Project no. 310806

RISKSUR

Providing a new generation of methodologies and tools for cost-effective risk-based animal health surveillance systems for the benefit of livestock producers, decision makers and consumers

Task 1.4: Development of an integrated evaluation matrix

Report for the deliverable No. 1.4

The EVA tool: an integrated approach for evaluation of animal health surveillance systems

KBBE
Collaborative Project
FP7-KBBE-2012-6

www.fp7-risksur.eu

Report prepared by (alphabetic order):

Clémentine Calba (CIRAD), Angus Cameron (AusVet), Flavie Goutard (CIRAD), Vladimir Grosbois (CIRAD), Barbara Haesler (RVC), Linda Hoinville (AHVLA), Ann Lindberg (SVA), Julio Pinto (FAO), Marisa Peyre (CIRAD), Dirk Pfeiffer (RVC), Daniel Traon (ARCADIA), Timothée Vergne (RVC).

Report reviewed internally by:

Katharina Stärk (SAFOSO) and Detlef Höreth-Böntgen (FLI)

Date: 16/12/2013
Type: Deliverable

Table of contents

Acronyms

AH	Animal Health
ASF	African Swine Fever
CEA	Cost-effective analysis
CSF	Classical Swine Fever
DALYs	Disability-adjusted life years
DM	Decision makers
EC	European Commission
EU	European Union
HPAI	Highly pathogenic avian influenza
MS	Member State
QALYs	Quality-Adjusted life years
SS	Surveillance Systems

SUMMARY

Surveillance systems need to be tailored to epidemiological systems which are driven by epidemiological, ecological, economic, social (including political, cultural) and environmental factors. Given the almost continuous changes occurring in this system, it is essential to regularly re-evaluate the effectiveness of the surveillance programme. This requires the design of comprehensive, practical, and affordable evaluation frameworks for timely assessment of not only the benefits and costs of a surveillance and control program but also the factors required for local acceptance, which itself is crucial for the effectiveness and sustainability of the system at national and international levels.

The use of economic evaluation in decision making for animal health surveillance policies has been limited so far. When considering the start, end or change of a surveillance programme, policy makers need to know if and how much surveillance is needed and what the most economical options are and how it best integrates with interventions. There are few specific guidelines available for economic evaluation of animal health surveillance, and only a limited number of empirical applications on the economic value of surveillance. It is critical to select appropriate economic efficiency criteria taking into account the evaluation context and viewpoint of the analysis. Priority setting, affordability, sustainability, social acceptance and communication are all issues that policy makers have to consider when designing and implementing disease management policy. These assessments are not that easy as the data available to quantify investments in surveillance and monitoring activities are very limited at both national and international levels. It is expected that MS would value access to a tool that would help them with technical and budgetary optimization of resource utilization when defining animal surveillance policy to as part of the national animal health policy.

Therefore one of the objectives of the RISKSUR project is to develop a practical tool to guide the decision makers in performing economic evaluation of their animal health surveillance systems. The first year of the project has been dedicated to the assessment of the challenges linked to this type of economic evaluation and decision support tool. A conceptual model of a decision support tool for economic evaluation of animal health surveillance system has been developed (EVA tool) based on several expert meetings and discussions. The EVA tool builds on existing evaluation framework, methods and tools and provides an integrated support guide for economic and epidemiological evaluation of animal health surveillance systems (SS). The aim of the EVA tool is to develop a standardized and internationally recognised approach to the evaluation process which provides flexibility to adapt to specific contexts. The objective is to develop a practical and easy stepwise guidance for the evaluation process along with a comprehensive toolkit. The users will have to provide inputs in relation to the evaluation question and the general context (epidemiological situation, surveillance objectives, and data availability) defining the context of the evaluation process. The tool then generates an optimum selection of evaluation attributes and corresponding measurement methods. The tool will provide detailed guidance on the application of the methods and the interpretation of the results. More details on the structure and logic of this tool are presented in this report together with the evaluation and development challenges yet to be addressed.

