human population
- The added values of diagnosing EG infection during meat inspection

2.3.1.2. Questionnaires

Nine questionnaires were received by July 31st. The Dutch expert on echinococcosis finally declined the participation, thus both questionnaires on *Echinococcus* were not filled in for NL. The last questionnaire (PRRS – FR) was received far after the deadline and thus was not considered for the analysis of responses.

A first analysis of the filled questionnaires highlighted similarities in the type of data mentioned as being not available by the MSs. The status of the disease in the EU logically influences its management in the different MSs. For example, few data are available on VEE in the three MSs, as being exotic to date and no specific disease surveillance has been implemented. Nevertheless, France implemented a surveillance network for equine diseases. PRRS is considered as endemic by two MSs (BE and NL) and in one region of FR (Brittany) but is not mandatory, so there is a lack of data in relation with surveillance as well.

It is interesting to notice similar trends between MSs, for all participating MSs (except slight differences). When data are available, their accessibility is not always easy, even for sanitary authorities and people involved in animal health such as the contact points in charge of filling the questionnaires. The problem of legitimacy of access to data should be underlined. In the questionnaires, data were specified as being accessible for the persons in charge of filling them that is

to say for professionals working for national sanitary authorities. Thus, most of them are probably not accessible to scientists. For example, data from NRL are available and accessible for sanitary authorities, but not for scientists for a question of data privacy. The availability and accessibility of data, as specified in questionnaires by MS's contact points, followed a similar trend for VEE and PRRS. Nevertheless, regarding *Echinococcus*, more data were mentioned as being available in BE.

2.3.2. Indirect survey: web searches

Web searches were performed between the 25th of June and the 8th of August, 2011. In order to ensure the standardization of the process (compilation of data resources), we fixed a10 minutes web search per data set. At least one data resource was gathered per data needed. The results were structured the results in several excel files related to the four MSs and the four case studies, thus obtaining sheets for data on disease general information, data on MSs general information [MS data] (climate, demography of animal populations, production systems, etc.) and data on MS disease specific situation (surveillance system, etc.).

Several types of data resources were included in the overview: books, courses, factsheets, guidelines, legislation, original articles, proceedings, reports, reviews and on-line data. Books, original articles and factsheets were the three most important contributors for Disease General Information (appendix D). Most MS data were found on websites (appendix D). For the case studies in the MS, websites and reports were the two major data resources (appendix D).

The status of the disease selected as case study greatly defines the availability of data. For example, few data were available on VEE, an exotic disease in EU. The fact that a disease is endemic and that no efficient control programme is implemented (or mandatory) could also go with the lack of data.

2.3.3. Data availability, accessibility and form: scoring system

In order to compare the results between MSs, a scoring system was used with reference to data **availability**, **accessibility** and **form** (the latter based on the opinion of three experts in data management with the view of applying the most appropriate classification in relation with data management) (see glossary). Figure 10 details the coding system process.

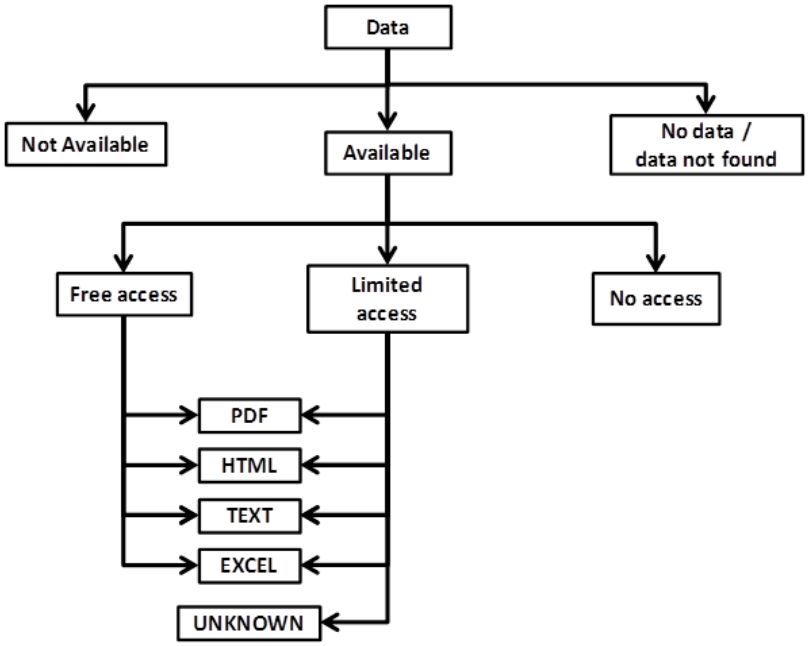


Figure 10: Assessment of the availability, accessibility and form of data

The estimations of data availability, accessibility and form are presented in appendix E (in terms of number of data). Regarding MS's data, between 25 and 35% of needed data did not exist or were not found within the limited time allowed to the web search (10 minutes). Depending of the MS considered, such range increases for VEE (16-85%), PRRS (15-76%), E. multilocularis (24-96%) and E. granulosus (30-58%). The difference of data availability was either in relation with the status of the disease targeted or with the different categories of data.

Once more, this approach reflected the parallelism existing between the status of the disease in a MS and the availability of its related data. Most of the time, when data were mentioned as being available, they also were accessible. Main forms of data resources were generally PDF and HTML files, as shown in table 7, although raw tabulated data are said to be more appropriate for risk assessment.

Table 7: Summary of availability, accessibility and form of data (all excel files considered – Disease General Information, MS_data, MS_disease)

Parameter	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	0 = no data or data not found	63	8	182	272	664	58	1,247
	1 = some data found	41	41	66	222	155	22	547
	2 = all data found	295	176	417	352	527	131	1,898
Accessibility	1 = limited access (restricted/charge)	37	22	92	60	67	10	288
	2 = free access	299	195	391	514	615	143	2,157
Form	0 = unknown	2	16	55	35	52	1	161
	1 = PDF	259	122	191	243	425	138	1,378
	2 = HTML	61	42	118	246	197	14	678
	3 = TEXT	14	14	30	48	8	0	114
	4 = EXCEL	0	23	89	2	0	0	114

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical epidemiology; PC = prevention and control; S = surveillance; PH = public health

The four excel files linked to data of each MS (MS data and MS diseases data) were the starting point of another approach which focused on the scores of availability, accessibility and form. For each of the five categories of data needed (DGI, DE, AE, PC, S and PH), the mean score of availability, accessibility were calculated, on the basis of the scores obtained for the N data found in each category. This process was repeated for MS data and for MS disease data (BE_PRRS, etc.). The mean scores could then be compared between the four MSs, both for MS data and for MS disease data. This approach was meant to identify any potential higher/lower score in one or another MS, for a same category/same disease data. An illustration of this approach is provided in appendix F, focusing on MS data. Very small variations were observed between MS. The lowest mean availability scores were recorded for prevention and control data, for all MSs. On the other hand, accessibility scores were very high in all MSs (with small SDs). NL presented the highest form score for the categories "Disease General Information" and "Descriptive Epidemiology", while the lowest scores were recorded for the "Prevention and Control" category (appendix F).

Such approach could also be performed for MS data related to a specific case study.

2.3.4. Overview of data resources

The overview included all resources gathered from the web searches, the experts' workshop and the questionnaires. A total of 2,445 sites and references composed the final overview. The same data resource was often reported for several data needed, thus reducing the total number of different data resources to 471.

To illustrate the overview of data resources, an example of a PHP/MySQL web application program was created (appendix G).

2.3.4.1. Possible mechanisms to collect surrogate data

Several data potentially useable as surrogate were proposed and were investigated during WP 4 (table 8).

Table 8: Some surrogates in case of absence of direct information on a specific disease

Category of data	Subcat. level 1	Subcat. level 2	Subcat. level 3	Data needed	Surrogate	Possible application in case studies
Descriptive Epidemiology	Morbidity			Morbidity rate	N veterinary consultations	VEE, PRRS
					Milk production	EG
					Milk somatic cell count	EG
					Trends in abortions	PRRS
					Sales of veterinary medicines	VEE, PRRS, EM, EG
	Mortality			Mortality rate	Data of rendering plants	VEE, PRRS, EG
	Spatio-temporal distribution			Annual incidence	N veterinary consultations	VEE, PRRS
					Sales of veterinary medicines	VEE, PRRS, EM, EG
					Trends in milk production	EG
					Trends in milk somatic cell counts	EG
					Trends in abortions	PRRS
					N samples sent to laboratories (clinical suspicion of the disease)	VEE, PRRS, EM, EG
	Demography of hosts			Populations of dogs	Sales of dog feed	EM, EG
Analytical Epidemiology	Factors of disease introduction	Importations /entries	Animals	N entries of domestic dogs	Sales of dog feed	EM, EG
			By-products	N entries of animal by-products	N commercial flights	VEE
			Humans	Human entries (arrivals) from areas at risk	N non-commercial flights	VEE
Surveillance	Surveillance network	Evaluation			N *pro justicia* involving actors of surveillance	VEE, PRRS, EM, EG
Public Health	Disease in humans	Spatio-temporal distribution		Incidence in humans	Sales of human medicines	VEE, EM, EG

3. WP3 - Facts and metadata model

In order to distinguish and clarify the different concepts used in the model, we used the terms **fact** and **metadata** to categorize AH experts' **data** needs, as identified in WP1 (see glossary). Thus, we named:

- **facts**, the values, also called datum, listed by the dataset for each of the variables (e.g. prevalence, mortality of a disease in a given region);
- **metadata,** the mean to describe the facts (e.g. how the prevalence was assessed, meaning the characteristics of the sampling frame, the case definition and the characteristics of the laboratory test, in terms of its sensitivity and specificity). They are also often presented in datasets;
- **data,** the fact (datum) and the metadata (datum about the datum) (WP1);
- **dataset**, a collection of data, usually presented in tabular form, where each column represents a particular variable and each row corresponds to its observations.

3.1. General concepts and methods

To ensure correct and proper use and interpretation of the datum, named from now facts, all users and owners of facts should have a common understanding of the meaning or semantics of the facts. To achieve this common understanding, a number of characteristics, or attributes of the facts have to be defined, also known as metadata. The metadata system enhances understanding of any given fact item within system – by managing the documentation of related concepts and definitions, its resource, the history of its values and the methodology used in its collection (e.g. design of surveys). Metadata is often defined as facts about facts. It is "structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage an information resource". Metadata should be structured and comprehensive. The development and construction of a metadata repository and its full integration with statistical facts are therefore upcoming key components to develop good metadata standards.
A metadata repository has its own facts' model. Indeed, a typical facts' model for database table contains metadata's fields like name of the table, location of the table, facts of table creation, systems accessing the table etc... We are proposing here to define a metadata model related to the identified AH experts' facts needs (WP1) and its priorisation. Depending on the context and the risk question, a fact can become a metadata and *vice versa*, what was considered here.

Facts and metadata, initially mixed in WP1's list of AH experts' data needs, were differentiated and related to each other, thus defining the model meta-dataset. The model was designed to provide experts' harmonised semantics to annotate existing facts as well as facts to be collected in the future. It gives experts clues on how to validate and assess the quality of available facts. To make it even more operational, we proposed facts **technical cards**, which could be used as an EFSA's form of MSs' facts' collection to prepare future AH opinions (annex WP3).
The model was as well constructed in order to facilitate the understanding and the access to different data resources, either for metadata or facts. Information from different data resources were integrated within the metadata model enabling a later retrieval of the underlying information items. As a consequence, storing facts identified or collected by EFSA-AHAW staffs and working group experts, would be available to AHAW experts' panel if relevant for new AHAW panel tasks.

In this section, the facts, the metadata related to the facts' quality assessment and the data resources (WP2) were differentiated and detailed. The resulting model aimed, through the technical cards, at linking a given risk question (WP1) to its needed facts along with its associated metadata. It was build, based on the approved Dublin Core Metadata Element Set (DCMES) (Dublin Core Metadata Initiative, 2010), as well as on existing EFSA and international data standards (appendix H).

To illustrate the relations that might exist between several metadata and a given fact, we used the Unified Modeling Language (UML). It allowed the specification, the mapping, the construction and the documentation of the different AH data model's components. The different metadata relationships created in the model (Sparks G. 2011) could either be (1) a "generalization relation" ("is a" relation) indicating that, between two related classes, one was the specialization form of the other (e.g. "an automobile is a type of vehicle" or the class "vector" is a specialized form of the class "organism"), (2) an "aggregation" ("has a" relation) when a class was a collection or container of other classes (e.g. A "addressee" is a part of an "email"), (3) an "association" when two classes were related (e.g. a "person" "subscribe" to a "magazine").

3.2. Metadata considered in the model for the facts validation and quality assessment

As shown in WP2, the resources of AH experts needed and used facts concerned both published facts (produced in scientific journals or scientific reports) and non published facts (collected and stored by governmental or non-governmental organizations).

Published facts are often related to scientific research activities where the needed facts are collected and analyzed with a defined protocol, which is normally described in the scientific paper or report. The outcome of a research activity is usually a summary (e.g. arithmetic mean, standard deviation, proportion, odds ratio, confidence intervals…) of the collected crude fact. It is rare that facts are tabulated and presented in the scientific paper or report. For transparency in scientific reports, such as the ones published by Food Safety authorities or agencies, crude facts are sometimes annexed or made available in an electronic format. For the purpose of risk assessment, crude facts are preferred but not necessary, unless the crude fact is summarized with the relevant statistics. As an example, the assessment of incubation period in the diseases contracted by biological causes (the duration between the entry of disease causing agent and the appearance of first sign of the disease is termed as incubation period), could be ascertained in scientific papers or reports on the values of average time incubation, minimum-maximum time of incubation, or the parameter of a stochastic model describing the variability between the observed animals. Depending of the type of risk assessment and the needed details for it, the format of the published results on the incubation period can be satisfactory or not. Because of published results' limitations risk assessors prefer the use of crude fact (i.e. in this case the individual incubation period observed for each animal) to perform properly uncertainty analysis The metadata used to validate and assess the quality of available facts, highly depend on the type of studies or resources used to produce these facts. Indeed, differences exist between the values of facts coming from experimental studies, from observational studies (epidemiological studies), from surveillance system, from animal production databases, from climatic recording systems etc…

Unpublished facts concern often management activities such as disease surveillance, prevention and intervention recording activities, animal movements, quarantine and veterinary inspection records, for which the protocols of facts' collection and gathering are not always directly available … In addition to AH facts, we included in this category of facts resource, other facts such as climate, land use, and feed production…

EFSA facts collection activity mainly concerns the second category of facts' resource (unpublished facts) because the published facts are generally collected and analyzed within the activity of the panel working groups, with the help of EFSA scientific staff or consortium, via article 36 projects. Moreover, EFSA scientific staffs recently proposed assistance on conducting systematic review. Because our retrospective analysis (WP1) pointed out the existence of different resources and mechanisms of facts collection by AHAW that were not always followed with success (lack of quantitative facts), we saw the need to build a metadata model allowing the systematisation, the structuring and the capitalization of experts facts collection. The result of it would be the insurance of the correct and proper use and interpretation of facts by all users and owners of facts, while assessing the risk related to animal diseases.

3.2.1. Metadata considered in the model for the facts extracted from published epidemiological studies

Epidemiological investigations consist of a series of type of studies whose ultimate aim is to provide relevant information for designing the control of disease:

- Descriptive epidemiological studies where the objective is to gain an understanding of the disease's characteristics in time and space, as it affects a certain population, and to identify the population involved. The methods used to describe the disease should aim to represent reality, according to the degree of accuracy and precision desired.

- Analytical epidemiological studies where the objective is to gain an understanding of the mechanisms governing the occurrence of a disease (the natural history of the disease). It consists of identify the different components of these mechanisms in order to understand how they function, allowing to explain them. Having found out what the nature of the disease agent is, it is important to know its resources, its targets (the hosts) and the modes of transmission it uses.

- Epidemiological evaluation studies where the objective is to evaluate the results of a control programme that has been implemented. Epidemiological evaluation provides the information necessary to steer the current control work, in the light of the changing situation, and for analysing the outcome of the work in relation to the cost of the control measures undertaken.

WP1 retrospective analysis pointed out the needs for AH experts to collect and validate published epidemiological facts, allowing their use with confidence while assessing the risk. The quality assessment and validation of published epidemiological studies are nowadays conducted following checklists specific to each type of epidemiological study and their objectives (EPA, 2006; EUROSTAT, 2007; STROBE, 2007). These checklists detail the metadata required for any assessment of facts' quality and proper use. Below are given some examples of facts' quality checklists that were, beside others, included in our metadata model.

- **Descriptive epidemiological study: measuring the disease frequency**
    - the way in which a case of disease is defined;
    - the definition of the unit of interest: an individual (animal) or a group of individuals (pen, herd, flock...);

- the definition of the target population;
- the definition with respect to time: incidence or prevalence;
- the definition with respect to space: the geographical area;
- the survey design including sampling, measurement and statistical analysis;
- the sampling type (simple random selection, stratified, cluster, etc...) for its adequacy to fit the purpose and the hypothesis of the study; e.g. design prevalence, homogeneity or heterogeneity of disease's geographical, demographic or during time distribution;
- the sample size in relation with its adequacy for the required level of precision and the selected sampling type;
- the characteristics of the screening test; e.g. its sensitivity, specificity (at the appropriate level, herd or individual);
- the actions made for standardisation of the test (so as to achieve the best possible repeatability and reproducibility) and the training of people responsible for making the measurements;
- the standardization of the questionnaire;
- the protocol made for facts' entry and facts' verification;
- the statistical procedures used (correct confidence intervals, standardisation of rates);
- the provisions made for preventing and analysing possible resources of bias and for avoiding any kind of error (facts' faults) as much as possible.

- **Analytical epidemiological study: measuring associations**

It is first crucial to check that a study is really analytical, which means that there is a clear defined hypothesis prior to the current study to be tested, and that the formulation of this hypothesis is valuable (is there a causal relationship, is there a relationship that can be measured?). Metadata to be collected could thus be listed as follow:

- presence of at least two groups to be compared;
- consistence of the overall approach for comparing the observed groups with the hypothesis;
- interpretation through deduction, based on the hypothesis formulated;
- characteristics of the groups selected for comparison; e.g. exposed and unexposed, cases and controls;
- are the groups comparable, both at the start and at the end of the study?

- case definition and its measurement (sensitivity and specificity);
- exposure definition and its measurement (sensitivity and specificity);
- statistical analysis: relative risk or odds ratio estimation, use of multivariate models, confidence intervals etc.;
- relationship between cause and effect (does the cause precede the effect? is the association strong enough? are there any potential confounders involved? are there any selection or measurement bias overestimating or underestimating the observed association?).

- **Evaluation epidemiological study: measuring disease control activities efficacy**

The objective of this type of evaluation is to measure the effect of a disease control activity on a population's health. In this case, what is being sought is a causal relationship between the control measure and the changes in the disease incidence or prevalence, in a given population. Thus, in a similar way to observational studies, following metadata should be collected:

- hypothesis of a statistically significant association between a control measure, A, and the frequency of the disease (A may only be one among many actions);
- time sequence involved;
- elimination of other possible resources of influence.

The different types of evaluative study are usually described as non-observational studies or experimental studies. Depending on the available facts, the following combinations are possible:

- a **'before-after'** comparison for a single group (the experimental group);
- only an **'after'** comparison based on two groups (experimental and control);
- a **'before-after'** comparison based on two groups.

These study designs can be implemented using either of two methods: (1) randomised trial: from which clear results usually emerge, because it is carried out under experimental conditions, but which is very difficult to reconcile with conditions in the field; (2) intervention study: which is more uncertain, leading to probable but not certain conclusions, is frequently carried out under practical conditions.

- The **randomised trial** takes advantage of being conducted under strictly controlled experimental conditions. It is theoretically possible to study a disease control activity by comparing a group, to which it was applied, with a control group. It is possible to control which individuals are chosen, the timing of the disease control measures, the conditions applying, etc. Furthermore, since individuals are randomly allocated either to the experimental or to the control group, the differences obtained at the end of the trial can be attributed entirely to the disease control activity. The randomised trial is the best

way of evaluating a health intervention. However, while it can be applied to animal diseases, for example to evaluate a treatment (groups of diseased animals and controls can be formed at will, while completely controlling all the parameters), it cannot be applied to large scale or complex animal health projects such as those combining vaccination, screening and slaughter of reactors, etc. In actual practice, intervention studies are the mainly used method.

- As in analytical epidemiology, the principle underlying **intervention studies** is the comparison of two situations, over time or in space, so as to be able to come to a conclusion about the impact of the disease control activity. The comparison over time consists of studying the disease's status in an area before and after a disease control activity was carried out (or at different times during the project): this is a before-after study.

The comparison in space consists of studying the disease's status in two areas, one where the disease control measure has been implemented (here) and one where it hasn't (there): this is a here-there study.

The underlying principle of before-after studies is simple. The reference point is the disease status before the project was implemented. Sufficiently precise information about it must be available for significant changes to be identified. After a given time period, which will vary according to the study, the health situation in the same area is reassessed using the selected indicators. The classical techniques should be used to determine if the values of these measures or indicators 'after' differ significantly from those 'before'. As was done for case/control and cohort studies, this should be done by checking for possible resources of error or bias (and in particular demographic changes and changes in environmental conditions). As was indicated above, finding a significant difference between the before and after situation does not mean that the whole of this difference can be attributed to the control activity carried out.

In here-there studies, the only latitude is in the choice of the two areas to be compared: an area where the disease control project is applied, and an area where it is not. Nevertheless, the choice is actually very restricted, since the two areas must be as comparable as possible. That is to say, the basic characteristics of the environment and of the farms, which can play a part in spontaneous changes in the disease status, must be as closely matched as possible; some such characteristics are: similar farm density, similar production systems (e.g. dairy or beef cows), similar average herd sizes, roughly the same type of livestock marketing, etc.

The same principles as in the previous case apply when coming to a conclusion about the difference observed. As in before-after studies, the fact that a significant difference is found between the two areas in favour of the one where the disease control activities were implemented, does not mean that the control activities alone account for the whole difference.

In fact, as has been mentioned before, one way of improving the efficiency of the control project's evaluation is to combine the two types of intervention studies that is the before-after with the here-there. A comparison of the two areas 'before' makes it possible to assess how well matched they are, and the 'after' comparison enables the effectiveness of the disease control activity to be studied.

After having given an overview of AH quality metadata to be considered at the time of collecting facts of published epidemiological studies, the same approach was used for facts extracted from

surveillance system. This analysis aimed at listing all the quality metadata to be included in the formal WP3 methodological framework for AH facts collection.

3.2.2. Metadata considered in the model for the facts extracted from surveillance system

Epidemiological surveillance systems of animal infectious diseases represent an important resource of facts for risk assessment. Different surveillance systems exist and can be characterized by:

- their type of objectives: to early detect the introduction of an exotic disease, estimate the frequency of various diseases found in a specific population to be able to establish priorities, to establish the importance of a specific disease (its incidence, prevalence, economic losses, etc.) and its spread, to assess the impact of a disease control programme, by monitoring decreases in the disease's incidence;
- the type of facts they collect: such as clinical (clinical observations) or biological (the presence of antibodies, isolation of an infectious agent);
- the way the system standardise the collected facts: for clinical facts (stating what clinical signs were observed and for how long), for meat inspection facts (the severity and stage of development for lesions found during meat inspections) and for biological facts (the used laboratory techniques, how the results are expressed, etc.);
- the frequency of the facts collection and transmission: facts collection can either be continuous (as by a veterinarian in the course of his daily routines) or can take place at intervals. In the same way, the frequency with which facts is to be sent to the central unit for processing must be set out, as well as how it is to be sent (as a printed document, electronic format);
- the way facts are validated, encoded, processed and stored;
- the ownership of the collected and produced information;
- the existence of periodic evaluation of the system by external and independent organisation;
- the existence of quality control process system in order to prevent errors and bias.

When extracting facts from any surveillance system the following question need to be answered:

- what types of facts are collected?
- who collect the facts?
- how are the facts collected?
- how are the facts transmitted?

- who process the facts?
- how are the facts stored?

In addition to the previous list of questions, a specific attention must be given to the issue of representativeness of the system. The observed population needs to be carefully described and existing constraints well documented (for example, voluntary participation or disease detection based on clinical signs by farmers), so as to be able to define precisely how these constraints will impair the extrapolation to the whole target population.

When the system is based on biological sample and laboratory analysis (test), the description of how the samples are taken, their transport and storage before the analysis is needed. The characteristics of the test including sensitivity, specificity, and limit of detection, repeatability and reproducibility must be provided.

3.3. Metadata elements considered in the model for AHAW panel data resources

Different types of data resources were identified and classified in WP2. Based on that, we enlarged the typology of the metadata model for epidemiological data resources (figure 11), taking into consideration the work done by Vieira da Silva Lopes' (Vieira da Silva Lopes, 2010) and the Dublin Core metadata standards for web resources (Dublin Core Metadata Initiative, 2010). This latter was also recognized as the standard ISO 15836:2009 (ISO, ISO 15836:2009 - Information and documentation - The Dublin *Core metadata element set*, Intec, 2009) (figure 12) (table 9).

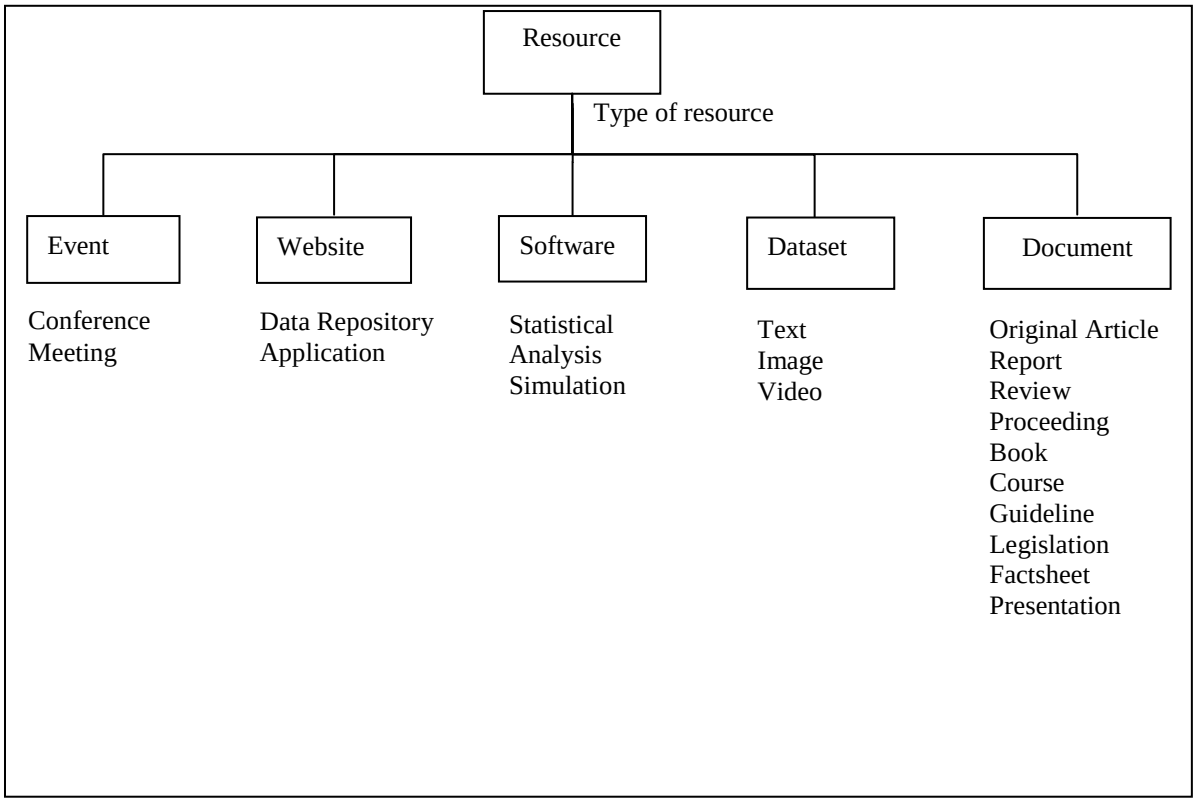


Figure 11: Types of resources managed by AHAW Panel

An "**Event**" was defined as either a conference or a meeting, a **"Website"** as either a data repository or an application, a "**Software**" as- either a statistical analysis or a simulation, a "**Dataset**" as either a text, an image or a video, and a "**Document**" as either a scientific paper, a scientific report, a proceedings, a book, a presentation, a guideline, a legislation or a factsheet. Whatever their type (e.g. document, software, dataset…), all data resources were said to share a common metadata set, named "**Resource**" (table 9).

Figure 12 presents the way AH data resources should be informed.

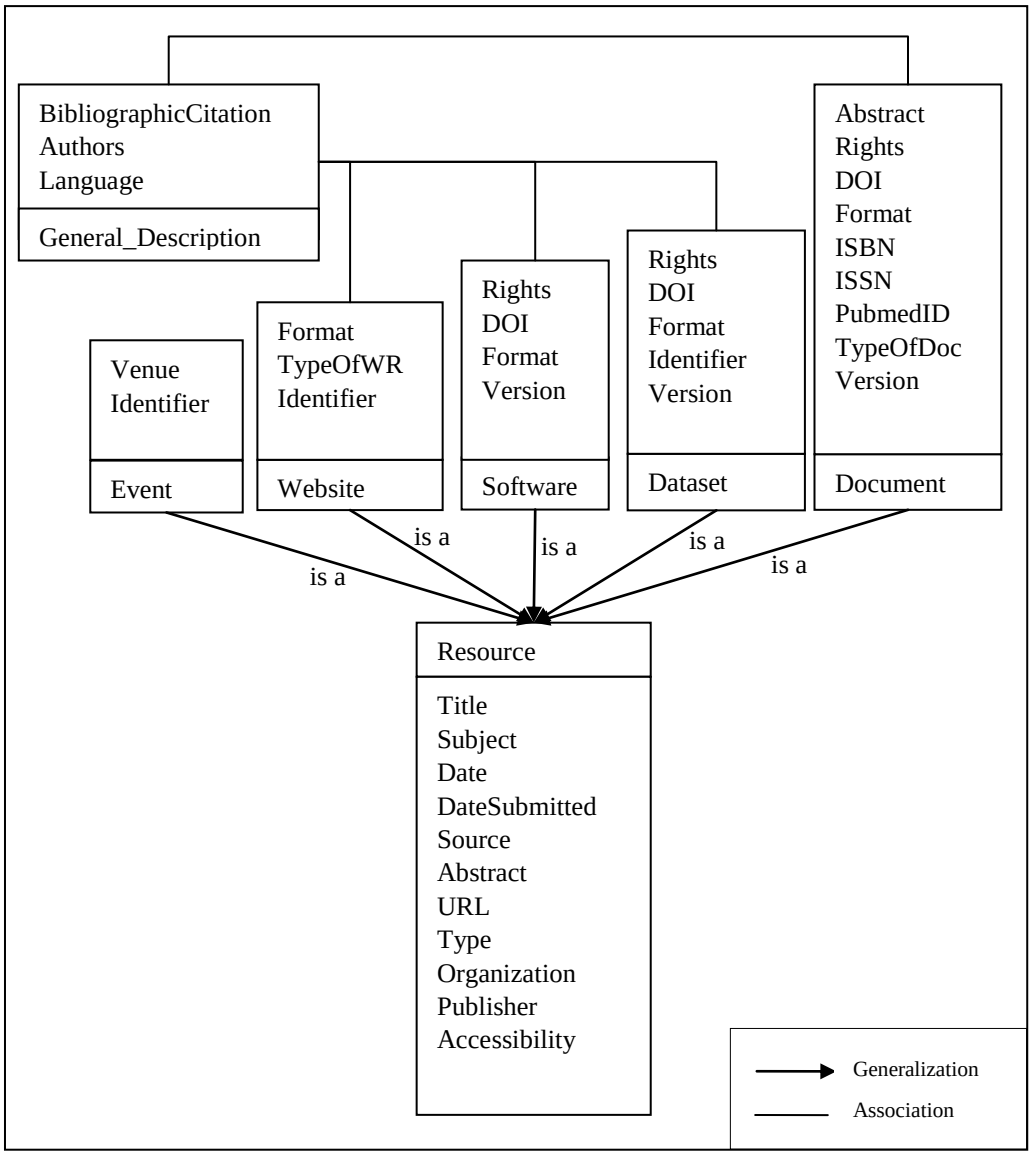


Figure 12: Metadata elements to be used to describe AHAW Panel types of resources

Table 9: Description of the DATASPEC metadata "Resource": list of the attributes

Resource	
Title	The title given to the resource
Subject	The topic of the content of the resource
Date	A date associated with an event in the life cycle of the resource. Should be the date that the dataset was created or last modified to assume the configuration in which it is submitted to the repository. Typically, the date will be associated with the creation or availability of the resource
Spatial	Spatial coverage of the resource
Temporal	Temporal characteristics of the resource
DateSubmitted	Submission date of the resource. Should be stored automatically by the program when the file is submitted and annotated, without need from user input
Source	A resource reference from which the present resource is derived (Description, Name, URL)
Abstract	A summary of the resource
Link	URL of the resource. This is applied to web resources that have a specific URL
Type	The nature or genre of the content of the resource
Organization	Set of elements to describe the organization responsible for the creation of a resource (Name, URL)
Publisher	An entity responsible for making the resource available. Example of publisher includes a person, an organization or a service (Name, Organization, URL)
Accessibility	Accessibility of the data
BibliographicCitation	A bibliographic reference for the resource (Reference [The citation of a reference for the resource], DOI [The DOI of a reference for the resource])
Authors	Name of the author(s)
Language	The language of the intellectual content of the resource
Format	The physical or digital manifestation of the resource: pdf, html, excel, text, word, database
DOI	The DOI of the resource
ISBN	The ISBN number of the resource
ISSN	The ISSN number of the resource
PubmedID	The ID of the pubmed repository
TypeOfDoc	Type of document resource (Book, Scientific paper...)
Version	Defines the resource version
TypeOfWR	Type of web resource
Venue	Specific location of an event, such an address
Rights	Information about rights held in and over the resource (Holder [A person or organization owning or managing rights over the resource], copyright, disclaimer)

3.4. Facts and associated metadata elements considered in the model

3.4.1. Review of EFSA existing standards

The consortium reviewed both EFSA and existing international data standards (e.g. OIE, FAOSTAT, EUROSTAT, USDAI, TRACE, STROBE, WHO) that could be used in the model, considering WP1 list of needed data. This latter was then completed and standards proposed by the consortium. The EFSA standards that were considered within this study were:

- the standard sample description for food and feed
- the food consumption database
- the zoonoses AMR reports and the zoonoses reports
- the food borne pathogens in foodstuffs
- the vector born disease AHAW catalogues

As for examples, we extracted from the **EFSA food and feed samples standards** the expiry date of a sample ("expiryY", "expiryM", "expiryD"), the origin type (farm or manufacture) of a feed or a food, the size of a sample lot ("lotSize"), the unit of a lot size ("lotSizeUnit"), the product's company manufacturer ("prodManuf") and the country of processing ("procCountry"). We used from the **EFSA vector born disease catalogue**, the standard "organism" considering both vector and host, as well as the different sample description attributes, such as: the altitude ("alt"), the coordinate type ("coordinates type"), the temperature ("temperature (°C)"), the humidity ("humidity"), the Coordinate Reference System (CRS), the sample tissues ("Sampling Tissue"), the sensitivity ("sensitivity") and the specificity ("specificity") of the analytical methods. We extracted from the **EFSA Zoonose reports** the sample stage attributes, detailed on: (1) Level1: Place or stage (e.g. at farm…), (2) Level2: Sample category (e.g. environmental…), (3) Level3: Sample type (e.g. dust, faeces…). Finally, we used from the **EFSA Zoonose guidance document on good practices for design of field surveys**, the standards developed on the hierarchies of the studied population: (1) the external population, meaning the total population that one would ideally like to extrapolate the results to, (2) the target population, meaning the immediate population which study results will be extrapolated to (the subject (items, animals, batches) included in the study is also derived from the target population), (3) the study population or experimental population, meaning the population of individuals (animals or groups of animals) selected to participate in the study (regardless whether or not they actually participate), and the sample size, considering the unit of the studied population, the materials used during the sampling, the material collection, the date of the collection, the storage of the material, the date of the statistical analysis, the description of the statistical analysis (e.g. the design of the analysis, the objectives of the survey, the number of the cases, the population at risk, the assumptions method, etc.).

3.4.2. List of the AHAW experts required epidemiological facts

Tables 10 to 14 present the different groups of facts by WP1 facts category with reference to WP1's list of experts' data needs. The tables include: (1) the facts extracted from the EFSA databases (mainly the EFSA Data Warehouse System (EFSA DWH, 2011)); (2) the facts extracted from international databases and proposed by the consortium.

Table 10: Description of the "disease general information" facts included in the AH model (labels named in bold)

Group of facts	Facts
Disease general information	
Human Status	Human_Epidemiology: Epidemic_Potential
Host Characteristics	Host_Receptivity: Pathogen_Resistance **(Host Pathogen Resistance)** Host_Susceptibility: Dose_Response_Effect, **(Dose response effect),** Intrinsic_Incubation_Period **(Intrinsic Incubation Period)** Immune_Response_facts **(Immune Response facts)**: Antibody_Level, Protection_Duration, Cross_Protection Disease_Infection_facts **(Disease Infection)**: Incubation_Period **(Incubation Period),**, Latency_Period **(Latency Period)**, Infection_Duration **(Infection duration)**, Infectious_Dose, Tissus_Pathogen_Load **(Tissus Pathogen Load)**, Excretion_Pathogen_Load **(Excretion Pathogen Load), Recovery_Rate (Recovery Rate)**
Vector Characteristics	Vector_Dispersion **(Vector Dispersion)**: Dispersion_Type {active:passive}, Dispersion_Description Vector_Capacity **(Vector Capacity)**: Vector_Relative_Density, Vector_Host_Bite_Frequency, Vector_Host_Blood_Index, Vector_Bites_Interval, Extrinsic_Incubation_Period, Vector_LifesSpan, Vector_Infectious_Load Vector_Competence **(Vector Competence)**: Overwintering
Chemical Resistance **(Chemical resistance)**	Chemical_Resistance: List of substance (e.g. insecticides) known for which the vector presents a resistance
Genetic Variations **(Genetic Variations)**	Evolutive_Character
Climatic facts	Climatic_facts **(Climatic Facts)**: Koppen-Geiger_Classification **(Climate Classification)**, Average_Temperature , Minimal_Temperature, Maximal_Temperature, Average_Relative_Humidity, Rainfalls, Sunlight, Wind_Speed, Prevaling_Direction, Altitude Drought **(Drought)**: Palmer_Drought_Severity_Index

Table 11: Description of the "descriptive epidemiology" facts included in the AH model (labels named in bold)

Groups of facts	Facts
Descriptive Epidemiology	
Disease Status	Disease_Status **(Disease Status)**: Disease_Status {Infected:NonInfected:Unknown} Status_at_the_end_of_the_period_facts **(Disease Status Facts)** (for Herd and/or Animal): N_status_Unknown / N_status_Not(OfficialIy)Free_lastCheck_Positive / N_status_Not(OfficialIy)Free_IastCheck_Negative / N_status_(officialIy)Free_Suspended / N_Status_Free –Type / N_Status_OfficialIy_Free
Spatio-Temporal Distribution	True_Annual_Prevalence **(Prevalence Facts)**, Observed_Prevalence **(Prevalence Facts)**, Incidence **(Incidence)**
Morbidity **(Disease's Mordidity)**	Morbidity_Rate
Mortality **(Disease's Mortality)**	Mortality_Rate
Case Fatality **(Disease's case Fatality)**	Case_Fatality_Rate
Demography facts **(Demography facts)**	Demography_facts: Density, Distribution

Table 12: Description of the "analytical epidemiology" facts included in the AH model (labels named in bold)

Groups of facts	Facts
Analytical Epidemiology	
Ecology facts **(Ecology facts)**	Ecology_facts: Land_Cover, Length_Pasture_Edges, Symbiotic_Factors
Pathogen Survival rate **(Pathogen Survival Rate)**	Pathogen_Survival_facts: Pathogen_Survival_Event_rate, Event_To_Survive (Processing, Cooked, Treatment...)
Physical Resistance **(Physical Resistance)**	Physical_Resistance: Minimal_Temperature, Maximal_Temperature, Optimal_Humidity, Optimal_pH, Optimal_Pressure, Maximal_Radiant_Flux
Population facts **(Population facts)**	Population_facts: N_Unit, N_Livestock_Per_Units
Commodity Description	Commodity_Pathogen_Load **(Commodity Pathogen Load)** Commodity_Description_facts **(Commodity Description facts)**: Related_Live_Animal_Unit, Pourcentage_Identified_Live_Animal Feed_Commodity_Description_facts **(Feed Commodity Description facts)**: N_Animal_Feed, N_Non_Animal_Feed
Preservation facts **(Preservation facts)**	Preservation_facts: Temperature, pH, Humidity
Production facts **(Production facts)**	Production_facts: Production_Volume_per_Unit
Movement	Movement_facts **(Movement facts)**: N_unit Transport_facts **(Transport facts)**: Volume_Per_Vehicle
Disease Spread	Disease_Form {epizootic:endemic:epidemic} **(Disease Form)**, Disease_Distribution **(Disease Distribution)**, Basic_Reproductive_Number_R0 **(Basic Reproductive Number R0),** Maximal_Distance_Of_Disease_Spread **(Maximal Distance of disease spread)**, Pathogen_Transmission **(Pathogen Transmission)**

Table 13: Description of the "surveillance" facts included in the AH model (labels named in bold)

Group of facts	Facts
Surveillance	
Surveillance Characteristics	Surveillance_facts **(Surveillance facts)**: N_Unit_Tested_Under_SurveilIance / N_Herd_Infected_Tested_Under_Surveillance / N_Unit_Tested_ByBulkMilk_Under_SurveilIance / N_Herd_Infected_Tested_ByBulkMilk_Under_SurveilIance, Surveillance_Sensitivity_Specificity **(Surveillance Sensitivity and Specificity)**: Sensitivity , Specificity
Outbreak **(Outbreak facts)**	Outbreak_facts: Contributory_Factor, Fact_Number_Outbreak, Fact_Cases, Facts_Death
Test Characteristics	Test_ Sensitivity_Specificity **(Test Sensitivity and Specificity)**: Specificity, Sensitivity Test_Distribution **(Test Distribution)**: Access_Delay, Distributor/Network_Actors, Test_Availability Test_Threshold **(Test Threshold)**: Criteria, Seuil, Description_Result
Test Result **(Test Result)**	Antimicrobial_Resistance_Result_facts: Fact_IZD_Value, Fact_Disc_Conc, Fact_Min_Value Measurement_facts: Fact_Res_Val (e.g. Substance_Concentration), Fact_ResVal_Rec, Fact_Res_ValUncertSD, Fact_Res_ValUnsert Subtyping: Fact_Total_Isolate, Fact_Isolates_Positive, Fact_isolates_Typed
Consumption **(Consumption facts)**	Consumption_facts: N_Consumer, Consumption_Date, Place_Of_Consumption, Fact_Amount, Meal, Food/Feed consommed
Network System **(Network System)**	Network_Communication_System: N_Awareness_Campaign, Feedback, Frequency Network_Early_Waring_System: Type_Early_Waring_System, Indicators, Website Network_Organization: Field {local, national, regional}, Figthing_Actors, N_Officier_Involved, N_Levels Commodity_Network **(Commodity Network)**: N_Intermediates_Between_Supplier_And_User, Sale_Points_List, Sales_Points_Control, Water_Supply_Type

Table 14: Description of the "prevention and control" facts included in the AH model (labels named in bold)

Groups of facts	Facts
Prevention and Control	
Biosecurity Measure **(Biosecurity Measure)**	Biosecurity _Measure: Existence_Of_Biosecurity, Biosecurity_Measure, Protocol
Control Measures **(Control Measure)**	Control_Measure: Existence_Of_Control, Control_Measure_Description
Program facts **(Program facts)**	Herd_facts: N_Herd_Under_Programme / N_Herd_Tested_Under_Programme / N_Positive_Herd / N_New_Positive_Herd / N_Herd_Depopulated / N_Positive_Herd_Depopulated / P_Herd_Covered_ind / P_Positive_Herd_ind / P_New_Positive_Herd_ind / P_Herd_Statuts_Officially_Free Animal_facts: N_Animal_Tested_Under_Programme / N_Animal_Tested / N_Animal_Tested_Individually / N_Animal_Positive / N_Animal_Positive_Slaughtered / N_Animal_Slaughtered / N_Animal_Covered_ind / P_Animal_Positive_ind Suspect_Case_Investigation_facts: N_Animal_Tested / N_Animal_ / N_Herd_Suspended / N_Animal_With_Suspecious_Lesions / N_Animal_examinated_positive
Event facts **(Event facts)**	Event_facts: Type_Of_Event (e.g. abortion), N_event, N_Isolation_For_Agent_Test, N_Event_Du_To_Agent, Duration
Inspection Disease Practices	Inspection_facts **(Inspection facts)**: N_Inspection_Points, Pourcentage_Visit_To_Purchase Quarantine_facts **(Quarantine facts)**: N_Animal_Quarantine, Quarantine_Duration_Period Slaughterhouse_facts **(Slaughterhouse facts):** N_Mandatory_Slaughters, Slaughters_Inspections_Sensitivity, N_Gross_Lesions_Inspected Inspection_Investigation_facts **(Inspection Investigation facts)**: Context (Routine, Suspicious), N_Unit_Tested, N_Unit_Positive, Corrective_Action (Quarantine, Rejected, Euthanasia, Destruction, Treatment, Slaugthered, Return_Of_Consignement), N_Unit_Corrected_Action Movement_Inspection_Point_facts **(Movement Inspection Point facts)**: N_BIP, N_Arrival_IP, N_Departure_IP
Processing Safeguard System	Process_Certification **(Process Certification)**: Good_Manufacturing_Practices_Protocol, GMP_Control, Iso_Standards, Iso_Satndard_Control Process_Treatment **(Process Treatment)**: Systematic_Treatment {Y:N}, Treatment_Type, Treatment_Pathogen_Survival_Rate, Supplement_Add {antibiotics…}

Awarness System **(Awarness System)**	Awarness_System: Awareness_Campaigns, Public_Targeted, Frequency
Program Implementation **(Program Implementation facts)**	Program_Implementation_facts: Evaluation_Type (internal/external), Indicators, Cheking_Frequency, Quality_Indicators_Of_Execution
Public Health Services **(Public Health Services)**	Public_Health_Professional_facts: N_Awareness_Practioners, Specialist_Ability, Competent_Specialists
Veterinary Services System **(Veterinary Services System)**	Veterinary_Services_Status: Level_Of_Authority, OIE_Standard_Conformity Veterinary_Services_Resource: Staff_Professional_Profile, N_Veterinarian_Involved_In_Epidemiosurveillance Veterinary_Services_Organization: Structure_Stability, Policies_Sustainability, Organization_Methods, Organization_Protocols Veterinary_Services_Legislation: Authority_Existence, Authority_Capability
Vaccination Practices **(Vaccination Practices)**	Vaccination_Practices_Status **(Vaccination Practices Status)**: Vaccination_Status {Mandatory:Factultative:Prohibed} Vaccination_Status **(Vaccination Status)**: Vaccination_Status Vaccination_Protocol **(Vaccination Protocol)**: Administration_Dose, Administration_Route, N_Dose_Regular_Protocol, N_Doses_PrimoVaccination, PrimoVaccination_Dose_Interval, PrimoVaccination_Booster_Interval, Dose_Frequency Vaccine_Conservation **(Vaccine Conservation):** Conservation_Duration, Conditions_Temperature, Percent_Dose_Really_Used, N_surplus Vaccine_Safety **(Vaccine Safety)**: Safety, Pathogen_Excretion_Risk Vaccine_Distribution **(Vaccine Distribution)**: Access_Delay, Vaccinologist, Network_Actors, Vaccine_Availability, Vaccine_Purity, Emergy_Vaccine_Bank Vaccine_Efficacy **(Vaccine Efficacy)**: Protection_Duration, Cross_Protection_Pathogen, Pathogen_Excretion_Risk Vaccination_Strategy_facts **(Vaccine Strategy)**: N_Unit_Vaccineated/Treated, N_Regions_Involved, Frequency
Treatment Practices **(Treatment Practices)**	Treatment_Status **(Treatment Status)**: Vaccination_Status Treatment_Protocol **(Treatment Protocol)**: Route_Exposure, Duration, Posology Substance_Conservation **(Substance Conservation)**: Conservation_Duration, Conditions_Temperature, Percent_Dose_Really_Used, N_surplus Substance_Safety **(Substance Safety)**: Safety, Pathogen_Excretion_Risk Treatment_Distribution **(Treatment Distribution)**: Access_Delay, Network_Actors, Substance_Availability, Substance_Purity Treatment_Strategy_facts **(Treatment Strategy)**: N_Unit_Vaccineated/Treated,

	N_Regions_Involved, Frequency
Traceability System (**Traceability System**)	Description, Identification_Method
Waste Management (**Waste Management)**	Existence of Waste Management, Protocol, Certification

Table 15 presents the model facts' categorisation. The terms "Existent", "Non Existent" and "Partially Existent" indicated the collection status of these facts' categories within EFSA; the term "Partially Existent" meant that some facts of the group were collected by EFSA and others not.