1 Challenges in Animal Health Surveillance Evaluation

1.1 Rationale for the evaluation of animal health surveillance systems

The need to set up efficient and sustainable surveillance systems (SS) has become a major concern particularly in recent years, following the SARS and H1N1 influenza pandemics which have also highlighted the importance to detect new diseases (Shahab 2009). In the field of animal health, SS provide information for effective disease control thereby improving productivity and food security, animal welfare, economic development and access to international trade. Moreover, around 75% of emerging infectious diseases in humans are zoonoses (Granger, 2011; Taylor et al., 2001). Therefore the capacity of SS to accurately describe patterns of animal diseases is of public health importance. Information about infectious diseases at a global scale relies on government surveillance systems. And yet the resources and reliability of these systems can vary considerably, especially in countries characterized by weak economies or political instability (Jebara, 2004). **To make best use of available resources, it is critical to perform timely and relevant evaluations of SS** (Shahab 2009). Evaluation is one essential step in the policy cycle (Jann and Wegrich, 2007). Evaluation which implies a judgment on the SS and recommendations for improvement is a critical part of surveillance in that it allows transparent interpretation of outputs, more objective decision making and resource allocation as well as improvements in system design and enhanced acceptance of system outputs by stakeholders (at local (e.g. farmers, veterinarians) and national levels (e.g. reference laboratory, veterinarians at central level)). This is particularly important given the knowledge gaps in our understanding of many diseases, which leads to varying degrees of uncertainty and bias in generated outputs. This in turn influences the assessment of the added value of new surveillance components such as risk-based surveillance which is based on the principle that the probability of danger occurring in the population is variable and that preferential surveillance is being performed in groups/zones that are being defined as " at risk" . The goal of these risk-based approaches is to allow for optimal resource allocation (Stärk et al., 2006).

Surveillance systems are complex, and need to be adapted to epidemiological systems driven by epidemiological, economic, social (including political, cultural) and environmental factors. To allow the design of cost-effective SSs, there is a need to design comprehensive, practical and affordable evaluation frameworks for timely evaluation of not only the benefits and costs of a surveillance and control program but also the factors required for local acceptance, which itself is crucial for the effectiveness and viability of the system at national and international scale (Antoine Moussiauxet al., 2011). Moreover, an assessment of system efficiency has to take account of each countries' specific needs and resources, and it has to be quantitative as much as possible to minimise the impact of subjectivity. In order to control diseases, institutional constraints must be considered together with technical aspects (FAO, 2009).

1.2 Limits of the existing evaluation frameworks

The review of the *evaluation methods of surveillance systems and current practices* performed as part of the RISKSUR project (D.1.2) has highlighted the importance and need to develop an integrated approach for epidemiological and economic evaluation of surveillance systems. This could be based on existing approaches but should also provide practical methods and tools for the assessment of attributes and cover not only the epidemiological aspects of the evaluation but also the social and economic aspects. Indeed, several authors have developed frameworks and guidelines for the evaluation of SS, each providing different levels of detail for implementation and usually targeting only partial aspects of the SS characteristics.

The review highlighted the need for standardisation of the terminology used in these evaluations and a requirement for the definition of gold standards, efficiency and effectiveness measures. The absence of these compromises the ability to compare different systems and therefore prevents the identification of systems that are most efficient. Indeed, several evaluation approaches are available and most of them have been developed and used on an *ad hoc* basis. The criteria considered by each approach are usually organised into a template structure, which controls the logical flow of the evaluation process. Various terms have been used to describe these processes (e.g. Guidelines, Method, Framework, Tool). However except for the ones classified as “tools” providing practical toolkits to be used in the evaluation including an assessment of attributes, no clear distinction could be made between the evaluation process templates described using different terms. In animal health there is no index of an effectiveness measure available as it exists in the human health system evaluation process (e.g. Disability-adjusted life years (DALYs) and Quality-Adjusted life years (QALYs)), (Dehove et al., 2012). Specific evaluation of surveillance (as opposed to the evaluation of disease interventions) has been performed only on limited occasions and a variety of approaches and methods are used without a generally agreed protocol (Drewe, 2011). Indeed more than 25 attributes have been described for the evaluation of animal health surveillance systems, making a complete evaluation – if all attributes are used – time-consuming and expensive. In some cases no methods have been described for the measurement of these attributes and only a fraction of these evaluation attributes have been included in the evaluation process templates and in the practical case studies (Drewe, 2011; Hoinville, 2011; Hoinville, 2013, RISKSUR D1.2). Economic evaluation activities currently focus mainly on disease control programmes and economic impact of diseases in populations (Rushton, 2009). The D1.2 review highlighted the relatively small number of papers published on the subject of economic evaluation of surveillance (RISKSUR D1.2).

The main recommendations to be developed from our review are i) the need for a standardized process template , allowing for structured evaluation of surveillance systems while still maintaining flexibility in the selection of evaluation attributes to allow adaptation to any particular epidemiological system context and providing guidance in relation to suitable methods for attribute assessment; ii) the need to design a glossary of evaluation terms (to complement the existing “surveillance glossary” (Hoinville, 2013) and iii) to develop a set of internationally recognised and standardised effectiveness metrics for economic evaluation of animal health surveillance.