Table 15: Existing EFSA standards for AHAW experts' needed data

Category of facts	Groups of facts	E	NE	PE
Human Status	Human_Disease_Status			X
	Human_Epidemiology		X	
Host Status	Host_Receptivity, Host_Susceptibility		X	
Vector Status	Vector_Dispersion, Vector_Capacity, Vector_Competence	X		
Organism Demography	Organism_Demography		X	
GeoClimatic	Climatic_Facts			X
	Ecology_Facts		X	
Organism Control	Organism_Control		X	
Vaccination Practices	Vaccination_Status	X		
	Vaccination_Practices_Status, Vaccination_Protocol, Vaccine_Conservation, Vaccin_Safety, Vaccine_Distribution, Vaccination_Strategy		X	
Commodity	Commodity_Facts (Pathogen_Load, Related_Live_Animal_Unit…)		X	
	Feed_Commodity (N_Animal_Feed, N_Non_Animal_Feed)		X	
	Process_Treatment		X	
Preservation	Preservation		X	
Production	Production_Facts		X	
Movement	Movement_Facts, Transport_Facts		X	
Safeguards System	Process_Certification, Safeguard_Inspection_Facts, Safeguard_Quarantine_Facts, Safeguard_Slaughterhouse_Facts, Awarness_System, Program_Implementation, Public_Health_Professional_Facts, Commodity_Network,		X	
	Network Systems: Communication, Early_Waring_System, Organization			
	Veterinary_Services: Status, resource, Organization, Legislation			
Population	Population_Facts	X		
Disease Status	Disease_Status_Fact (Prevalence, Incidence, Disease_Status)			X
	Disease_Spread (Disease_Form, Disease_Distribution, R0)		X	

	Status_at_the_end_of_the_period_Facts, Herd_Facts, Animal_Facts, Surveillance_Facts, Suspect_Case_Investigation, Surrogates_Facts	X		
Outbreak	Outbreak_Facts	X		
Antimicrobial Resistance	Antimicrobial_Resistance	X		
Measurement	Measurement_Facts (e.g. Food concentration)	X		
Subtyping	Subtyping_Fact	X		
Consumption	Consumption_Facts	X		

Legend: E = Existent, NE = Non Existent, PE = Partially Existent

3.4.3. Metadata associated with AHAW experts required facts

Tables 16 to 18 describe the AH DATASPEC metadata-set. OIE, FAO, USDA, WHO and UN international standards were the major ones taken as references to complete the model and helped precising the definitions of the different attributes in a given metadata-set.

Table 16: Description of the metadata and example of associated facts or group of facts (1/3)

Metadata	Description	Example of associated categories or group of facts
Epidemiological Unit	The unit which the facts or the group of facts represent, which could be considered either infected (contaminated) or not for example.	Population, Prevalence, Disease_Status

Organism
• Animal
  ○ Host
• Vector
• Human
  ○ Individual
  ○ Professional
	Describe the organism that is studied and that the data collection represents. This metadata contains the classification of the organism and its nature (Wild, Pet or Domestic). The type of the organism considered could be precise by the sub-metadata corresponding: a vector, an animal or a human.
	An animal could be considered like a host. A human could be considered like an individual participant in a food consumption survey or a professional working in any system (e.g. veterinarian, employee).
		Population, Movement, Production, Host Status, Vector Status, Human Status, Organism Demography, Disease Status, Organism Control, Vaccination_Status, Consumption

Organism Description	The "Organism Description" metadata contains attributes which describe the life stage and the sex of an organism or a group of organism.	Population, Production, Outbreak
Organism Treatment	Describes any treatment which could be ministered to an organism.	Production, Movement
Resistance Characteristic	Describes, for a vector or a biological agent, the chemical resistance, the physical resistance and the genetic variations.	Disease_Spread, Vector Status
Vector Characteristics	This metadata contains the vector characteristics including the habitat, the activity and the cycle of the vector.	Disease_Spread
Population	The population metadata describe a group of organism, including the description of the study, the target and the external population.	Population, Movement, Disease Status
Parameter • Disease • Substance • Biological Agent	The parameter metadata describe the parameter that the outcome values represent and contains the identifier specific to each parameter. The type of the parameter is precised by the sub-metadata "Disease", "Substance" or "Biological Agent".	Disease Status, Antimicrobial Resistance, Outbreak, Vaccination Practices
Disease History	This metadata regroup all elements describing the history of the disease, including the immune response, the existent treatments, the existent vaccines and their protocol, the description of the infection, the clinical signs and the different stage of the disease.	Disease_Spread
Disease Transmission	Informing on the different pathways of a disease transmission, including the context (naturally or provoked experimentally), the mode (direct or indirect), the entry of the pathogen and the vehicle by which the transmission occurs.	Disease_Spread
Case Definition	The case definition metadata contains the set of criteria (clinical, epidemiological or biological) used or standardized to fully identify a case of a particular disease.	Prevalence, Disease_Status, Outbreak
Outbreak	Describes the type of the outbreak (primary or secondary), the circumstance of detection, the level of evidence and the resource.	Outbreak

Table 17: Description of the metadata and example of associated facts or group of facts (2/3)

Metadata	Description	Example of associated categories or group of facts
Food/Feed	Describes and identifies the food or feed item.	Consumption, Commodity, Antimicrobial Resistance, Movement
Program • Surveillance • Prevention • Control	Describes the program in which the facts will be collected. It includes, among others, the type of program (sub metadata Surveillance, Prevention or control), the frequency of testing and the scale of time.	Program Implementation, Organism Control, Disease_Status

Study
• Analytical
  ○ Observational
  ○ Experimental
• Descriptive
	The study metadata describes the type of study used for the program. Different types of study are defined in the analytical and descriptive sub-metadata.	Disease Status, Measurements Fact, Antimicrobial Resistance, Vaccination Practices, Organism Control

Data Collection Design	The Data Collection Design describe the purpose of the planned procedure used in a study including the legislation, the level in which the study is design and implemented, the disease context, the sampling strategy (based on the EUROSTAT typology) and the reporting year.	Disease Status, Measurements Fact, Antimicrobial Resistance, Vaccination Practices, Organism Control
Sample	The sample metadata describes the attributes required to fully describe the sample tested (an animal, an individual, a food or feed or an environmental elements) in laboratory or a population under a specific surveillance program.	Disease Status, Measurements Fact, Antimicrobial Resistance, Vaccination Practices, Organism Control
Sample Matrix	Describes the matrix where the sample is taken.	Prevalence, Measurements Fact, Antimicrobial Resistance
Sampling	The sampling characteristics describing the sample, including the temporal and the spatial elements.	Prevalence, Measurements Fact, Antimicrobial Resistance
Sampling Point	Describe the point in the food chain were the sample was taken.	Prevalence, Measurements Fact, Antimicrobial Resistance
Analytical Method	The analytical method identifies the method used in the laboratory (e.g. To measure the parameters in a sample) or a statistical method used.	Prevalence, Measurement fact, Outbreak
Test	Describes the characteristics of the test: the status (standard, complementary), the type of test (diagnosis, agent identification…), the sensitivity and specificity, the criteria threshold values, and the material used to perform the test.	Prevalence, Measurement fact, Outbreak
Laboratory	Contains attributes to describe the laboratory and its quality criteria while performing the analytical method.	Measurement fact, Outbreak

Table 18: Description of the metadata and example of associated facts or group of facts (3/3)

Metadata	Description	Example of associated categories or group of facts
Organisation	Describes the organization which was responsible for collecting and submitting the dataset to EFSA or which was implicated in any system or program.	All data collected need these metadata information
Geography	Contains the attributes that can be used to provide a spatial context.	
Date	Contains the attributes that can be used to provide a temporal context	
Fact	Describe the fact, like the aggregation level, the fact representation or the measurement unit…	
Commodity	Contains the attributes to describe the commodity considered, including the type of commodity (animals, animal product, animal genetic material…) and other descriptive elements.	Commodity, Movement, Production
Movement	This metadata allow a description of the movement that may be involved in a disease release, exposure or spreading, including the subject (animal, commodity or population), the type (trade, show, visit inspection, wildlife…), the scale (national, European or International) and the pathway description. The movement is composed by different step (describing in the "Movement Step" metadata).	Movement, Disease_Spread
Movement Step • Transport • Transit	The step of a movement could be a transport type (the transport sub-metadata) or a transit step (Transit sub-metadata)	Movement, Disease_Spread
Facility	The facility represents any premise implicated in a system like a farm or an industry.	Production, Movement, Commodity
Production System	Describe the system that produces an organism or a food or feed item.	Production, Population, Commodity

Safeguard System
• Biosecurity
  ○ Bioconfinement
  ○ Bioexclusion
  ○ Biocompartimentation
	Describe the different safeguard system setting out in any facility (biosecurity measures), or in a specific country (Veterinary Services, Quarantine measure…)	Production, Movement, Commodity, Disease_Spread, Safeguards System

A full description of the different metadata along with their attributes is given hereafter.

3.4.3.1. Description of the AH DATASPEC metadata: "Epidemiological Unit", "Fact", "Organization", "Geography", "Date" and "Events"

- "Epidemiological Unit"

The metadata "Epidemiological Unit" (table 19) represents a prerequisite for all epidemiological studies. It refers to the EFSA DWH attribute "SampleUnit" (RASA INTERNAL TECHNICAL Report of EFSA on Logical data model proposal for the EFSA Data Warehouse).

Table 19: Description of the attribute of the metadata "Epidemiological Unit"

Epidemiological Unit	
Unit	Specimens or facts collected; could either consider infected (contaminated) unit or not (e.g. Animal, flock, Herd, Holding, Slaughter batch, Environmental…)

An "Epidemiological Unit" means: "a group of animals with a defined epidemiological relationship that share approximately the same likelihood of exposure to a pathogen. This may be because they share a common environment (e.g. animals in a pen), or because of common management practices. Usually, this is a herd or a flock. However, an epidemiological unit may also refer to groups such as animals belonging to residents of a village, or animals sharing a communal animal handling facility. The epidemiological relationship may differ from disease to disease or even strain to strain of the pathogen" (OIE Territorial Animal Health Code, 2011).

- "Fact"

The metadata "Fact" describes the "facts" (table 20). The attribute named "Fact_Level" replaces the DWH "Parameter_Type". Other attributes directly characterizing the quantitative data of the facts, taken from the DWH System, were also included in the conceptual model (e.g. "Measurement_Units").

Table 20: Description of the attributes of the metadata "Fact"

Fact	
Fact_Level	Whether the result is individual or aggregated data
Fact_Type	Type of the data (e.g. a statistical data, a laboratory data, Quantity, Weight,)
Fact_Presentation	Graph, Histogram, Table, etc. (Frodsham A., 2007)
Result_Type	Type of the result obtained: LOD, LOQ, VAL, BIN, etc.
Expression_Result	Expression of the result: wet weight, dry weight, fat weight
Units	Units of the measurement fact (Kg, m^2)

Figure 13 illustrates the relationships existing between the different attributes of the metadata "Fact".

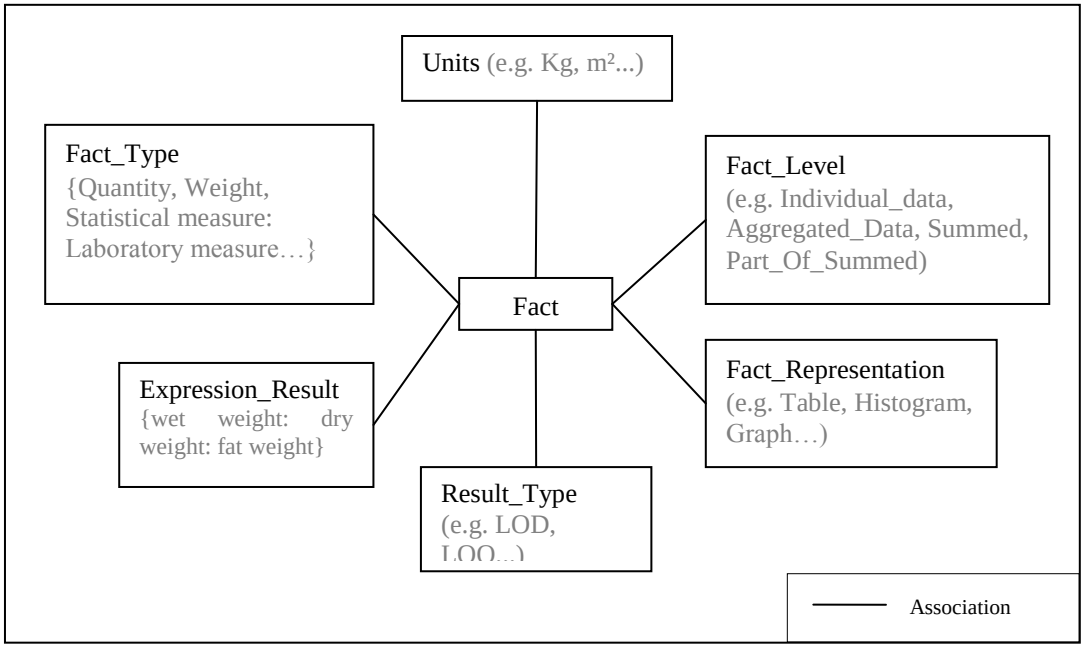


Figure 13: Class diagram of the metadata "Fact"

- "Organisation"

The metadata "Organisation" represents the entity responsible for collecting, collating and submitting the monitoring (and/or survey) facts to EFSA. It also considers any organisations (e.g. producer) that are needed to carry out a given animal risk assessment (table 21).

Table 21: Description of the attributes of the metadata "Organisation"

Organisation	
OrganisationName	Organisation providing the data or organization like producer.
OrganisationCountry	Country of the organisation
OrganisationType	Type of the organisation

- "Geography"

The metadata "Geography" provides a spatial context for the outcome value (table 22).

Table 22: Description of the attributes of the metadata "Geography"

Geography	
Zone	Zone in which the area is located
RegionL0	Country or other top classification level
RegionL1	NUTS, HASC, FAO level one region
RegionL2	NUTS, HASC, FAO level two region
RegionL3	NUTS, HASC, FAO level three region

- "Date"

The metadata "Date" provides a temporal context of the fact (table 23).

Table 23: Description of the attributes of the metadata "Date"

Date	
Date	Reporting date of the fact
Year	Year
Quarter	Financial quarter, season
Month	Month
Week	Number of the week (According to ISO-8601)
Day	Day

- "Event"

The metadata "Event" was created, representing all the events characterising the fact (table 24).

Table 24: Description of the attributes of the metadata "Event"

Event	
Event_Type	Biological (Abortion, Death, Pregnancy...) or Physical (Cooked, Treatment, Processing...)
Event_Description	Name of the event (Abortion, Death, Treatment...)

3.4.3.2. Description of the AH DATASPEC metadata: "Organism" (Animal, Vector, Human, Individual, Professional), "Organism Description", "Host" and "Population"

- "Organism"

The metadata "Organism", presented in the DWH system, was precised at different levels before being included in the AH DATASPEC model. Indeed two attributes, named "Organism_Classification" and "Organism_Type", were created (table 25).

Table 25: Description of the attributes of the metadata "Organism"

Organism	
Organism_Classification	Classification of the organism, from kingdom to species and strain if necessary.
Organism_Type	Nature of the organism: Wild, Pet, Domestic

The metadata "Organism" includes all kind of hosts constituting a population and distributed among vectors, humans or animal species. Figure 14 details the relationships that exist between the metadata "Organism" and its related components, as well considered as metadata and defines the different attributes of each metadata.

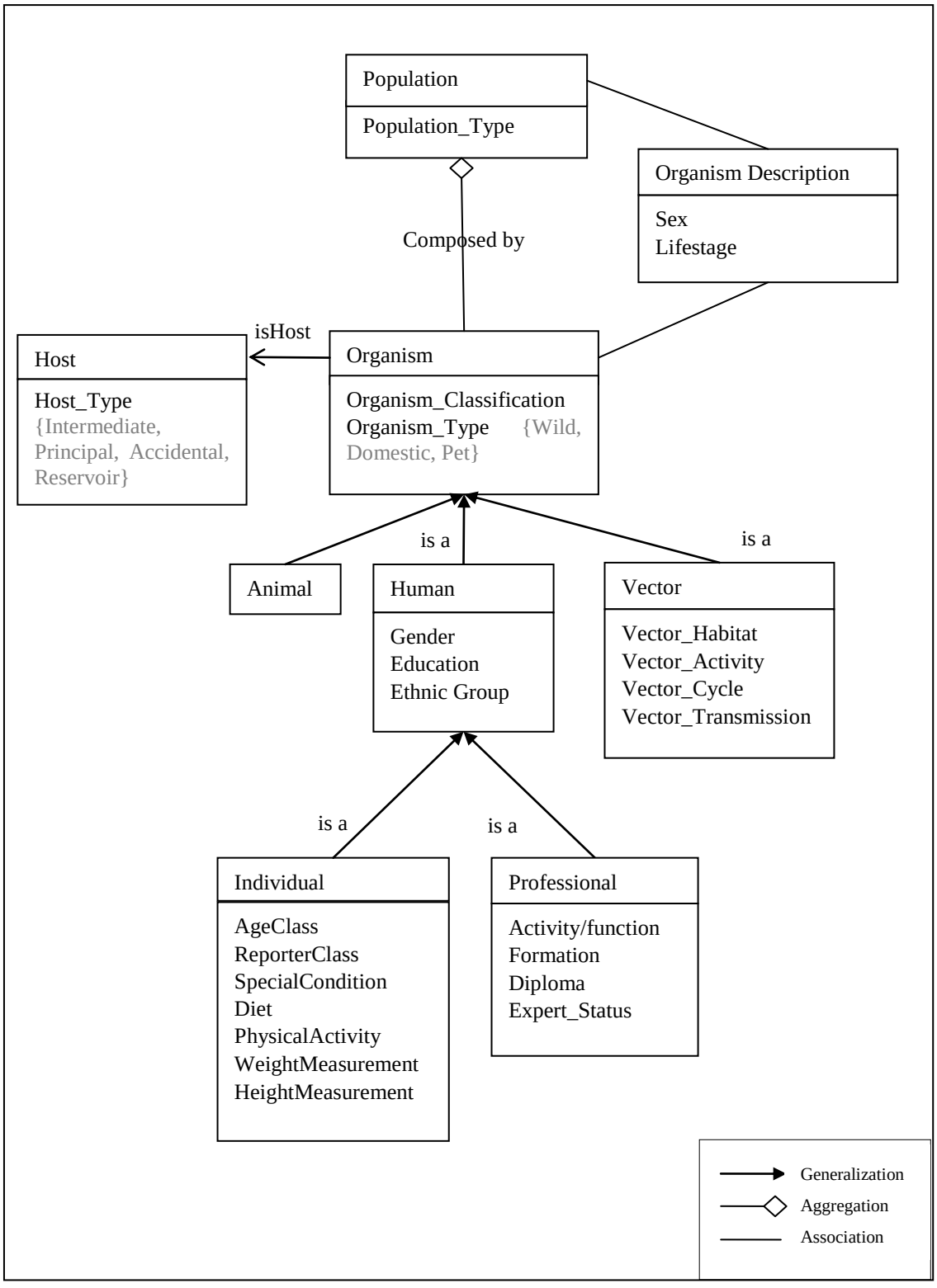


Figure 14: Class diagram of the metadata "Organism" considering its related components

The metadata "Animal" is a description of an individual animal. This could be a mammal, a bird or a bee (OIE Territorial Animal Health Code, 2011). A vector is defined as an insect or any living carrier that transports (passively or actively) an infectious agent from an infected individual to a susceptible individual, its food or its immediate surroundings (OIE Territorial Animal Health Code, 2011). The metadata "Vector" uses existent EFSA standards as well as outputs from WP1 analysis (table 26). It is linked to the metadata "Organism" for all the data concerning the vector's taxonomic classification (figure 14).

Table 26: Description of the attributes of the metadata "Vector"

Vector	
Vector_Habitat	Type, Optimal_Temperature, Relative_Humidity, Wind_Speed, Altitude
Vector_Activity	Type {indoor:outdoor}, Period{day:night}, Seasonality, Overwintering
Vector_Cycle	Duration, Lifestage
Vector_Transmission	Transmission_Type {Trans-ovarial:Trans-Stadial}

The metadata "Organism Description" (figure 14) describes the facts needed for the risk assessment that are related to the individual organisms or populations considered in the model. The attribute "Sex" is a proportion of female or male, in a case of a population.

The metadata "Individual" contains, in the DWH, characteristics that are mainly used to describe an individual participating in a food consumption survey. In order to fit with AH Panel data needs (WP1), a more general individual class, named "Human" was created and linked to the existent "Individual" one (figure 14). Indeed, the metadata "Human" is defined by attributes from an individual participating to a food consumption study or another epidemiological study, as well as attributes, detailed in the metadata "Professional" (table 27). This new metadata "Professional" allows the collection of data about professionals, responsible for Member States' prevention and surveillance programs, thus allowing a better description and evaluation of these programs.

Table 27: Description of the attributes of the metadata "Professional"

Professional	
Activity/Function	Professional activity (e.g. Veterinarians, farmers, pet owner, doctor, employee)
Formation	Continued formation followed and frequency
Diploma	Diploma obtained
Expert_Status	Expert status or not

The metadata "Population" was extracted from the EFSA Zoonose report. A "Population" means a group of units sharing common characteristics (OIE Territorial Animal Health Code, 2011). It includes the description of the target population, the external population and the population studied (table 28).

Table 28: Description of the attributes of the metadata "Population"

Population	
Population_Type	Study_Population: Population of individuals (animals or groups of animals) selected to participate in the study (regardless of whether or not they actually participate). Target_Population: Immediate population to which the study results will be extrapolated. The subject (items, animals, batches) included in the study would be derived from the target population External_Population: The total population that one would ideally like to be able to extrapolate results to. It might vary depending on the perspective of the individuals interpreting the result of the study.

3.4.3.3. Description of the AH DATASPEC metadata: "Food/Feed"

The metadata "Food/Feed" contains attributes to describe food or feed items that are sampled or surveyed. In order to precise the composition of the feed and/or the food with regards to AH risk assessments, attributes were added in the AH DATASPEC model, such as the "Type" of the feed or the food, the "Animal_Origin" and the "Processing" (table 29).

The EFSA DWH attributes "Facet_Preservation" was considered in the AH DATASPEC model as a group of facts describing the preservation conditions and renamed "Preservation".

Table 29: Description of the attributes of the metadata "Food/Feed"

Food/Feed	
Food/Feed item	Identifier of the food or feed of the classification chosen
Synonyms	Synonyms used to name the food or feed
Scientific Name	Scientific names to name the food or feed
Ingredients	List of ingredient composing the food or feed
Packaging	Packaging of the food or feed
Type	Pasture vs concentrates
Animal_Origin	Animal vs non-animal product {Y:N}
Nutritional characteristics	Nutritional characteristics and additive
Processing	Indicates if the food or feed is processed or not {Y:N}
Treatment	Indicates if the food or feed is subject to a treatment {Y:N}
Origin_Geo	Place of origin of the product
Manufacturing_Date	Date of production of the food or feed

As shown in figure 15, the metadata "Food/Feed" may be linked to the metadata "Organism" (table 25) when the food or the feed item has an "Animal_Origin".

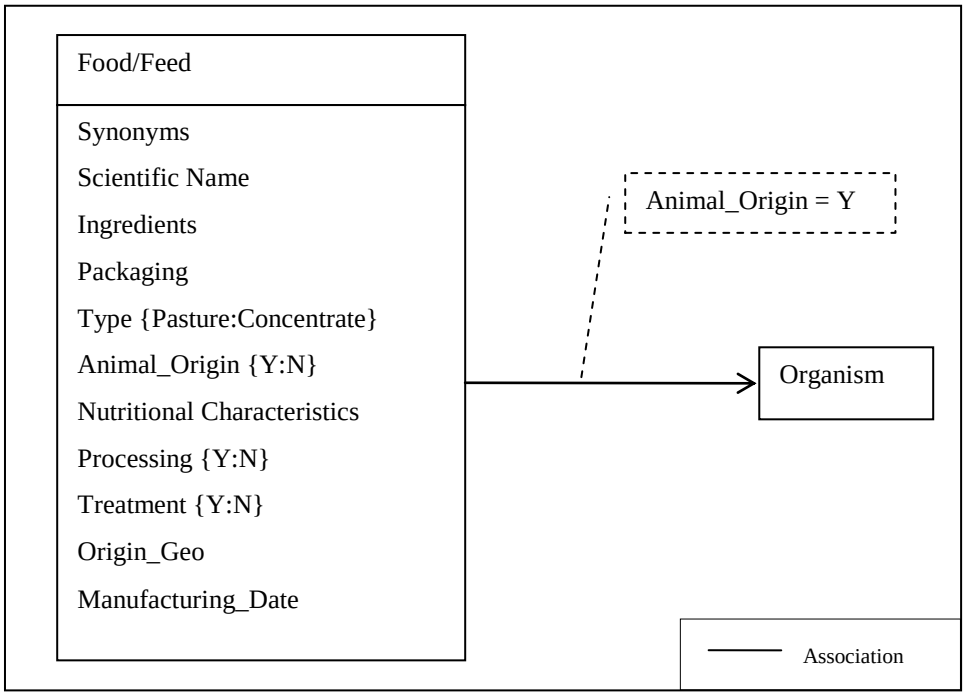


Figure 15: Class diagram of the "Food/feed" metadata considering its related metadata

3.4.3.4. Description of the AH DATASPEC metadata: “Parameter” (Disease, Biological Agent, Substance), “Case Definition”, “Disease Transmission”, “Disease Treatment, Vaccine, Immune Response and Clinical Sign” and “Outbreak”

Data about the disease, including its biological agent, its host and its natural history were also considered, while developing the AH DATASPEC model (WP1). Indeed there is the need of specifying which disease’s characteristics are to be considered at the time of assessing the risk.

Figure 16 and 17 link the needed information taken from the specific metadata sets that were created by the working group, considering: (1) the AH panel metadata needs (WP1), (2) the existing metadata standards and (3) the specific link existing between these metadata and the metadata “Disease”.

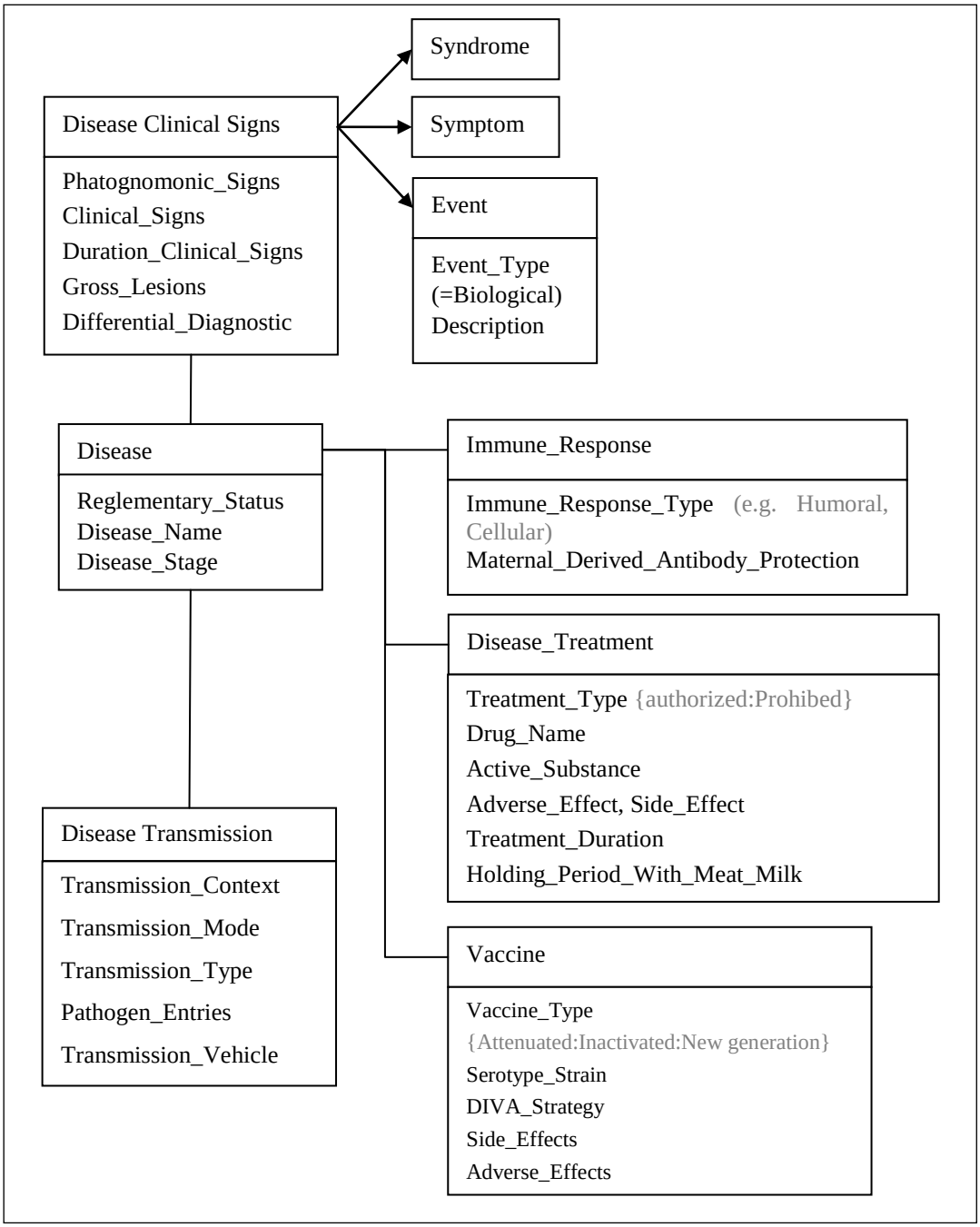


Figure 16: Class diagram of the metadata "Disease" considering its related disease natural history metadata (1/2)

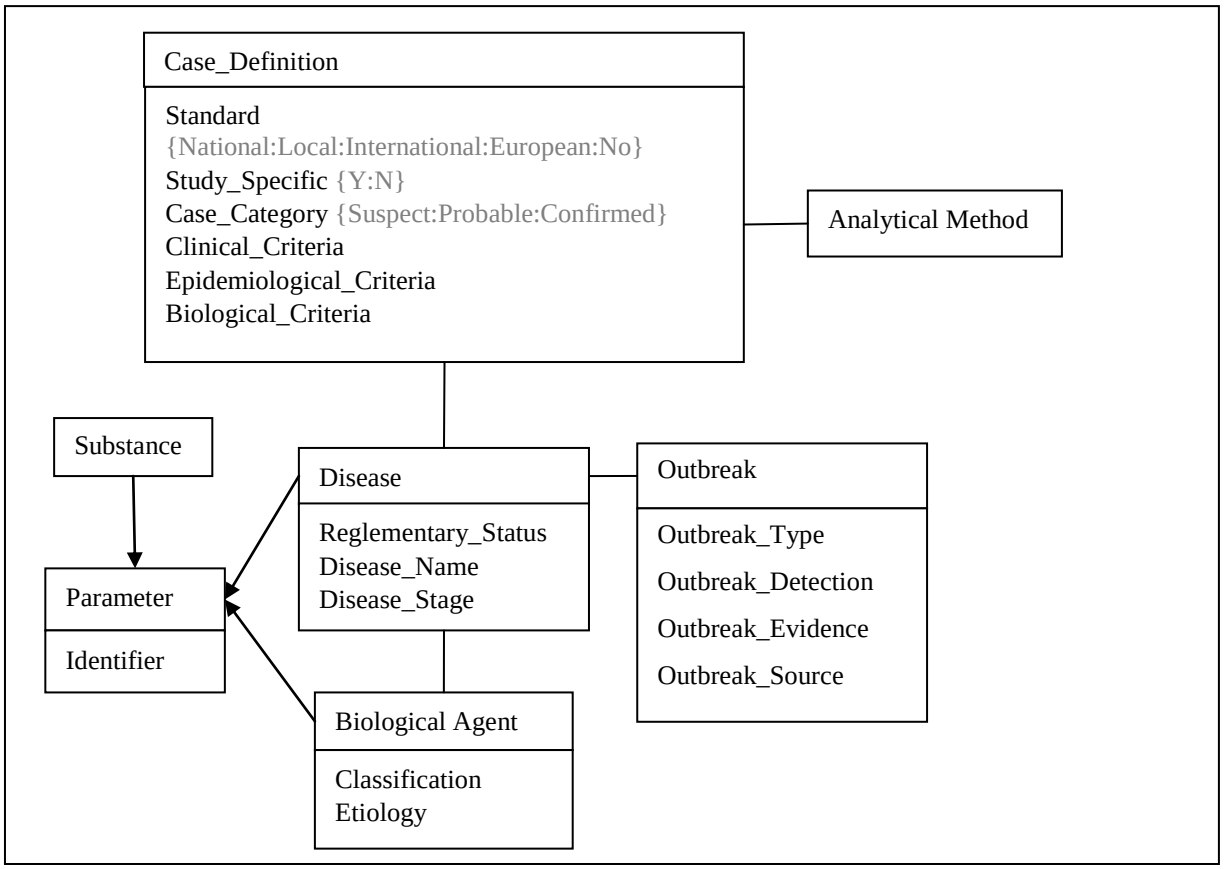


Figure 17: Class diagram of the metadata "Disease" considering its related monitoring and survey metadata (2/2)

The following tables (30 to 35) detail and define some metadata sets presented in the previous figures.

Table 30: Description of the attributes of the metadata "Disease Transmission"

Disease Transmission	
Transmission_Context	If the transmission is arrived naturally or was induced by an experimentation {Natural:Experimental}
Transmission_Mode	Whether the transmission is direct or indirect {Direct:Indirect}
Transmission_Type	If the transmission is from one individual to another in the same generation (Horizontal) or is from one parent to his offspring (Vertical or mother-to-child transmission) {Vertical, Horizontal}
Pathogen_Entries	Entries by which the transmission occurred (e.g. Insemination, Vector)
Transmission_Vehicle	Support of the transmission (e.g. Food, Blood, Air)

Any program for animal diseases' surveillance, prevention or control, relies on the disease case definition, which can vary from Member State to Member State. This information is a prerequisite for the good evaluation of the disease epidemiological situation reported by member States and consequently for any AH risk assessment. This is the reason why, a metadata "Case Definition" was created by the DATASPEC consortium (figure 17 and table 31).

Table 31: Description of the attributes of the metadata "Case Definition"

Case Definition	
Standard	Whether the case definition is a standardized definition and its level {National:Local:International:European} or not {no}
Study_Specific	Whether the case definition has been realized during a specific study {Y:N}
Case_Category	Classification of case, function of levels of certainty. {Suspect:Probable:Confirmed}
Clinical_Criteria	Clinical sign or lesions which allow confirming type case → **metadata Disease_Clinical_Signs**
Epidemiological_Criteria	Epidemiological information allowing determining the type of the case.
Biology_Criteria	Analytical method and validation criteria used which allow confirming type case → **metadata Analytical Method**

The data said to be required to value the disease's case definition of a Member State were identified taking the WHO, OIE and USDA definitions as a reference (USDA, 2006; WHO, 2011; OIE Terrestrial Animal Health Code, 2011).

An outbreak is defined as the occurrence of one or more cases in an epidemiological unit; a case being an individual animal infected by a pathogenic agent, with or without clinical signs (OIE Terrestrial Animal Health code, 2011). Table 32 lists the definition of the different terms, needed by AH Panel to correctly value Member States outbreaks reported data and consequently run the risk assessment models.

Table 32: Description of the attributes of the metadata "Outbreak"

Outbreak	
Outbreak_Type	Type of outbreak: Primary or Annex
Outbreak_Detection	Circumstances of the outbreak detection (e.g. Inspection in Slaughterhouse)
Outbreak_Evidence	Level of outbreak evidence
Outbreak_Source	Detection entity of the biological agent (e.g. Organism, Feed/food…)

3.4.3.5. Description of the AH DATASPEC metadata: "Sample", "Sample Matrix", "Sample Point" and "Sampling"

Tables 33 to 35 define the metadata "Sample", "Sample Matrix", "Sampling Point" (the point where sample was taken) and "Sampling" that were considered in the AH DATASPEC model. EFSA existing standards on sampling elements were merged in one unique metadata, named "Sampling" (table 35).

Table 33: Description of the attributes of the metadata "Sample"

Sample	
Sample	Sample tested in the laboratory or Population that the sample represents.
SampleMethod	Method for selecting or collecting sampling units
SampleSize	Size of the sample (e.g. number of sample taken in a population unit, weight of an individual sample)
SampleFrame	List of the sampling units in the sample lot or population.

Table 34: Description of the attributes of the metadata "Sample Matrix"

Sample Matrix	
Sample_Matrix	Sampled entity (e.g. Blood, Faeces, Dust, Milk, etc.)
Matrix_Type	Type of the matrix {Live animal sample, animal product, animal genetic material, feedstuffs, biological products, pathological material, plant, Environmental sample}

Table 35: Description of the "Sampling" and "Sampling Point" attributes

Sampling	
Latitude	Latitude of sampling
Longitude	Longitude of sampling
Altitude	Altitude of sampling
CoordinateType	Coordinate Type of sampling
Temperature	Temperature during sampling
CRS	Indicates the reference of coordinate system.

Sampling Point	
SamplingPointL1	Highest level to describe the sampling point
Sample point	Description of the sampling point

Figure 18 presents the relations existing between the metadata "Sample" and the others related. In order to precise the metadata "Sample Point" with regards to the definition given by the DWH (EFSA DWH, 2011), we considered that a sample could either be (1) an individual sample tested in a laboratory (e.g. an individual, a food or a feed, an animal or an environmental element such as dust) or (2) a population of any organism, thus describing the population from which the animal sampled came from or describing the population under which a specific surveillance program was settled. The metadata "Sample" was completed with the attributes "SampleSize" and a "SampleFrame" (USDA, 2006). The relation with the metadata "Epidemiological Unit" indicates which unit was considered in the sampling. Depending on the risk assessment context, it is worse noting that the metadata "Sample Point" could be linked to the metadata "Facility" and the "Geography" metadata.

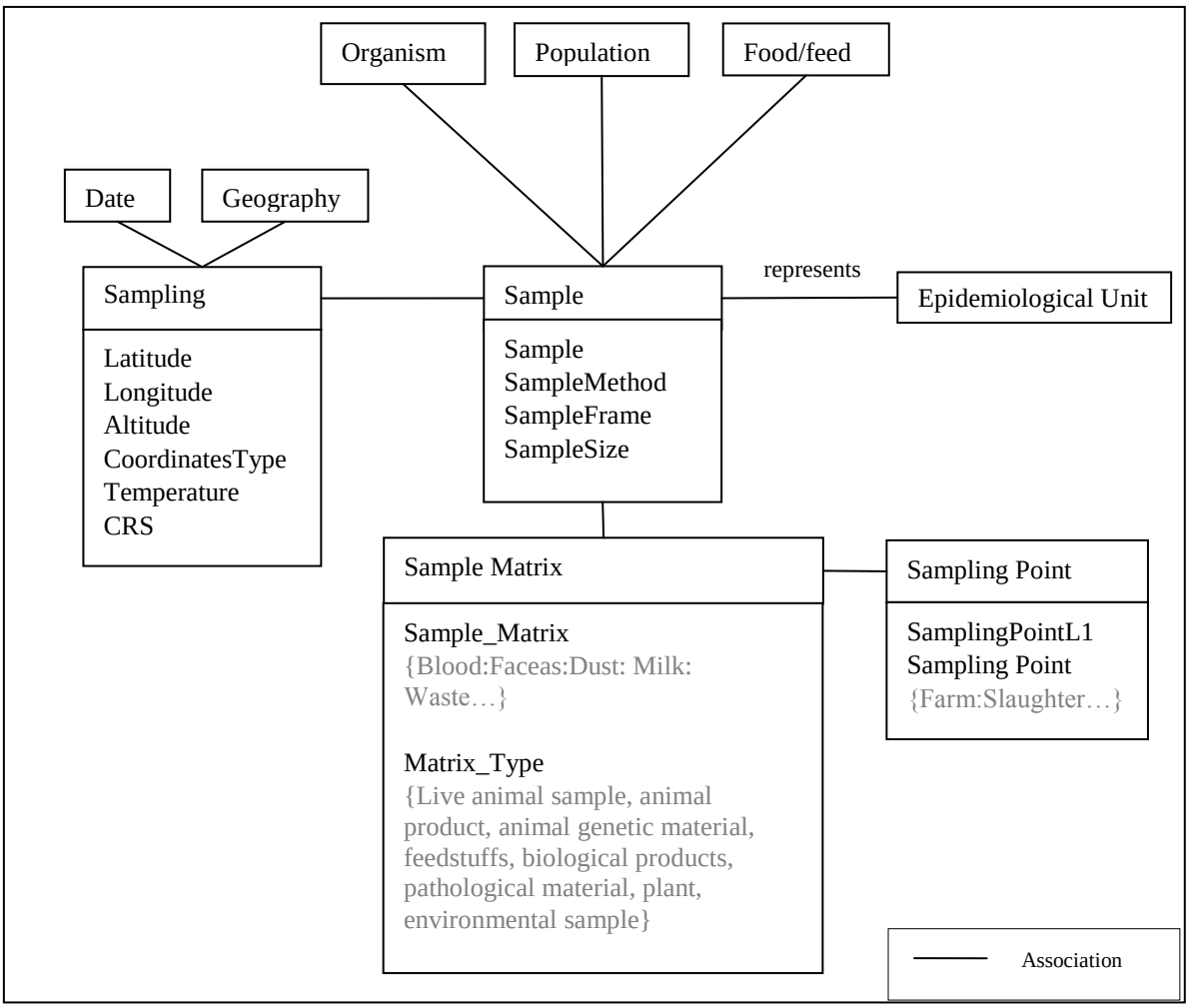


Figure 18: Class diagram of the metadata "Sample" considering its related metadata

3.4.3.6. Description of the AH DATASPEC metadata: "Analytical Method", "Test", "Analysis Date" and "Laboratory"

The metadata named "Analytical Method" is defined in the AH DATASPEC model not only by a laboratory measurement method, like in the DWH System, but also by a statistical method.

As shown in the figure 19, each "Test" is characterized by a "Sample", an "Analytical Method" and a "Date". A "Test" is defined as "a procedure used to classify a unit either as positive, negative or suspect, with respect to a disease or an infection" (OIE Terrestrial Manual 2008). This definition easily illustrates the need for AH experts to also collect, for a given test, information about the laboratory in charge of this test, justifying the creation of a metadata "Laboratory".

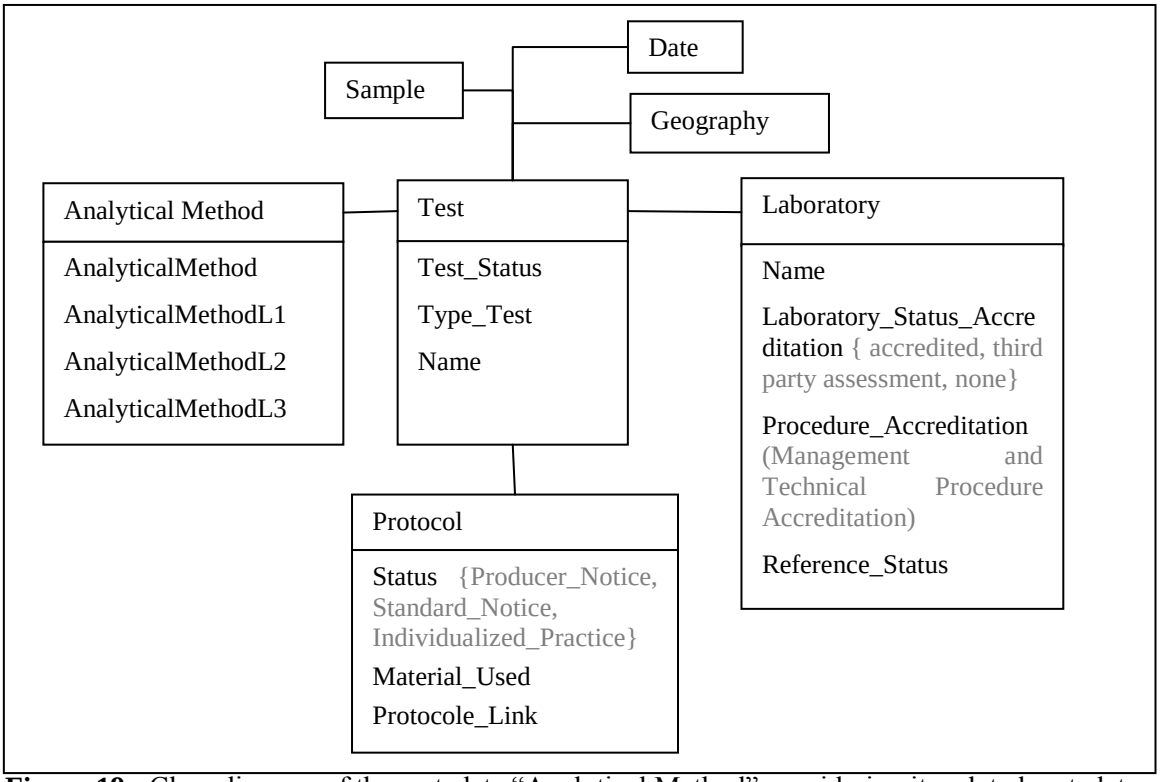


Figure 19: Class diagram of the metadata "Analytical Method" considering its related metadata

Table 36 details the definitions of the attributes constituting the metadata exposed figure 19.

Table 36: Description of the attributes of the metadata "Analytical Method" and "Test"

Analytical Method	
AnalyticalMethod	Description of the analytical method (e.g. GC-MS-MS, Student's test)
AnalyticalMethodL1	First level of the analytical method description (Biological, Chemical or Statistical)
AnalyticalMethodL2	Second level of the analytical method description (e.g. Chromatographic Test, Atomic Spectrography, Test of Association)
AnalyticalMethodL3	Third level of the analytical method description (e.g. Gas Chromatography, Liquid Chromatography)
Test	
Test_Status	Status of the diagnostic test: Standard, Complementary, Unknown
Type_Test	Type of the test considered (e.g. Diagnosis Test, Agent Identification)
Name	The commercial name of the test
Protocol	
Status	The status of the protocol applied: The protocol describes in the notice, a specific protocol given by OIE or legislation, or a specific protocol of user.
Material_Used	The description of the material used like the reference of the serum or the type of instrument used.
Protocol_Link	The complete description of each step of the protocol (e.g. Direct link toward the protocol sheet)

A laboratory "means a properly equipped institution staffed by technically competent personnel under the control of a specialist in veterinary diagnostic methods, who is responsible for the validity of the results. The Veterinary Authority approves and monitors such laboratories with regard to the diagnostic tests required for international trade" (OIE Terrestrial Manual 2008). Therefore, the AH model metadata "Laboratory" contains attributes describing the laboratory and its quality criteria, while performing the analytical method (e.g. Technical_Procedure Accreditation), allowing part of the collected data quality assessment by AH experts (table 37)

Table 37: Description of the attributes of the metadata "Laboratory"

Laboratory	
Name	The name of the laboratory
Laboratory_Status_Accreditation	Accreditation status of the laboratory performing the analytical method {accredited, third party assessment, none}
Procedure_Accreditation	Management Accreditation procedure within the laboratory, operation and effectiveness of the quality management system within the laboratory (ISO/IEC 17025, third party assessment, internally validated, not validated) / Technical requirement for the analytical method used in the laboratory, factors which determines the correctness and reliability of the tests and calibrations performed in laboratory (ISO/IEC 17025, third party assessment, internally validated, not validated)
Reference_Status	Allow to determine if this laboratory is a referenced laboratory listed by OIE.

.

3.4.3.7. Description of the AH DATASPEC metadata: “Program”, “Study” and “Data Collection Design”

The DWH’s metadata “Program” was clarified in order to distinguish attributes related to a program *sensu stricto* (e.g. inspection, surveillance or control programs) from attributes related to a study (e.g. descriptive, analytical studies); the latter being done during a program.

Figure 20 exposes the data relations included in the AH DATASPEC model. New metadata (“Program”, “Study” “Data Collection Design”) were created, using when possible attributes already existing in the DWH System (e.g. “Legislation”, “Level”).

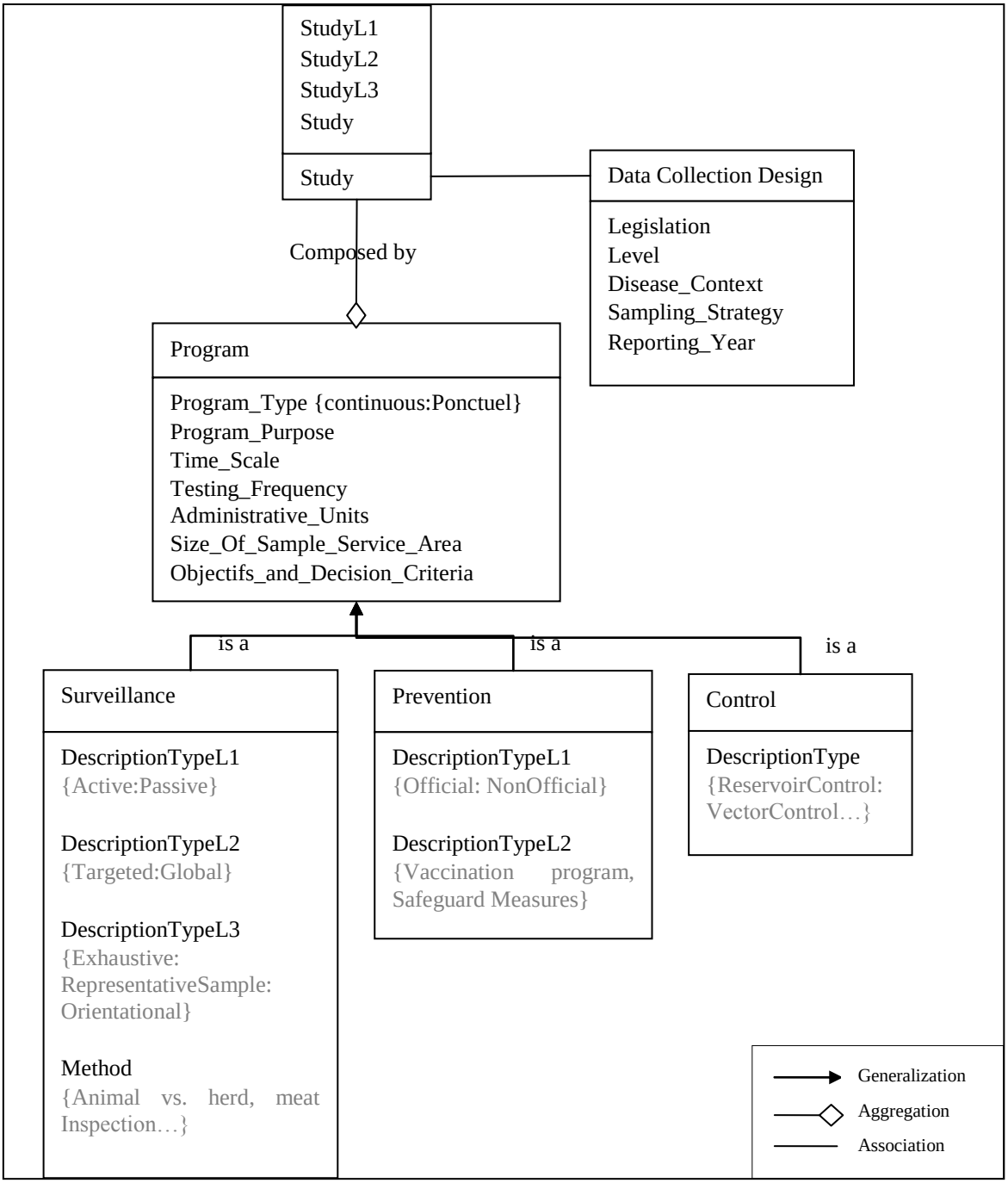


Figure 20: Class diagram of the metadata "Program" considering its related metadata

Tables 38 and 39 detail the different definitions of the metadata attributes.

Table 38: Description of the attributes of the metadata "Program"

Program	
Program_Type	Whether the program is continuous or punctual
Program_Purpose	Objectives and purposes of the program
Testing_Frequency	Frequency of the testing within the time scale
Time_Scale	Time scale of the program
Administrative_Units	Units included in the program: States, regions, zones, country, Zip Code Areas, statistical reporting units, sample grid references, neighbourhoods, parcel (USDA, 2006)
Size_Of_Sample_Service_Area	Number of reporting units (e.g., labs, clinics, slaughter plants), of geographic area per unit sampled, of eligible units per reporting unit (USDA, 2006)

Table 39: Description of the attributes of the metadata "Study" and "Data Collection Design"

Study	
StudyL1	Descriptive or Analytical study {(e.g. Analytical Study)
StudyL2	Second level of the study description with regards to StudyL1 (e.g. Experimental_Study)
StudyL3	Third level of the study description with regards to StudyL2 (e.g. With_Comparaison Randomized Controlled Trial)
Study	Study (e.g. Randomization)
Data Collection Design	
Legislation	Legislation frame of the program
Level	Level at which the program is designed and implemented (EU, National, Industry, Research)
Disease_Context	Sample infectious phase and/or disease context
Sampling_Strategy	EUROSTAT typology of sampling strategy (see appendix H)

As described table 39, different hierarchical study levels were considered while developing the AH DATASPEC methodological framework (Frodsham A., 2007): firstly, **descriptive studies** that describe AH-related phenomena, without doing any comparison between study groups (e.g. exposed versus non-exposed or treated versus not treated), thus avoiding any conclusion about associations between exposure and outcomes (Dohoo Ian, 2003). Secondly, **analytical studies** those make the comparisons between groups of study subjects. It allows investigator to make inferences about

relationships between exposures of interest (e.g. risk factors, treatments etc) and outcomes of interest (e.g. disease occurrence, productivity effects etc) (figure 21).

Depending on the study from which the data are collected, different criteria of quality are to be considered by experts. Figure 21 details the different studies that were considered in the AH DATASPEC, in order to specify the related metadata collection. This study categorization aims at facilitating the assessment of the suitable use of the collected data from a given study.

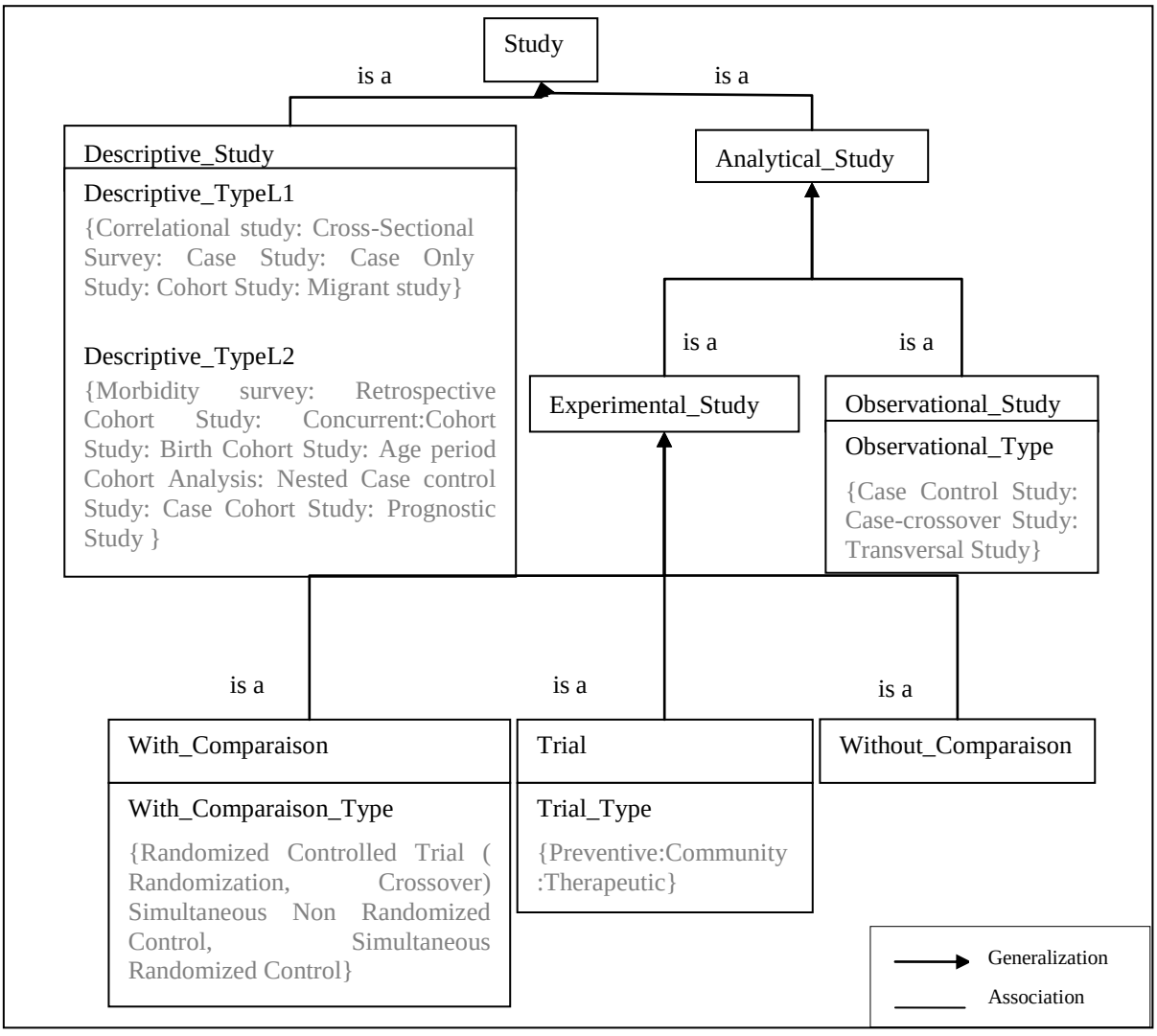


Figure 21: Class diagram of the hierarchy considered in the AH DATASPEC for the existing types of "Studies"

3.4.3.8. Description of the AH DATASPEC metadata: "Commodity", "Production System", "Facility"

The commodity refers to the risk of disease's release or exposure. A complete list possible commodities is presented in the United Nations Commodity Trade Statistics Database (UN, Commodity List), where each commodity's name and description are given. Table 40 lists the different attributes considered in the AH DATASPEC model for the metadata commodity.

Table 40: Description of the "Commodity" attributes

Commodity	
Commodity_Name	Name of the commodity (UN, Commodity List)
Commodity_Description	Description of the commodity (UN, Commodity List)
Commodity_Type	Type of the commodity: animals, animal product, animal genetic material, feedstuffs, biological products, pathological material, (OIE Terrestrial Animal Health Code 2011) and plant
Production_Chain	Type of the production chain: Food, Insemination, Pharmaceutical
Commodity_Level	Whether the commodity is a Primary Product or a by Product
Animal_Origin	Possible animal origin of the commodity {Y:N}

Due to the high diversity of the "Commodity_Type" and its related production systems, a metadata Production_System (table 41) and a metadata Facility (table 42) were also defined in the AH DATASPEC model.

Table 41: Description of the attributes of the metadata "Production System"

Production System	
ProductionSystemL1	Type of the production system (Dairy, Meat, Fight, Reproduction system, Egg, Wool, farming…)
ProductionSystemL2	Whether the production system is intensive or extensive {Intensive:Extensive}
ProductionSystemL3	Outdoor system (pasture system), Indoor System, With outdoor environment access
Housing System	Description of the housing system: condition in which the animal are conserved (Cages [Battery cages / Furnished cages], Littered floor, Compartimentation.)

Table 42: Description of the attributes of the metadata "Facility"

Facility	
Facility_Type	Activity of the facility (e.g. industry, laboratory, farm, veterinary clinic, insemination center, pet shop, riding school, border post, zoo, slaughter house)
Facility_Material	Material used

The links existing between the different metadata are presented figure 22, taking into account the fact that the metadata "Production System" describes the system that produces an organism or a Food/feed item (Sorensen J.T and all, 2006).

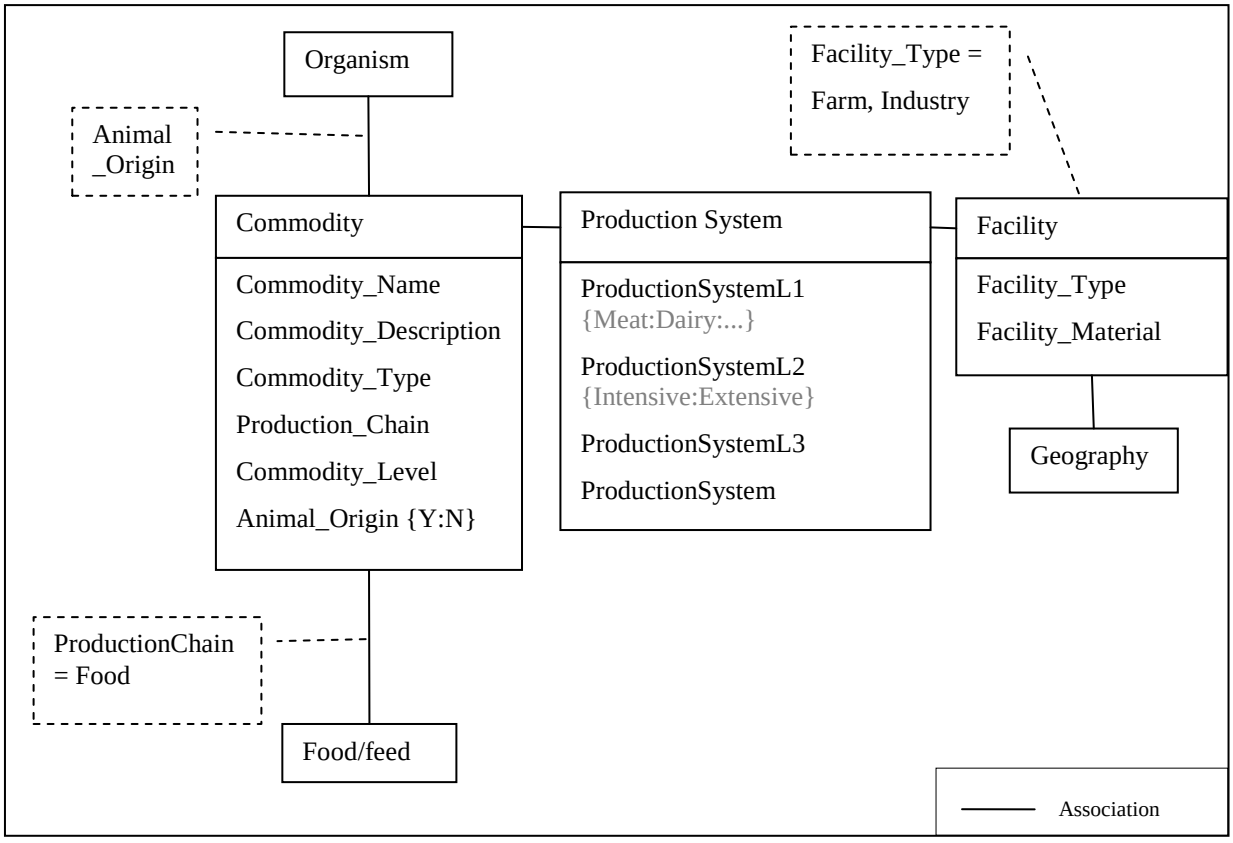


Figure 22: Class diagram of the metadata "Production System" and its related metadata

3.4.3.9. Description of the AH DATASPEC metadata: "Movement"

The metadata "Movement" describes all the different movements that may be involved in the release, exposure or spreading of a disease, such as the trade (legal or illegal), shows, visit inspections, wildlife's movement, farming (e.g. transhumance) (table 43).

Table 43: Description of the attributes of the metadata "Movement", "Movement Step", "Transport" and "Vehicle"

Movement	
Movement_TypeL1	Trade, Event, Inspection, Wild, Farming, or Tourism
Movement_TypeL2	Movement direction, geographical or temporal movement context (e.g. Importation, Exportation, Seasonal Migration, Intra_Area...)
Movement_Scale	National, European or International
Movement_Pathway	Global pathway, including departure zone, zones crossed and arrival zone
Movement_Duration	Duration of the movement.
Movement Step	
Step_Duration	Duration of the step
Transport	
Departure	The zone or the facility from which the transport start
Arrival	The final destination (Geographical area or the facility) of the transport
Crossed_Zone	The crossed zone (Geographical area or the facility) during the transport
Transit	
Place	Facility or a geographical area
Vehicle	
Identification	Matriculation number of the vehicle
Vehicle_Type	Conveyance, including train, truck, aircraft or ship (OIE Terrestrial Animal Health Code 2011)
Vehicle_Capacity	Capacity of the vehicle
Vehicle_Speed	Speed of the vehicle
Vehicle_Condition	Condition in which the vehicle is (e.g. Closed, Frozen, Chilled)
Authorization_Status	Transporter_Authorization, Certificate...

Figure 23 presents the different links considered in the AH DATAPEC model, in relation with the metadata commodity.

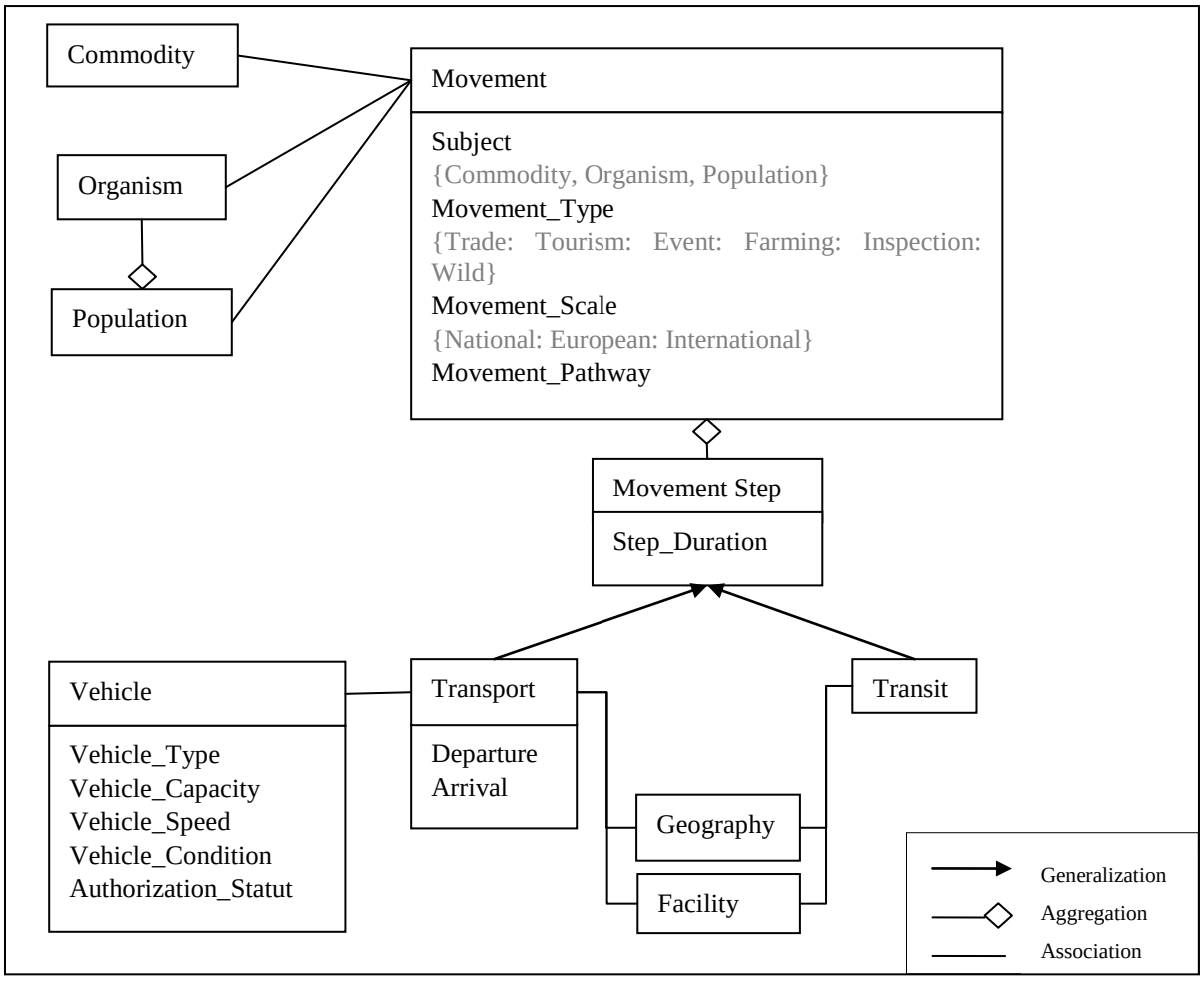


Figure 23: Class diagram of the metadata "Movement" and its related metadata

3.4.3.10. Description of the AH DATASPEC metadata: "Safeguard System"

In order for AH experts to collect information, regarding the factors influencing the disease transmission, a metadata "Safeguards System" was created (table 44). This metadata is defined by several "Safeguard_Type" for which specific descriptive elements (attributes) may exist. As an example, table 45 presents the description of one "Safeguard_Type" ("Biosecurity") and figure 24 the corresponding data-links considered in the AH DATASPEC model.

Table 44: Description of the attributes of the metadata "Safeguard System"

Safeguard System	
Safeguard_Type	Considered safeguard system (e.g. Veterinary_Services:Quarantine:Biosecurity:Inspection:Public_Health_Services)
Safeguard_Quality	Protocol quality criteria

Table 45: Description of the attributes of the metadata "Biosecurity"

Biosecurity	
Biosecurity_Level	Bio exclusion, Bio-confinement, Bio-compartimentation
Biosecurity_Type	The type of bio-security system, function of its role. (e.g. Waste Management,

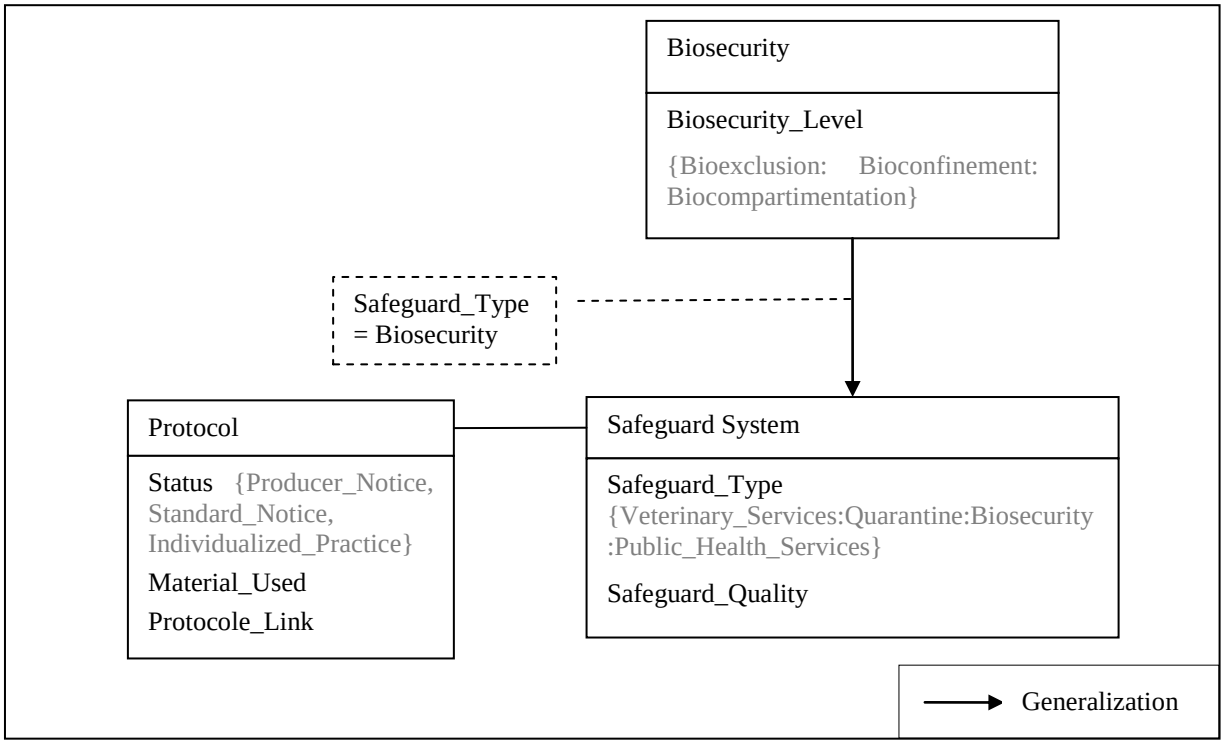


Figure 24: Class diagram of the metadata "Safeguard System" and one of its related metadata

3.4.4. Link between risk questions (WP1), facts and associated metadata (WP3)

A proposition of linkage was done between the identified AH most recurrent risk questions (WP1) and the associated needed facts (table 46). Besides, for each identified fact of our model, a technical card that detailed its identified needed metadata was created. Metadata needs were prioritised, distinguishing: (1) the **structural metadata** that describes the nature of the fact; (2) the **mandatory documentation metadata** that are necessary for primary use and understanding of the fact; (3) the **expected documentation metadata** that precise the understanding of the fact; (4) the possible data resources; and (5) the possible data surrogate. All technical cards were detailed in annex WP3 (annex WP3- appendix A).

As a consequence, the resulting DATASPEC model could link a given risk question (WP1) to its needed facts along with its associated metadata, making the data collection process even more operational for AH experts.

Table 46: Facts by type of risk question

WP1 identified risk request on "Disease Status"	
Facts	**Description**
Disease Status	Disease status either of the geographical area (presence / absence) or of the host population (infected / non infected) of interest
Demography data	Density and distribution of each disease susceptible host (animal [wildlife, domestic or pet], vector and human) of interest
Disease's Incidence	Disease incidence for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Disease's Prevalence	Disease prevalence for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Disease's mortality rate	Disease mortality for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Outbreak data	Outbreak notification in the geographical area of interest
Surveillance data	Data of the disease surveillance program in place in the geographical area of interest
Vaccination Status/ Vaccination Practices	Vaccination status of each disease susceptible host and vaccination protocols in place in the geographical area of interest
WP1 identified risk request: "Host/Pathogen characteristic"	
Facts	**Description**

Demography data	Density and distribution of each disease susceptible host (animal [wildlife, domestic or pet], vector and human) of interest
Disease's Prevalence	Disease prevalence for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Disease's Morbidity rate	Disease morbidity rate for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Disease's Mortality rate	Disease mortality rate for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Host receptivity, Infection data, Immune Response	For each susceptible host, the dose-response effect, the intrinsic incubation period, the disease immunity (e.g. antibody level/protection duration/ cross protection) and the characteristics of the infection (e.g. incubation period, latency period, infection duration, infectious dose, tissue pathogen load, excretion pathogen load)
Production data	Volume of the commodities by commodity in susceptible host production system
Vector capacity / Vector competence / Vector dispersion	Vector characteristics including its disease capacity and competence and its dispersion
Vaccination status	Vaccination status of each disease susceptible host

WP1 identified risk request: "Potential risk of spreading to susceptible population / Pathways of transmission & speed of the spread"

Facts	**Description**
Disease status	Disease status either of the geographical area (presence / absence) or of the host population (infected / non infected) of interest
Climatic data/ Ecology data	Climatic and ecological characteristics of the geographical area of interest
Control measures	Control measures applied in the geographical area of interest
Movement data	Description of the susceptible host/commodity movement that could induce the spreading of the disease into the geographical area of interest
Outbreak data	Outbreak notification in the geographical area of interest
Population data	List of the susceptible population living in the geographical area of interest
Production data	Volume of the commodities by commodity in susceptible host

	production system
Transport data	Volume of the susceptible host/commodity transported in a given vehicle
Treatment practices	The complete description of the susceptible host/commodity treatment practices in the geography area of interest, including the protocol, the substance conservation, the substance safety, the treatment distribution and the treatment strategy
Vector capacity / Vector competence / Vector dispersion	Vector characteristics including its disease capacity and competence and its dispersion
Disease Spread: Disease form / Basic reproductive number (Ro) / Pathogen Transmission / Maximal Distance Of Disease Spread,	Form (e.g. epizootic) and potential spread (e.g. basic reproductive number, transmission pathway, maximal distance of spread) of the disease
Vaccination status	Vaccination status of each disease susceptible host

WP1 identified risk request: "Risk of (re)introduction"

Facts	**Description**
Demography data	Density and distribution of each disease susceptible host (animal [wildlife, domestic or pet], vector and human) of interest
Disease Prevalence	Disease prevalence for each susceptible host (animal [wildlife, domestic or pet], vector and human) in the area of interest
Ecology data/ Climatic data	Climatic and ecological characteristics of the geographical area of interest
Resistance characteristics	Physical and chemical resistance of a given vector or biological agent
Intrinsic Incubation Period	
Movement data	Description of the susceptible host/commodity movement that could induce the spreading of the disease into the geographical area of interest
Transport data	Volume of the susceptible host/commodity transported in a given vehicle
Movement Inspection data	Control/inspection done during susceptible host/commodity movement (inspection border point, inspection before departure)

Inspection investigation data	Inspection investigations of each inspection point (e.g. number of unit tested during inspection)
Quarantine data	Quarantine practices (e.g. length)
Slaughterhouse data	Slaughterhouse practices (e.g. inspection)
Test sensibility and specificity	Specificity and sensitivity of the tests
Production data	Volume of the commodities by commodity in susceptible host production system
Specific commodity data: Pathogen load	Pathogen load observed in each commodity of interest
Program implementation data	Indicators and quality of the surveillance program
Vaccination status	Vaccination status of each disease susceptible host
Vaccination strategy	Strategy of the hosts vaccination in the geographical area of interest (e.g. host vaccinated, N_Unit_Vaccineated/Treated, N_Regions_Involved, frequency)
Biosecurity measures	Biosecurity measures applied in each facility (farm…), production system and vehicle, in the geographical area of interest
Traceability system	Traceability system of a susceptible host/ commodity

WP1 identified risk request: "Risk of establishment"

Facts	**Description**
Movement data	Description of the susceptible host/commodity movement that could induce the spreading of the disease into the geographical area of interest
Biosecurity Measures	Biosecurity measures applied in each facility (farm…), production system and vehicle, in the geographical area of interest
Awareness System	
Control Measures	Control measures applied in the geographical area of interest
Vaccination strategy	Strategy of the hosts vaccination in the geographical area of interest (e.g. host vaccinated, N_Unit_Vaccineated/Treated, N_Regions_Involved, frequency)
Veterinary Services	

Public Health Services	
Network system	
Process Treatment	
Inspection data	Number of inspection point existing in the geographical area of interest
Quarantines data	Quarantine practices (e.g. length)
Slaughterhouse data	Slaughterhouse practices (e.g. inspection)
Traceability system	Traceability system of a susceptible host/ commodity

WP1 identified risk request: "Effectiveness of surveillance measures"

Facts	**Description**
Surveillance data	Data of the disease surveillance program in place in the geographical area of interest
Event data	Event (Death, Pregnancy, Symptoms…) considered in the surveillance program
Inspection data	Number of inspection point existing in the geographical area of interest
Inspection investigation data	Inspection investigations of each inspection point (e.g. number of unit tested during inspection)
Program Implementation data	Indicators and quality of the surveillance program
Outbreak data	Outbreak notification in the geographical area of interest

WP1 identified risk request: "Effectiveness of biosecurity measures"

Facts	**Description**
Biosecurity measures	Biosecurity measures applied in each facility (farm…), production system and vehicle, in the geographical area of interest

WP1 identified risk request: "Effectiveness of prevention tools"

Facts	**Description**
Inspection data	Number of inspection point existing in the geographical area of interest

Quarantines data	Quarantine practices (e.g. length)
Slaughterhouse data	Slaughterhouse practices (e.g. inspection)
Inspection investigation data	Inspection investigations of each inspection point (e.g. number of unit tested during inspection)
Movement Inspection points	Control/inspection done during susceptible host/commodity movement (inspection border point, inspection before departure)

WP1 identified risk request: "Effectiveness of control measures"

Facts	**Description**
Disease status	Disease status either of the geographical area (presence / absence) or of the host population (infected / non infected) of interest
Control measures	Control measures applied in the geographical area of interest
Program data	Strategy and result of the susceptible host/commodity control program in the geographical area of interest
Treatment practices	The complete description of the susceptible host/commodity treatment practices in the geography area of interest, including the protocol, the substance conservation, the substance safety, the treatment distribution and the treatment strategy
Vaccination practices	Vaccination protocols in place in the geographical area of interest
Disease Inspection practices	Inspection point in place in the geographical area (inspection data, movement inspection data, quarantine data, slaughterhouse data) and results of the inspection (inspection investigation)
Pathogen Survival data	Pathogen survival rate following the control measures

WP1 identified risk request: "Diagnostic tool availability and efficiency"

Facts	**Description**
Test distribution	Distribution of the test of interest.
Test specificity and sensitivity / Test Threshold	Specificity and sensitivity of the tests and threshold used for the analysis
Test result	Specific value obtained for a given test analysis

WP1 identified risk request: "Treatment availability and efficiency"

Facts	Description
Treatment practices	The complete description of the susceptible host/commodity treatment practices in the geography area of interest, including the protocol, the substance conservation, the substance safety, the treatment distribution and the treatment strategy
Program implementation data	Indicators and quality of the surveillance program
Pathogen Survival data	Pathogen survival rate following the control measures

WP1 identified risk request: "Vaccine availability and efficiency"

Facts	Description
Vaccination Practices	Vaccination protocols in place in the geographical area of interest
Program Implementation data	Indicators and quality of the surveillance program

4. WP4 - Methodological framework

4.1. Objectives of WP4

The risk questions typology, presented in section 1, showed that a full risk assessment is not always needed. Furthermore, EFSA AHAW opinions and reports incorporate simultaneously several related questions that are addressed using different methodologies and tools. EFSA AHAW approach starts always with the production of scientific reports reviewing available scientific information on the considered animal health issues. In some circumstances, conclusions drawn from scientific review are judged sufficient to answer to the mandate terms of reference and no additional work is conducted. For more complex animal health issues, models are used based on available facts and information. Different modelling approaches are possible. The specific objective, or risk question, will guide whether a particular approach is more or less appropriate.

4.2. General methodological framework

The proposed general methodological framework (figure 25) starts first with the questions. Knowing the type of the question a list of data needs can be derived from the tables produced in WP1. The availability and accessibility of the needed data are then evaluated as in WP2. Once the data resources are identified, data specification is conducted (including attributes of the needed facts and their corresponding metadata) thanks to the tables produced in WP3. Data specifications are used to extract data, either facts or metadata, from available resources or to build form and questionnaire to collect these data directly from member states authorities, scientific or professional organisations. In parallel, in regard to the nature of the available data and the type of the risk question the more suitable methods (Statistical methods, linear probabilistic risk assessment or Mechanistic or dynamic models) are selected.

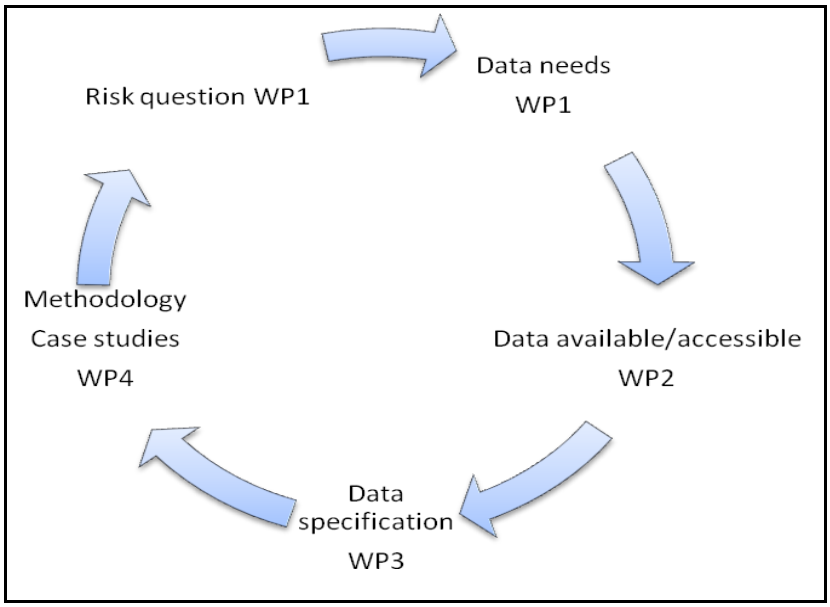


Figure 25: Integration of the fourth work packages

In its development, the general framework was thoroughly tested in the course of three case studies selected in agreement with EFSA-AHAW unit: 1) *Echinococcus granulosus*: what is the added value of meat inspection with respect to disease prevention and control in the definite host (dog) and intermediate hosts (sheep, goat, cattle and swine)? 2) *Echinococcus multilocularis*: surveillance on domestic and wild canids to demonstrate freedom; 3) Porcine Reproductive and Respiratory Syndrome: what is the effectiveness of intervention measures at reducing the prevalence of PRRS in France? Intervention measures to be considered should include: vaccination, herd management, and biosecurity; and 4) Venezuelan Equine Encephalitis: what is the risk of introduction of the virus into the EU, taking into account risk reduction measures in place?

4.3. Risk assessment methodologies and tools

The specific objective, or risk question, will guide whether a particular approach is more or less appropriate. Models may be used either for more comprehensive description of collected facts, or aims at a systematic understanding, or the prediction of future events. For simplicity, three main types of methods can be distinguished:

- **Statistical methods**: that may be for example suitable for endemic diseases, where it is needed to assess the disease frequency (Prevalence/Incidence) and to assess possible growing or declining of the disease frequency. Statistical methods are also useful to assess specific correlation or association between exposure and occurrence of diseases or the effect of an intervention measure at the point of its application.

- **Linear probabilistic risk assessment**: that may be for example appropriate to assess the probability of an exotic agent entrance.

- **Mechanistic or dynamic models**: that may be for example suitable to describe and assess the spread of a disease in a certain population.

4.3.1. Statistical methods

Statistical methods can be used for describing a collection of facts (descriptive statistics) or for drawing inferences about the process or population being studied (Inferential statistics). Statistical inferences may take the form of: answering yes/no questions about the facts (hypothesis testing), estimating numerical characteristics of the facts (estimation), describing associations within the facts (correlation) and modelling relationships within the facts (for example, using regression analysis). Inference can extend to prediction and estimation of unobserved values either in or associated with the population being studied; it can include extrapolation and interpolation of time series or spatial facts, and can also include facts mining. Statistical methods focus on the uncertainty of the observed events and do not study usually the mechanism of the occurrence of the events.

Statistical methods may be used to answer directly to the term of references or be used as an intermediate analysis to assess key elements or inputs for a risk assessment model. As an example, the first situation could be the case where the mandate is asking for the comparison between different screening tests. A statistical analysis may be conducted to compare the sensitivities of the candidate tests. The output of the analysis will be used directly to provide a conclusion about the more sensitive tests. Considering the second situation, a statistical model can be used to estimate the sensitivity of a

serological test that may be used as an input to assess the risk of introduction of a disease to a disease free herd where the purchased animal are subject to a serological test before their entrance to the herd.

Beyond basic techniques, there are more complex analytical methods used to analyze facts from experimental and observational studies. For descriptive purposes, factor analysis may be used to examine the correlations or dependency among different study variables with the intent of creating index measures for deeper analysis. Regression techniques may be used to examine how particular variables of interest affect a particular outcome and to test possible interactions.

Broad categories of methods are identified:

- methods based on the normal distribution,
- methods following transformation of facts,
- single-distribution generalized linear models (GLMs),
- parametric models based on skewed distributions outside the exponential family,
- models based on mixtures of parametric distributions,
- two (or multi)-part and Tobit models,
- survival methods,
- methods based on averaging across models, and
- non-parametric methods,

The choice between the different categories of statistical methods depends on the nature of collected facts and the objective of the statistical analysis. Description of statistical methods could be found easily in different text books and it is not necessary to develop them in this report. However, it was considered important to describe some examples using statistical methods. Numerous statistical models exist, to evaluate control strategies (e.g. Hadorn & Stärk, 2008). Several examples can be found in the scientific literature that illustrates the usefulness of modelling in evaluating and adapting strategies for animal disease control.

4.3.2. Linear Probabilistic risk assessment

It starts with listing all steps required for the risk to occur, on differentiating release, exposure and consequences (table 47). At each stage facts needed are described.

Table 47: Example of risk question: "What is the risk for the introduction of virus X through migratory birds into wild bird population in country C?"

Risk assessment steps	***Outcomes***	**Stage in the conceptual model**	**Facts**
Release	*Probability of Entrance*	Migratory bird infected Migratory bird enter C	Prevalence of infection Flyways of migratory birds
Assessment	*Probability of target population being exposed*	Infected migratory bird in resting sites Contact with local wild bird in resting sites	Flyways and resting sites Local birds density in resting sites and type of contacts
Consequences	*Magnitude of consequences and their probability*	Local wild bird infected local wild bird spread disease among the target population	Mode of transmission, host susceptibility…

Using the example of table X, the overall risk could be assessed using a simple model:

$$Risk = P(En) \times P(Ex|En) \times P(Cce|Ex) \times Cce$$

Where P(En) is the probability of entrance of the agent to the country A, P(Ex|En) is the conditional probability of exposure knowing the entrance of the agent X, P(Cce|Ex) is the conditional probability of observing consequence Cce knowing the exposure and Cce is the magnitude of the consequence.

Probability event tree may be used to accurately identify the different probabilities that need to be considered for release, exposure and consequences assessment. It describes the sequences of events necessary for disease entrance, exposure of local populations and their consequences.

The risk assessment starts with the hypothesis that disease or pathogenic agent X is present or circulating in location A. The pathogenic agent could be moved from location A to location B where the disease or the pathogenic agent X is absent (release or entrance). The Release of X may occur through the movement of animal or animal products (for example meat, milk, hide, exhaled air, vesicle fluid) carrying or containing the pathogenic agent under consideration, or through other vehicles such as biological vectors, mechanical vectors or fomites... The release assessment may consider one or several scenarios of entrance that could lead to one or several scenarios of exposure and subsequently to one or several scenarios of outbreaks and consequences (figure 26). The general linear risk formula will take the following form:

$$Risk = \sum_i P(En_i) \times P(Ex_i|En_i) \times P(Cce_i|Ex_i) \times Cce_i$$

Where the subscript "i" represents the different scenarios combinations.

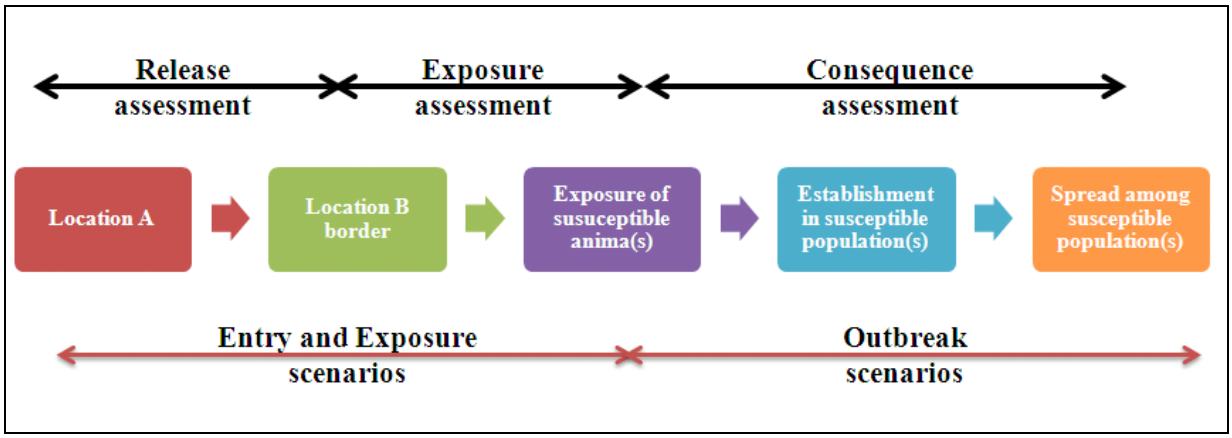


Figure 26: Generic linear risk assessment model

Figure 26 represents a simple example of release assessment model. In this example, the agent is introduced to a location B trough the importation of live animals from another location A and it is assumed that all imported animals are tested before their entrance to the location B. The model requires inputs parameters that may include the probability that location A is free, number of declared outbreaks, sensitivity and specificity of the surveillance system in order to assess the true prevalence, number of herds in location A, number of animals in infected and in non infected herds, number of imported animals and sensitivity of the test used before exporting the animals.

The input parameters may be subdivided as:

Constant inputs such as the total number of animals in location A;

Variable inputs such as the annual number of declared outbreaks (annual/seasonal variability), number of animals per herd (between herd variability), intra-herd prevalence (between herd variability and/or between season variability), number of imported animals (annual/seasonal variability);

Uncertain inputs such as the sensitivity of the test used before exporting animals (e.g. sensitivity is assessed in an experimental study including small number of infected animals).

In addition to the input parameters, the model can be build accepting a number of hypothesis or assumptions, e.g.:

- the characteristics of imported animals do not affect their risk to be infected at the time of their movement;
- an infected animal remains infectious during the period of time between its movement and arrival to the importing country and survive to the infection;
- the probability of observing false positive result does not depend on the number of days separating the onset of infection and the day of testing (no difference between newly and early infected animals).

To obtain a more comprehensive model, facts about the disease natural history are needed. A description of the progression of animal's condition from the onset of exposure to the causal agent up to the recovery or even death of the animal is needed. Basically, each disease has different natural history. As an example, individual in the population may be assigned to different subgroups, each representing a specific stage of the disease: Susceptible, Exposed, Infectious, Removed with immunity. The transition rates from one subgroup to another (Effective Contact Rate, Average Death Rate, Average Latent Period, Average Infectious Period, Average Loss of Immunity Rate of Recovered Individuals, Average Temporary Immunity Period...) need to be assessed using epidemiological or experimental facts. This type of model sophistication cannot be in general achieved using simple probability event tree analysis. Mechanistic or dynamic model will therefore be more suitable (WP4).

4.3.2.1. Stochastic versus deterministic models

Models that contain no random variables are classified as deterministic. Deterministic models have a known set of inputs which will result in a unique set of outputs. A stochastic simulation model has one or more random variables as inputs. The randomness is modelled using probability distributions.

An important step in stochastic probabilistic risk assessment is to select the most appropriate probability distributions to represent the factors that have a significant influence on the risk estimates. This step in the construction of a Monte Carlo model can be very demanding and resource intensive. Defining probability distributions for every factor in a probabilistic risk assessment may not generally be necessary. If the sensitivity analysis indicates that a particular factor does not contribute significantly to the overall variability or uncertainty of the risk estimate, then this factor may be represented as a point estimate. Different approaches to sensitivity analysis may be applied.

A probability distribution is a function that describes all the possible values and likelihoods that a random variable can take within a given range. This range is between the minimum and maximum statistically possible values, but where the possible value is likely to be plotted on the probability distribution depends on a number of factors, including the distributions mean, standard deviation, skewness and kurtosis. Probability distributions may be used to characterise variability or uncertainty.

For example, if a lognormal distribution provides a good fit to a large facts set of number of animal per herd. Therefore, the distribution type (lognormal) and associated parameters (mean and standard deviation) fully describes the probability distribution of number of animal per herd, from which other statistics of interest can be calculated (e.g., median, and 95th percentile). Alternatively, a probability distribution may be specified to characterise parameter uncertainty. For example, the sample mean (m) is generally an uncertain estimate of the population mean (μ) due to measurement error, small sample sizes, and other issues regarding representativeness. A Probability distribution can be used to represent the distribution of possible values for the true, but unknown parameter.

Understanding whether uncertainty or variability is being represented by a probability distribution is essential to determine how the distribution and parameters should be specified and used in a probabilistic risk assessment.

We distinguish four approaches for probability distribution specification:

- **Large facts set are available:** When large facts set do exist two options are possible: i) use the observed distribution as a probability distribution to describe the factor, ii) fit probability distribution to facts.

- **No fact is available**: When there is uncertainty associated with a factor, such as a facts gap, expert judgment may be appropriate for obtaining distributions. Note that distributions elicited from experts reflect individual or group inferences, rather than empirical evidence. Distributions based on expert judgment can serve as Bayesian priors in a decision-analytic framework. The distributions and Bayesian priors can be modified as new empirical facts become available.

- **Use of statistical theories to derive a probability distribution**: for example, central limit theorem states that the sampling distribution of any statistic will be normal or nearly normal, if the sample size is large enough.

- **Mechanistic approach:** There may be mechanistic reasons depending on known physical or biological processes that dictate the shape or the nature of the distribution. For example, normal distributions result from processes that sum random variables whereas lognormal distributions result from multiplication of random variables. Another example, Poisson distribution can be used to describe the frequency of discrete events, where the events are independent of one another, and randomly distributed in space or time.

4.3.2.2. Monte-Carlo simulation models

A point estimate of risk of introduction may be obtained directly using probability calculations. As an example, the risk of pathogen entrance could be derived as:

$$P(E) = 1 - \left(1 - HP \times AP \times (1 - Se)\right)^{n}$$

Where HP is Herd Prevalence, AP is Animal Prevalence, Se is Sensitivity of the test applied before animal movement, and n is the number of imported animals.

When dealing with variable and/or uncertain factors the probability calculations is complicated (figure 27). Monte Carlo simulation consists on repeating random samples from a set of independent probability distributions to compute their results as with a point estimate of the risk. The output of a probabilistic risk assessment is a distribution of all the possible values of the risk. The total uncertainty (unpredictability) of the derived risk is the combination of two components, namely, uncertainty and variability. A second-order Monte-Carlo Simulation (MCS), also known as two phases or two dimensional MCS may be used to separately propagate uncertainty and variability (Pouillot et al, 2003). The second-order Monte-Carlo Simulation involves a double looping (or nesting) procedure consisting of multiple realisations of model parameters and iterations of input variables. The outcome is a collection of cumulative distributions that simultaneously display the uncertainty and variability in the results.

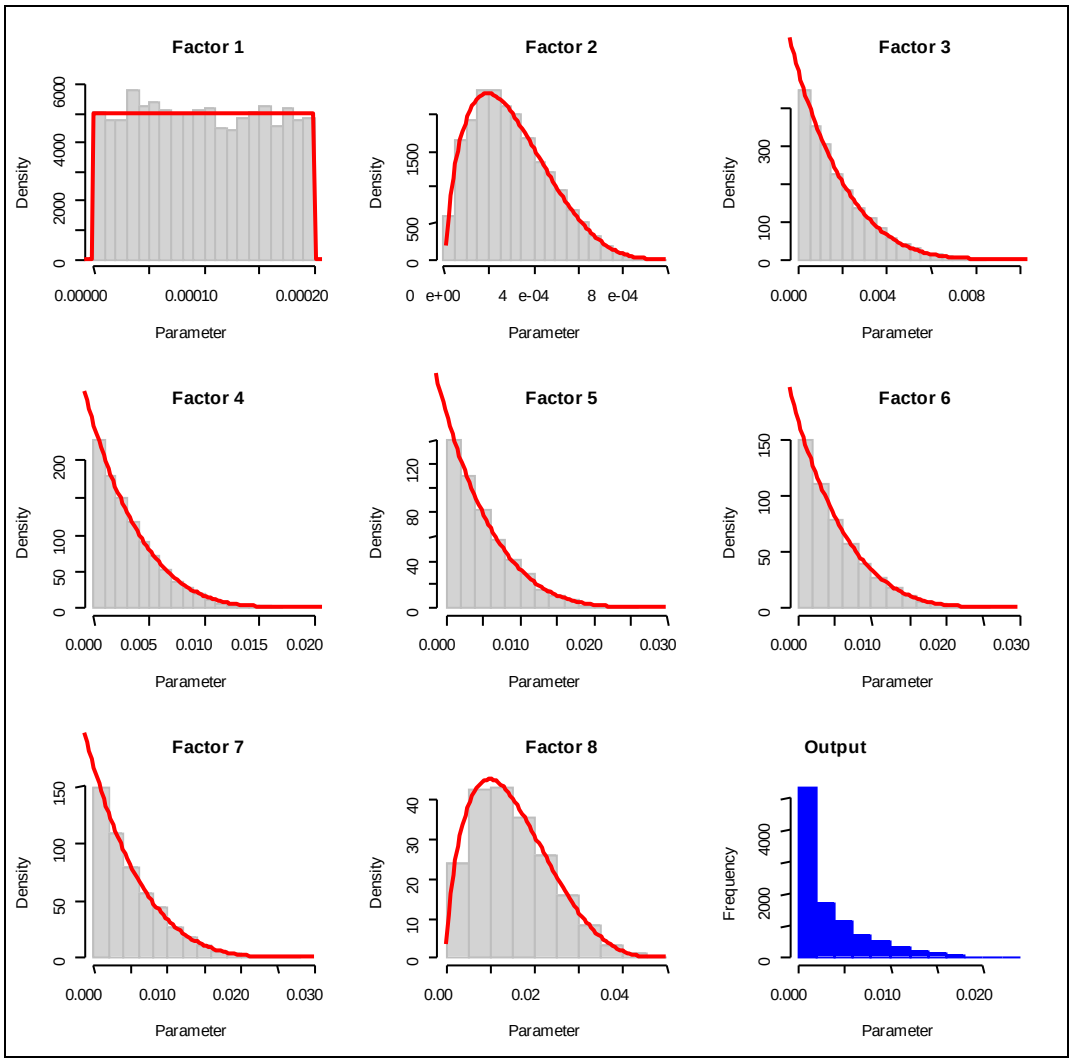


Figure 27: Monte-Carlo simulation

4.3.3. Mechanistic dynamic models

Under mechanistic dynamic models, many different types of models are incorporated, which all have the common denominator of dealing with the non-linearity of an epidemic stage of infectious disease transmission. All these methods tend to require experienced modellers with good analytical mathematical knowledge.

Generally, dynamic methods can be separated into the detailed or "black box" methods and the parameter sparse:

The **black box methods** tend to include all the known quantitative information on the infection, the host and their interaction. These methods tend to work with individual based stochastic simulation models. They take a long time to develop, but if available and if made user-friendly, they can be very helpful. The main disadvantage is that it is difficult to see how uncertainty and model assumptions

influences the outcome, hence the name "black box" models. This also makes it rather risky to extrapolate these models to a different situation or to address a different question with them. Furthermore, unknown parameters values for the model tend to be quantified based on expert judgement, while the evaluation of the impact of such highly uncertain values tend to be minimal and offers little insight, due to the complexity of the model.

The **parameter sparse methods** usually aim at pattern analysis and evaluation and comparison of control measures. They deal with global patterns and ignore some known detail of the system, because their impact is limited in specific stages of an epidemic, especially the exponential growth stages. Thus the main patterns of the infection, which have a major impact, can thus be evaluated better. Technically, this category still contains a wide range of different methods, each with their own pros and cons:

Theoretically, the simplest method is the **R0 method**, which reduces the non-linear process of epidemic (exponential) growth into a summarizing term, named the basic reproduction ration (R0). R0 gets to be expressed into basic parameters and variables, which is rather straight forward for many infections, while it requires special expertise to do this for slow infections and vector-borne infections. Infections with reactivation or persistence mechanisms within a host are also tricky. Many definitions of the reproduction ratio for specific systems have been published and can safely be taken from literature. If the reproduction ratio is well defined, it can be used to evaluate the impact of control measures if well quantified information on the control measures is available. It can also be applied in the reverse, how much reduction of transmission is required for e.g. eradication, and which control measures are needed to get to that level. Although this method is theoretically simple, because the non-linearity of the system is removed by studying the (mostly linear) effect of the per generation reproduction, in practice this method tends to require scientists, experienced in working with R0.

Parameter sparse deterministic models offer good insight into the dynamic behaviour of an epidemic. The growth rate and final size of the infections are important parameters which help in the evaluation of an epidemic with these tools. Although these methods are straightforward as such, they require some expertise in mathematics and/or good knowledge of the relevant literature. If well applied, they can be valuable in evaluating the impact of future epidemics, or in quantifying parameters of an experienced epidemic, which can subsequently be used to evaluate (new) control measures. The deterministic methods can occasionally be extended to probabilistic methods, when probability density functions replace one or more parameters. In some cases such methods can lead to exact answers regarding uncertainty and variability, without needing simulation. Thus, an exact answer can be obtained immediately for all relevant scenarios. This surely needs experts in mathematical models to read and explain these results

Parameter sparse stochastic methods are valuable especially if there is a very high probability of extinction of the infection in an early stage, and to analyse the initial and final stages of an epidemic. In these early of late stages, when few infectious animals are around, stochasticity can have a large impact on the rate of growth and decline (fade out) of the epidemic. This method can therefore help in evaluating the probability that an epidemic will actually take of under specific conditions. It can also help in determining the probability distribution of the time to extinction, which unfortunately tends to be a long tailed distribution.

Spatial analysis using dynamic methods uses two main tools: (1) the speed of the wave front, i.e. how fast does the epidemic move through space and (2) a spatial transmission kernel, which determines the probability that a herd gets infected, given that it is located a specific distance from a resource. Both methods require special expertise and good computer tools to be useful. Both methods can be valuable

in determining the size of a zone with control measures to contain the infection. The latter method is also valuable when control measures aim at reducing transmission between farms. However, quantifying such models is generally not easy, and requires a lot of facts.

All of the above methods for evaluation of the non-linear dynamics in a system require statistical models that are based on these dynamic models, to quantify parameters from field facts of epidemics and from experimental facts. Dynamic models are often part of a complete risk assessment for infectious diseases with epidemic behaviour. The results of the dynamic models are subsequently used in a further risk analysis, where probabilistic methods are the most common tool.

4.3.3.1. Suitability of the dynamic models to the question type

Dynamic models are required when the impact of the exponential growth phase of an epidemic has a large impact on the answer. Thus, it is an essential tool in questions regarding invasion, prevention, eradication and control of infectious (epidemic) diseases, especially when optimisation questions arise.

4.3.3.2. Facts requirements

Generally, the parameter sparse methods lead to generic answers, with limited facts needs. Often, they can already answer a few basic questions, without much good facts. Summarized parameter estimates on the growth rate of the infection in a population from literature have proven to be valuable also under different conditions.

However, for all dynamic methods, it is valid that detailed answers, especially answers regarding optimisation of for example a vaccination or surveillance program, require good facts input. Some facts from similar species and infection strains are needed in order to obtain results with reasonable quality (low uncertainty). Furthermore, original facts from epidemics are valuable, but analysis of such facts generally takes time. If the facts have not been analysed with a statistical method based on an underlying dynamic model, the results can usually not be used. Thus, usually only analyses with parameter estimations aiming at dynamic models can be used. Fortunately, such analyses are published more and more frequently now.

4.3.3.3. Uncertainty and variability

Black box methods should offer a good overview of the uncertainties and variability's in their result. Generally, the variability (stochasticity) in the result is a direct output of the model. The way uncertainty influences the results can mostly be found in one of the first publications on the model, or should at least be found in a report. Without such an analysis, the model is rather useless. In a black box type model, the impact of uncertainty is difficult to oversee and therefore requires thorough analysis.

Parameter sparse methods are generally sufficiently simple (i.e. little complexity in the number of parameters and variables) to allow direct insight in the impact of uncertainty on the model results. Occasionally, these are still supported with an analysis of the uncertainty, because the model ended up with more parameters and more complexity than can be overseen quickly. In those cases, the applied tool is generally stochastic modelling, and they can approach the complexity of the black box models.

There is no exact division between these two methods; they appear on a continuous scale of complexity, with the two examples given above as the extremes. The variability of such models is often evaluated using Latin hypercube sampling.

4.4. Case studies

4.4.1. Echinococcus granulosus

4.4.1.1. WP4 – Risk question as formulated by EFSA

What is the added value of meat inspection with respect to disease prevention and control in the definite host (dog) and intermediate hosts (sheep, goat, cattle and swine)?

The main objective of *E. granulosus* surveillance (meat inspection) is to assess the prevalence (and incidence) of the disease. The estimation of prevalence will rely on a stochastic model. As carcasses/organs presenting cysts are not further submitted to lab diagnosis, it is not possible to know the exact prevalence. Nevertheless, the FASFC (Federal Agency for the Safety of the Food Chain) centralizes all data related to carcasses seized for hydatid cysts (de Smedt, personal communication). Thus it will be possible to estimate the prevalence of the disease when considering the annual number of animals slaughtered, per species and per age category.

4.4.1.2. Background

Echinococcosis is a zoonotic infection caused by adult or larval (metacestode) stages of cestodes belonging to the genus *Echinococcus* and the family *Taeniidae* (Eckert et al., 2002). To date, four species of *Echinococcus* are recognised, namely *E. granulosus*, *E. multilocularis*, *E. oligarthrus* and *E. vogeli*.

- Strains and life cycle

Within the species *E. granulosus*, several 'strains' have been identified, because of genetic heterogeneity as shown in table 48.

Table 48: Strains of *E. granulosus* (Adapted from Eckert and Deplazes, 2004)

Strain/isolate (G: genotype)	Intermediate hosts	Infectivity for humans	Definitive hosts	Probable geographic distribution
Common sheep strain (G1)	Sheep, cattle, pig, camel, goat, macropods	Yes	Dog, fox, dingo, jackal, hyena,	Europe, Middle East, Africa, Asia (Iran, India, Nepal, China), Russia, Australian mainland, Tasmania, New Zealand, USA
Tasmanian sheep strain (G2)	Sheep, cattle?, man	Yes	Dog, fox	Tasmania, Argentina
Buffalo strain (?) (G3)	Buffalo, cattle?	?	Dog (fox?)	Asia
Horse strain (G4)	Horses and other equines	No/?	Dog	Europe, Middle East, South Africa (New Zealand?, USA?)
Cattle strain (G5)	Cattle, buffalo, sheep, goat	Yes	Dog	Europe, South Africa, India, Nepal, Sri Lanka, Russian Federation, South America?
Camel strain (G6)	Camel, goat, cattle	Yes	Dog	Middle East, Iran, Africa, China, Nepal, Argentina
Pig strain (G7)	Pig	Yes	Dog	Poland, Slovakia, Ukraine, Russia, Argentina
Cervid strain (G8)	Cervids	Yes	Wolf, dog	North America, Eurasia
Lion strain (G9?)	Zebra, wildebeest, warthog, bush pig, buffalo, various antelopes, giraffe? Hippopotamus?	?	Lion	Africa

Legend: "?" = Unclear status

The life-cycle of *E. granulosus* is presented in figure 28. The most important cycle involves domestic dogs and sheep. *Echinococcus* requires two mammalian hosts for completing its cycle. Definitive hosts of *E. granulosus* are generally carnivores (mainly dogs) while intermediate hosts which

perpetuate the cycle are generally ruminants (originally cervids). Aberrant hosts (accidental hosts), in which the infection does not lead to further capacity of transmission, include humans.

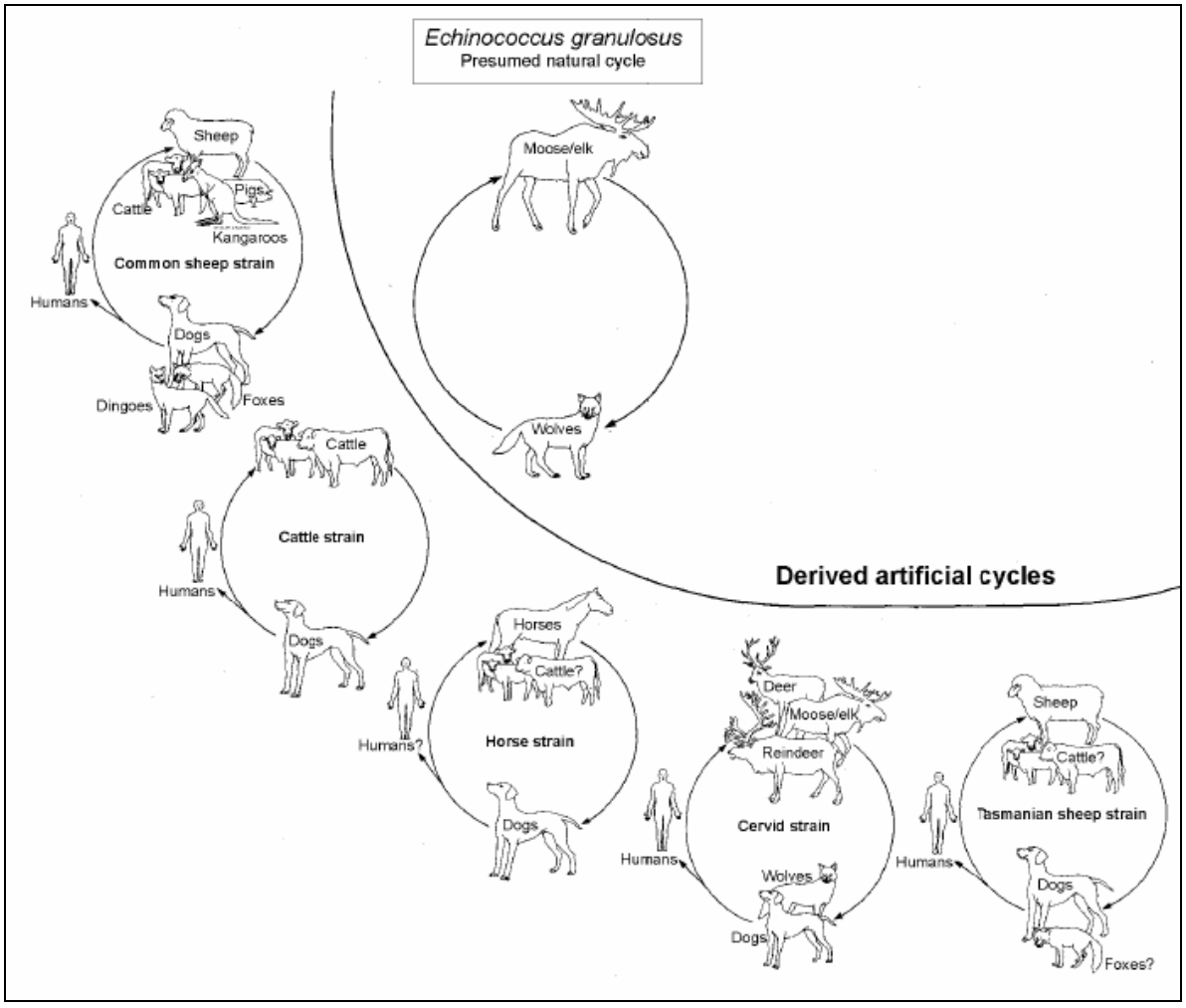


Figure 28: Life-cycle of Echinococcus showing the presumed natural cycle and derived artificial cycles (Adapted from the Eckert et al., 2002)

- Clinical aspects
  - Definitive hosts

No significant pathology is generally observed in definitive hosts.

  - Intermediate and aberrant hosts

It usually takes several years for hydatid cysts to develop to a size that may cause disease in animals (Eckert et al., 2002). In intermediate host species, cysts occur most frequently in the liver and lungs, as shown in figure 29, but sometimes they develop in other organs (central nervous system, skeletal muscles and bone marrow). The cysts of *E. granulosus* are typically unilocular, but sometimes

multilocular. They vary in size and shape. The strain of *E. granulosus* can influence the location of cysts and their morphology. Usually host and *Echinococcus* coexist well. Following initial infection, there is a cellular response and a fibrous capsule develops around the parasite.

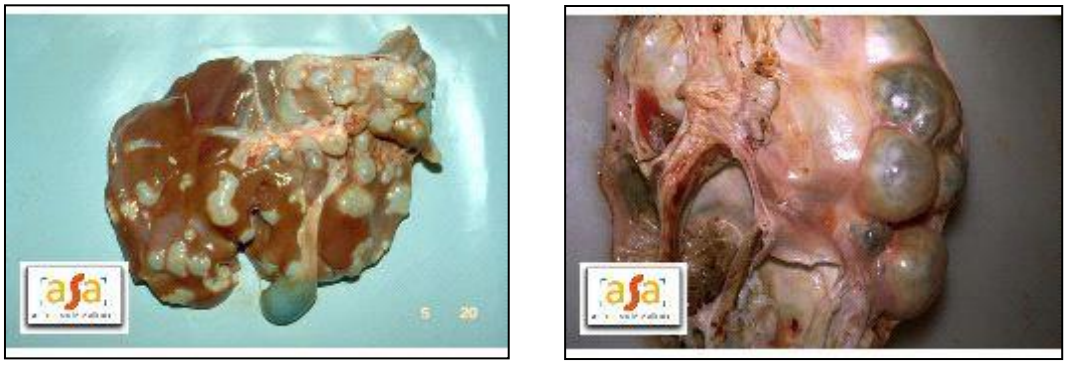


Figure 29: Hepatic (A) and pulmonary (B) hydatid cysts in sheep (Asadia – online atlas of slaughterhouses lesions[1])

- Diagnosis (Eckert et al., 2002)
  - Definitive hosts (dogs)
    - Living animals:

Detection of eggs and proglottids on faecal samples (routine flotation technique) or on perianal skin (adhesive tape).

Arecoline purging: standard method long time used for surveys in dog populations. After application of arecoline to dogs (a parasympathomimetic drug acting on smooth muscle of small intestine and causing worm paralysis), the faecal material discharged is examined. Sensitivity is not very high but specificity reaches 100%.

Immunodiagnosis: (i) **Coprantigen detection** (ELISA): specificity around 97% and sensitivity between 63 and 77%. This method is now preferred to arecoline purging as a screening test in dogs for routine surveillance. (ii) **Serum antibody detection** (ELISA – IgA, IgE, IgA): sensitivity is low (between 35 and 73%) and specificity reaches 70 to 95%. This method can be used as a screening method to assess the presence of *E. granulosus* in a dog population but the main constraints are: Antibodies persist after the worm burden is eliminated; Low sensitivity and specificity? Lack of correlation with the worm burden; Risk of false positives if infection with *Taenia* species.

Copro-DNA recognition method (PCR): confirmatory test (laborious and expensive) but 100% Specificity (EFSA, 2010).

- Necropsy

Direct examination of intestine (microscope).

Sedimentation and counting technique (SCT) – 100% Sensitivity and 100% Specificity (Eckert et al., 2002).

Intestinal wall scraping technique (IST) – 78% Sensitivity and 100% Specificity (EFSA 2010).

Shaking in a vessel technique – 96% Sensitivity and 100% Specificity (EFSA 2010).

- Intermediate hosts

Meat inspection (age-dependent prevalence) – Sensitivity and Specificity?

Immunological tests – poorly sensitive and not very specific in livestock (cannot replace necropsy) (Eckert et al., 2002): (i) Serum antibody detection: poor sensitivity and specificity; (ii) Detection of circulating antigens: not useful; (iii) DNA technology.

- Differential diagnosis in intermediate hosts

The differential diagnosis of hydatid cysts in livestock includes retention cysts in kidneys, liver cysts, granulomatous lesions, *Cysticercus tenuicolis* and tuberculosis (FAO, 2004). The most frequent confusion is often with hepatoperitoneal cysticercosis (*Tenia hydategina*) (Ministère de l'Agriculture, de l'Alimentation, de la pêche, de la ruralité et de l'aménagement du territoire, 2011).

- Current E. granulosus surveillance in Belgium

Currently, according to the European legislation (European Parliament and Council, 2004), the surveillance of *E. granulosus* relies on the detection of cysts in intermediate hosts at the slaughterhouse (sheep, cattle, swine, goats and horses). It could be qualified as a passive exhaustive surveillance. Whole carcasses or parts of them (organs) are rejected if *E. granulosus* cysts are found (Working group on foodborne infections and intoxications, 2003). A suspicion of cysticercosis will entail carcasses are kept frozen for a specific duration of time. No further lab confirmation of cysticercosis or any other clinically similar disease (such as hydatid disease) is routinely performed (Claes L., ITG and personal communication). Thus, the diagnosis only relies on meat inspection.

4.4.1.3. Methodology to apply to answer the risk question

On the basis of available data, it is to say the annual number of carcasses/organs seized after observing hydatid cysts, it will be possible to estimate the prevalence of cysts among slaughtered animals, per species and category of age (apparent prevalence). Knowing the characteristics of meat inspection (sensitivity and specificity), we will be able to estimate the true prevalence in the different livestock populations. The formula to apply is the following:

$$Pr = \frac{Pa + (Sp - 1)}{Se + (Sp - 1)}$$

Where Pr = true prevalence, Pa = apparent prevalence, Sp = test specificity and Se = test sensitivity.

Nevertheless, it is important to keep in mind that, even if meat inspection is highly specific (100%), its sensitivity is very poor (Dorny et al. 2004, Geysen et al 2007).

The trends in true prevalence will allow estimating the effectiveness of meat inspection in identifying positive carcasses.

1. Assumptions:

Horses will not be considered in the following model, as EFSA specified it is not relevant to monitor them because of the non-zoonotic risk (EFSA 2010).

2. Three approaches could be applied:

- to use former Belgian data on echinococcosis (apparent prevalence)

- to adapt prevalence data from other MSs, where prevalence surveys have been performed (table below).

- to work on the basis of cysticercosis data (surrogate procedure)

- **First approach: based on former echinococcosis prevalence at slaughterhouses in Belgium**

There is a real lack of data after 2004, as illustrated in table 49, so this methodology would not be the most suitable.

Table 49: Number of carcasses identified as *Echinococcus* + at meat inspection (FASFC)

	Cattle*		Pigs		Sheep		Goats		Solipeds	
	N	N +	N	N +	N	N+	N	N+	N	N+
2002	641,292	171	11,200,914	1	89,114	3	2,733	-	15,672	-
2003	570,099	200	11,609,933	-	83,112	3	2,514	-	12,304	-
2004	564,266	48	11,229,149	0	87,119	2	3,814	0	11,655	-
2005	523,795		10,861,234		112,771	34	2,585		11,542	
2006	496,181		10,794,757		151,803[£]				10,728	
2007	493,222		11,536,172		137,491[£]				8,939	
2008	522,557		11,588,072		139,555[£]				9,253	
2009	480,068	0	11,677,883		141,214[£]				8,910	
2010	503,277		11,924,052		151,158[£]				8,970	

*Cattle only include adult animals, not calves; [£]from 2006, sheep and goats are counted together.

- **Second approach: based on prevalence estimated in other MSs**

Numerous studies have been carried out to estimate *Echinococcus granulosus* prevalence in other MSs and European countries. The literature review performed during the indirect survey of WP2 (web searches) allowed identifying several of these prevalence studies (table 50).

Table 50: Prevalence of *Echinococcus granulosus* reported in different MSs

Member State	Year(s)	Targeted species	Prevalence rate (%)	Reference
Bulgaria	1983-1995	Sheep	32	Todorov et Boeva 1997
		Cattle	19	
		Pigs	1.5	
Cyprus (North)	2004	Sheep	1.53	Dakkak 2010
		Cattle	6.61	
		Goats	0.13	
Czech Republic	1994-1995	Sheep	0.73	Kolarova 1997
		Cattle	0.003	
		Pigs	0.005	
France	1989	Small ruminants	0.42	Soulé et al. 1995
		Cattle	0.13	
		Pigs	0.009	
Midi-Pyrénées	1994-1996	Livestock	2.5	Bichet et Dorchies, 1998
Greece				
Peloponnese (South)	2005	Sheep	30.4	Varcasia et al. 2007
		Goats	14.7	
	2002-2006	Sheep	39.3	Christodoulopoulos et al. 2008
Thessaly	2010	Sheep	23.0	
North		Goats	7.6	Sotiraki et Chaligiannis 2010
Italy				
North of the country	2003-2005	livestock	<1%	Garippa 2006
		Sheep	20.2	
Abruzzo		Cattle	15.3	
		Cattle	14.8	
Campania		Cattle	67.1	
Sicily		Sheep	57.6	
		Sheep	75.3	
Sardinia		Cattle	41.5	
		Pigs	9.4	
Poland	1997	Sheep and goats	18.7	Derylo et Szilman 1998
		Cattle	0.007	
		Pigs	4.5	
Romania	2001-2004	Sheep	12.65	Neghina et al. 2010
		Cattle	18.98	
		Swine	3.81	
Spain	2005	Sheep	0.57	Carmena et al. 2008
		Cattle	0.7	
		Goats	0.03	
UK (Wales)	1984-1998	Sheep	4.3 to 6.0	Lloyd et al 1998, Palmer et al 1995

- **Third approach: based on actualised cysticercosis data – surrogate approach.**

Two possibilities could be envisaged:

1. to rely on 2002-2004 data in order to estimate the proportional lesions diagnosed as *Echinococcus* compared to cysticercosis. The ratio would then be applied to estimate the apparent real prevalence of *Echinococcus* for the period between 2005 and 2010, starting from the cysticercosis data. Indeed, data are available on the number of carcasses/organs discarded for cysticercosis cysts (localised *vs*. generalised cysticercosis), as shown in the table 51. Cysticercosis data could be considered as surrogate data.

Table 51: Number of carcasses presenting cysticercosis cysts (localised *vs*. generalised cysticercosis) between 1998 and 2010, in Belgium.

	Cattle			**Calves**			**Pigs**		
	N slaugth.	**N Loc.**	**N Gen.**	**N slaugth.**	**N Loc.**	**N Gen.**	**N slaugth.**	**N Loc.**	**N Gen.**
1998		1,644							
1999		1,331							
2000		1,531	9						
2001		1,982	18						
2002	641,292	3,336	29	309,023			11,200,914		
2003	570,099	3,849	25	317,000	10	1	11,609,933	9	0
2004	564,266	2,981	21	317,269			11,229,149		
2005	523,795	2,374	15	313,115	2	1	10,861,234		
2006	496,181	1,796	28	327,391			10,794,757		
2007	493,222	1,527		306,315			11,536,172		
2008	522,557	2,356	18	301,102			11,588,072		
2009	480,068	1,811	9	319,188			11,677,883		
2010	503,277	1,756	10	334,013			11,924,052		

N slaugth. = N animals slaughtered; N Loc. = N carcasses with localised cysticercosis and N. Gen = N carcasses with generalised cysticercosis.

2. to rely on experts' opinion within the frameworks of a Bayesian approach: starting from the differential diagnosis of echinococcosis cysts, the following model could be applied:

$$P_r = \sum_{i=1}^{5} x \quad P_i \quad x \quad P(k/i)$$

Where Pr = true prevalence, Pi = probability of other disease (cysticercosis, tuberculosis, etc.) and P(k/i) = probability the cyst is related to the disease. Experts will then have to fill in a table (table 52) and specify the frequency of a disease compared to another (diseases compared two by two).

Table 52: Example of table to be filled in by experts, compiling the frequency of one disease compared to another. Each cell is filled with a coefficient depending on the importance of the disease compared to another (e.g. cysticercosis *vs*. tuberculosis: 7)

	Echinococcosis	**Cysticercosis**	**Other parasites**	**Granulomatous lesions**	**Tuberculosis**
Echinococcosis	1				
Cysticercosis		1			
Other parasites			1		
Granulomatous lesions				1	
Tuberculosis					1

Coefficients from 1 to 9 can be used, respecting the following scale: 1 = not more important; 3 = moderately more important; 5 = strongly more important; 7 = very strongly more important; 9 = extremely more important

A frequency for each disease can then be estimated. Data on other diseases may also be gathered through literature review. It will be necessary to work by species and by category of age (e.g. adult cattle *vs*. calves), as echinococcosis prevalence is age-dependant as already mentioned above.

Afterwards, a matrix is used to estimate the frequency of each disease. When working by category of age, it is necessary to specify the proportion of animals by category of age.

4.4.1.4. Data needed

Most needed data will be related with some characteristics of the parasite and diagnosis at meat inspection. To take into account the risk represented by imported animals, all data related to importation of livestock species from countries where *E. granulosus* is known to be endemic must be considered also. To be complete, all data linked to the surveillance programme implemented should be considered as well, as it can influence the prevalence of cysts observed. As no further diagnostic test is performed, apart from meat inspection, all data related to lab diagnosis are not required. All data needed are summarized in table 53 (on the same frame as the one presented in the second interim report focusing on WP2).

Table 53: Data needed (starting from the exhaustive list of data needed elaborated in WP2)

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	Data needed
Disease General Information	Disease information	Hosts	Intermediate	List of intermediate hosts
		Pathogenicity	Infection	Incubation period
				Latency period - **Mean duration for the development of a viable cyst in intermediate hosts (per species)**
				Gross lesions
		Diagnosis	Clinical diagnosis	Pathognomonic signs
				Differential diagnosis
			Meat inspection	**Sensitivity**
				Specificity
Descriptive Epidemiology	Morbidity			Morbidity rate (**N carcasses/organs with hydatid cysts, per livestock species and category of age**)
	Demography of hosts	Intermediate hosts	Livestock	Populations and subpopulations of cattle, sheep, goats, swine N livestock herds/flocks
Analytical Epidemiology	Factors of disease introduction	Importations	Animals	N importations of cattle, sheep, goats, swine
	Factors of disease spreading and establishment	Production system	Industries	Cattle, sheep, goat and swine production systems N cattle, sheep, goat and swine heads in meat industry + N
			Farming systems	N cattle, sheep, goat and swine heads under intensive farming system + N herds/flocks
		Animal movement and traceability		Identification of live animals (cattle, sheep, goats, swine and dogs)
Surveillance	Surveillance network	Organisation		Data collection
				Detection of the cases
				Reports of the cases
				Action based on surveillance findings
				N officers involved in the surveillance system
				N levels of the surveillance system
		Characteristics		Field of surveillance networks

			Type of surveillance
			Situation of the disease surveyed
			Population(s) surveyed
			Mode of data collection
			Dependence towards fighting actors
		Evaluation	Existing evaluation
			Type of evaluation (internal/external)
			List of performance indicators
			Frequency
	Objectives and decision criteria	Objectives	Description of the purpose and rationale for surveillance (estimate the magnitude and baseline status of a problem, facilitate planning of national control or eradication programs and strategies, evaluate control measures and intervention efforts)
	Study design	Type of surveillance	Type of the surveillance (active *vs*. passive)
			Surveillance at slaughterhouses
	Population description and characteristics	Sampling units	Simple units or aggregated units, geographical or spatial measure included, time constraints
		Target population	Population about which statistical inference will be made
			Size of target population
		Study population	Population from which the sample is to be drawn
			Size of study population
			Sample frame (list of units to be sampled)
		Targeted population	Population defined by specific diseases variables inherent to the disease in question
		Administrative units	Which units are included in the surveillance system (states, province, sample grid reference, ...)

		Size of sample	Number of reporting units, should include geographic area serviced per unit sampled, number of eligible units served by reporting unit (per unit of geographic area being serviced)
		Animal and group type	Species, breed and type (if applicable) of animals should be evident; include breeds and crosses, define the animal by appropriate production phase concept, age categories, including all appropriate categories pertinent to the surveillance objectives
	Case definition	Clinical description and case definition	List of criteria for positive case, negative case and others as applicable
		Case classification	Definition of the so called suspect, probable and confirmed case categories
			Levels of the classification certainty

Data in bold are additional data needed, specific to the case study and the related risk question, compared to the initial exhaustive and generic list of data needed.

4.4.1.5. Availability of data

Table 54 presents the overview of the data needed to answer the risk question; data availability is as well detailed.

Nevertheless, more specific additional data are also required:

- N animals slaughtered per species and per category of age (prevalence is strongly age-dependent (Eckert et al., 2002)
- N carcasses/organs discarded for hepatic cysts, per species (cattle, sheep, goat and swine)
- N carcasses/organs discarded for pulmonary cysts, per species (cattle, sheep, goat and swine)
- N carcasses/organs discarded for cysticercosis (localised *vs*. generalised) per species and category of age – surrogate data

Compared to the initial exhaustive list of data needed (generic list), some data had to be adapted for the specific case study (characteristics of meat inspection for example). On the other hand, additional data will also be required if willing to apply all three approaches proposed in the methodology section.

Table 54: Availability of data needed, as specified in the inventory of data resources (WP2)

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	Data needed
Disease General Information	Disease information	Hosts	Intermediate	List of intermediate hosts
		Pathogenicity	Infection	Incubation period
				Latency period - Mean duration for the development of a viable cyst in intermediate hosts (per species)
				Gross lesions
		Diagnosis	Clinical diagnosis	Pathognomonic signs
				Differential diagnosis
			Meat inspection	**Sensitivity**
				Specificity
Descriptive Epidemiology	Morbidity			*Morbidity rate (N carcasses/organs with hydatid cysts, per livestock species and category of age)*
	Demography of hosts	Intermediate hosts	Livestock	Populations and subpopulations of cattle, sheep, goats, swine
Analytical Epidemiology	Factors of disease introduction	Importations	Animals	N importations of cattle, sheep, goats, swine
	Factors of disease spreading and establishment	Production system	Industries	Cattle, sheep, goat and swine production systems N cattle, sheep, goat and swine heads in meat industry + N herds/flocks per species N cattle, sheep and goat heads in dairy industry + N herds/flocks per species N swine in breeding industry
			Farming systems	N cattle, sheep, goat and swine heads under intensive farming system + N herds/flocks N cattle, sheep, goat and swine heads under extensive farming system + N herds/flocks
		Animal movement and traceability		Identification of live animals (cattle, sheep, goats, swine)
Surveillance	Surveillance network	Organisation		Data collection
				Detection of the cases
				Reports of the cases
				Action based on surveillance findings

		N officers involved in the surveillance system
		N levels of the surveillance system
	Characteristics	Field of surveillance networks
		Type of surveillance
		Situation of the disease surveyed
		Population(s) surveyed
		Mode of data collection
		Dependence towards fighting actors
	Evaluation	**Existing evaluation**
		Type of evaluation (internal/external)
		List of performance indicators
		Frequency
Objectives and decision criteria	Objectives	Description of the purpose and rationale for surveillance (estimate the magnitude and baseline status of a problem, facilitate planning of national control or eradication programs and strategies, evaluate control measures and intervention efforts)
Study design	Type of surveillance	Type of the surveillance (active *vs.* passive)
		Surveillance at slaughterhouses
Population description and characteristics	Sampling units	Simple units or aggregated units, geographical or spatial measure included, time constraints
	Target population	Population about which statistical inference will be made
		Size of target population
	Study population	Population from which the sample is to be drawn
		Size of study population
		Sample frame (list of units to be sampled)
	Targeted population	Population defined by specific diseases variables inherent to the disease in question
	Administrative units	Which units are included in the surveillance system (states, province, sample grid reference, ...)

		Size of sample	Number of reporting units, should include geographic area serviced per unit sampled, number of eligible units served by reporting unit (per unit of geographic area being serviced)
		Animal and group type	Species, breed and type (if applicable) of animals should be evident; include breeds and crosses, define the animal by appropriate production phase concept, age categories, including all appropriate categories pertinent to the surveillance objectives
	Case definition	Clinical description and case definition	**List of criteria for positive case, negative case and others as applicable**
		Case classification	**Definition of the so called suspect, probable and confirmed case categories**
			Levels of the classification certainty

In bold, data specified as being not available (or not found) in the overview of data resources (WP2) ; in italic : some data are available (Overview of data resources – WP2).

4.4.1.6. How to collect data mentioned as being not available (not available/not found)

Additional literature/web searches

Data related to the characteristics of meat inspection could be gathered through additional literature review and web searches. Indeed, a specific time was dedicated to web searches when building the inventory of data resources (10 minutes). A longer search time might solve the problem and lead to the identification of additional data resources.

Contacts with the Federal Agency for the Safety of the Food Chain

Missing data related to surveillance (evaluation of the surveillance system and case definition) could be gathered through the consultation of experts of the National Food Safety authority.

4.4.1.7. How to answer the risk question if no time to collect data?

Meat inspection should be useful only when all or a part of carcasses/organs presenting cysts are further submitted to lab diagnosis.

4.4.2. Echinococcus multilocularis

4.4.2.1. WP4 – Risk question as formulated by EFSA

What are the surveillance measures to adopt on domestic and wild canids to demonstrate freedom?

The main objective to model the answer to the risk question will be based on the estimation of the true prevalence in foxes and domestic dogs. To demonstrate freedom, it is necessary to investigate a representative sample of correct size in order to estimate prevalence under a specific threshold fixed to qualify a MS as free.

4.4.2.2. Background

Echinococcus multilocularis is responsible of the so-called alveolar echinococcosis in humans. It is currently endemic in several countries from Western and Central Europe, Asia and North America.

- Life cycle

In Europe, the definitive hosts for *E. multilocularis* are foxes, raccoon dogs and wolves (domestic dogs may act as definitive hosts occasionally also) and *Felidae* to a lesser extent (lynx, wildcats and domestic cats), as shown in figure 29 (Eckert *et al*., 2002).

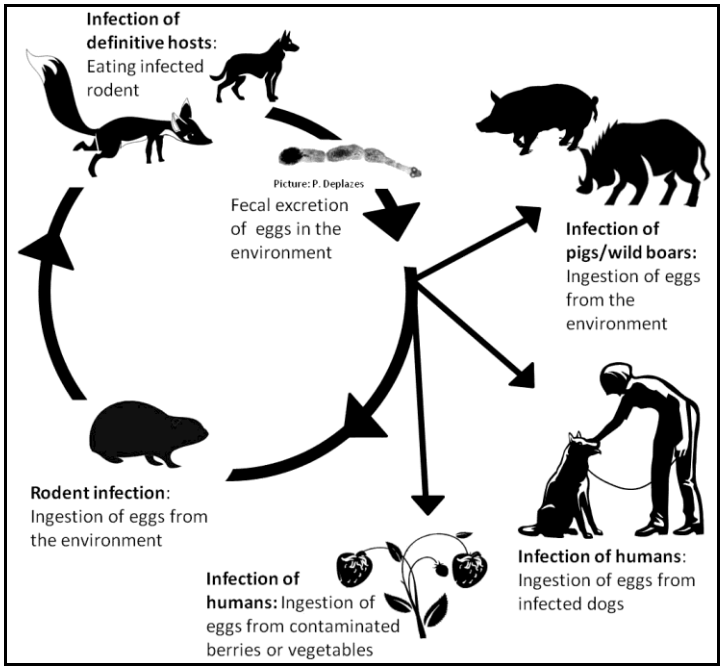


Figure 30: Echinococcus multilocularis life cycle (Wahlström et al., 2011)

Intermediate hosts of *E. multilocularis* are mainly rodents. Humans and various animal species, e.g. wild boar, are considered as accidental hosts (Eckert *et al*., 2002).

- Clinical aspects

In rodents, the main clinical signs are related to cellular infiltration of the liver, peritoneal cavity, other abdominal organs and sometimes lungs (leading to death). Just like *E. granulosus*, *E. multilocularis* does not cause clinical signs in definitive hosts.

- Diagnosis
  - Definitive hosts (foxes, raccoon dogs, domestic dogs and cats)
    - Live animals:

Detection of eggs and proglottids on faecal samples (routine flotation technique) or on perianal skin (adhesive tape).

Arecoline purging: standard method long time used for surveys in dog populations. After application of arecoline to dogs (a parasympathomimetic drug acting on smooth muscle of small intestine and causing worm paralysis), the faecal material discharged after purging is examined. Sensitivity is not very high but specificity reaches 100%.

Immunodiagnosis: (i) Coprantigen detection: ELISA using a monoclonal Ab EmA9 raised against adult *E. multilocularis* somatic antigens; sensitivity is about 61% if worm burden is below 100, but may reach 95% for animals harbouring more than 100 worms. Specificity reaches 95 to 99% (Eckert *et al*., 2002), (ii) Detection of circulating antibodies – sensitivity of 12 to 60%. There is a persistence of circulating antibodies after elimination of the parasite.

Detection of copro-DNA: **PCR** – sensitivity about 94% and specificity of 100% (but laborious, expensive); used as confirmatory test (Eckert *et al*., 2002; OIE, 2008).

- After necropsy: To date, examination of the small intestine at necropsy is the most reliable form of diagnosis in definitive hosts (OIE, 2008).

Direct examination of intestine (microscope).

Intestinal scraping technique (IST) – sensitivity reaches 76 to 78% and specificity is around 99% (Eckert *et al*., 2002). This technique has been widely used for studies on the prevalence in foxes.

Sedimentation and counting technique (SCT) – sensitivity higher than IST; specificity around 99% (Eckert *et al*., 2002).

Shaking in a vessel technique – 96% sensitivity and 100% specificity (EFSA, 2010).

- Current surveillance of *E. multilocularis* in Belgium

Currently, no active surveillance is implemented for *E. multilocularis* in definitive hosts.

In the past, several studies have estimated the prevalence in foxes in different regions of the country (Losson *et al*., 2003; Vervaeke *et al*., 2003; Brochier *et al*., 2007; Hanosset *et al*. 2008). Neither in dogs nor in cats has the prevalence ever been estimated in Belgium. On the other hand, *E. multilocularis* is now endemic, at least in the south of the country. Most prevalence studies were carried out by scientific teams (Universities, etc.).

4.4.2.3. Methodology to apply to answer the risk question

The main objective is the estimation of the true prevalence (foxes and dogs). The specific surveillance programme shall be designed to detect per epidemiologically geographical unit in the MS a prevalence of not more than 1 % at confidence level of at least 95 % (European Commission, 2011). Testing a country free only implies that the prevalence is below a certain limit (Bødker *et al*., 2006).

The methodology to answer the risk question will rely on the estimation of true prevalence, according to a **Bayesian approach** (WinBUGS), with convergence checking based on modified Gelman-Rubin analysis (Saegerman et al., 2006). The Bayesian approach allows integrating field data and expert opinions in a probability model (Branscum *et al*., 2005).

To estimate the true prevalence, it is necessary to estimate three parameters: apparent prevalence (number of positive test results divided by the number of animals tested), sensitivity and specificity. Data (tests results) contain only information about the first apparent prevalence. Test sensitivity and specificity must come from other resources (other studies, etc.) or from expert opinion (Lesaffre *et al*., 2007). The Bayesian analysis will allow the combination of external information (prior information)

with data. Prior information refers to the knowledge about the parameter of interest (probability that an animal is infected). The results of previous prevalence studies performed in Belgium are presented table 55.

Table 55: Previous studies of *E. multilocularis* prevalence performed in foxes in Belgium

Region	Year	Prevalence in foxes (%)	N animals sampled	Reference
Wallonia	1998-2002	20.2	709	Losson *et al*., 2003
- Ardenne		33.1	320	
- Belgian Lorraine		23.1	26	
- Fagne-Famenne		17.2	29	
- Condroz		12.6	190	
- Hesbaye		1.6	126	
Wallonia	2003-2004	24.55	990	Hanosset *et al*., 2008
- Ardenne		40.84	213	
- Belgian Lorraine		33.8	142	
- Fagne-Famenne		61.8	34	
- Condroz		24.86	181	
- Hesbaye		10.0	420	
Flanders	2002	1.7	236	Vervaeke *et al*., 2003
Flanders and Brussels	2007-2008	0.0	187	Van Gucht *et al*., 2010

Some restrictions on the parameter estimates must be fixed to estimate the prevalence:

- Deterministic: setting Se (or Sp) to a particular value (assumption of conditional independence)

- Probabilistic: based on a Bayesian framework when expert knowledge is available (Berkvens *et al*., 2006). It can consist in specifying a prior distribution for a parameter or for a function of parameters.

In order to validate the Bayesian analysis, several criteria will be used (Praet *et al*., 2006):

- Bayesian p-value: indicates if experts' opinions are contradictory to field data

- Deviance information criterion (DIC): allows verifying the adequacy between experts' opinions and experimental results, relying on the credibility of observations.

- P_D parameter: number of parameters really estimated (must be lower or equal to the maximal number of estimable parameters).

- Modified Gelman-Rubin convergence analysis

1. Assumptions:

Felidae, especially domestic cats, will not be considered as they are of lower zoonotic significance because of a slower development and reduced egg production (Deplazes *et al*., 2004).

No clinical signs are observed in dogs (even if it already happened that dogs are intermediate hosts; in such case, there is no risk of contamination because eggs are not shed).

The treatment of dogs is delivered correctly (100% efficacious at eliminating the parasite)

2. Sensitivity of the surveillance system:

To assess the value of surveillance data, it will be necessary to estimate the sensitivity of the surveillance system. A surveillance system includes several components (SSC), each of them related to a separate data resource (derived from a separate surveillance process or data collection system). A stochastic scenario tree model used to describe each SSC can be elaborated to estimate the sensitivity of each SSC (Martin *et al*., 2007; Wahlström *et al*., 2011). The overall sensitivity is then calculated from the sensitivities of the different SCC. It is assumed specificity of the surveillance system is 100% (Martin *et al*., 2007). SCC are rarely independent, thus it is necessary to account for this lack of independence.

The scenario tree divides the population in smaller groups, in which each individual has the same probability of being detected as diseased. At each branch of the tree, probabilities are estimated for each possible outcome. Estimating the SSC sensitivity involves multiplying the probabilities down each limb of the tree and summing those that give a positive outcome (disease detected) (Martin *et al*., 2007).

SSC sensitivity depends on the prevalence of the disease in the population (design prevalence or minimum expected prevalence or threshold prevalence). An example of complex scenario tree is presented figure 31.

The probability of each branch can be estimated from a quantitative, semi-quantitative or qualitative estimation point of view, according to available data. The uncertainty to the analysis should be considered as possible also (stochastic modelling). If necessary, e.g. when data are not available, probabilities could be estimated using expert opinion.

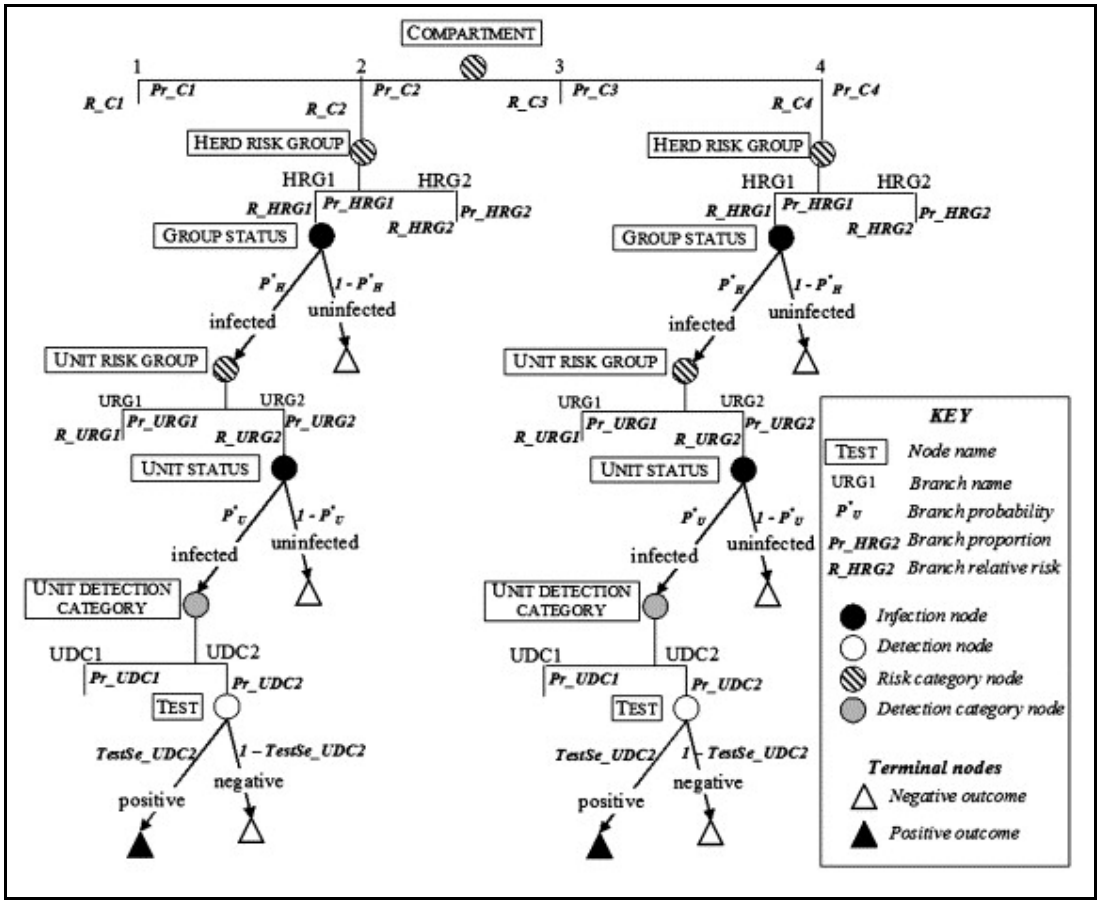


Figure 31: Example of a complex stylised scenario tree for surveillance process active in four compartments within a country, showing risk category nodes at both the group and unit levels, and a detection factor node (Martin et al., 2007)

4.4.2.4. Data needed

Starting from the exhaustive list of data needed elaborated in WP2, a list of data needed specific to the case study and the risk question were selected (table 56).

In order to describe the sampling strategy in wildlife, all data related to surveillance will be required. As no official surveillance system and no surveillance programme are implemented, all data related to the surveillance network are not required. Nevertheless, data related to the design of (regional) prevalence studies are pertinent. Also prevention and control data are interesting to collect within the frameworks of a free status denomination.

Table 56: List of data needed for E. multilocularis case study

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	Data needed
Disease General information	Disease information	Hosts	Definitive	List of definitive host species
		Diagnosis	Lab diagnosis	Sample type
				Test used
				Test sensitivity
				Test specificity
				Test predictive values
Descriptive epidemiology	Spatio-temporal distribution			Annual incidence
				Form of the disease
	Demography of hosts			Populations of dogs
				Populations of foxes
				Populations of raccoon dogs
				Density of rodents
Analytical epidemiology	Factors of disease introduction	Entries	Animals	N entries of domestic dogs
				Chemotherapy status of entering dogs
Prevention and control	Disease inspection practices	Quarantine		**N dogs quarantined**
				Average duration of quarantine
		Animal movement and traceability		**Identification of dogs**
				Follow-up of dog movements
	Reservoir control	Reservoir		**Existence of a control of foxes**
	Treatment			**Chemotherapy of dogs**
				Chemotherapy of wildlife
Surveillance	Objectives and decision criteria	Objectives		Description of the purpose and rationale for surveillance (estimate the magnitude and baseline status of a problem, facilitate planning of national control or eradication programs and strategies)
	Study design	Type of surveillance		Type of the surveillance (active *vs*. passive)

			Proportion of active vs. passive surveillance
			Surveillance of wildlife
	Population description and characteristics	Sampling units	Simple units or aggregated units, geographical or spatial measure included, time constraints
		Target population	Population about which statistical inference will be made
			Size of target population
		Study population	Population from which the sample is to be
			Size of study population
			Sample frame (list of units to be sampled)
		Targeted population	Population defined by specific diseases variables inherent to the disease in question
		Administrative units	Which units are included in the surveillance system (states, province, sample grid reference, ...)
		Size of sample	Number of reporting units, should include geographic area serviced per unit sampled, number of eligible units served by reporting unit (per unit of geographic area being serviced)
		Animal and group type	Species, breed and type (if applicable) of animals should be evident; include breeds and crosses, define the animal by appropriate production phase concept, age categories, including all appropriate categories pertinent to the surveillance objectives
	Case definition	Clinical description and case definition	List of criteria for positive case, negative case and others as applicable
		Laboratory criteria	Diagnostic test(s) sensitivity
			Diagnostic test(s) specificity

			Identification of limitations of the test(s) used for the disease confirmation
			type of diagnostic test(s) and cut-off points or dilution used to define categories of cases
		Case classification	Definition of the so called suspect, probable and confirmed case categories
			Levels of the classification certainty
	Sampling methods		Description of the field and laboratory data collection techniques
			Level of detection
			Statistical level of confidence
			Diagnostic test sensitivity of sampling
			Predictive value
			Time intervals of data collection
			Frequency of data collection
			Geographic extent of the study area under
			Methods of data collection and handling (how raw data are gathered from the field, sample handling protocol, cold chain measures, sample degradation factors)
			Resource(s) of potential bias
			Trigger for data collection
	Laboratory	Diagnostic test	Date of sampling collection and lab test
			Test(s) used
			Sensitivity
			Specificity
		Accreditation	Existence of accreditation programs
			Existence of proficiency testing
			Frequency of proficiency testing

Data in bold are additional data needed, compared to the initial exhaustive and generic list of data needed.

Apart from these data extracted from the generic exhaustive list of data, additional data appear to be necessary to answer the risk question.

- Geographic origin of animals ('rural' *vs*. urban areas): the risk is higher for rural dogs (Torgerson and Craig, 2009).
- Number of dogs with access to wooden areas (increased risk of rodent predation?)
- Transmission:
    - Number of eggs released per gram by the definitive host
    - Duration of pre-patent period
    - Delay before production of eggs
- Differential diagnosis in dogs, foxes and raccoon dogs
- Routine anthelminthic treatment of dogs (passport or vaccination card including information on deworming frequency and status)
- Existence of a treatment in wildlife (Praziquantel-containing oral baits)?
- Sensitivity of the surveillance system (Wahlström *et al*., 2011)

On the other hand, the following surrogate data could be used to help answering the risk question:

- Density of rodents: common voles are the most important intermediate hosts in the MS of interest. A high density of rodents could have, as a consequence, a higher density of foxes, and thus represent areas at higher risk of alveolar echinococcosis.
- Density of foxes: in a recent study carried out in Switzerland, Schweiger and collaborators (2007) reported an increase in human cases of alveolar echinococcosis following an increased density of foxes.
- Human annual incidence of alveolar echinoccosis cases: as sampling in wildlife is not always easy to implement, it could be suggested to consider the trends in human cases of alveolar echinococcosis as reflecting the situation in wildlife. Nevertheless, the disease develops only 5 to 15 years after infestation in humans, it is thus not considered as an early indicator!).
- Human laboratory results (serology): a serological survey of human populations could be less binding than sampling wildlife, as, to date, the best diagnostic test performed in wildlife requires the death of the animal.

4.4.2.5. Availability of data

Table 57 summarizes the availability of data needed to answer the risk question.

Only some data are available on the annual incidence of the disease (prevalence studies for wild species). No data are available for incidence in domestic dogs. Regarding entries of domestic dogs, only commercial movements of dogs are registered. No data are available for non-commercial movements. No follow-up of non-commercial movements is performed, so no data will be available. Data on the chemotherapy status of dogs is not required on the identification document for intra-EU movements, so this data will not be available either.

Table 57: Availability of data needed, as specified in the inventory of data resources (WP2)

Category	Sub-category level 1	Sub-category level 2	Sub-category level 3	Data needed
Disease General information	Disease information	Hosts	Definitive	List of definitive host species
		Diagnosis	Lab diagnosis	Sample type
				Test used
				Test sensitivity
				Test specificity
				Test predictive values
Descriptive epidemiology	Spatio-temporal distribution			*Annual incidence*
				Form of the disease
	Demography of hosts			Populations of dogs
				Populations of foxes
				Populations of raccoon dogs
Analytical epidemiology	Factors of disease introduction	Entries	Animals	*N entries of domestic dogs*
				Chemotherapy status of entering dogs
Prevention and control	Disease inspection practices	Quarantine		**N dogs quarantined**
				Average duration of quarantine
		Animal movement and traceability		**Identification of dogs**
				Follow-up of dog movements
	Reservoir control	Reservoir		Existence of a control of foxes
	Treatment			Chemotherapy of dogs
				Chemotherapy of wildlife
Surveillance	Objectives and decision criteria	Objectives		Description of the purpose and rationale for surveillance (estimate the magnitude and baseline status of a problem, facilitate planning of national control or eradication programs and strategies)
	Study design	Type of surveillance		Type of the surveillance (active *vs.* passive)
				Proportion of active vs. passive surveillance

			Surveillance of wildlife
	Population description and characteristics	Sampling units	Simple units or aggregated units, geographical or spatial measure included, time constraints
		Target population	Population about which statistical inference will be made
			Size of target population
		Study population	Population from which the sample is to be drawn
			Size of study population
			Sample frame (list of units to be sampled)
		Targeted population	Population defined by specific diseases variables inherent to the disease in question
		Administrative units	Which units are included in the surveillance system (states, province, sample grid reference, ...)
		Size of sample	Number of reporting units, should include geographic area serviced per unit sampled, number of eligible units served by reporting unit (per unit of geographic area being serviced)
		Animal and group type	Species, breed and type (if applicable) of animals should be evident; include breeds and crosses, define the animal by appropriate production phase concept, age categories, including all appropriate categories pertinent to the surveillance objectives
	Case definition	Clinical description and case definition	List of criteria for positive case, negative case and others as applicable
		Laboratory criteria	Diagnostic test(s) sensitivity
			Diagnostic test(s) specificity
			Identification of limitations of the test(s) used for the disease confirmation

			type of diagnostic test(s) and cut-off points or dilution used to define categories of cases
		Case classification	**Definition of the so called suspect, probable and confirmed case categories**
			Levels of the classification certainty
	Sampling methods		Description of the field and laboratory data collection techniques
			Level of detection
			Statistical level of confidence
			Diagnostic test sensitivity of sampling
			Predictive value
			Time intervals of data collection
			Frequency of data collection
			Geographic extent of the study area under surveillance
			Methods of data collection and handling (how raw data are gathered from the field, sample handling protocol, cold chain measures, sample degradation factors)
			Source(s) of potential bias
			Trigger for data collection
	Laboratory	Diagnostic test	**Date of sampling collection and lab test**
			Test(s) used
			Sensitivity
			Specificity
		Accreditation	Existence of accreditation programs
			Existence of proficiency testing
			Frequency of proficiency testing

In bold, data mentioned as not available or not found in the overview of data resources; in italic: data partly available

4.4.2.6. How to collect data mentioned as not being available (not available/not found)

Contacts with the Federal Agency for the Safety of the Food Chain

Data on quarantine of dogs could be investigated by consulting the sanitary authorities in charge of border inspection points.

Contacts with research units

Specific data on prevalence studies (sampling programmes, case definition, etc.) could be gathered by consulting the scientific teams who have performed the prevalence studies.

Contacts with laboratories

Regarding missing data on diagnostic and proficiency testing, information will be sought in laboratories in charge of the diagnosis (National Reference Laboratory and research laboratories).

4.4.3. Porcine Reproductive and Respiratory Syndrome

4.4.3.1. WP4 – Risk question as formulated by EFSA

What is the effectiveness of intervention measures at reducing the prevalence of PRRS in France? Intervention measures to be considered should include: (1) vaccination, (2) herd management, and (3) biosecurity.

4.4.3.2. Background

Porcine Reproductive and Respiratory Syndrome (PRRS) is caused by a virus classified as an arterivirus. PRRS represents an important cause of reproductive disease in sows, high mortality in pre-weaning piglets infected in utero and respiratory disease in pigs infected post-weaning (Evans et al., 2010).

The between-herds' transmission of PRRSv can take place when animals are traded from infected herds, semen from newly infected boars is used or mechanical vectors, such as human, transmit the virus to susceptible animals. Airborne transmission was suggested, however, its importance is still uncertain. Because purchasing practices, isolation facilities, herd size and pig density mainly explain the PRRSv prevalence variability between herds, both fade out (spontaneous extinction) and reintroduction are important characteristics of PRRSv transmission dynamic. Table 58 presents the possible contamination sources (Evans et al., 2010; Pitkin et al., 2009).

Table 58: Possible contamination sources of PRRSv in between herds' transmission and their related measures of prevention control

Contamination sources	Circumstances	Prevention
Pig introduction	Traded from infected herds	Purchase animals from non infected herds, Isolate purchased animals before their introduction into the herd (quarantine)
Semen	Newly infected boars	Insurance that semen from insemination's centers offers sanitary guarantee of PRRS virus free status.
Transport vehicles	Vehicles moving from infected to free herds	Combine litter removal, disinfection, and drying…
Fomites	Shoes, clothing, material and equipment can allow mechanical transmission of the virus…	Wash and disinfect inanimate objects before their introduction to the farm…
Aerosol	Air transmission	Use air filters
Humans	Humans can act as mechanical vectors.	Biosecurity protocols: including change of boots and clothes, showering… Avoid visitors
Other animals and insects	Flies picking up the virus from PRRS-infected animals and transport it…	Use insecticides, insect bait and manage the site (grass cutting, weeds removal…)

PRRS virus possesses a complex pattern of transmission, including efficient horizontal and vertical transmission. It can be transmitted among pigs through nose secretion, manure, and semen, as well as *in utero* for piglets born from infected sows. The incubation period varies from 3 to 28 days but is most likely between 4 and 7 days. The virus is usually present in the blood during a period of 4-6 weeks, during which it can be excreted through nose secretion, manure, semen, and urine. After 4.5-20 months post infection, pigs might become again susceptible, due to a loss of protective immunity. The infection is often maintained in the herd by young stock being infected. It has been shown experimentally that infection of sows on or after week 12 of gestation causes late abortions, and the birth of stillborn and mummified pigs; before this period the utero transmission of the virus is still uncertain (Evans et al., 2010). Passive immunity against PRRS virus is maintained until 4-6 weeks of age.

Following its introduction, the spread of the virus within a herd is mainly determined by the type of the pig production system, the intensity and the frequency of contacts between pigs (Lurette et al., 2008). Its persistence and fade-out within a herd are mainly determined by: (1) the birth rate of the piglets', (2) the rate of breeding sows' replacement, and (3) the size of the herd (Nodelijk et al., 2000). In herds where the weaning unit is not separated, the weaned pigs are infected in this unit, whereas weaners in separated weaning units are usually not infected until they are mixed with others through continuous transfer.

4.4.3.3. Methodology to apply to answer the risk question

In order to assess PRRS preventive measures, both between-herds and within-herd transmission dynamics of PRRSv are to be investigated, using a Susceptible-Infected-Recovered-Susceptible (SIRS) compartmental model. The three basic parameters determining the natural history of infection (rate of loss of passive immunity, rate of recovery, rate of loss of protective immunity) are collected from the scientific literature. The persistence and fade out of the virus are investigated taking into account different prevention and control measures, especially vaccination, herd management (e.g. contact structure, herd size, frequency of animal introduction) and biosecurity procedures (e.g. quarantine...).

4.4.3.4. Needed data

Table 59 list the needed data.

Table 59: List of the data needed to answer the risk questions (The bold text indicates the title of the matching technical card that was created within WP3 in order to link a given data to its required metadata set)

Data	Description	Related Technical Card (WP3)
Data on disease and General Information		
Disease transmission	The description of all pathogen transmission pathway	Pathogen Transmission
Climatic data	The climatic data in France	Climatic data
The test sensibility and specificity	The sensibility and specificity of the diagnostic test used during inspection and surveillance.	Test sensitivity and specificity
Data on host		
Disease infection data	Incubation period, excretion pathogen load, Recovery rate, Tissues Pathogen Load	Disease Infection data, Incubation Period, Excretion Pathogen Load, Recovery Rate
The demography data of the host	The density and distribution of the host in Belgium	Demography data
Data on the Production System Information in the region of interest		
Population of each system	Counting of live animals in each production system, and each step in this system	Population data
The production	The volume of commodity produced by each production system	Production data
Commodity pathogen load	The commodity pathogen load in each commodity.	Commodity pathogen Load
Ecology data	Ecology data in France: Land cover surrounding the farm, existence of lagoon effluent.	Ecology data
Disease Status	Disease status of importing host and herd, as well as from the region of interest	
Data of surveillance program if	The description of surveillance: N Herd Under Programme / N Herd Tested Under	Herd data

existing	Programme / N Positive Herd / N New Positive Herd / N Herd Depopulated / N Positive Herd Depopulated / Percent Herd Covered individually / Percent Positive Herd individually / Percent New Positive Herd individually / Percent Herd Status Officially Free	

Data of factors of interest to reduce PRRS prevalence within a region and/or farm

Importation data	Introduction of external host and movement in the production system	Movement data
Biosecurity measures	Description of the biosecurity measures applied in the production system	Biosecurity Measures, Waste Management
Inspection Point data	N Inspection Points, Percentage Visit To Purchase	Inspection Point data
Quarantine data	N Animal Quarantine, Quarantine Duration Period	Quarantine data
Inspection Investigation data	Description of investigation during inspection**:** Context (Routine, Suspicious), N Unit Tested, N Unit With Suspicious Lesions, N Unit Positive, Corrective Action (Quarantine, Rejected, Euthanasia, Destruction, Treatment, Slaughtered, Return Of Consignment, Suspended), N Unit Corrected Action	Inspection Investigation data
Control Measures	Control Measures applied in the production system	Control Measures
Vaccination Status	Vaccination Status of importing host and livestock applied in the area	Vaccination Status
Vaccination strategy	Vaccination strategy applied in the area and in production system	Vaccination Strategy
Chemical and Physical resistance of virus	Chemical and Physical resistance of virus and the pathogen survival rate	Chemical Resistance, Physical Resistance, Pathogen Survival rate
Virus Resistance at treatment	Use of treatment, its efficiency and safety (Safety, Pathogen Excretion Risk) and the pathogen survival rate to treatment	Treatment practices, Substance Safety, Pathogen Survival rate
Vector resistance to insecticide	Description of the use of insecticide in France	Treatment practices

Parameters used to value the impact of measures on PRRS prevalence and control

Immune response data	Description of immunity response of host (passive immunity, protective...)	Immune Response data
Prevalence	Prevalence before and after the prevention measures	Disease's Prevalence

4.4.3.5. Data availability

Thanks to WP2 overview of PRRS data availability, especially within the French data collection system, some data were said not to be available:

- Data on PRRS pathognomonic signs
- Data on PRRS differential diagnosis of the disease
- Data on industry and management factors affecting disease transmission and spread: confinement operation, biosecurity practices, industry awareness
- Data on PRRS annual incidence
- Data on PRRS vaccination status of arriving animals, DIVA strategy, duration of vaccinal protection, doses of vaccine used at the national level, N animals vaccinated at the national level and frequency of vaccination

4.4.3.6. How to collect data mentioned as being not available (not available/not found)

In order to collect the not available data, associated technical cards as presented in annex WP3 could be used to create forms including the needed facts and their metadata.

4.4.4. Venezuelan Equine Encephalitis

4.4.4.1. WP4 – Risk question as formulated by EFSA

What is the risk of introduction of the virus into the EU, taking into account risk reduction measures in place?

The assessment includes introduction of VEE virus into the EU and one local host becoming infected. Spread and establishment of the virus is not in this risk assessment.

4.4.4.2. Background

VEE (Venezuelan Equine Encephalitis) virus is a virus belonging to the family of the Togaviridae, genus Alphavirus. VEE is not present in the EU at the moment (it can be considered exotic). There are 6 antigenic subtypes of the virus (I-VI), each divided by antigenic variants. Antigenic variants I-AB and I-C are associated with epizootic/epidemic activity (central and South America) in equids and humans. Variants I-D, I-E, I-F and type II-VI are considered to circulate in enzootic cycles (southern USA, central and South America) (figure 32).

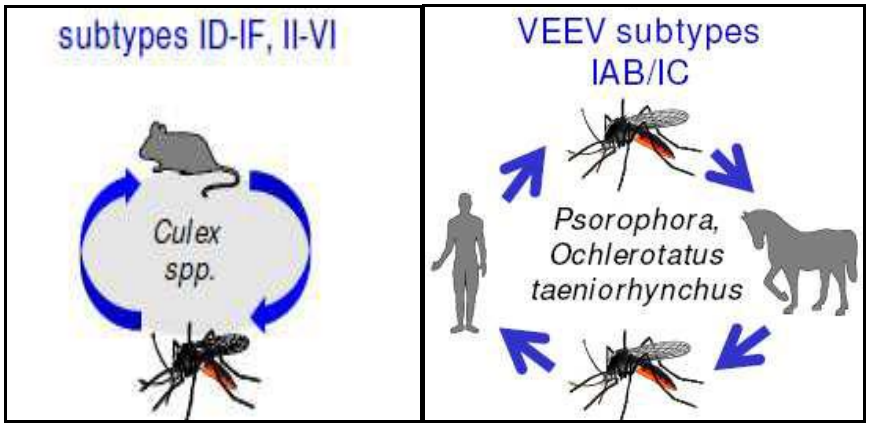


Figure 32: VEEv enzootic and epizootic forms (from left to right)

The VEE main reservoir is constituted by wild rodents. Susceptible species are horses but also humans. Some species of birds (herons) also develop high and prolonged viremias and can infect blood-sticking mosquitoes.

4.4.4.3. Methodology to apply to answer the risk question

We aim for a qualitative risk assessment here. Therefore, we have to think of the possible pathways for introduction of the VEE virus in the EU. For each pathway the likelihood that this pathway can happen should be estimated. To estimate the likelihood we need data.

1. Theoretical pathways for introduction of VEE in EU

Step 1 = Import VEE infected animal:

- Import of horses
- Trade
- Temporarily (sport, breeding)
- Passing through EU for trade
- Import after temporarily export
- Import of donkeys
- Import of zebras
- For example for breeding in zoo
- Import of birds (also own movement of birds from neighbouring countries)
- Import of rodents

Step 2 = Import VEE infected vector

- Vector on living animals from risk countries
- Vector on plants/flowers from risk countries
- Vector introduced by natural movement of vector

Step 3 = Import VEE virus

- Import of biologicals (ova/embryo's/sperm/serum/plasma)
- Import of meat
- Import of vaccines

Step 4 = "Import" of infected human

- The following tables (60 to 63) present for each pathway the data needed to estimate the likelihood of this pathway.

Table 60: Data needed to estimate the likelihood of the VEEv introduction in relation with the import VEE infected animal (equids/ rodents and birds are presented in different tables)

Import of VEE infected horse, donkey, zebra		
Pathway step 1	**Likelihood depends on**	**Data needed to estimate likelihood**
Animal must be infected	- Number of horses imported from risk countries - Number of donkeys and zebras imported from risk countries - Likelihood of being infected	- Risk countries for VEE - Number of horses, donkeys, zebras imported from risk countries to EU (split up in number imported for trade; temporarily imported for competitions; import after temporarily export for competition; passing through EU for trade) - Prevalence of infection of horses, donkeys, zebras in risk countries - Vaccination status of horses from risk countries - Are horses, donkeys, zebras tested for VEE before export from risk countries? - Available tests for horses/donkeys/zebras - If yes, sensitivity/specificity of test.
Infection is not detected in country of origin	- Showing symptoms of disease - Tests performed - Incubation period	- Knowledge about clinical disease, do horses, zebras, donkeys show symptoms? Which symptoms? - Are symptoms distinguishable from other diseases symptoms? - Are horses tested for being infected with VEE before transport in country of origin? - Sensitivity/specificity of test
Infection is not detected during transport	- Showing symptoms of disease - Incubation period	- Knowledge about clinical disease; do horses, zebras, donkeys show symptoms? Which symptoms - Length of incubation period
Animal is still infected after transport	- Length of transport - Length of incubation period, disease, recovery, immunity - Is incubation period longer than transport time and quarantine time? - Testing/quarantine at border	- Length of incubation period, viremic period - Are horses tested for VEE at border? - Which tests are available? - Sensitivity/specificity of test - Are horses put in quarantine at border when importing in EU from risk countries for VEE? How long? - Surveillance system in importing country? - Likelihood of detection of infection by surveillance system
Animal is bitten by a vector	- Availability of competent vectors in country of import (EU) - Number of competent vectors	- Which vectors are known to be able to transmit VEE? - Are these present in EU? - Vector abundance

	- Season (are vectors active?) - Probability of transmission from vector to animal	- Where? - Season/temperature in which these vectors are active - Probability of transmission of virus to animal per bite
Vector survives EIP and time to next blood meal	- Season	- Season/temperature in which the vectors are active/can survive
Vector bites susceptible host	- Availability of susceptible hosts in country of destination	- Which animals can be a susceptible host? - Numbers of host animals in country of destination
Host is infected		
Import of VEE infected rodent into EU		
Pathway step 1	**Likelihood depends on**	**Data needed to estimate likelihood**
Rodent must be infected	- Number of rodents imported from VEE risk countries - Likelihood of being infected - Species of rodents that can be host for VEE - Does season play a role in cycle of VEE in rodent?	- Number of rodents imported from VEE risk countries to EU - Number of rodents imported by accident (in feed etc) from VEE risk countries to EU - Prevalence of rodents with VEE in risk countries - Which rodent species can be host for VEE? - Does season play a role in cycle of VEE in rodent?
Rodent is still infectious after travel to EU	- Length of incubation period, viremic period	- Length of incubation period - Do viremic birds show symptoms?
Rodent is bitten by a vector	See above	See above
Vector survives EIP and time to next blood meal		
Vector bites susceptible host		
Host is infected		
Infected bird comes into EU		
Pathway step 1	**Likelihood depends on**	**Data needed to estimate likelihood**
Bird must be infected	- Number of birds migrating to EU from risk countries - Likelihood of being infected - Number of birds migrating to EU from neighbouring countries - Possible travel distance for herons (is it possible for a heron to fly on its own from a risk country to EU)	- Number of birds (herons) that migrate from VEE risk countries to EU - Prevalence of birds with VEE in risk countries - Distance that a heron can fly/travel on its own. Seasonal movements of herons during the year. - Viremic period of VEE in herons - Symptoms of VEE infected herons

	- Viremic period in herons	
Bird is still infectious after travel to EU	- Length of incubation period, viremic period	- Length of incubation period in herons - Do viremic herons show symptoms?
Bird is bitten by a vector	See above	See above
Vector survives EIP and time to next blood meal		
Vector bites susceptible host		
Host is infected		

Table 61: Data needed to estimate the likelihood of the VEEv introduction in relation with the import of VEE infected vector

Pathway step 2	Likelihood depends on	Data needed to estimate likelihood
Import VEE infected vector - Vector on living animals from risk countries - Vector on plants/flowers from risk countries - Vector introduced by natural movement of vector		
Infected vector enters EU	- Is vector still infected after travel? - How many vectors are imported into the EU on living animals from risk countries? - How many vectors are imported in the EU on living plants/flowers from risk countries? - Is it possible for vectors to travel independently to the EU? How? Chances for survival? - Does a vector live longer than the transport to EU? - Is vertical transmission of VEE in vectors possible?	- Which insects can be vectors for VEE? - Which countries are risk countries for VEE? In which countries is VEE endemic/epidemic? - Prevalence of VEE in vectors in risk countries - Numbers of animals from risk countries imported to EU that theoretically can carry vectors for VEE. - Amounts of plants/flowers from risk countries imported to EU that theoretically can carry vectors for VEE. - Number of potential VEE vectors that are imported by plant/animal import to the EU. - Distance form risk countries to EU. - Distance that vectors of VEE can travel independently. - Life expectancy of potential vectors - Life cycle of VEE in vectors (vertical transmission possible?) - Climate conditions needed for vectors of VEE. - Climate conditions in EU, comparable to needed conditions?
Vector bites susceptible host	- Availability of susceptible hosts in country of destination	- Which animals can be a susceptible host? - Numbers of host animals in country of destination
Host becomes infected		

Table 62: Data needed to estimate the likelihood of the VEEv introduction in relation with the import of VEEv by biologicals, meat and vaccines (presented in three tables)

Pathway step 3	Likelihood depends on	Data needed to estimate likelihood
Import VEE virus by import of biologicals (ova/embryo's/sperm/serum/plasma/cell culture/tissue, etc), from all possible host animals/humans		
Biological with VEE virus enters EU	- Amount of biologicals imported into the EU from VEE risk countries - Can these biologicals contain VEE virus?	- Number/amount of all biological products from humans/horses/rodents that are imported into the EU from VEE risk countries - Are measures taken in country of origin to make sure that biological is virus free? (testing biological, testing animals, only producing from animals that are free of disease, etc) - Can VEE virus survive in these products? - Are these products tested before export? - Are these products tested before use?
Horse is inseminated with material, or Serum/plasma is injected in a horse/human/host animal, or Other use of biological in horse/human/host animal		- Use of biologicals from risk countries in host animals and humans in EU (which products, which animals, how often, etc)
Host animal develops infection		- Infectivity of virus in the different biological products - virus dose needed for infection in horse - Virus dose needed for infection in humans - Can this dose be present in biological product? - Can VEE be transferred by insemination with infected sperm or infected ovum/embryo? - Can VEE be transferred by injecting serum/plasma that contains VEE?

Pathway step 3	Likelihood depends on	Data needed to estimate likelihood
Import VEE virus by import of (horse) meat		
Infected horse meat is imported to EU	- Is horse meat from VEE risk countries imported into the EU? - How much horse meat in imported? - In which form? (fresh meat/frozen/pet food, etc) - Prevalence of VEE in horses in countries from which meat is imported. - Survival of VEE virus in meat/meat products	- Amount of horse meat from VEE risk countries imported into the EU (fresh/frozen/processed products, pet food, etc) - Transport conditions of imported horse meat (frozen, chilled, etc) - Prevalence of VEE in horses in countries from which meat is imported - Can VEE survive in horse meat? For how long?
Meat is eaten by host animal/human - Consumption by human - Consumption by host animal - Consumption of waste by host animal		- Amount of horse meat consumption per person in EU - Percentage of this horse meat consumption which is originating from VEE risk countries - Amount of horse meat that is fed to VEE host animals in the EU. - Percentage of this horse meat consumption which is originating from VEE risk countries
Host animal develops infection		- Is infection transferrable by eating meat? - Dose of VEE virus necessary for infecting host - Is this dose present in meat? - Percentage of horse meat that is eaten raw - Survival of VEE virus in cooked meat - Survival of VEE virus in animal feed (during processing etc)
Host animal is bitten by vector	See above	
Etc	See above	
Pathway step 3	**Likelihood depends on**	**Data needed to estimate likelihood**
Import VEE virus by import of vaccines	- If a live vaccine is available	- Is a vaccine against VEE available? - Details about the vaccine (live, modified, inactivated, virus strain) - Number of VEE vaccines imported into/produced in the EU
Horses are vaccinated with the vaccine		- Number of horses vaccinated against VEE in EU
Vaccine virus mutates in horse in virulent strain	- Type of vaccine	- Type of vaccine - Virus strain in vaccine - Historic data about mutation of virus strain to virulent strain

Horse is bitten by a vector	
Etc	

Table 63: Data needed to estimate the likelihood of the VEEv introduction in relation with the "import" of VEEv infected human

Pathway step 4	Likelihood depends on	Data needed to estimate likelihood
Entry of VEE infected human in EU		- Number of people travelling from VEE risk countries to EU - Prevalence of VEE in persons from VEE risk countries
Infection is not detected in country of origin	- Showing symptoms of disease - Tests performed - Incubation period	- Knowledge about clinical disease, do humans show symptoms? Which symptoms? - Are symptoms distinguishable from other diseases symptoms?
Human is infected after travel	- Duration of travel - Length of incubation period, disease, recovery, immunity - Is incubation period longer than travel time and quarantine time?	- Length of incubation period, viremic period in humans - Likelihood of detection of infection of VEE in humans in Europe - Knowledge of doctors in EU about VEE/human encephalitis - Which tests are available? - Sensitivity/specificity of test - Surveillance system in importing country?
Human is bitten by a competent vector	- Availability of competent vectors in country of import (EU) - Number of competent vectors - Season (are vectors active?)	- Which vectors are known to be able to transmit VEE? - Are these present in EU? - Where? - Season/temperature in which these vectors are active
Etc, see above		

CONCLUSIONS

WP1: Typology of risk questions and identification of data needs

Twelve types of questions were identified.

The top three risk question types were, by decreasing order, concern: the risk of (re)introduction (21 opinions), the risk of potential spread to susceptible population, the pathways of transmission and speed of the spread (20 opinions), and the effectiveness of control measures (17 opinions).

Scientific opinions are dealing in general with more than two types of questions.

We notice absence of a formal working procedure that helps to achieve adequacy between the used methodologies, the questions and the available data.

The conducted retrospective analysis of AHAW opinions pointed out several difficulties encountered by AHAW experts to address quantitatively the risk questions. The main difficulties are associated to data availability. Data gaps are in general recognized. However, needs of collecting new data are not prioritised in regard to their added value on facilitating the answer to risk managers questions.

Data needs were classified in 5 categories: disease general information, descriptive epidemiological data, analytical epidemiological data, prevention and control data, and disease surveillance data. Within each category, three levels of subcategories were defined. This classification permits the establishment of a comprehensive list of data needs independently to the potential risk questions.

WP2: Data availability

The overview of these three different animal health data collection systems illustrates how similar their organisation is. Their apparent differences mainly rely on the different maturity stages of the systems, the management of collection systems (unique vs. multiple) and the level of integration of animal health with other compartments of the food chain. By order of increased system maturity, one could indeed propose the following countries: France, Netherland and Belgium. Except for Belgium (multiple databases which are continuously updated, interconnected and well centralized), other systems aim at getting centralized without centralizing their collection of data itself. Despite the decentralization of data collection gives to countries the possibility to collect wide range of specific and precise field data, it sometimes leads to growing polemics about data ownership that often hamper the easy access of available data for animal health professionals. Consultation and collaboration of all actors and stakeholders involved in data collection systems should be encouraged to increase the maturity of all systems and to ensure useful risk assessment and reactivity (adaptability) to the discovery of new diseases.

Differences of data availability either in relation with the status of the disease targeted or with the different categories of data are observed (e.g. descriptive epidemiology vs. public health). For example, few data are available on VEE, an exotic disease. When a disease is endemic (PRRS) or of limited interest due to its limited impact on public health (few annual human cases such as Echinoccocosis), it is poorly surveyed if not at all.

Numerous types of data resources were identified. Nevertheless, for the vast majority of diseases, the disease-specific information will be found in textbooks, papers, scientific literature, etc. Websites represented the majority of resources regarding Member States data.

The main forms of data resources were PDF and HTML files. Although raw tabulated data are more appropriate for risk assessment, these resources are not often available and sometimes difficult to access (e.g. restricted or paying access). In addition, a restricted or paying access may represent a constraint to conduct a risk assessment in a short timeframe. Few data were in a form directly exploitable. Indeed, the access to Oracle or other databases compiling data directly exploitable is probably more often restricted or not directly available through a classical search. Furthermore, the delay spent to perform web searches was limited to 10 minutes per data. A longer time would potentially provide additional data resources. Furthermore, a look at some of the data resources particularly those identified in all diseases, would have allowed a better understanding and metadata specification.

The lack of availability was mainly observed for prevention and control, surveillance and public health. The accessibility of data is generally related to their availability.

The lack of availability was mainly observed for prevention and control, surveillance and public health.

Data resources gathered in the inventory will be useful for future EFSA and/or national risk assessments, among others, to prepare future mandates.

Because the same data resource can be reported for several data, a total of 471 different data resources were inventoried.

WP3: Data specification, validation and management

Data collection and metadata elements needed for animal health scientific and risk assessment were defined based on approved Dublin Core Metadata and existing EFSA data standards.

In order to facilitate the data specification facts and group of facts and their corresponding metadata were defined.

To link WP3 with WP1 and WP2 tables associating facts with type of risk questions and data categories and subcategories were created.

The link between all the created tables (data categories, type of questions, facts, and metadata) enable different types of queries starting from the type of questions, ending with the needed data or more specifically facts or group of facts, their associated metadata and their possible resources (WP2).

WP4: Methodological framework

Three main types of methods were distinguished: 1) Statistical methods: that may be for example suitable for endemic diseases, where it is needed to assess the disease frequency (Prevalence/Incidence) and to assess possible growing or declining of the disease frequency. Statistical methods are also useful to assess specific correlation or association between exposure and occurrence of diseases or the effect of an intervention measure at the point of its application; 2) Linear probabilistic risk assessment: that may be for example appropriate to assess the probability of an exotic agent entrance; and 3) Mechanistic or dynamic models: that may be for example suitable to describe and assess the spread of a disease in a certain population.

For each type of methods data needs and their suitability to answer the risk questions were provided.

PERSPECTIVES

Thanks to the review of the different opinions published by EFSA from 2004 to 2010 (WP1), only twelve risk questions are found to be the most recurrent. Whatever the disease and its epidemiological context, we see that these questions don't vary that much, confirming the possibility for EFSA to **anticipate** its **data and methodological needs**. The review of its opinions also allows AHAW to identify its current **data gaps**. In order to reinforce the anticipation of its panel needs and orientate its future MSs' data collection calls, AHAW should promote the analysis of **AHAW risk questions' real-time evolution and the inventory of data availability**, directly related to the EU AH new trends and focus' changes. The analysis of the four case studies shows how the DATASPEC general methodological approach could lead to more efficient preparatory data-collection, therefore enabling more rapid response to new risk managers questions.

The different WPs show what are AHAW **facts**' needs and their possible **resources** and suggest a **metadata model** to structure and organise it. There is indeed a need for AHAW to house and organised the data yet- and subsequently-collected to facilitate its **reuse** for future opinions. The model indeed aims at creating a global database of AH knowledge, insuring the anticipation of future opinions needs and both information's **traceability** and **capitalization**. This database should be regularly updated, therefore providing an **information platform** for the integration, management and sharing of heterogeneous data, information and knowledge resources needed for animal disease risk assessment.

The WP3 is to develop a methodology for data collection including the definition of metadata standards for outcomes values to support data **validation and quality assessment**. After identification of data resources, protocols and methodology are needed in order to extract properly the data needed for the risk assessment question. Currently, there is not metadata model publicly available for the specific annotation of animal infectious disease knowledge.

As planned in the proposal of this project, EFSA need to define mechanisms and protocols allowing data collection from Member states. The use of term 'collection' is ambiguous, because what is expected from MS is not to collect the data specifically for EFSA but to extract EFSA required data from existing systems. It is more an activity of data assembly and transfer. The developed **facts technical cards**, which detail for each given needed fact, the type, the related metadata, the format and the informatics channels and vehicles that could be used to transfer them correctly to EFSA, could therefore be used as back-up material, but would need to be made operational and validate by MSs.

REFERENCES

Bauer C, 2009. Is the cat a source of human echinococcosis? 7th Merial European Symposium on Parasitosis and Arthropod-Borne Diseases, Athens, Greece, May 6-8.

Belgian federal agency for the safety of the food chain (FASFC) Scientific Committee, 2007. Scientific exploitation of databases within the frameworks of food safety risk assessment. Sci. Com. Workshop- http://www.favv.be/comitescientifique/publications/_documents/2007_WS_SciCom_en.pdf [consulted on July 20, 2011].

Berkvens D, Speybroeck N, Praet N, Adel A, Lesaffre E, 2006. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. Epidemiology 17, 145-53.

Bichet H, Dorchies P, 1998. Prevalence of hydatidosis in cattle in the Midi-Pyrenees. Parasite 5, 61-68.

Bødker, R, Maddox-Hyttel, C, Vigre H (Institute for Infektiøse Husdyrsygdomme), 2006. Evaluation of the risk of *Echinococcus multilocularis* spreading from Denmark to Sweden. J. nr. 55017-0005, 34 pp.

Branscum AJ, Gardner IA, Johnson WO, 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev Vet Med. 68, 145-63.

Breiman L, Friedman JH, Olshen RA, Stone CJ, 1984. Classification and regression trees. Wadsworth, Pacific Grove, California, United States.

Brochier B, De Blander H, Hanosset R, Berkvens D, Losson B, Saegerman C, 2007. Echinococcus multilocularis and Toxocara canis in urban red foxes (Vulpes vulpes) in Brussels, Belgium. Prev Vet Med. 80, 65-73.

Cardoen S, Berkvens D, Claes L, Van Gucht S, Dewulf J, De Zutter L, Saegerman C, 2010. Qualitative and quantitative risk evaluation of the risks associated with Trichinella in Belgium: current status and prospects. Epidémiol. et Santé Anim. 57, 31-43.

Carmena D, Sánchez-Serrano LP, Barbero-Martínez I, 2008. *Echinococcus granulosus* infection in Spain. Zoonoses Public Health. 55, 156-65.

Christodoulopoulos G, Theodoropoulos G, Petrakos G, 2008. Epidemiological survey of cestode-larva disease in Greek sheep flocks. Vet Parasitol. 153, 368-373.

Dakkak A, 2010. Echinococcosis/hydatidosis: a severe threat in Mediterranean countries. Vet Parasitol. 174, 2-11.

Deplazes P, Hegglin D, Gloor S, Romig T, 2004. Wilderness in the city: the urbanization of *E. multilocularis*. Trends in Parasitology 20, 77-84.

Derylo A, Szilman P, 1998. Prevalence of echinococcosis in swine, beef-cattle, sheep and goats in Poland during 1995-1997. Wiadom. Parazytol. 44, 413.

Dohoo Ian, Wayne Martin, Henrik Stryhn, 2003. Veterinary Epidemiologic Research.

Dórea FC, Sanchez J, Revie CW, 2011. Veterinary syndromic surveillance: current initiatives and potential for development. Prev. Vet. Med. 101, 1-17.

Dorny P, Phiri IK, Vercruysse J, Gabriel S, Willingham AL III, Brandt J, Victor B, Speybroeck N, Berkvens D, 2004. A Bayesian approach for estimating values for prevalence and diagnostic test characteristics of porcine cysticercosis. Int. J. Parasitol. 34, 569-576.

Dublin Core Metadata Initiative, 2010. http://dublincore.org/ [consulted on September 2012].

Dubois MA, 2005. Modélisation en épidémiologie : objectifs et méthodes. Épidémiol. Santé anim. 47, 1-13.

Eckert J, Deplazes P, 2004. Biological, epidemiological, and clinical aspects of echinococcosis, a zoonosis of increasing concern. Clin Microbiol Rev. 17, 107-35.

Eckert J, Gemmell MA, Meslin FX, Pawlowski ZS, Eds. – WHO-OIE, 2002. WHO/OIE Manual on Echinococcosis in Humans and Animals: a Public Health Problem of Global Concern, 286 pp. http://whqlibdoc.who.int/publications/2001/929044522X.pdf [consulted on January 23rd, 2012].

EPA, 2006. Guidance on systematic planning using the data quality objectives processes. Office of environmental information. http://www.epa.gov/quality/qs-docs/g4-final.pdf [consulted on August 2012].

European Commission, 2007. Handbook on Data Quality Assessment Methods and Tools. http://unstats.un.org/unsd/dnss/docs-nqaf/Eurostat-HANDBOOK%20ON%20DATA%20QUALITY%20ASSESSMENT%20METHODS%20AND%20TOOLS%20%20I.pdf [consulted on July 20, 2011].

European Commission, Directorate-General for Health and Consumers, 2010. Animal Disease Notification System (ADNS) – Annual Report 2009, 3. http://ec.europa.eu/food/animal/diseases/adns/adns_report2009_en.pdf [consulted on July 20, 2011].

European Commission Delegated Regulation (EU) No 1152-2011 of 14 July 2011 supplementing Regulation (EC) No 998/2003 of the European Parliament and of the Council as regards preventive health measures for the control of Echinococcus multilocularis infection in dogs. http://eur-lex.europa.eu/Notice.do?mode=dbl&lng1=en,pt&lang=&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=626836:cs&page=&hwords=null [consulted on February 2, 2012].

EFSA, 2010. Development of harmonized schemes for the monitoring and reporting of *Echinococcus* in animals and foodstuffs in the EU – Question No EFSA-Q-2009-01071.

EFSA, 2010. Käsbohrer A., Tenhagen B-A., Appel B. and Fetsch A. Development of harmonized survey methods for food-borne pathogens in foodstuffs in the European Union. Scientific Report. Publication on 22 november 2010.

EFSA, 2010. Standard sample description for food and feed. EFSA Journal 2010;8(1):1457 [54 pp.] doi:10.2903/j.efsa.2010.1457. Available online: www.efsa.europa.eu.

EFSA, 2011. Manuel for reporting on Zoonoses, Zoonotic Agents and Antimicrobial Resistance in the framework of Directive 2003/99/EC and some other pathogenic microbiological agents for information derived from the year 2010. Supporting Publications 2011:135 [119 pp.]. Available online: www.efsa.europa.eu.

EFSA DWH, 2011. European Food Safety Authority: Logical data model proposal for the EFSA Data Warehouse (DWH). SAS Internal Technical Report. 2011.

European Parliament and Council, 2004. Regulation No 854/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption. *Off. J. Eur. Union*, L 155/206.

EUROSTAT, 2007. Handbook on quality Assessment Methods and Tools. http://unstats.un.org/unsd/dnss/docs-nqaf/Eurostat-HANDBOOK%20ON%20DATA%20QUALITY%20ASSESSMENT%20METHODS%20AND%20TOOLS%20%20I.pdf. [consulted on August 2012].

EUROSTAT European Commision, 2010. Typology of sampling strategies. Item 4.4.b. of the Agenda Updated Methodological information. Working Group "Food Safety Statistic". Doc. ESTAT/F5/ES/201.

Evans, C.M., Medley, G.F., Creasey, S.J. and Green L.E., 2010. A stochastic mathematical model of the within-herd transmission dynamics of porcine reproductive and respiratory syndrome virus (PRRSV): Fade-out and persistence. Prev. Vet. Med, 93: 248-257.

Food and Agriculture Organisation (FAO), 2004. Section 6: Ante-mortem inspection. In. Good practices for the meat industry, 322 pp.

Frodsham A., 2007. Guideline for the Epidemiological Ontology, Epidemiology Ontology-Web version I 18/07/2007 http://www.hugenet.org.uk/resources/informatics/Ontology_Version_1.pdf [consulted on September 2012].

Geysen D, Kanobana K, Victor B, Rodriguez-Hidalgo R, de Borchgrave J, Brandt J, Dorny P, 2007. Validation of meat inspection results for *Taenia saginata* cysticercosis by PCR-Restriction fragment length polymorphism. J Food Prot. 70, 236-40.

Guatteo R, Seegers H, Taurel AF, Joly A., Beaudeau F, 2011. Prevalence of *Coxiella burnetii* infection in domestic ruminants: a critical review. Vet. Microbiol. 149, 1-16.

Hadorn DC, Stärk KDC, 2008. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. Vet. Res. 39, 57.

Hanosset R, Saegerman C, Adant S, Massart L, Losson B, 2008. Echinococcus multilocularis in Belgium: prevalence in red foxes (Vulpes vulpes) and in different species of potential intermediate hosts. Vet Parasitol. 151, 212-7.

Humblet M-F, Boschiroli M-L, Saegerman C, 2009. Classification of worldwide bovine tuberculosis risk factors in cattle: a stratified approach. Vet. Res. 40, 50.

Humblet M-F, Gilbert M, Govaerts M, Fauville-Dufaux M, Walravens K, Saegerman C, 2010. New assessment of bovine tuberculosis risk factors in Belgium based on nationwide molecular epidemiology. J. Clin. Microbiol. 48, 2802-2808.

Humblet M-F, Moyen J-L, Bardoux P, Boschiroli M-L, Saegerman C, 2011. Importance of awareness for veterinarians involved in cattle tuberculosis skin testing. Transbound. Emerg. Dis. 58, 531-6.

Humblet M-F, Moyen J-L, Bardoux P, Boschiroli ML, Saegerman C, 2011. The Importance of Awareness for Veterinarians Involved in Cattle Tuberculosis Skin Testing. Transbound. Emerg. Dis. doi: 10.1111/j.1865-1682.2011.01228.x.

Khosrowpour M, 2007. Dictionary of information science and technology. Ed. Idea Group Inc (IGI), London, 761 pp.

Kirch W, 2008. Encyclopaedia of Public Health. Ed. Springer Science, New York, 1061 pp.

Kolarova L, 1997. Die Echinokokkosen in der Tschechischen Repulik. Mitt. Österr. Ges. Tropenmed. Parasitol. 19, 65-70.

Lesaffre E, Speybroeck N, Berkvens D, 2007. Bayes and diagnostic testing. Vet Parasitol. 148, 58-61.

Lloyd S, Walters TMH, Craig PS, 1998. Use of sentinel lambs to survey the effect of an education programme on control of transmission of *Echinococcus granulosus* in south Powys, Wales. Bull. WHO 76, 469-473.

Loshin D, 2008. Monitoring Data Quality Performance Using Data Quality Metrics. Informatica, 22 pp.

http://www.informationweek.com/whitepaper/Business_Intelligence/wp100380?articleID=100380 [consulted on October 5, 2011].

Losson B, Kervyn T, Detry J, Pastoret PP, Mignon B, Brochier B, 2003. Prevalence of *Echinococcus multilocularis* in the red fox (*Vulpes vulpes*) in southern Belgium. Vet Parasitol. 117, 23-28.

Lurette, A., Belloc, C., Touzeau, S., Hoch, T., Seegers, H., Fourichon, C., 2008. Modelling batch farrowing management within a farrow-to-finish pig herd: influence of management on contact structure and pig delivery to the slaughterhouse. Animal 2, 105–116.

Martin PAJ, Carmeron AR, Greiner M, 2007. Demonstrating freedom from disease using multiple complex data sources 1: A new methodology based on scenario trees. Prev Vet Med 79, 71-97.

Merianos A, 2007. Surveillance and response to disease emergence. Curr. Top. Microbiol. Immunol. 315, 477-508.

Méroc E, Herr C, Verheyden B, Hooyberghs J, Houdart P, Raemaekers M, Vandenbussche F, De Clercq K, Mintiens K, 2009. Bluetongue in Belgium: episode II. Transbound. Emerg. Dis. 56, 39-48.

Ministère de l'Agriculture, de l'alimentation, de la pêche, de la ruralité et de l'aménagement du territoire, Direction générale de l'alimentation (DGAL), service de l'alimentation, sous-direction de la sécurité sanitaire des aliments, bureau des zoonoses et de la microbiologie alimentaires (Décembre 2011) Plan de surveillance de la contamination des viscères d'animaux de boucherie par Echinococcus granulosus au stade de l'abattoir – 2012. Note de service DGAL/SDSSA/N2011-8280. http://agriculture.gouv.fr/IMG/pdf/DGALN20118280Z.pdf [Consulted on February 1st, 2012].

Morgan ER, 2008. *Echinococcus multilocularis* in veterinary practice in Europe. EJCAP 18, 255-258.

Morris RS and Pfeiffer DU, 1995. Directions and issues in bovine tuberculosis epidemiology and control in New Zealand. N. Z. Vet. J. 43, 256-265.

Neghina R, Neghina AM, Marincu I, Iacobiciu I, 2010. Epidemiology and epizootology of cystic echinococcosis in Romania 1862-2007. Foodborne Pathog Dis. 7, 613-8.

Nodelijk, G., De Jong, M.C.M., Van Nes, A., Vernooy, J.C.M., Van Leengoed, L.A.M.G., Pol, J.M.A., Verheijden, J.H.M., 2000. Introduction, persistence and fade-out of porcine reproductive and respiratory syndrome virus in a Dutch breeding herd: a mathematical analysis. Epidemiology and Infection 124, 173–182.

OECD, 2008. OECD Glossary of Statistical Terms. http://www.oecd-ilibrary.org/docserver/download/fulltext/3008121e.pdf?expires=1324585706&id=id&accname=guest&checksum=D7FA0A575B73177052FAB7E8E10694E6 [consulted on December 22, 2011].

OIE, 1998. World animal health in 1997. Part 2. OIE, Paris, 422-754.

OIE, 2008. Chapter 2.1.4 – Echinococcosis/hydatidosis. In OIE Terrestrial Manual 2008. pp. 175-189. http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.01.04_ECHINOCOCCOSIS.pdf [Consulted on January 23rd, 2012].

OIE Terrestrial Manual 2009, Glossary of Terms, Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2009. http://web.oie.int/eng/normes/MMANUAL/A_Index.htm [consulted on October 3rd, 2012].

OIE Terrestrial Animal Health Code 2011, Glossary of the Terrestrial Animal Health Code (2011). http://www.oie.int/index.php?id=169&L=0&htmfile=glossaire.htm#sous-chapitre-2 [consulted on October 3rd, 2012].

Olson JE, 2003. Chapter 2: Definition of Accurate Data. In: Data quality: the accuracy dimension. Ed. Morgan Kaufmann, 293 pp.

Olteanu C, Panaitescu D, Ghernian I, 1997. Echinococcosis/hydatidosis (E/H) in man and animals in Romania. Arch. Int. Hydatid. 32, 292-293.

OMG, 2012. Object Management Group, Inc., Object Management Group Terms and Acronyms.

Palmer SR, Biffin AH, Craig PS, Walters TM, 1995. Control of hydatid disease in Wales. Siriraj Hosp. Gaz. 47, 674-675.

Parola P, Raoult D, 2001. Ticks and tick-borne bacterial diseases in humans: an emerging infectious threat. Clin. Infect. Dis. 32, 897-928.

Petrie A, Watson P, 2006. Statistics for veterinary and animal science. Blackwell Publishing, Oxford, 299 pp.

Pfeffer, M and Dobler, G., 2010. Emergence of zoonotic arboviruses by animal trade and migration. Parasites &Vectors 2010, 3:35.

Pitkin, A., Otake, S., and Dee. S., 2008. Biosecurity protocols for the prevention of spread of porcine reproductive and respiratory syndrome virus. Swine Disease Eradication Center. University of Minnesota College of veterinary Medicine. http://www.aasv.org/aasv/PRRSV_BiosecurityManual.pdf [accessed on April 16th].

Porta M, 2008. A dictionary of Epidemiology – Fifth edition. Oxford University Press, UK. 316 pp.

Porter RS, Leblond A, Lecollinet S, Tritz P, Cantile C, Kutasi O, Zientara S, Pradier S, van Galen G, Speybroek N, Saegerman C, 2011. Clinical diagnosis of West Nile fever in equids by classification and regression tree (CART) analysis and comparative study of clinical appearance in three European countries. Transbound. emerg. Dis. 58, 197-205.

Praet N, Dorny P, Saegerman C, Marcotty T, Berkvens D, 2006. Estimation de la prévalence d'une maladie et des caractéristiques des tests diagnostiques par une approche bayésienne. Epidémiol. Santé anim. 49, 113-130.

Report of Task Force on Zoonoses Data Collection on Guidance Document on Good Practices for Design of Field Surveys. *The EFSA Journal* (2006), 93, 1-24.

Saegerman C, De Blander H, Hanosset R, Berkvens D, Losson B, Brochier B, 2006. Evaluation des risques liés à la présence d'*Echinococcus multilocularis* et de *Toxocara canis* dans la population vulpine en région Bruxelloise. Epidémiol. et santé anim. 50, 97-104.

Saegerman C, Mellor P, Uyttenhoef A, Hanon J-B, Kirschvink N, Haubruge E, Delcroix P, Houtain J-Y, Pourquier P, Vandenbussche F, Verheyden B, De Clercq K, Czaplicki G, 2010. The most likely time and place of introduction of BTV8 into Belgian ruminants. PloS ONE 5:e9405.

Saegerman C, Speybroeck N, Roels S, Vanopdenbosch E, Thiry E, Berkvens D, 2004. Decision support tools for clinical diagnosis of disease in cows with suspected bovine spongiform encephalopathy. J. clin. Microbiol. 42, 172-178.

Saegerman C, Speybroeck N, Vanopdenbosch E, Wilesmith JW, Berkvens D, 2006. Trends in age at detection in cases of bovine spongiform encephalopathy in Belgium: an indicator of the epidemic curve. Vet. Rec. 159, 583-587.

Schweiger A, Ammann RW, Candinas D, Clavien PA, Eckert J, Gottstein B, Halkic N, Muellhaupt B, Prinz BM, Reichen J, Tarr PE, Torgerson PR, Deplazes P, 2007. Human alveolar echinococcosis after fox population increase, Switzerland. Emerg. Infect. Dis. 13, 878-882.

Singer A, 2010. EFSA- External report reviewing the previous opinions of the Panel on Animal Health and Welfare concerning the application of quantitative tools, in the sequence of the current self mandate on "Good Practice in Conducting Scientific Assessments in Animal Health Using Modelling". Publication on 4 march 2010.

Sorensen J.T., Edwards S., Noordhuizen J., Gunnarsson S., 2006. Animal production Systems in the industrialised world. Rev. Sci. Tech. Off. Int. Epiz;, 2006, 25(2), 493-503.

Sotiraki S, Chaligiannis I, 2010. Cystic echinococcosis in Greece. Past and present. Parasite 17, 205-10.

Soulé C, Fabien JF, Maillot E, 1995. Animal echinococcosis in France: national survey in slaughterhouses in 1989. Epidémiol. Santé anim. 16, 131-141.

Sparks G. 2011. "Database Modelling in UML". Retrieved 8 September 2011.

Speybroeck N, Berkvens D, Mfoukou-Ntsakala A, Aerts M, Hens N, Huylenbroeck GV, Thys E, 2004. Classification trees versus multinomial models in the analysis of urban farming systems in central Africa. Agric. Syst. 80, 133-149.

Stachowiak H, 1973. Allgemeine Modelltheorie. Springer, Vienna, New York.

STROBE, 2007. Strobe Statement-Strengthening the reporting of observational studies in epidemiology. http://www.strobe-statement.org/index.php?id=available-checklists [consulted on August 2012].

Todorov T & Boeva V, 1997. Epidemiology of echinococcosis in Bulgaria – a comparative study. Arch. Int. Hydatid. 32, 232-233.

Toma B, Dufour B, Sanaa M, Benet J-J, Shaw A, Moutou F, Louza A, 2001. Epidémiologie appliquée à la lutte collective contre les maladies animales transmissibles majeures. 2ème Edition. AEEMA, Maison-Alfort, France.

Torgerson PR and Craig PS, 2009. Risk assessment of importation of dogs infected with Echinococcus multilocularis in to the UK. Vet Rec. 165, 366-8.

UN, Commodity List. United Nations Commodity Trade Statistics Database http://comtrade.un.org/db/mr/rfCommoditiesList.aspx [Consulted on January 2012].

USDA, 2006. Surveillance and Data Standards for USDA/APHIS/Veterinary Services, v 1.0, July 11, 2006. Prepared by the Centers for Epidemiology and Animal Health.

Van Gucht S, Van Den Berge K, Quataert P, Veschelde P, Le Roux I, 2010. No emergence of *Echinococcus multilocularis* in foxes in Flanders and Brussels Anno 2007-2008. Zoonoses Public Health 57, e65-e70.

Vandeweerd J-M, Saegerman C, 2009. Guide pratique de médecine factuelle vétérinaire. Les Editions du Point Vétérinaire, Paris, 193 pp.

VanEngelsdorp D, Evans JD, Saegerman C, Mullin C, Haubruge E, Nguyen BK, Frazier M, Frazier J, Cox-Foster D, Chen Y, Underwood R, Tarpy DR, Pettis JS, 2009. Colony collapse disorder: a descriptive study. PloS ONE 4, e6481.

VanEngelsdorp D, Speybroeck N, Evans J, Nguyen BK, Mullin C, Frazier M, Frazier J, Cox-Foster D, Chen Y, Tarpy DR, Haubruge H, Pettis JS, Saegerman C, 2010. Weighing risk factors associated with bee colony collapse disorder by classification and regression tree analysis. J. econ. Entomol. 103, 1517-1523.

Varcasia A, Canu S, Kogkos A, Pipia AP, Scala A, Garippa G, Seimenis A, 2007. Molecular characterization of *Echinococcus granulosus* in sheep and goats of Peloponnesus, Greece. Parasitol. Res. 101, 1135-1139.

Vervaeke M, Dorny P, Vercammen F, Geerts S, Brandt J, Van Den Berge K, Verhagen R, 2003. *Echinococcus multilocularis* (Cestoda, Taeniidae) in Red foxes (*Vulpes vulpes*) in northern Belgium. Vet Parasitol. 115, 257-263.

Vieira da Silva Lopez.P., 2010. A metadata model for the annotation of epidemiological data. Lisbon University, Faculdade de Ciências, Departamento de Informatica. Mestrado em Tecnologias da Informação Aplicadas às Ciências Biológicas e Médicas.

W3C, 2011. "Web Services Glossary". W3C. February 11, 2004. Retrieved 2011-04-22.

Wahlström H, Isomursu M, Hallgren G, Christensson D, Cedersmyg M, Wallensten A, Hjertqvist M, Davidson RK, Uhlhorn H, Hopp P, 2011. Combining information from surveys of several species to estimate the probability of freedom from *Echinococcus multilocularis* in Sweden, Finland and mainland Norway. Acta Vet Scand. 53, 9.

WHO, 2011. World Health Organization. Health action in crises. Definitions: Emergencies. http://www.who.int/hac/about/definitions/en/index.html [consulted on December, 2011].

Wilesmith J, Morris RD, Stevenson MA, Cannon R, Prattley D, Benard HJ, 2004. Development of a method for evaluation of national surveillance data and optimization of national surveillance strategies for bovine spongiform encephalopathy. Report submitted to DG SANCO, Brussels, 34 pp.

Working group on foodborne infections and intoxications, 2003. Report on zoonotic agents in Belgium. http://www.favv.be/publications-en/_documents/2003_Report-on-zoonotic-agents_en.pdf [Consulted on January 23rd, 2012].

APPENDICES

A. LIST OF THE REVIEWED AHAW OPINIONS SORTED BY EFSA QUESTION NUMBER

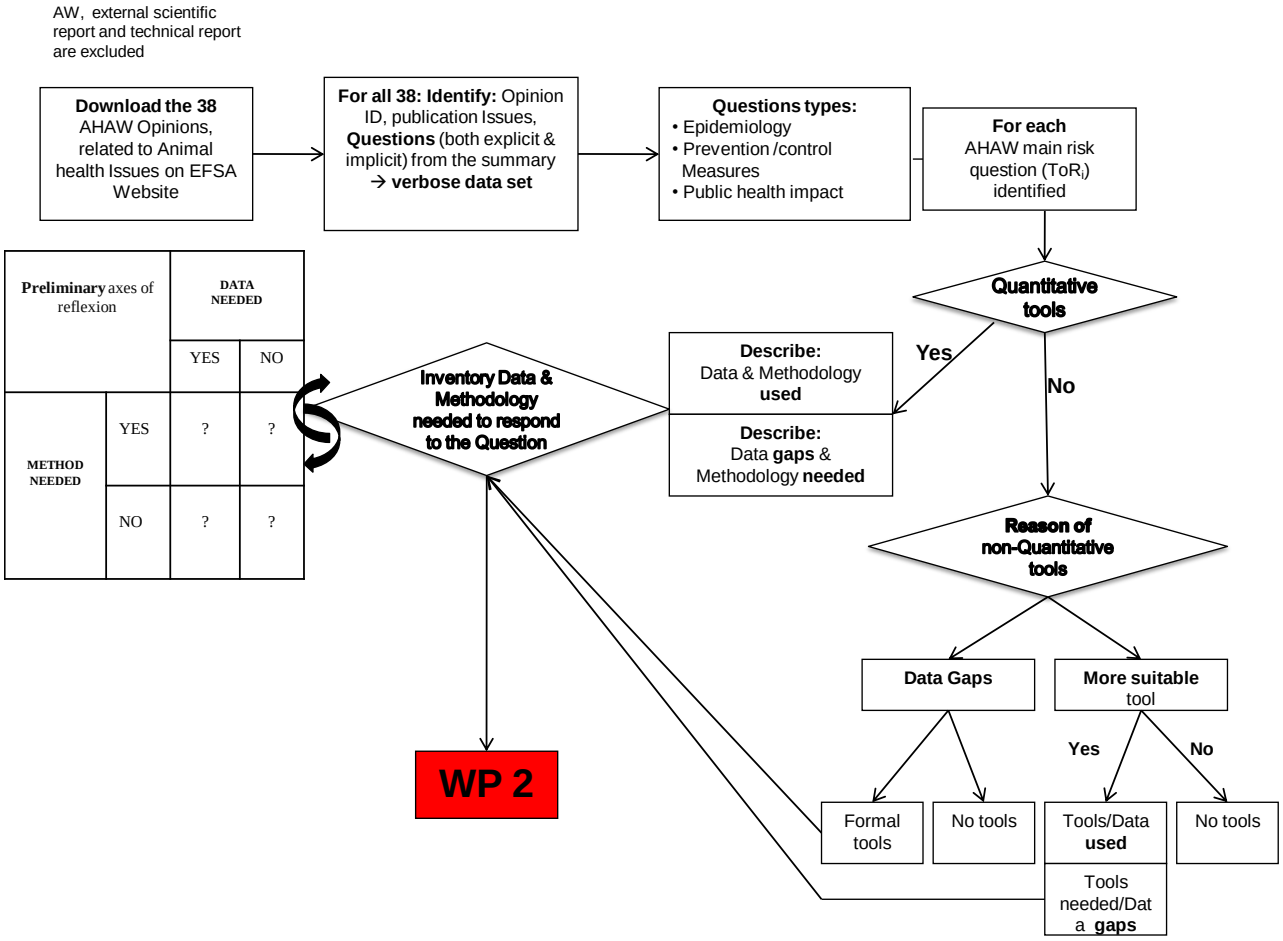
ID	EFSA question	Document
1	EFSA-Q-2004-005	- The EFSA Journal (2004) 110, 1-59, the risk of transmission of Mycobacterium avium subsp. paratuberculosis via bovine semen
2	EFSA-Q-2004-050	- The EFSA Journal (2005) 238, 1-128. The risk of a Rift Valley fever incursion and its persistence within the Community
3	EFSA-Q-2004-075	- The EFSA Journal (2005) 266, 1-21. Animal health and welfare aspects of Avian Influenza - Annex to The EFSA Journal (2005) 266, 1-21; Animal health and welfare aspects of Avian Influenza
4	EFSA-Q-2004-100	- The EFSA Journal (2005) 239, 1-85. The probability of transmission of Porcine Reproductive and Respiratory Syndrome virus (PRRSv) to naive pigs via fresh meat
5	EFSA-Q-2004-113	- The EFSA Journal (2006) 313, 1-34, Risk Assessment on Foot and Mouth Disease - The EFSA Journal (2006) 313, Risk Assessment on Foot and Mouth Disease: Part 1 - The EFSA Journal (2006) 313, Risk Assessment for Foot and Mouth Disease: Part 2 - The EFSA Journal (2006) 313, Risk Assessment on Foot and Mouth Disease: Part 3 - The EFSA Journal (2006) 313, Risk Assessment for Foot and Mouth Disease: References
6	EFSA-Q-2004-161	- The EFSA Journal (2006) 347, 1-21, "Animal health risks of feeding animals with ready-to-use dairy products without further treatment" - Annex to the EFSA Journal (2006) 347, 1-21. Animal health risks of feeding animals with ready to use dairy products without further treatment.
7	EFSA-Q-2005-018	- The EFSA Journal (2006) 311, 1-20 - Opinion on the "Definition of a BoHV-1-free animal and a BoHV-1-free holding, and the procedures to verify and maintain this status" - Annex to the EFSA Journal (2006) 311, 1- 65; Definition of a BoHV-1-free animal and a BoHV-1- free holding, and the procedures to verify and maintain this status.
8	EFSA-Q-2005-057	- The EFSA Journal (2006) 410, 1-55, Scientific Opinion on "Animal health and welfare risks associated with the import of wild birds other than poultry into the European Union" - Annex to the EFSA Journal (2006) 410, 1-55, "Animal health and welfare risks associated with the import of wild birds other than poultry into the European Union"
9	EFSA-Q-2005-060	- The EFSA Journal (2006) 432, 1-44 Scientific Opinion on "Performance of Brucellosis Diagnostic Methods for Bovines, Sheep, and Goats" - Annex to the EFSA Journal (2006) 432, 1-44, Scientific Opinion on "Performance of Brucellosis Diagnostic Methods for Bovines, Sheep, and Goats"
10	EFSA-Q-2005-243	- The EFSA Journal (2006) 357, 1-46, Opinion on "Migratory birds and their possible role in the spread of highly pathogenic Avian Influenza" - Annex to The EFSA Journal (2006) 357, 1-46, "Migratory birds and their possible role in the spread of highly pathogenic Avian Influenza" - Annex to The EFSA Journal (2006) 357, 1-18, Addendum to the Scientific Opinion on "Migratory birds and their possible role in the spread of highly pathogenic Avian Influenza (EFSA-Q-2005-243)" - Annex with figures (2005-243 ahaw_op_ej357_migratorybirds_1_annex.pdf)

11	EFSA-Q-2006-014	- The EFSA Journal (2006) 436 1-54, "Assessment of the risk of rabies introduction into the UK, Ireland, Sweden, Malta, as a consequence of abandoning the serological test measuring protective antibodies to rabies" - Excel-worksheet presenting the model used in "Assessment of the risk of rabies introduction into the UK, Ireland, Sweden, Malta, as a consequence of abandoning the serological test measuring protective antibodies to rabies"
12	EFSA-Q-2006-050/051	-The EFSA Journal (2006) 403 1-62, " Review of the Community Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Antimicrobial Resistance in the European Union in 2004"
13	EFSA-Q-2006-112	- The EFSA Journal (2006) 441, 1-54, "Assessment of the risk of Echinococcosis introduction into the UK, Ireland, Sweden, Malta and Finland as a consequence of abandoning national rules"
14	EFSA-Q-2006-145	- The EFSA Journal (2007) 477, 1-25, Opinion of the Scientific Panel on Animal Health and Animal Welfare regarding a request from the European Commission to review ND focusing on vaccination worldwide in order to determine its optimal use for disease control purposes - Annex to the EFSA Journal (2007) 477, 1-24. Opinion of the Scientific Panel on Animal Health and Animal Welfare: "Review on Newcastle disease focusing on vaccination worldwide in order to determine its optimal use for disease control purposes."
15	EFSA-Q-2006-156	- The EFSA Journal (2007) 450, Scientific Opinion on "Vaccination against avian influenza of H5 and H7 subtypes as a preventive measure carried out in Member States in birds kept in zoos under Community approved programs"
16	EFSA-Q-2006-179	- The EFSA Journal (2008) 645, 1-34, Opinion of the Scientific Panel on Animal Health and Animal Welfare: "Tuberculosis testing in deer" - Annex to the EFSA Journal (2008) 645, 1-34; Opinion of the Scientific Panel on Animal Health and Animal Welfare: "Tuberculosis testing in deer"
17	EFSA-Q-2006-309	- The EFSA Journal (2007) 489, Scientific Opinion on "Vaccination against avian influenza of H5 and H7 subtypes in domestic poultry and captive birds"
18	EFSA-Q-2006-311	- The EFSA Journal (2007) 479, 1-29, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on request from the European Commission on bluetongue vectors and vaccines. - The EFSA Journal (2007) 479, 1-29 and The EFSA Journal (2007) 480, 1-20, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA-Q-2007-063) on bluetongue - Annex I - Summary of the results of the questionnaire on BT vaccines sent to the vaccine companies - The EFSA Journal (2007) 479 1-29 and The EFSA Journal (2007) 480 1-20, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA-Q-2007-063) on bluetongue - Annex II - Summary of the results of the questionnaire on bluetongue sent to CVOs - The EFSA Journal (2007) 479, 1-29 and The EFSA Journal (2007) 480, 1-20, Scientific Report of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA Q-2007-063) on bluetongue
19	EFSA-Q-2006-326	- The EFSA Journal (2007) 469, 1-102, "Assessment of the risk of tick introduction into the UK, Ireland, and Malta as a consequence of abandoning the national rules"
20	EFSA-Q-2007-044	- The EFSA Journal (2007)584, 1- 163, For citation purposes: Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission on possible vector species and live stages of susceptible species not transmitting disease as regards certain fish diseases.
21	EFSA-Q-2007-061	- The EFSA Journal (2007) 597, 1-116, Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission on possible vector species and live stages of susceptible species not transmitting disease as regards certain mollusc diseases

22	EFSA-Q-2007-062	- The EFSA Journal (2007) 598, 1-91, Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission on possible vector species and live stages of susceptible species not transmitting disease as regards certain crustacean diseases
23	EFSA-Q-2007-063	- The EFSA Journal (2007) 480, 1-20, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on the EFSA Self mandate on bluetongue origin and occurrence - The EFSA Journal (2007) 479, 1-29 and The EFSA Journal (2007) 480, 1-20, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA-Q-2007-063) on bluetongue - Annex I - Summary of the results of the questionnaire on BT vaccines sent to the vaccine companies - The EFSA Journal (2007) 479 1-29 and The EFSA Journal (2007) 480 1-20, Scientific Opinion of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA-Q-2007-063) on bluetongue - Annex II - Summary of the results of the questionnaire on bluetongue sent to CVOs - The EFSA Journal (2007) 479, 1-29 and The EFSA Journal (2007) 480, 1-20, Scientific Report of the Scientific Panel on Animal Health and Welfare on request from the Commission (EFSA-Q-2006-311) and EFSA Self mandate (EFSA Q-2007-063) on bluetongue
24	EFSA-Q-2007-179	- The EFSA Journal (2008) 715, 1-161, Scientific Opinion of the Panel on Animal Health and Welfare on a request from The European Commission on Animal health and welfare aspects of avian influenza and the risk of its introduction into the EU poultry holdings.
25	EFSA-Q-2007-200	- The EFSA Journal (2009) 932, 1-18, Scientific Opinion / Statement / Guidance of the Panel on AHAW on a request from Commission on “Control and eradication of Classic Swine Fever in wild boar” - Annex to The EFSA Journal (2008) 932, 1-18 and 933, 1-16, Scientific Report - Annex to The EFSA Journal (2009) 932, 1-18 and 933, 1-16, ANNEX A – DATA COLLECTION ON WILD BOAR - Annex to The EFSA Journal (2009) 932, 1-16 and 933, 1-18, ANNEX B - TECHNICAL DESCRIPTION OF THE MODELS
26	EFSA-Q-2007-201	- The EFSA Journal (2008) 735, 1-70, Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission (DG SANCO) on Bluetongue vectors and insecticides
27	EFSA-Q-2008-074	- The EFSA Journal (2008) 808, 1-144, Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission (DG SANCO) on aquatic species susceptible to diseases listed in Directive 2006/88/EC
28	EFSA-Q-2008-427	- The EFSA Journal (2009) 933, 1-16, Scientific Opinion of the Panel on AHAW on a request from Commission on “Animal health safety of fresh meat derived from pigs vaccinated against Classic Swine Fever” - Annex to The EFSA Journal (2008) 932, 1-18 and 933, 1-16, Scientific Report - Annex to The EFSA Journal (2009) 932, 1-18 and 933, 1-16, ANNEX A – DATA COLLECTION ON WILD BOAR - Annex to The EFSA Journal (2009) 932, 1-16 and 933, 1-18, ANNEX B - TECHNICAL DESCRIPTION OF THE MODELS
29	EFSA-Q-2008-436	- The EFSA Journal (2008) 795, 1-56, Scientific Opinion of the Panel on Animal Health and Welfare on a request from the European Commission (DG SANCO) on Risk of Bluetongue Transmission in Animal Transit
30	EFSA-Q-2008-665	- The EFSA Journal (2009) 1144, 1-112, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on porcine brucellosis (Brucella suis)
31	EFSA-Q-2009-00011	- The EFSA Journal (2010) 8(11): 1894, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on a scientific opinion on the increased mortality events in Pacific Oysters, *Crassostrea gigas*

32	EFSA-Q-2009-00503	- The EFSA Journal (2010) 8(8): 1703, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on Epizootic Hemorrhagic Disease
33	EFSA-Q-2009-00506	- The EFSA Journal (2010) 8(3): 1556, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on African Swine Fever
34	EFSA-Q-2009-00594	- The EFSA Journal (2010) 8(8): 1703, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on the role of tick vectors in the epidemiology of Crimean-Congo Hemorrhagic Fever and African Swine Fever in Eurasia
35	EFSA-Q-2009-00595	- The EFSA Journal (2010) 8(9): 1723, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on geographical distribution of tick-borne infection and their vectors in Europe and the other regions of the Mediterranean Basin
36	EFSA-Q-2009-00879	- The EFSA Journal (2010) 8(2): 1499, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on Besnoitiosis: an emerging disease in Europe
37	EFSA-Q-2009-00935	- The EFSA Journal (2010) 8(10): 1770, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on a request from the Commission on the pandemic (H1N1) 2009 influenza and its potential implications for animal health
38	EFSA-Q-2010-00010	The EFSA Journal (2010) 8(5): 1595, Scientific Opinion of the Panel on Animal Health and Welfare (AHAW) on Q Fever

B. WP1'S WORKFLOW



C. RETROSPECTIVE ANALYSIS FILLING FORM

QUESTIONS	RESPONSES			
AHAW Opinion's ID				
Number of Question type identified in the Opinion				
Type Title	Choose between: EPIDEMIOLOGY/ PREVENTION&CONTROL/ PUBLIC HEALTH			
Q1: Are sub questions identified in the mandate?	**If YES:** go to Q2 and Q3 / **If NO**: go directly to Q4			
Q2: Number of AHAW opinion main risk questions (ToRi)				
Q3: Identification of the ToR_i	ToRi titel:……			
NATURE OF THE TOOLS				
Q4: Quantitative Approach	ToR_i: NO/YES……	**If YES**: go to "Use of Quantitative Approach"		
Q5: Qualitative Approach	ToR_i: NO/YES……	**If YES**: go to "Use of Qualitative Approach"		
Q6: Narrative Approach	ToR_i: NO/YES……	**If YES**: go to "Use of Narrative Approach"		
QUANTITATIVE APPROACH				
Q6: METHODOLOGY description	**USED**	**RELEVANT**	**NEEDED**	
ToRi:				
Q5: DATA description	**USED**	**RELEVANT**	**GAPS**	**SOURCES**
ToRi:				
QUALITATIVE APPROACH:				
Q6: METHODOLOGY description	**USED**	**RELEVANT**	**REASONS OF QUALITATIVE APPROACH**	**NEEDED**
ToRi:				
Q5: DATA description	**USED**	**RELEVANT**	**GAPS**	**SOURCES**
ToRi:				
NARRATIVE APPROACH:				
Q6: METHODOLOGY description	**USED**	**RELEVANT**	**REASONS OF NARRATIVE APPROACH**	**NEEDED**
ToRi:				
Q5: DATA description	**USED**	**RELEVANT**	**GAPS**	**SOURCES**
ToRi:				

D. RESULTS OF WP2 WEB-SEARCHES (INDIRECT SURVEY)

Results of web searches classified by type of data source (N sources) – Disease General Information

Resource	VEE	PRRS	*Echinococcus multilocularis*	*Echinococcus granulosus*	TOTAL
Book	7	8	37	34	86
Factsheet	22	2	4	7	35
Original Article	10	13	8	7	38
Proceedings	0	10	0	0	10
Report	0	16	0	0	16
Review	13	1	1	2	17
Websites	16	11	1	1	29
Not found	19	5	7	7	38
NA	348	369	377	377	1,471
TOTAL	**435**	**435**	**435**	**435**	**1,740**

Results of web searches classified by type of data source (N sources) – MSs' data

Resource	BE	FR	NL	SP	TOTAL
Book	0	3	0	0	3
Factsheet	0	3	0	0	3
Guidelines	0	1	0	0	1
Legislation	3	6	0	3	12
Original article	1	1	1	2	5
Proceedings	0	1	4	8	13
Report	0	7	0	0	7
Websites	133	111	128	91	463
Not found	75	80	79	108	342
NA	223	222	223	223	891
TOTAL	**435**	**435**	**435**	**435**	**1,740**

Results of web searches classified by type of data source (N sources) – VEE

Resource	BE	FR	NL	SP	TOTAL
Book	0	1	0	0	1
Factsheet	0	0	0	1	1
Original Article	3	8	9	0	20
Report	5	0	41	0	46
Review	0	2	0	2	4
Websites	13	65	2	0	80
Not found	7	34	36	18	95
NA	407	325	346	414	1,492
TOTAL	**435**	**435**	**435**	**435**	**1,740**

Results of web searches classified by type of data source (N sources) – PRRS

Resource	BE	FR	NL	SP	TOTAL
Course	0	0	0	5	5
Factsheet	0	0	0	9	9
Legislation	0	10	0	0	10
Original article	9	17	6	10	42
Report	2	2	7	0	11
Review	0	1	0	0	1
Websites	26	26	26	5	83
Not found	18	70	87	102	277
NA	380	309	309	304	1,302
TOTAL	**435**	**435**	**435**	**435**	**1,740**

Results of web searches classified by type of data source (N sources) – *Echinococcus multilocularis*

Resource	BE	FR	NL	SP	TOTAL
Factsheet	0	3	4	3	10
Original article	22	3	9	0	34
Report	5	39	18	0	62
Review	0	9	0	0	9
Websites	32	25	20	1	78
Not found	54	38	63	93	248
NA	322	318	321	338	1,299
TOTAL	**435**	**435**	**435**	**435**	**1,740**

Results of web searches classified by type of data source (N sources) – *Echinococcus granulosus*

Resource	BE	FR	NL	SP	TOTAL
Factsheet	1	6	4	3	14
Legislation	17	28	19	19	83
Original article	0	0	1	5	6
Report	9	15	13	12	49
Review	0	5	5	1	11
Websites	45	5	6	2	58
Not found	43	57	71	76	247
NA	320	319	316	317	1,272
TOTAL	**435**	**435**	**435**	**435**	**1,740**

E. WP2 DATA AVAILABILITY, ACCESSIBILITY AND FORM COMPARISON

Comparison of data availability, accessibility and form – Disease General Information

Parameter		Codification	DGI	DE	AE	PC	S	PH	Total
Availability	VEE	0 = no data or data not found	12		5	3			20
		1 = some data found							0
		2 = all data found	66	9	10	20	13	11	129
		TOTAL	78	9	15	23	13	11	149
	PRRS	0 = no data or data not found	3		2				5
		1 = some data found					8		8
		2 = all data found	54	3	8	38	84		187
		TOTAL	57	3	10	38	92		200
	E. multilocularis	0 = no data or data not found	4		3				7
		1 = some data found							0
		2 = all data found	56	2	11	13	12	12	106
		TOTAL	60	2	14	13	12	12	113
	E. granulosus	0 = no data or data not found	4		3				7
		1 = some data found							
		2 = all data found	62	2	12	16	20	13	125
		TOTAL	66	2	15	16	20	13	132
Accessibility	VEE	1 = limited access (restricted or charge)	9			3			12
		2 = free access	57	9	10	17	13	11	117
		TOTAL	66	9	10	20	13	11	129
	PRRS	1 = limited access (restricted or charge)	7			1			8
		2 = free access	47	3	8	37	8		103
		TOTAL	54	3	8	38	8	0	111
	E. multilocularis	1 = limited access (restricted or charge)	5		1	1			7
		2 = free access	51	2	10	12	12	12	99
		TOTAL	55	2	11	13	12	12	106
	E. granulosus	1 = limited access (restricted or charge)	6		2	3			11
		2 = free access	56	2	10	13	20	13	114
		TOTAL	62	2	12	16	20	13	125
Form	VEE	0 = unknown							0
		1 = PDF	56	9	8	14	7	11	105
		2 = HTML	9		2	6	6		23
		3 = TEXT	1						1
		4 = EXCEL							0
		TOTAL	66	9	10	20	13	11	129
	PRRS	0 = unknown							0
		1 = PDF	49	3	7	26	8		93
		2 = HTML	5		1	12			18
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	54	3	8	38	8	0	111
	E. multilocularis	0 = unknown							0
		1 = PDF	52	2	9	13	12	12	100
		2 = HTML	3		2				5
		3 = TEXT	1						1
		4 = EXCEL							0
		TOTAL	56	2	11	13	12	12	106
	E. granulosus	0 = unknown							0
		1 = PDF	54	2	10	13	20	13	112
		2 = HTML	8		2	3			13
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	62	2	12	16	20	13	125

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

Comparison of data availability, accessibility and form – MS's data

Parameter	Member State	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	Belgium	0 = no data or data not found	1		32	42			75
		1 = some data found		10	31	68			109
		2 = all data found	8	21	57	23			109
		TOTAL	9	31	120	133	0	0	293
	France	0 = no data or data not found	1	1	22	56			80
		1 = some data found		3	1	61			65
		2 = all data found	8	27	96	36			167
		TOTAL	9	31	119	153	0	0	312
	The Netherlands	0 = no data or data not found	1	1	34	43			79
		1 = some data found			9	60			69
		2 = all data found	8	35	82	23			148
		TOTAL	9	36	125	126	0	0	296
	Spain	0 = no data or data not found	1		27	80			108
		1 = some data found		2	16	15			33
		2 = all data found	9	27	112	28			176
		TOTAL	10	29	155	123	0	0	317
Accessibility	Belgium	1 = limited access (restricted or charge)	1	3	18	7			29
		2 = free access	7	28	70	84			189
		TOTAL	8	31	88	91	0	0	218
	France	1 = limited access (restricted or charge)		1	23	26			50
		2 = free access	8	29	74	71			182
		TOTAL	8	30	97	97	0	0	232
	The Netherlands	1 = limited access (restricted or charge)	1	6	20	7			34
		2 = free access	7	29	71	76			183
		TOTAL	8	35	91	83	0	0	217
	Spain	1 = limited access (restricted or charge)		4	20	11			35
		2 = free access	9	25	108	32			174
		TOTAL	9	29	128	43	0	0	209
Form	Belgium	0 = unknown							
		1 = PDF	1	3	15	5			24
		2 = HTML	1	11	21	37			70
		3 = TEXT	6	14	30	47			97
		4 = EXCEL		3	22	2			27
		TOTAL	8	31	88	91	0	0	218
	France	0 = unknown		1	15	20			36
		1 = PDF	1	20	38	41			100
		2 = HTML	7	1	19	36			63
		3 = TEXT							0
		4 = EXCEL		8	25				33
		TOTAL	8	30	97	97	0	0	232
	The Netherlands	0 = unknown	1	6	20	5			32
		1 = PDF	1	8	18	32			59
		2 = HTML		12	33	46			91
		3 = TEXT	6						6
		4 = EXCEL		9	20				29
		TOTAL	8	35	91	83	0	0	217
	Spain	0 = unknown		4	20	10			34
		1 = PDF	1	11	50	17			79
		2 = HTML	8	11	36	16			71
		3 = TEXT							0
		4 = EXCEL		3	22				25
		TOTAL	9	29	128	43	0	0	209

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

Comparison of data availability, accessibility and form between the four MSs – VEE

Parameter	Member State	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	Belgium	0 = no data or data not found	1		6				7
		1 = some data found	23			1	2		26
		2 = all data found	2				9		11
		TOTAL	26	0	6	1	11	0	44
	France	0 = no data or data not found	1		7	1	25		34
		1 = some data found	2				7		9
		2 = all data found	14			4	70		88
		TOTAL	17	0	7	5	102	0	131
	The Netherlands	0 = no data or data not found	2		4		30		36
		1 = some data found	11						11
		2 = all data found	5		4	1	47		57
		TOTAL	18	0	8	1	77	0	104
	Spain	0 = no data or data not found	11		7				18
		1 = some data found			1				1
		2 = all data found				1		1	2
		TOTAL	11	0	8	1	0	1	21
Accessibility	Belgium	1 = limited access (restricted or charge)					1		1
		2 = free access	25			1	10		36
		TOTAL	25	0	0	1	11	0	37
	France	1 = limited access (restricted or charge)	5				7		12
		2 = free access	11			4	70		85
		TOTAL	16	0	0	4	77	0	97
	The Netherlands	1 = limited access (restricted or charge)	2						2
		2 = free access	14		4	1	47		66
		TOTAL	16	0	4	1	47	0	68
	Spain	1 = limited access (restricted or charge)							0
		2 = free access			1	1		1	3
		TOTAL	0	0	1	1	0	1	3
Form	Belgium	0 = unknown					1		1
		1 = PDF	13						13
		2 = HTML	12			1	10		23
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	25	0	0	1	11	0	37
	France	0 = unknown					7		7
		1 = PDF	14			3	14		31
		2 = HTML	2			1	56		59
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	16	0	0	4	77	0	97
	The Netherlands	0 = unknown							0
		1 = PDF	13		4	1	46		64
		2 = HTML	3				1		4
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	16	0	4	1	47	0	68
	Spain	0 = unknown							0
		1 = PDF			1			1	2
		2 = HTML							0
		3 = TEXT				1			1
		4 = EXCEL							0
		TOTAL	0	0	1	1	0	1	3

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

Comparison of data availability, accessibility and form between the four MSs – PRRS

Parameter	Member State	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	Belgium	0 = no data or data not found	1	3	1	8	5		18
		1 = some data found	1	6	2	3	14		26
		2 = all data found	1			52	22		75
		TOTAL	3	9	3	63	41	0	119
	France	0 = no data or data not found	1	1	4	9	55		70
		1 = some data found		1			13		14
		2 = all data found	2	6	2	18	32		60
		TOTAL	3	8	6	27	100	0	144
	The Netherlands	0 = no data or data not found	1		1	9	76		87
		1 = some data found	4	2	3	8	11		28
		2 = all data found		6	2	20	1		29
		TOTAL	5	8	6	37	88	0	144
	Spain	0 = no data or data not found	2	1	4	11	84		102
		1 = some data found			1	3			4
		2 = all data found		5	2	13	8		28
		TOTAL	2	6	7	27	92	0	134
Accessibility	Belgium	1 = limited access (restricted or charge)		1			1		2
		2 = free access	2	5	2	55	35		99
		TOTAL	2	6	2	55	36	0	101
	France	1 = limited access (restricted or charge)	1	2	1		10		14
		2 = free access	1	5	1	18	35		60
		TOTAL	2	7	2	18	45	0	74
	The Netherlands	1 = limited access (restricted or charge)				1			1
		2 = free access	4	8	5	27	12		56
		TOTAL	4	8	5	28	12	0	57
	Spain	1 = limited access (restricted or charge)							0
		2 = free access		5	3	16	8		32
		TOTAL	0	5	3	16	8	0	32
Form	Belgium	0 = unknown							0
		1 = PDF		6	2	25	8		41
		2 = HTML	2			30	28		60
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	2	6	2	55	36	0	101
	France	0 = unknown	1						1
		1 = PDF	1	7	1	5	27		41
		2 = HTML			1	13	18		32
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	2	7	2	18	45	0	74
	The Netherlands	0 = unknown							0
		1 = PDF	3	8	5	17	12		45
		2 = HTML	1			11			12
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	4	8	5	28	12	0	57
	Spain	0 = unknown							0
		1 = PDF		1	3	6	7		17
		2 = HTML		4		10	1		15
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	5	3	16	8	0	32

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

Comparison of data availability, accessibility and form between the four MSs – *Echinococcus multilocularis*

Parameter	Member State	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	Belgium	0 = no data or data not found	2			1	41	9	53
		1 = some data found		6		1	31	4	42
		2 = all data found		6	2	12	36	12	68
		TOTAL	2	12	2	14	108	25	163
	France	0 = no data or data not found	2		1	1	29	5	38
		1 = some data found					1		1
		2 = all data found		7	7	10	60	31	115
		TOTAL	2	7	8	11	90	36	154
	The Netherlands	0 = no data or data not found	2			4	51	6	63
		1 = some data found		3			42	10	55
		2 = all data found		6	5	4	15	8	38
		TOTAL	2	9	5	8	108	24	156
	Spain	0 = no data or data not found	2		4		87		93
		1 = some data found							0
		2 = all data found				4			4
		TOTAL	2	0	4	4	87	0	97
Accessibility	Belgium	1 = limited access (restricted or charge)		3			13		16
		2 = free access	0	9	2	13	54	16	94
		TOTAL	0	12	2	13	67	16	110
	France	1 = limited access (restricted or charge)			4		8	1	13
		2 = free access		7	3	10	53	30	103
		TOTAL	0	7	7	10	61	31	116
	The Netherlands	1 = limited access (restricted or charge)			2			4	6
		2 = free access		9	3	4	57	14	87
		TOTAL	0	9	5	4	57	18	93
	Spain	1 = limited access (restricted or charge)							0
		2 = free access				4			4
		TOTAL	0	0	0	4	0	0	4
Form	Belgium	0 = unknown		3			12		15
		1 = PDF		9	2	6	35	14	66
		2 = HTML				7	20	2	29
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	12	2	13	67	16	110
	France	0 = unknown					5		5
		1 = PDF		4	6	4	30	26	70
		2 = HTML		3	1	6	26	5	41
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	7	7	10	61	31	116
	The Netherlands	0 = unknown							0
		1 = PDF		9	5	4	47	13	78
		2 = HTML					10	5	15
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	9	5	4	57	18	93
	Spain	0 = unknown							0
		1 = PDF				3			3
		2 = HTML				1			1
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	0	0	4	0	0	4

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

Comparison of data availability, accessibility and form between the four MSs – *Echinococcus granulosus*

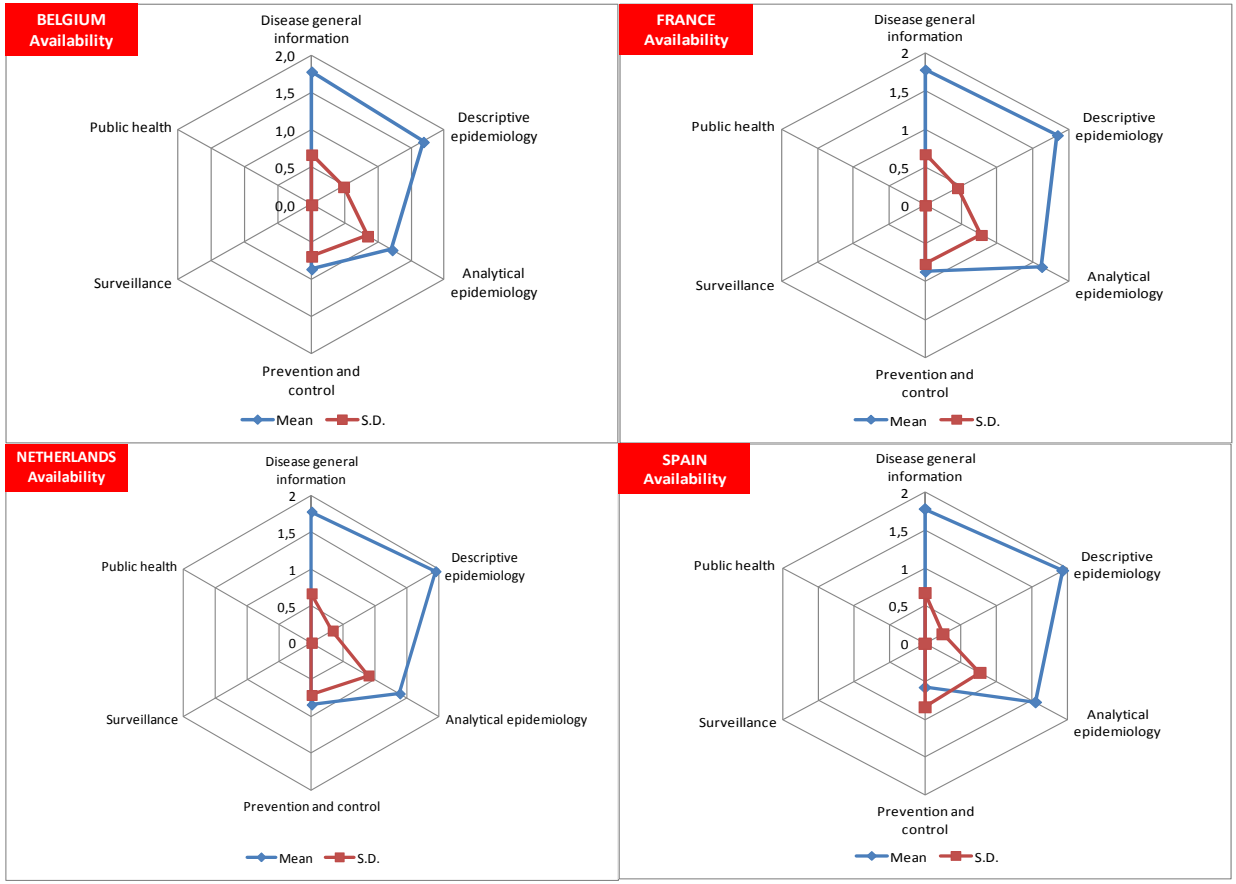
Parameter	Member State	Codification	DGI	DE	AE	PC	S	PH	Total
Availability	Belgium	0 = no data or data not found	2		4		26	11	43
		1 = some data found		3		2	21	2	28
		2 = all data found		3	1	3	54	7	68
		TOTAL	2	6	5	5	101	20	139
	The Netherlands	0 = no data or data not found	2		3		56	10	71
		1 = some data found		4	2			5	11
		2 = all data found		2	1	4	35	7	49
		TOTAL	2	6	6	4	91	22	131
	France	0 = no data or data not found	2	1	4		44	6	57
		1 = some data found					13		13
		2 = all data found		3	1	5	56	19	84
		TOTAL	2	4	5	5	113	25	154
	Spain	0 = no data or data not found	2		4	4	55	11	76
		1 = some data found		1				1	2
		2 = all data found		6	2	4	84	10	51
		TOTAL	2	7	6	8	22	22	129
Accessibility	Belgium	1 = limited access (restricted or charge)		2			18	1	21
		2 = free access		4	1	5	57	8	75
		TOTAL	0	6	1	5	75	9	96
	The Netherlands	1 = limited access (restricted or charge)			1			1	2
		2 = free access		6	2	4	35	11	58
		TOTAL	0	6	3	4	35	12	60
	France	1 = limited access (restricted or charge)					9	2	11
		2 = free access		3	1	5	60	17	86
		TOTAL	0	3	1	5	69	19	97
	Spain	1 = limited access (restricted or charge)						1	1
		2 = free access		7	2	4	29	10	52
		TOTAL	0	7	2	4	29	11	53
Form	Belgium	0 = unknown		2			18	1	21
		1 = PDF		4	1	1	32	7	45
		2 = HTML				4	20	1	25
		3 = TEXT					5		5
		4 = EXCEL							0
		TOTAL	0	6	1	5	75	9	96
	The Netherlands	0 = unknown							0
		1 = PDF		6	3	4	35	11	59
		2 = HTML						1	1
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	6	3	4	35	12	60
	France	0 = unknown					9		9
		1 = PDF		3	1		56	19	79
		2 = HTML				5	1		6
		3 = TEXT					3		3
		4 = EXCEL							0
		TOTAL	0	3	1	5	69	19	97
	Spain	0 = unknown							0
		1 = PDF		7	2	3	29	11	52
		2 = HTML				1			1
		3 = TEXT							0
		4 = EXCEL							0
		TOTAL	0	7	2	4	29	11	53

DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health

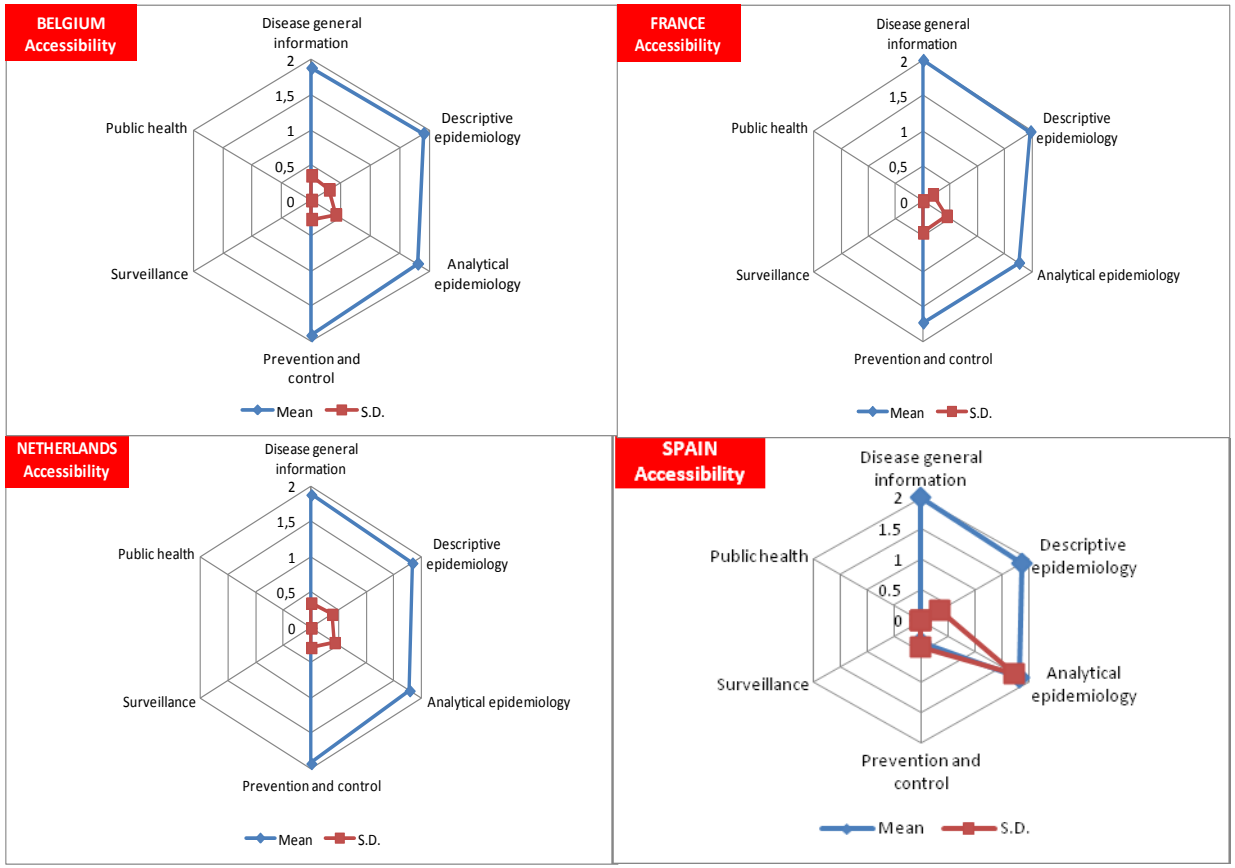
F. COMPARISON OF DATA AVAILABILITY, ACCESSIBILITY AND FORM BETWEEN THE FOUR MSS FOR MSS DATA (MEAN, MIN. AND MAX. SCORES WITH SD)

		Availability						Accessibility						Form of data					
MS	**Parameter**	**DGI**	**DE**	**AE**	**PC**	**S**	**PH**	**DGI**	**DE**	**AE**	**PC**	**S**	**PH**	**DGI**	**DE**	**AE**	**PC**	**S**	**PH**
BE	Number	9	31	120	133	0	0	8	31	88	91	0	0	8	31	88	91	0	0
	Min	0.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Max	2.00	2.00	2.00	2.00	0.00	0.00	2.00	2.00	2.00	2.00	0.00	0.00	2.00	4.00	4.00	4.00	0.00	0.00
	Mean	**1.78**	**1.68**	**1.21**	**0.86**	-	-	**1.88**	**1.90**	**1.80**	**1.92**	-	-	**1.63**	**1.65**	**1.92**	**1.53**	-	-
	SD	0.67	0.48	0.84	0.69	-	-	0.35	0.30	0.41	0.27	-	-	0.74	1.02	1.39	0.70	-	-
FR	Number	9	31	119	153	0	0	8	30	97	97	0	0	8	30	97	97	0	0
	Min	0.00	0.00	0.00	0.00	0.00	0.00	2.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	Max	2.00	2.00	2.00	2.00	0.00	0.00	2.00	2.00	2.00	2.00	0.00	0.00	2.00	4.00	4.00	2.00	0.00	0.00
	Mean	**1.78**	**1.84**	**1.62**	**0.87**	-	-	**2.00**	**1.97**	**1.76**	**1.73**	-	-	**1.88**	**1.8**	**1.81**	**1.16**	-	-
	SD	0.67	0.45	0.78	0.77	-	-	0.00	0.18	0.43	0.45	-	-	0.35	1.37	1.42	0.75	-	-
NL	Number	9	36	125	126	0	0	8	35	91	83	0	0	8	35	91	83	0	0
	Min	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Max	2.00	2.00	2.00	2.00	0.00	0.00	2.00	2.00	2.00	2.00	0.00	0.00	3.00	4.00	4.00	2.00	0.00	0.00
	Mean	**1.78**	**1.94**	**1.38**	**0.84**	-	-	**1.88**	**1.83**	**1.78**	**1.92**	-	-	**2.38**	**1.94**	**1.80**	**1.49**	-	-
	SD	0.67	0.33	0.89	0.71	-	-	0.35	0.38	0.42	0.28	-	-	1.19	1.41	1.39	0.61	-	-
SP	Number	9	30	155	123	0	0	8	30	128	43	0	0	8	30	128	43	0	0
	Min	0.00	1.00	0.00	0.00	0.00	0.00	2.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	Max	2.00	2.00	2.00	2.00	0.00	0.00	2.00	2.00	2.00	2.00	0.00	0.00	2.00	4.00	4.00	2.00	0.00	0.00
	Mean	**1.78**	**1.93**	**1.55**	**0.58**	-	-	**2.00**	**1.87**	**1.84**	**1.74**	-	-	**1.88**	**1.57**	**1.64**	**1.14**	-	-
	SD	0.67	0.25	0.77	0.84	-	-	0.00	0.35	0.36	0.44	-	-	0.35	1.07	1.26	0.77	-	-

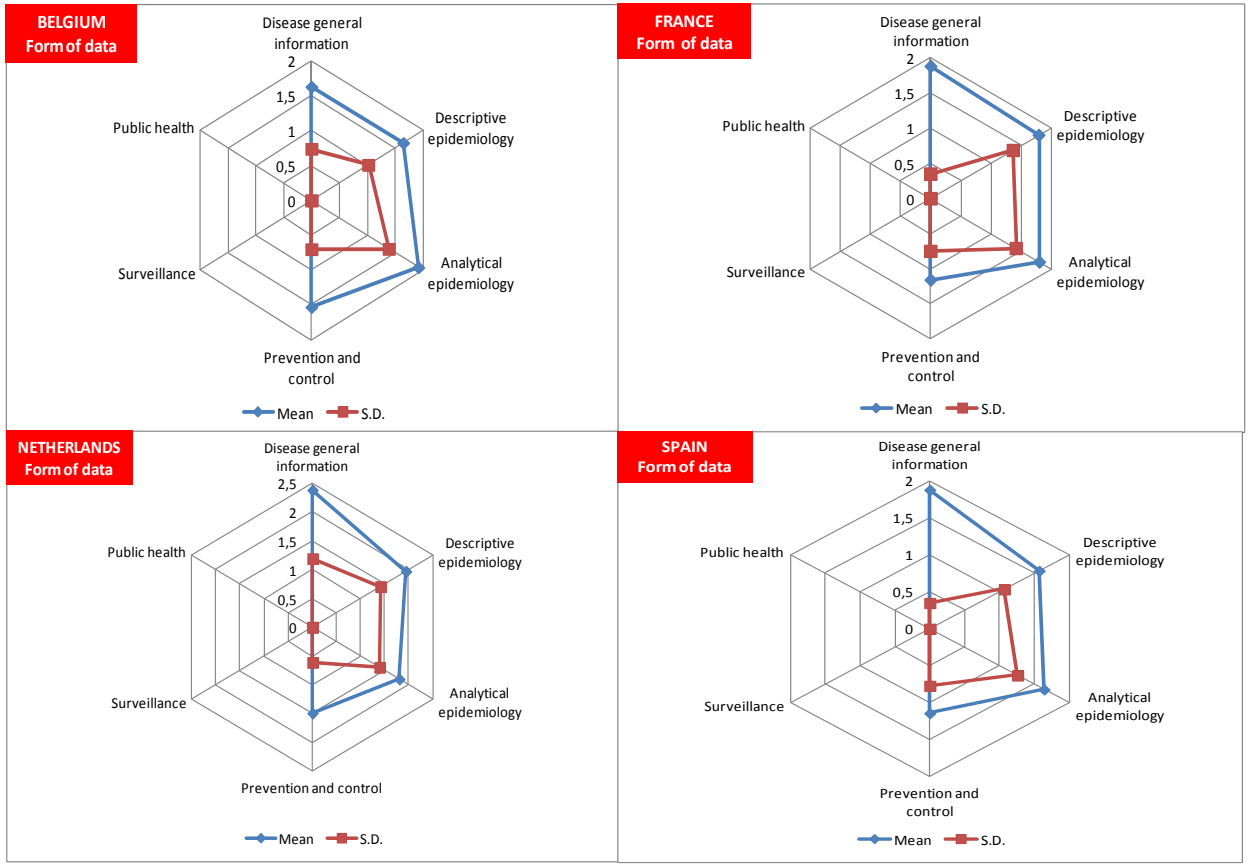
DGI = Disease General Information; DE = Descriptive Epidemiology; AE = Analytical Epidemiology; PC = Prevention and Control; S = Surveillance; PH = Public Health



Comparison of means between the four MSs – data availability –MSs data



Comparison of means between the four MSs – data accessibility – MSs data

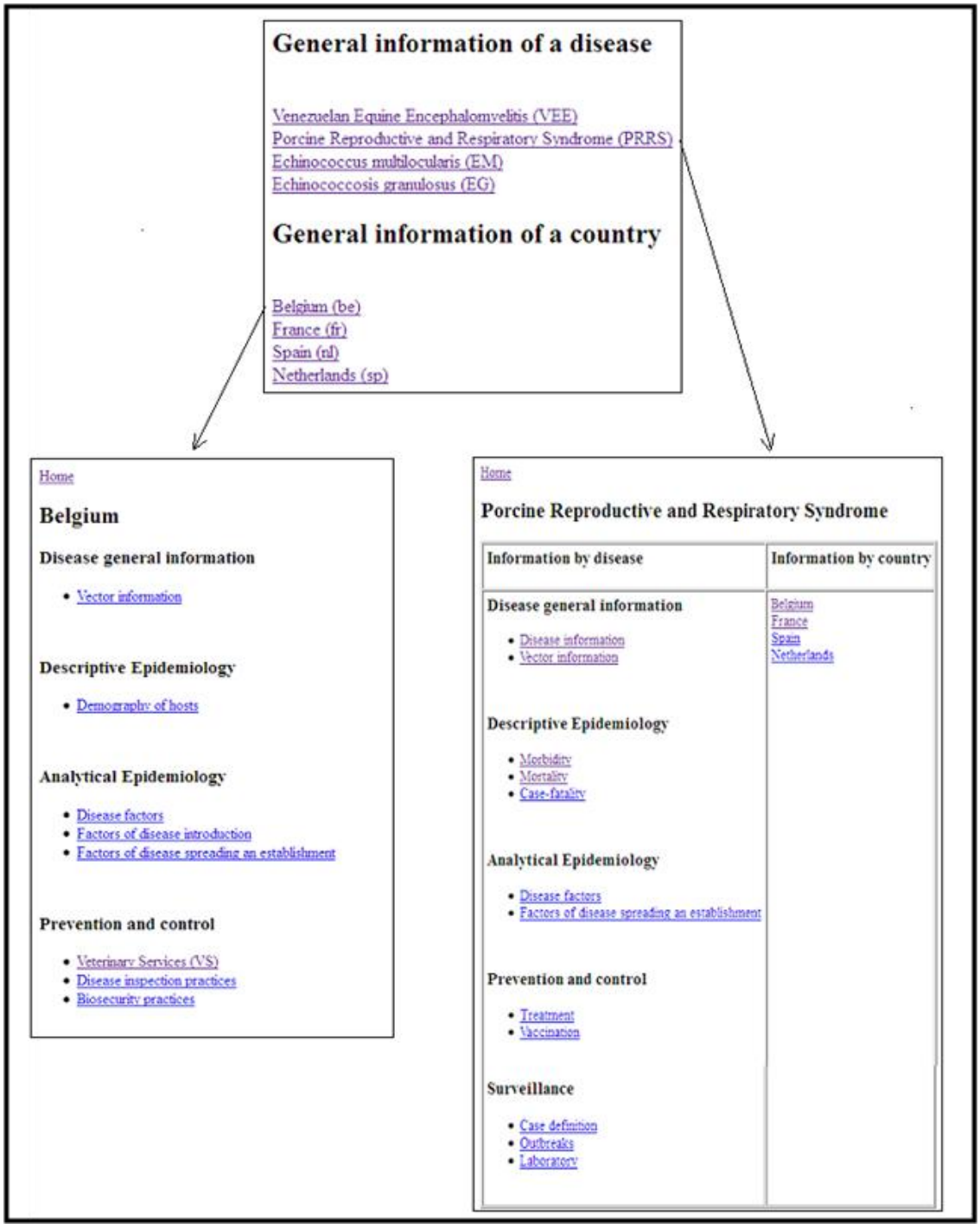


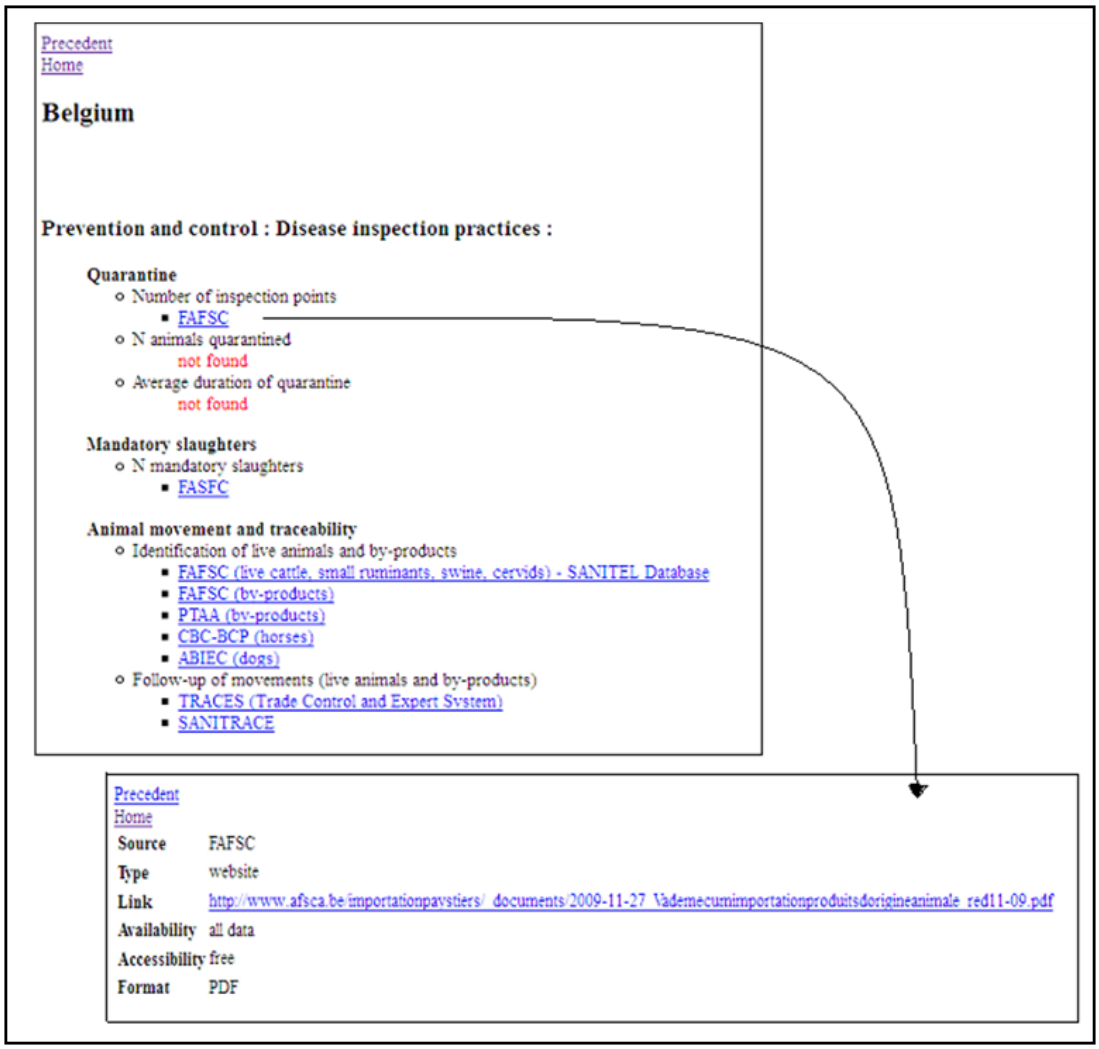
Comparison of means between the four MSs – Form of data – MSs data

G. ILLUSTRATION OF THE DATASPEC DATA SOURCES INVENTORY APPLICATION

The following figures detail a proposed web application that link for each AHAW experts' query the related WP2 identified source(s)

Home Page





To complete the identification card, another tool is used, especially for the sources which are present in the Pubmed database. Indeed, a service web, named "Entrez Utilities Web Service" or "eutils" (Bookshelf ID: NBK25500, E-utilities Quick Start, Eric Sayers, PhD. NCBI sayers@ncbi.nlm.nih.gov Created: December 12, 2008; Last Update: December 14, 2011.), was created, allowing the interrogation of its database and the sources' metadata fetching through various references (see below).

Precedent
Home

Pubmed metadata

1: Annu Rev Entomol. 2004;49:141-74.

Venezuelan equine encephalitis.

Weaver SC, Ferro C, Barrera R, Boshell J, Navarro JC.

Center for Biodefense and Emerging Infectious Diseases and Department of Pathology, University of Texas Medical Branch, Galveston, Texas 77555-0609, USA. sweaver@utmb.edu

Venezuelan equine encephalitis virus (VEEV) remains a naturally emerging disease threat as well as a highly developed biological weapon. Recently, progress has been made in understanding the complex ecological and viral genetic mechanisms that coincide in time and space to generate outbreaks. Enzootic, equine avirulent, serotype ID VEEV strains appear to alter their serotype to IAB or IC, and their vertebrate and mosquito host range, to mediate repeated VEE emergence via mutations in the E2 envelope glycoprotein that represent convergent evolution. Adaptation to equines results in highly efficient amplification, which results in human disease. Although epizootic VEEV strains are opportunistic in their use of mosquito vectors, the most widespread outbreaks appear to involve specific adaptation to Ochlerotatus taeniorhynchus, the most common vector in many coastal areas. In contrast, enzootic VEEV strains are highly specialized and appear to utilize vectors exclusively in the Spissipes section of the Culex (Melanoconion) subgenus.

Publication Types:
    Research Support, Non-U.S. Gov't
    Research Support, U.S. Gov't, P.H.S.
    Review

MeSH Terms:
    Aedes/virology
    Animals
    Disease Outbreaks
    Encephalitis Virus, Venezuelan Equine/genetics
    Encephalitis Virus, Venezuelan Equine/pathogenicity*
    Encephalomyelitis, Venezuelan Equine/epidemiology
    Encephalomyelitis, Venezuelan Equine/transmission*
    Equidae
    Horse Diseases/epidemiology
    Horse Diseases/transmission
    Horses
    Host-Parasite Interactions
    Humans
    Insect Vectors/virology*
    Mutation
    Viral Envelope Proteins/genetics*
    Viral Vaccines
    Virulence/genetics

Substances:
    Viral Envelope Proteins
    Viral Vaccines
    glycoprotein E2, equine encephalitis virus

Grant Support:
    AI-25489/AI/NIAID NIH HHS
    AI48807-01/AI/NIAID NIH HHS
    AI49725-01/AI/NIAID NIH HHS
    TW5919/TW/FIC NIH HHS

PMID: 14651460 [PubMed - indexed for MEDLINE]

Source	Weaver SC, Ferro C, Barrera R, Boshell J, Navarro JC. (2004) VENEZUELAN EQUINE ENCEPHALITIS. Annu. Rev. Entomol. 49:141–74
Type	Review
Link	http://www.annualreviews.org/doi/pdf/10.1146/annurev.ento.49.061802.123422
Availability	all data
Accessibility	free
Format	PDF

H. DESCRIPTION OF THE EXISTING INTERNATIONAL STANDARDS USED IN THE DATASPEC MODEL - EFSA DATA WAREHOUSE (DWH) DIMENSIONS

Dimension	Description and Attributes	
Individual	An individual participating in a food consumption survey	
Individual attributes:	IndividualIdentifier	Unique identifier for each individual included in dimension
	Gender	Sex of the participant
	AgeClass	Age classification of individual
	ReporterClass	Subject identified as under or over reporter
	SpecialCondition	Physiological or health condition of subject
	Diet	Dietary classification of individual
	Education	Level of education described using international Standard classification of education 1997
	EthnicGroup	Description of ethnic group of individual, this is distinct of nationality
	PhysicalActivity	Level of physical activity of the individual
	WeightMeasurement	Method used to measure body weight
	HeightMeasurement	Method used to measure height
Organism	Organism that have been sampled in the course of the survey, monitoring or scientific studies	
Organism attributes:	OrganismIdentifier	Unique identifier for each sample unit included in the dimension.
	Taxonomic Classification	Kingdom: Kingdom of the organism Phylum: Phylum of the organism Class: Class of the organism Order: Order of the Organism Family: Family of the organism Genus: Genus of the organism Species: Species of the organism Strain: Strain, variety, breed, or genetic event of organism tested
	Sex	Sex of organism
	LifeStage	Life stage of organism
	ProductionMethod	Production method revelant for animals sampled from the agricultural environnements
	Treatment	Treatment applied to the Organism (Control, Dose, Vaccination)
	RouteExposure	Route of exposure or application method of treatment (Oral: feed, water, intramuscular…)
	TraetmentDuration	Duration of treatment or exposure
Food/feed	Food or feed item sampled for purpose of survey and monitoring	

Food/feed attribute and facets:	Food/feed item	Description of the food or feed item at the lowest base term. That could be describe by more standard with different levels of hierarchy: EFSAConciseClass (Three level classification to describe food item according to the EFSA concise food consumption system), FoodExClass (Three level classification of the base terms in FoodEx), GEMSClass (Three level classification according to the Worlds health organisation Global Environment Monitoring system), ZoonosesClass (Four Level classification of food items to support Zoonose annual report)
	Facets	Synonyms, Scientific name, Ingredients, alcohol, cooked, fat, part, packaging, ProductionMethod, Preservation, Packaging, treatment, IntendedUse
Programme	Describe the purpose of the survey or monitoring program, including where relevant the European legislation for the program	
Program attributes:	Programme	Program used to collect fact described in terms of legislation, level, study type and sampling strategy.
	Legislation	Legislation that describe the program
	Level	Level at with the program is design and implemented (EU, National, Industry, Research…)
	Study	Type of study design used for the program
	Strategy	Strategy of sampling. EUROSTAT typology of sampling strategy
Sample	Describe the sample taken for laboratory analysis.	
Sample attributes:	Sample	Each individual sample tested in the laboratory
	SampleMethod	Method of selecting / collecting sampling units
	SampleUnit	The unit which the specimens taken and which is considered either infected (contaminated) or not, based on the analyses result
Sampling Point	The point in the food chain where sample was taken.	
Sample Point Attributes:	SamplingPointL1	Highest level to describe the sampling point
	Sample point	Description of the sampling point
Analytical Method	Method used in the laboratory to measure the parameter reported in the survey or monitoring results.	
Analytical Method Attributes:	AnalyticalMethodL1	Highest level to describe the analytical method used (Biological, chemical…)
	AnalyticalMethodL2	Level 2 to describe the analytical method used (Chromatography tests, Atomic spectroscopy…)
	AnalyticalMethodL3	Level 3 to describe the analytical method used (Gas chromatography, Liquid chromatography…)
	AnalyticalMethod	Description of the analytical method (GC-MS-MS…)

Parameter	Parameter that the outcome values represent	
Parameter Attribute:	Parameter	Full description of parameter at the lowest level. That could be describe by more standard with different levels of hierarchy: OrganicClass (describing organic compounds), ToxinClass (describing toxins, poisonous substance produced by living cells or organisms), MetalsClass (describing metallic elements in order to account mettalic compounds and transition metals), PesticidesClass (describing Pesticides to discriminate between pesticide residues definitions and pesticide residues), ZoonosesClass (describing zoonotics agents and other pathogens)
Document	Describe the attributes related to the resource of the information for result extracted from scientific literature.	
Document Attributes:	Document	Full description of the document include Title, author and resource
	DocumentType	Type of document using OECD classification with extension to describe EFSA outputs
	Year	Year of publication
Organisation	Organisation which was responsible for collating and submitting the monitoring or survey dataset to EFSA	
Organization Attributes:	Organisation	Organisation providing the facts
	OrganisationCountry	Country of the organisation
	OrganisationType	Type of organisation
Geography	Provide a spatial context for the survey or monitoring outcome value	
Geography attributes:	Zone	Zone in which the area is located
	RegionL0	Country or other top classification level
	RegionL1	NUTS, HASC, FAO level one region
	RegionL2	NUTS, HASC, FAO level two region
	RegionL3	NUTS, HASC, FAO level three region
Date	Provide a temporal context for the survey or monitoring outcome value	
Date attributes:	Date	Date is required to allow reporting of temporal information where the full date is not available
	Year	Year
	Quarter	Financial quarter, season
	Month	Month
	Week	Number of the week (According to ISO-8601)
	Day	Day

- EFSA DATA WAREHOUSE FACTS TABLE

Fact	Description	Dimensions	Lookup
Substance Concentration	Result of a survey and monitoring programs using laboratory methods to measure the concentration of a substance in a food or feed item	Organisation, date, geography, Food/feed, Program, sample, sampling point, parameter, analytical method	Parameter_Type, Expression_Result, Result_Type, Laboratory_Accreditation, Procedure_Accreditation, LegalLimit
Food Consumption	Result from food consumption surveys in term of the amount of each food item consumed per individuals as part of a meal	Organisation, date, Geography, Food/Feed, individual, Program, Analytical method	Meal, Exception_Day, Home_Consumption
Food Composition	Information about the composition of food items consumed in Europe	Organisation, document, date, geography, Food/feed, program, sample, sampling point, parameter, analytical method	Units, quality, Result_Type
Prevalence	Result of survey and monitoring programs testing for presence of a microbiological agent.	Organisation, date, geography, food/feed or organism, program, sample, sampling point, parameter, analytical method	Microbiological_Limit
Subtyping	Similar information to that presented in the prevalence table however, in this case, the number of isolates tested and the number of isolates positive is recorded where the laboratory has undertaken further testing on biological agents isolated from positive samples. Results are presented at the subspecies/serovar/phagetype level.	Organisation, Date, Geography, Food/Feed or Organism, Program, Sample, Sampling Point, Parameter, Analytical Method	
Animal Population	Information on a susceptible animal population in the regions supplying prevalence facts.	Organisation, date, geography, food/feed or Organism, program, sample	
Disease status	Information on EU co financed programs to control brucellosis in bovines and brucellosis in bovine, caprine and ovine animals.	Organisation, Date, Geography, Food/feed or organism, parameter	

Food borne outbreaks Summary	All outbreaks for foodborne pathogens	Organisation, date, geography, program, parameter	
Food borne outbreaks detail	Outbreak where the food vehicle has been identified	Organisation, date, geography, food/feed program, parameter, sample, sampling point	Contributory_Factor, Outbreak_Evidence, Outbreak_Type
Antimicrobial Resistance	Results of the monitoring program testing for zoonotic pathogens which exhibit resistance to antimicrobial substance, in particular antibiotics used to treat illness in animals and humans.	Organisation, date, geography, food/feed or organism, program, sample, sampling point, parameter, analytical method	Laboratory_Accreditation, Procedure_Accreditation, Result_Type

- EXISTING METADATA ON SAMPLING STRATEGY AND CORRESPONDING METADATA (EUROSTAT EUROPEAN COMMISION, 2010; IAN DOHOO, 2003)

Strategy	Definition	
Objective sampling	Strategy based on the selection of a *random sample* from a population on which the facts are reported.	→ Simple random sample: every element in the target population has an equal probability of being included.A complete list of the target population is required and a formal random process is used.
Selective sampling	Strategy based on the selection of a random sample from a subpopulation (or more frequently from subpopulations) of a population on which the facts are reported. The subpopulations are determined on a risk basis or not. The sampling from each subpopulation is not proportional: the sample size is proportionally bigger for instance in subpopulations considered at high risk. This sampling includes also the case when the facts reported refer to censuses on subpopulations.	→ Systematic random sample: A complete list of the population to be sampled is not required provided an estimate of the total number of animal sis available and all of the animals are sequentially available. → Stratified random sample: Prior to sampling, the population study is divided into mutually strata based on factors likely to affect the outcome. → Cluster sampling: in the cluster sample, the primary sampling unit is larger than the unit of concern. → Multistage sampling:
Census	When the totality of a population, on which the facts are reported, is controlled.	
Suspect sampling	Selection of an individual product or establishment in order to confirm or reject a suspicion of non-conformity. It's a not random sampling. The facts reported refer themselves to suspect units of the population.	
Convenient sampling	Strategy based on the selection of a sample for which units are selected only on the basis of feasibility or ease of facts collection. It's a not random sampling. The facts reported refer themselves to units selected according to this strategy. This "new" typology was added as included in sampling strategies used in facts collected by EFSA.	
Other sampling strategies	In order to document the Controls database, further typologies of sampling strategies can be assigned if none of the previous ones has been used for the selection of the sample units: More than one sampling strategy: the facts reported refer to sample units selected according more than one sampling strategy, for example: some units are selected according "Objective sampling" and other units according "Suspect sampling". Other: the facts reported refer to sample units selected according a strategy not included in the previous ones. Not specified: the facts reported refer to sample units selected according a strategy not specified.	

Metadata	Description and Attributes
Epidemiological Unit	
Unit	Specimens or facts collected; could either consider infected (contaminated) unit or not (e.g. Animal, flock, Herd, Holding, Slaughter batch, Environmental…)
Fact	
Fact_Level	Whether the result is individual or aggregated facts
Fact_Type	Type of the fact (e.g. a statistical fact, a laboratory fact, Quantity, Weight...)
Fact_Presentation	Graph, Histogram, Table, etc. (Frodsham A., 2007)
Result_Type	Type of the result obtained: LOD, LOQ, VAL, BIN, etc.
Expression_Result	Expression of the result: wet weight, dry weight, fat weight
Units	Units of the measurement fact (Kg, m^2)
Organisation	Organisation which was responsible for collating and submitting the monitoring or survey dataset to EFSA
OrganisationName	Organisation providing the facts
OrganisationCountry	Country of the organisation
OrganisationType	Type of organisation
Geography	Provide a spatial context for the survey or monitoring outcome value
Zone	Zone in which the area is located
RegionL0	Country or other top classification level
RegionL1	NUTS, HASC, FAO level one region
RegionL2	NUTS, HASC, FAO level two region
RegionL3	NUTS, HASC, FAO level three region
Date	Provide a temporal context for the survey or monitoring outcome value
Date	Date is required to allow reporting of temporal information where the full date is not available
Year	Year
Quarter	Financial quarter, season
Month	Month
Week	Number of the week (According to ISO-8601)

Day	Day
Event	
Event_Type	Biological (Abortion, Death, Pregnancy...) or Physical (Cooked, Treatment, Processing...)
Event_Description	Name of the event (Abortion, Death, Treatment...)
Organism	Organism that have been sampled in the course of the survey, monitoring or scientific studies
OrganismIdentifier	Unique identifier for each sample unit included in the dimension.
Taxonomic Classification	Kingdom: Kingdom of the organism Phylum: Phylum of the organism Class: Class of the organism Order: Order of the Organism Family: Family of the organism Genus: Genus of the organism Species: Species of the organism Strain: Strain, variety, breed, or genetic event of organism tested
Organism_Type	Nature of the organism: Wild, Pet, Domestic
Organism Description	
Sex	Sex of organism
LifeStage	Life stage of organism
Vector	
Vector_Habitat	Type, Optimal_Temperature, Relative_Humidity, Wind_Speed, Altitude
Vector_Activity	Type {indoor:outdoor}, Period{day:night}, Seasonality, Overwintering
Vector_Cycle	Duration, Lifestage
Vector_Transmission	Transmission_Type {Trans-ovarial:Trans-Stadial}
Individual	An individual participating in a food consumption survey.
IndividualIdentifier	Unique identifier for each individual included in dimension
Gender	Sex of the participant
AgeClass	Age classification of individual
Education	Level of education described using international Standard classification of education 1997
EthnicGroup	Description of ethnic group of individual, this is distinct of nationality
Professional	

Activity/Function	Professional activity (e.g. Veterinarians, farmers, pet owner, doctor, employee)
Formation	Continued formation followed and frequency
Diploma	Diploma obtained
Expert_Status	Expert status or not
Population	
Population_Type	Study_Population: Population of individuals (animals or groups of animals) selected to participate in the study (regardless of whether or not they actually participate). Target_Population: Immediate population to which the study results will be extrapolated. The subject (items, animals, batches) included in the study would be derived from the target population External_Population: The total population that one would ideally like to be able to extrapolate results to. It might vary depending on the perspective of the individuals interpreting the result of the study.
Food/feed	Food or feed item sampled for purpose of survey and monitoring
Food/Feed	
Food/Feed item	Identifier of the food or feed of the classification chosen
Synonyms	Synonyms used to name the food or feed
Scientific Name	Scientific names to name the food or feed
Ingredients	List of ingredient composing the food or feed
Packaging	Packaging of the food or feed
Type	Pasture vs concentrates
Animal_Origin	Animal vs non-animal product {Y:N}
Nutritional characteristics	Nutritional characteristics and additive
Processing	Indicates if the food or feed is processed or not {Y:N}
Treatment	Indicates if the food or feed is subject to a treatment {Y:N}
Origin_Geo	Place of origin of the product
Manufacturing_Date	Date of production of the food or feed
Parameter	Parameter that the outcome values represent
Parameter	Full description of parameter at the lowest level. That could be describe by more standard with different levels of hierarchy: OrganicClass (describing organic compounds), ToxinClass (describing toxins, poisonous substance produced by living cells or organisms), MetalsClass (describing metallic elements in order to account mettalic compounds and transition metals), PesticidesClass (describing Pesticides to discriminate between pesticide residues definitions and pesticide residues), ZoonosesClass (describing zoonotics agents and other pathogens)
Disease Transmission	

Transmission_Context	If the transmission is arrived naturally or was induced by an experimentation {Natural:Experimental}
Transmission_Mode	Whether the transmission is direct or indirect {Direct:Indirect}
Transmission_Type	If the transmission is from one individual to another in the same generation (Horizontal) or is from one parent to his offspring (Vertical or mother-to-child transmission) {Vertical, Horizontal}
Pathogen_Entries	Entries by which the transmission occurred (e.g. Insemination, Vector)
Transmission_Vehicle	Support of the transmission (e.g. Food, Blood, Air)
Case Definition	
Standard	Whether the case definition is a standardized definition and its level {National:Local:International:European} or not {no}
Study_Specific	Whether the case definition has been realized during a specific study {Y:N}
Case_Category	Classification of case, function of levels of certainty. {Suspect:Probable:Confirmed}
Clinical_Criteria	Clinical sign or lesions which allow confirming type case → **metadata Disease_Clinical_Signs**
Epidemiological_Criteria	Epidemiological information allowing determining the type of the case.
Biology_Criteria	Analytical method and validation criteria used which allow confirming type case → **metadata Analytical Method**
Outbreak	
Outbreak_Type	Type of outbreak: Primary or Annex
Outbreak_Detection	Circumstances of the outbreak detection (e.g. Inspection in Slaughterhouse)
Outbreak_Evidence	Level of outbreak evidence
Outbreak_Source	Detection entity of the biological agent (e.g. Organism, Feed/food…)
Sample	
Sample	Sample tested in the laboratory or Population that the sample represents.
SampleMethod	Method for selecting or collecting sampling units
SampleSize	Size of the sample (e.g. number of sample taken in a population unit, weight of an individual sample)
SampleFrame	List of the sampling units in the sample lot or population.
Sample Matrix	
Sample_Matrix	Sampled entity (e.g. Blood, Faeces, Dust, Milk, etc.)
Matrix_Type	Type of the matrix {Live animal sample, animal product, animal genetic material, feedstuffs, biological products, pathological material, plant, Environmental sample}
Sampling	

Latitude	Latitude of sampling
Longitude	Longitude of sampling
Altitude	Altitude of sampling
CoordinateType	Coordinate Type of sampling
Temperature	Temperature during sampling
CRS	Indicates the reference of coordinate system.
Sampling Point	
SamplingPointL1	Highest level to describe the sampling point
Sample point	Description of the sampling point
Analytical Method	
AnalyticalMethod	Description of the analytical method (e.g. GC-MS-MS, Student's test)
AnalyticalMethodL1	First level of the analytical method description (Biological, Chemical or Statistical)
AnalyticalMethodL2	Second level of the analytical method description (e.g. Chromatographic Test, Atomic Spectrography, Test of Association)
AnalyticalMethodL3	Third level of the analytical method description (e.g. Gas Chromatography, Liquid Chromatography)
Test	
Test_Status	Status of the diagnostic test: Standard, Complementary, Unknown
Type_Test	Type of the test considered (e.g. Diagnosis Test, Agent Identification)
Name	The commercial name of the test
Protocol	
Status	The status of the protocol applied: The protocol describes in the notice, a specific protocol given by OIE or legislation, or a specific protocol of user.
Material_Used	The description of the material used like the reference of the serum or the type of instrument used.
Protocol_Link	The complete description of each step of the protocol (e.g. Direct link toward the protocol sheet)
Laboratory	
Name	The name of the laboratory
Laboratory_Status_Accreditation	Accreditation status of the laboratory performing the analytical method {accredited, third party assessment, none}
Procedure_Accreditation	Management Accreditation procedure within the laboratory, operation and effectiveness of the quality management system within the

	laboratory (ISO/IEC 17025, third party assessment, internally validated, not validated) / Technical requirement for the analytical method used in the laboratory, factors which determines the correctness and reliability of the tests and calibrations performed in laboratory (ISO/IEC 17025, third party assessment, internally validated, not validated)
Reference_Status	Allow to determine if this laboratory is a referenced laboratory listed by OIE.
Program	
Program_Type	Whether the program is continuous or punctual
Program_Purpose	Objectives and purposes of the program
Testing_Frequency	Frequency of the testing within the time scale
Time_Scale	Time scale of the program
Administrative_Units	Units included in the program: States, regions, zones, country, Zip Code Areas, statistical reporting units, sample grid references, neighbourhoods, parcel (USDA, 2006)
Size_Of_Sample_Service_Area	Number of reporting units (e.g., labs, clinics, slaughter plants), of geographic area per unit sampled, of eligible units per reporting unit (USDA, 2006)
Study	
StudyL1	Descriptive or Analytical study {(e.g. Analytical Study)
StudyL2	Second level of the study description with regards to StudyL1 (e.g. Experimental_Study)
StudyL3	Third level of the study description with regards to StudyL2 (e.g. With_Comparaison Randomized Controlled Trial)
Study	Study (e.g. Randomization)
Data Collection Design	
Legislation	Legislation frame of the program
Level	Level at which the program is designed and implemented (EU, National, Industry, Research)
Disease_Context	Sample infectious phase and/or disease context
Sampling_Strategy	EUROSTAT typology of sampling strategy (see appendix H)
Commodity	
Commodity_Name	Name of the commodity (UN, Commodity List)
Commodity_Description	Description of the commodity (UN, Commodity List)
Commodity_Type	Type of the commodity: animals, animal product, animal genetic material, feedstuffs, biological products, pathological material, (OIE Terrestrial Animal Health Code 2011) and plant
Production_Chain	Type of the production chain: Food, Insemination, Pharmaceutical
Commodity_Level	Whether the commodity is a Primary Product or a by Product
Animal_Origin	Possible animal origin of the commodity {Y:N}

Production System	
ProductionSystemL1	Type of the production system (Dairy, Meat, Fight, Reproduction system, Egg, Wool, farming…)
ProductionSystemL2	Whether the production system is intensive or extensive {Intensive:Extensive}
ProductionSystemL3	Outdoor system (pasture system), Indoor System, With outdoor environment access
Housing System	Description of the housing system: condition in which the animal are conserved (Cages [Battery cages / Furnished cages], Littered floor, Compartimentation.)
Facility	
Facility_Type	Activity of the facility (e.g. industry, laboratory, farm, veterinary clinic, insemination center, pet shop, riding school, border post, zoo, slaughter house)
Facility_Material	Material used
Movement	
Movement_TypeL1	Trade, Event, Inspection, Wild, Farming, or Tourism
Movement_TypeL2	Movement direction, geographical or temporal movement context (e.g. Importation, Exportation, Seasonal Migration, Intra_Area...)
Movement_Scale	National, European or International
Movement_Pathway	Global pathway, including departure zone, zones crossed and arrival zone
Movement_Duration	Duration of the movement.
Movement Step	
Step_Duration	Duration of the step
Transport	
Departure	The zone or the facility from which the transport start
Arrival	The final destination (Geographical area or the facility) of the transport
Crossed_Zone	The crossed zone (Geographical area or the facility) during the transport
Transit	
Place	Facility or a geographical area
Authorization_Status	Transporter_Authorization, Certificate...

Vehicle	
Identification	Matriculation number of the vehicle
Vehicle_Type	Conveyance, including train, truck, aircraft or ship (OIE Terrestrial Animal Health Code 2011)
Vehicle_Capacity	Capacity of the vehicle
Vehicle_Speed	Speed of the vehicle
Vehicle_Condition	Condition in which the vehicle is (e.g. Closed, Frozen, Chilled)
Safeguard System	
Safeguard_Type	Considered safeguard system (e.g.Veterinary_Services:Quarantine:Biosecurity:Inspection:Public_Health_Service)
Safeguard_Quality	Protocol quality criteria
Biosecurity	
Biosecurity_Level	Bio exclusion, Bio-confinement, Bio-compartimentation
Biosecurity_Type	The type of bio-security system, function of its role. (e.g. Waste Management,

GLOSSARY [AND/OR] ABBREVIATIONS

Accidental host: a host for an infection that is not part of the normal ecology of the infectious agent/parasite. Most accidental hosts are also 'dead end' hosts, i.e. they do not transmit the infection further (Reproductive rate ratio (R_0<<1)). Some, however, may maintain the infection for a few generations of transmission, or even pass it on to a different host species (liaison hosts).

Accuracy of data: data are accurate if they meet characteristics proposed by 'gold standards'. This criterion includes data validity – the data capability of a data item to measure what it is meant to, and reliability – the capability of a data item to measure what it is meant to when the measurement is repeated (Kirch, 2008). It is indeed a criterion of data quality.

Aggregation: is a form of UML association that implies the collection of one class of objects within another (Sparks G. 2011).

AHAW opinions most recurrent risk questions: both AHAW opinions explicit (ToR) and implicit risk questions that are required in order to properly run the risk model.

Animal: means a mammal, bird or bee (OIE Terrestrial Animal Health Code, 2011).

Animal product: means product which are animal origin. The pathological material means samples obtained from live or dead animals, containing or suspected of containing infectious or parasitic agents, to be sent to a laboratory (OIE Terrestrial Animal Health Code, 2011).

Antimicrobial Resistance: the ability of microorganisms of certain species to survive or even to grow in the presence of a given concentration of an antimicrobial agent, that is usually sufficient to inhibit or kill microorganisms of the sample species (Dir. 2003/99/EC). Resistance against an antimicrobial is considered to be present if the Minimum Inhibitory Concentration (MIC) exceeds the breakpoint or the epidemiological cut-off value (EFSA DWH).

Association: is a UML relationship between 2 classes indicating that at least one side of the relationship knows about and somehow uses or manipulates the other side. This relationship may by functional (do something for me) or structural (be something for me). For example, structural relationship: an Address class may be associated with a Person class (Sparks G 2011).

Biological products: are those products derived from living organisms which are manufactured and distributed in accordance with the requirements of appropriate national authorities, which may have special licensing requirements, and are used either for prevention, treatment, or diagnosis of disease in humans or animals, or for development, experimental or investigational purposes related thereto. (CDC) (OIE Harmonisation and improvement of registration, distribution and quality control of vaccines in the Middle East).

Case: means an individual animal infected by a pathogenic agent, with or without clinical signs (OIE Terrestrial Animal Health Code, 2011).

Categorization of data: one way of organizing information on objects or ideas. The process of recognition, differentiation, and understanding of objects by grouping into categories, usually for

statistical analysis or graphic representation. Category is from late Latin categoria — a division within a system of classification.

Clarity of data: clarity refers to the data's information environment: whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality also available (including limitation in use) and the extent to which additional assistance is provided by National Statistical Institutes (OECD, 2008).

Class Diagram: a UML diagram shows a collection of declarative (static) UML model elements such as classes and types, with their contents and relationships (OMG 2012).

Coherence of data: coherence of statistics is their adequacy to be reliably combined in different ways and for various uses (OECD, 2008). Coherence of data reflects the degree to which the data from a single statistical program, and data brought together across data sets are logically connected and complete. Fully coherent data are logically consistent – internally, over time, and across products and programs.

Commodity: means live animals, products of animal origin, animal genetic material, biological products, pathological material (OIE Terrestrial Animal Health Code, 2011) and Plant.

Comparability of data: comparability is the extent to which differences between data from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the data (OECD, 2008). It measures the impact of differences when data are compared between geographical areas, non-geographical domains, or over time (European Commission, 2007).

Correctness of data: the correctness of data can be defined as a measure of the proximity of a data value to some other data value that is considered correct.

Crude data: data not processed or subjected to analysis (= raw data). Forms of crude data are generally Excel, Access, Text files, etc.

Data: a collection of items of information. A data is specific and generally measurable. One data is related to one specific parameter. It is important to distinguish the terms ‘information’ and ‘data’.

Data accessibility: physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro- or macro-data, various formats (paper, files, CD-ROM, Internet, etc.), etc. (European Commission, 2007). In the present project, three categories of accessibility were considered, with the scoring system applied in brackets: (1) **No access:** no data or data not found (score = 0); (2) **Limited access:** the access to data is limited if restricted to some categories of persons (e.g. sanitary authorities) or if a charge is required to access data (score = 1); (3) **Free access** (score = 2).

Data availability: refers to the degree to which data can be instantly accessed. Data are available if not only they exist, but also can be accessed easily. Data availability is also a criterion of data quality (Kirch, 2008). The following scoring system was applied when handling data: (1) No data available or data not found (score = 0); (2) Some data available (score = 1); (3) All data available (score = 2).

Data collection: process of gathering data from various sources (Kirch, 2008).

Data completeness: extent to which all data that are needed are available. It is usually described as a measure of the amount of available data from a statistical system compared to the amount that was expected to be obtained (European Commission, 2007).

Data consistency: consistency refers to data values in one data set being consistent with values in another data set. A strict definition of consistency specifies that two data values drawn from separate data sets must not conflict with each other, although consistency does not necessarily imply correctness (Loshin, 2006).

Data gaps: the lacking data of the report (either numerical or not) said to be necessary, either by experts or by reviewers, to answer opinions' risk questions. Reviewers created for each mandate, an exhaustive list of "data gaps" related to a specific ToR without paying attention, neither on data availability nor on data accessibility. This assessment was done taking into account the more idealistic context of risk assessment studies. The requirements specified by experts, at the time of the mandate, could differ from the ones mentioned by reviewers, at the time of this study, because of a different scientific background or contexts.

Data integrity: data integrity is the quality or condition of being whole and unaltered, and it refers to the consistency, accuracy and correctness of data (Khosrowpour 2007).

Data not found: within the frameworks of the indirect survey (web searches), only a specific time (10 minutes) was assigned to each parameter to investigate. When no data were found after that specific timeframe, we specified 'data not found'. It does not mean they do not exist, but only that they were not found within the timeframe assigned to the search.

Data quality: data has quality if it satisfies the requirements of its intended use (Olson, 2003). Data quality depends as much on the intended use as it does on the data itself. To satisfy the intended use, several attributes characterise collectively the quality of data: relevance, accuracy, timeliness, punctuality, comparability, coherence, accessibility and clarity (European Commission, 2007).

Data source (named **resource** in the model): a specific data set, metadata set, database or metadata repository from where data or metadata are available. The source of data is often used as a synonym for the term 'data provider' (OECD, 2008)

Data used: the data (either numerical or not) used by AHAW Scientific Panel in the opinions in order to answer the different opinions' risk questions. Data used were therefore characterized by their "Source(s)" (e.g. scientific literature data, data from quarantine centers…).

Data Warehouse: the EFSA data warehouse (DWH) system is a subject-oriented, integrated, time-variant and non-volatile repository of data designed to support reporting and analysis (EFSA DWH).

Dataset: represents the data that are collected by EFSA from each Member State. It includes the "facts" (data from MS directly collected by EFSA) and the "metadata" of these facts.

Definitive host: the host in which sexual maturation of a parasite occurs (Porta, 2008).

Degree of precision: the quality of being sharply defined or stated. One measure of precision is the number of distinguishable alternatives from which a measurement was selected, sometimes indicated by the number of significant digits in the measurement (Porta, 2008).

Dimension (DWH): the dimensions are object that contains descriptive attributes (or fields) that provide information about outcome value stored in the data warehouse. (EFSA DWH).

Disease: means the clinical and/or pathological manifestation of infection (OIE Terrestrial Animal Health Code, 2011).

Disease general information: compiles all information and data related to the natural history of the disease such as incubation period, duration of clinical signs. These data are not affected by local/national contexts.

DOI number: A DOI ("digital object identifier") is a unique identification code that allows an article to be easily tracked by many different archival and research programs on the Internet. Every article that is accepted for publication in an APS Journal is assigned a DOI number. It's the copyeditor's responsibility to make sure that this number is formed correctly. The DOI number is composed with the publisher code (first 6 digits), the journal abbreviation (string of alphabetic characters), and the article number (last nine digits of the DOI).

Early warning system: in disease surveillance, a specific procedure to detect as early as possible any departure from usual or normally observed frequency of phenomena.

Epidemiological unit: means a group of animals with a defined epidemiological relationship that share approximately the same likelihood of exposure to a pathogen. This may be because they share a common environment (e.g. animals in a pen), or because of common management practices. Usually, this is a herd or a flock. However, an epidemiological unit may also refer to groups such as animals belonging to residents of a village, or animals sharing a communal animal handling facility. The epidemiological relationship may differ from disease to disease, or even strain to strain of the pathogen (OIE Terrestrial Animal Health Code, 2011).

Fact (DWH): 1) A measurement value, often numeric and typically aggregated, stored in a data warehouse. (Named "**facts**" in this report). 2) A schema object representating a column in a data warehouse table and containing basic or aggregated numbers (Named "Fact table" in this report). (EFSA DWH).

Facts (AHAW DATASPEC Model): see Member State data.

Form of data: the informatics format of data. Several forms were considered and a scoring system was also applied as follows (in an increasing order of utility for data handling): 0 = not found; 1 = unknown; 2 = PDF; 3 = HTML; 4 = TEXT; 5 = EXCEL.

Hidden data (acronym = 'underground' data): data that are not directly accessible through a classical way; data for which the existence is not always known by potential users.

Implicit ToR: opinions' implicit ToR could either be identified by experts, at the time of the study and reported by reviewers, or identified by reviewers themselves, at the time of the review. Gathered implicit questions were either based on structural features in the text (e.g. chapter headings) or by evaluating textual descriptions in the opinions.

Incidence: means the number of new cases or outbreaks of a disease that occur in a population at risk in a particular geographical area within a defined time interval (OIE Terrestrial Animal Health Code, 2011).

Incubation period: means the longest period which elapses between the introduction of the pathogen into the animal and the occurrence of the first clinical signs of the disease (OIE Terrestrial Animal Health Code, 2011).

Infection: means the entry and development or multiplication of an infectious agent in the body of humans or animals (OIE Terrestrial Animal Health Code, 2011).

Information: facts that have been arranged and/or transformed to provide the basis for interpretation and conversion into knowledge (Porta, 2008). In the context of the present project, the term 'information' is also used to combine data related to a specific aspect of the disease, e.g. information on vectors gathers data on the type of vector, the geographical distribution, etc. It is important to distinguish information from data. Information can include several related data. For example, data on the number of dairy and meat herds provide information on the production system.

Intermediate host: in parasitology, the host in which asexual forms of the parasite develop.

Inventory of data sources: the complete list of data sources compiled thanks to the direct (questionnaires and workshop) and indirect (web searches) surveys.

ISBN number: (International Standard Book Number). An ISBN is assigned to each edition and variation (except reprinting) of a book. The ISBN is 13 digits long if assigned after January 1, 2007, and 10 digits long if assigned before 2007.

Literature Review (LR): descriptive review of the different publications on a specific topic over a large period of time.

Logical data model: is a representation of an organization's data, organized in terms entities and relationships and is independant of any particular data management technology (DWH).

Lookup (DWH): look up list represent possible value terminology associated with attributes and required for the correct interpretation.

Member State data: data related to the Member State that will be useful for all the case-studies, e.g., animal species demography, trade data, etc. In the context of the present project, the Member State Data are considered as the "facts" of the AHAW model.

Metadata: data that defines and describes other data and processes. This means that metadata are data that describe other data, and data become metadata when they are used in this way. This happens under particular circumstances and for particular purposes, as no data are always metadata. The set of circumstances and purposes (or perspective) for which some data are used as metadata is called the context. So, metadata are data about data in some context (OECD, 2008).

Narrative approach: the narrative approach was any reports that presented only a descriptive review of the different papers published on a specific topic over a large period of time.

Natural history of the disease: the course of a disease from pathological onset or inception to resolution. Many diseases have certain relatively well-defined stages that, taken all together, are referred to as the 'natural history of the disease' in question. These stages are as follows (Porta, 2008):

1. Stages of pathological onset. They are constantly being changed in many diseases; "onset," in particular, tends to be redefined in increasingly smaller microbiological (e.g., molecular and genetic) terms.

2. Pre-symptomatic stage: from initiation of disease to the first appearance of symptoms and/or signs.

3. Clinically manifest disease, which may progress inexorably to a fatal termination, be subject to remissions and relapses, or regress spontaneously, leading to recovery.

Needed data: data that were defined as necessary, by risk assessors and Consortium members, after reading the risk questions selected in WP1.

No data: within the frameworks of the indirect survey (web searches), the terminology 'no data' was used when no data exist on the specific parameter (e.g. no data exist on the possible reservoir of Venezuelan equine encephalitis in France, as the disease is still exotic to data and has not been investigated so far).

Ontology: ontology is used to mean different things, e.g. glossaries & data dictionaries, thesauri & taxonomies, schemas & data models, and formal ontology & inference. A formal ontology is a controlled vocabulary expressed in an ontology representation language. This language has a grammar for using vocabulary terms to express something meaningful within a specified domain of interest.

Outbreak: means the occurrence of one or more cases in an epidemiological unit (OIE Terrestrial Animal Health Code, 2011).

Parameter: parameter that the outcome value represents. In the context of the present project, it's could be a disease, a biological agent or a substance.

Pathological material: means samples obtained from live or dead animals, containing or suspected of containing infectious or parasitic agents, to be sent to a laboratory (OIE Terrestrial Animal Health Code, 2011).

Population: means a group of units sharing a common defined characteristic (OIE Terrestrial Animal Health Code, 2011).

Prevalence: means the total number of cases or outbreak of a disease that are present in a population at risk, in a particular geographical area, at one specified time or during a given period (OIE Terrestrial Animal Health Code, 2011).

Public access: unrestricted access to data.

Punctuality of data: punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar or previously agreed among partners (OECD, 2008).

Qualitative approach: the qualitative approach (QL) was any tool that was not clearly identified as quantitative in the opinion or that did not clearly use quantitative data (even semi-quantitative approaches).

Quantitative approach: the quantitative approach (QN) was any tool that was used by AHAW experts to estimate the risk based on quantitative data.

Reference Laboratory: laboratory of recognised scientific and diagnostic expertise for a particular animal disease and/or testing methodology; includes capability for characterising and assigning values to reference reagents and samples (OIE Terrestrial Manual 2008).

Relevance: degree to which data meet current and potential users' needs. IT refers to whether all data that are needed are produced and the extent to which concepts used (definitions, classifications, etc.) reflects user needs.

Reservoir: any person, animal, arthropod, plant, soil, or substance, or combination of these in which an infectious agent normally lives and multiplies, on which it depends primarily for survival, and where it reproduces itself in such a manner that it can be transmitted to a susceptible host (Porta, 2008).

Restricted access: imposing conditions on access to data. Access of users to the protected data is restricted and they can only have a certain number of outputs (e.g. tables) or maybe only outputs of a certain structure (OECD, 2008). In the present project, a restricted access is used to specify an access authorized for certain categories of persons (e.g. animal health authorities), but not for the scientific community in general.

Sample: material that is derived from a specimen and used for testing purposes (OIE Terrestrial Manual 2008).

Sample frame: means complete list of all sampling units in the target population, from which the sample can be selected. It should have the property that every element in the population has some chance to being selected in the sample by whatever method is used to select elements from the sampling frame (Dohoo et al., 2003, USDA, 2006).

Sensitivity (analytical): synonymous with 'Limit of Detection', smallest detectable amount of analyte that can be measured with a defined certainty; analyte may include antibodies, antigens, nucleic acids or live organisms (OIE Terrestrial Manual 2008).

Sensitivity (diagnostic): proportion of known infected reference animals that test positive in the assay; infected animals that test negative are considered to have false-negative results (OIE Terrestrial Manual 2008).

Slaughterhouse: means premises, including facilities for moving or lair aging animals, used for the slaughter of animals to produce animal products (OIE Terrestrial Animal Health Code, 2011).

Specificity (analytical): degree to which the assay distinguishes between the target analyte and other components in the sample matrix; the higher the analytical specificity, the lower the level of false-positives (OIE Terrestrial Manual 2008)

Specificity (diagnostic): proportion of known uninfected reference animals that test negative in the assay; uninfected reference animals that test positive are considered to have false-positive results (OIE Terrestrial Manual 2008)

Spillover (host): in a spillover host, infection will not persist indefinitely unless there is re-infection from another species. Infection in spillover hosts will disappear progressively if the disease is

eliminated or reduced in the species which is acting as the source. A spillover host may be a dead-end host if it plays no significant role in the onward transmission of infection or at the other extreme may be an amplifying host which increases the prevalence of infection in domestic stock or expands the number of species affected by the disease (Morris and Pfeiffer, 1995).

Surrogate: when no direct data are available, indirect related data can be generated.

Surveillance: represents an extension of monitoring and consist of the close and continuous observation of the occurrence of infection for the purpose of active control (Noordhuizen et al. 2001, Report of the guidance on good practices for Design of field survey)

Susceptible species: a species presenting the capacity to clinically exhibit the disease after being in contact with a pathogen agent (Toma et al., 2001)

Syndromic surveillance: syndromic Surveillance uses individual and population health indicators that are available before confirmed diagnoses or laboratory confirmation to identify outbreaks or health events and monitor the health status of a community (Centers for Disease Control and Prevention[12])

Terms of Reference (ToR): represent the AHAW opinion most recurrent risk questions that are theoretically the questions asked to AHAW Panel by the Commission. These most recurrent risk questions represented the main requests of the different Opinions that AHAW scientific panel had to answer to.

Timeliness of data: length of time between data availability and the event/phenomenon it describes (European Commission, 2007).

Type of data sources: the following categories of data sources were considered: (1) **Book**: online book or book for which extracts can be consulted online if not owned in a paper format; (2) **Course**: online course (University); (3) **Factsheet:** sheet compiling complete information on the disease considered (e.g. epidemiology, clinical disease, control, prevention, etc.); (4) **Guidelines**: information intended to advise people on how something should be done or what something should be; (5) **Legislation**: texts retrieved from the European or National legal texts; (6) **Original Articles**: Scientific publications submitted to a lecture committee; (7) **Proceedings:** abstracts from booklets related to (inter-)national congresses or conferences; (8) **Report:** document compiling information and results related to a specific thematic and emitted by a scientific team or an (inter-)national organism; (9) **Review:** a review article compiling information gathered from original articles previously published; (10) **Website:** data only present on the website page (online data source).

UML modeling: unified Modeling Language is the standard language for analysis and design of applications, specifying the structure and behaviour of systems. UML is defined as an underlying abstract syntax and an overlying graphical concrete representation (OMG, 2012).

Vaccination: means the successful immunisation of susceptible animals through the administration, according to the manufacturer's instructions and the Terrestrial Manual, where relevant, of a vaccine comprising antigens appropriate to the disease to be controlled (OIE Terrestrial Animal Health Code, 2011).

[12] http://www.cdc.gov/EHRmeaningfuluse/Syndromic.html

Vector: means an insect or any living carrier that transports an infectious agent from an infected individual to a susceptible individual or its food or immediate surroundings. The organism may or may not pass through a development cycle within the vector (OIE Terrestrial Animal Health Code, 2011).

Vehicle: means any means of conveyance including train, truck, aircraft or ship (OIE Terrestrial Animal Health Code, 2011).

Veterinary Services: means the governmental and non-governmental organisations that implement animal health and welfare measures and other standards and recommendations in the Terrestrial Code and the OIE Aquatic Animal Health Code in the territory. The Veterinary Services are under the overall control and direction of the Veterinary Authority. Private sector organisations, veterinarians, veterinary paraprofessionals or aquatic animal health professionals are normally accredited or approved by the Veterinary Authority to deliver the delegated functions (OIE Terrestrial Animal Health Code, 2011).

Web Service: a Web service as a software system designed to support interoperable machine-to-machine interaction over a network (W3C, 2011).

Zoonosis: means any disease or infection which is naturally transmissible from animals to humans. (OIE Terrestrial Animal Health Code, 2011).

ABBREVIATIONS

ANMV = *Agence Nationale du Médicament Vétérinaire* (France) – National Agency of Veterinary Medicines

ARSIA = *Association Régionale de Santé et d'Identification Animale* (Wallonia, Belgium) – Regional Association for Animal Health and Identification

AWE = *Association Wallonne de l'Elevage* (Wallonia, Belgium) – Walloon Breeding Agency

BDNI = *Base de données Nationale d'identification* (France) – National cattle base register

CNEV = *Centre National d'Expertise sur les Vecteurs* (France) **–** National Centre of Expertise on vectors

CNIEL = *Centre National Interprofessionnel de l'Economie Laitière* (France) – National Inter-professional Centre of Dairy Economy

Coop de France = *Entreprise Coopérative Agricole* (France) **-** French Agricultural Cooperative

CRL = Community Reference Laboratory

CVI = Central Veterinary Institute – Wageningen UR (Netherlands)

DDCSPP = *Direction Départementale de la Cohésion Sociale et de la Protection des Populations* (France) – Departmental Direction of Social Cohesion and Population Protection.

DGAL = *Direction Générale de l'Alimentation* (France) **–** General Directorate for Food of the French Ministry of Agriculture, Food, Fisheries and Rural Affairs

DGS = *Direction Générale de la Santé* (France) **–** General Directorate of Health

DGZ = *Dierengezondheidszorg Vlaanderen* (Flanders, Belgium) – Flemish Animal Health Care

EDI-Sasha = *échange de données informatisées de laboratoire* (France) **–** Harmonized laboratory electronic data transfer

EDI-Span = *échange de données informatisées des clos d'équarrissage* (France) – Harmonized rendering plants electronic data transfer

EL&I = *Ministerie van Economische Zaken, Landbouw en Innovatie* – Dutch Ministry of Agriculture

FAMPH = Federal Agency for Medicines and Health Products (Belgium)

FASFC = Federal Agency for the Safety of the Food Chain (Belgium)

GD = *Gezondheidsdienst voor Dieren* (Netherlands) – Animal Health Services

GDS = *Groupements de Défense Sanitaire* (France) – Sanitary Defense Groupings.

NERGAL = Slaughterhouse condemnation database (France)

ONCFS = *Office National de la Chasse et de la Faune Sauvage* (France) – National Office for Wildlife and Hunting.

OSP = *Organisation de Sélection Porcine* (France) – Swine Selection Organization

RESPE = Réseau d'Epidemio-Surveillance en Pathologie Equine (France) - Equine disease epidemiological surveillance Network

RIVM = *Rijksinstituut voor Volksgezondheid en Milieu* (Netherlands) – National Institute for Public Health and the Environment

SAGIR = *Réseau de surveillance sanitaire de la faune sauvage en France* - Wildlife epidemiological surveillance system

SIGAL = *Système d'Information de la DGAL* (France) - General Directorate for Food of the French ministry of Agriculture, Food, Fisheries and Rural Affairs Information System

UNCEIA = *Union Nationale des Coopératives d'Elevage et d'Insémination Animale* (France) – National Union of Farming and Animal Insemination Agricultural Cooperatives

VAR = Veterinary and Agrochemical Research Center (Belgium)

VIC = *Veterinair Incidenten- en Crisiscentrum* – Database on notifiable diseases (Netherlands)

VWA = *Voedsel en Waren Autoriteit* (Netherlands) - Agency for Safety of Food and Consumer Products

VWS = *Ministerie van Volksgezondheid, Welzijn en Sport* (Netherlands) – Dutch Ministry of Public Health