1.3 The need for an effectiveness measure and the benefit issue

When considering the start, end or change of a surveillance programme, policy makers need to know if and how much surveillance is worth it. Animal disease creates two sources of lost well-being for people. First, the monetary value of losses represented caused by the negative effects of disease itself; second, the additional resource costs incurred in the attempt to offset those output value losses, when those resources could have been used to generate other outputs people value elsewhere in the economy. In assessing the rationality of any resource-usage decision, the key criterion is whether the value of recovered outputs is sufficient to at least cover the additional resource costs (Mclnerney, 1992).

Surveillance provides information for decisions regarding the choice of interventions. Intervention is the process of applying specific measures targeted at reducing the risk of disease or infection, i.e. by either reducing the probability of disease/infection and/or reducing the consequences. Jointly surveillance and intervention achieve loss avoidance through the process of reducing the impact of disease by avoiding, containing, reducing or removing it – the outcome which decision-makers are ultimately interested in (Häsler *et al.* 2011).

In this three variable relationship (surveillance – intervention - loss avoidance) surveillance and intervention can be economic complements or substitutes. If surveillance and intervention resources are complements, it means they are used in a given ratio and must be treated as one input. This is the case, for example, in testing and culling strategies. If they are economic substitutes, using more of one input will require the use of less resource for the other, for example when surveillance is used to detect disease early and thus saves intervention resources. **In practice, these considerations mean that surveillance cannot be properly evaluated without simultaneously considering intervention** (Howe et al.*,* 2013).

The losses avoided (e.g. reduced losses due to mortality, abortion, prolonged calving interval, premature culling, drop in milk yield, meat or wool, human disease in the case of zoonosis, and market access) can be calculated by comparing a situation with the mitigation policy under discussion (a combination of surveillance and intervention) with the situation in the absence of this policy (the counterfactual). Both the counterfactual and the situation with the mitigation policy under discussion can be highly dynamic depending on the disease in question. Therefore, epidemiological simulation models are often necessary to simulate proxies of loss avoidance over time, such as disease prevalence or incidence. These disease frequency reductions can then be translated into economic values by, for example, multiplying the number of animals that are not dying due to disease by their market value. **These monetary benefits could then be compared to the costs of surveillance and intervention in a cost-benefit analysis to assess whether an investment is worthwhile**.

In some instances, decision makers may not require explicit quantification of the (monetary) benefit, but may be interested in the costs of the surveillance options related to selected performance indicators, which can act as proxies for a benefit. In an economic analysis, these performance indicators (i.e. effectiveness measures) would be compared to the costs of surveillance in a cost-effectiveness analysis (CEA). Cost-effectiveness analysis, commonly used to assess human health interventions, has rarely been applied to animal health decision-making problems (Babo Martins and Rushton, in press). In human health economics, the effectiveness often refers to the avoidance of illness or death, but the outcome of any objective can – in theory - be measured in various technical terms, for example reduction of abortions or detection of cases of disease. However, it is important that the value of the effect in question reflects a (non-monetary or monetary) benefit.

Unlike in health economics, where attempts have been made to harmonise CEA methodologies and encourage comparability of studies[1] (Murray, 2000), **there are no specific guidelines available yet for its application in animal health. The key is to select effectiveness measures that are meaningful; otherwise they will not inform the allocation of scarce resources.** Ideally, they can be compared against some pre-defined standards or values that have been established in studies in the past. For example, if previous studies established that each day of earlier detection of a highly pathogenic avian influenza (HPAI) outbreak resulted in the avoidance of losses worth £100,000, a cost-effectiveness ratio of a surveillance system to early detect HPAI expressed as costs/days of earlier detection can be easily interpreted. However, without this information, effectiveness measures like time of introduction of disease until detection or the probability of detecting an outbreak are not informative in a CEA.

Further, surveillance generates other benefits that cannot be easily measured, but nonetheless have a value. Examples include consumer confidence, reputation, feelings of safety, contentment, peace of mind – all these have perceived values that are generally not converted to monetary values by the price system of the market. Therefore, indirect methods of valuation such as willingness-to-pay or stated choice preference approaches need to be adopted (Antoine-Moussiaux, 2012).

1.4 The decision-makers point of view

Current world trade patterns and the globalization in general terms favour the rapid movement of people, animals and material resulting in increased risks of introduction of new diseases on a given territory. Animal health surveillance is an essential tool to detect disease or infection, to monitor disease trends, to facilitate the control of disease or infection, to support claims for freedom from disease or infection, to provide data for use in risk assessment, for animal and/or public health purposes, and to substantiate the rationale for sanitary measures. Today, animal health surveillance programs play a key role in the EU animal health policy. They are fully integrated into most livestock industries, food production systems and rural economies.

These elements all lead to the necessity of a public intervention that has been initiated for several decades at multiple levels (international-OIE, regional-EU and national levels-MS in the EU, other countries e.g. Australia, etc.). The type of surveillance applied depends on the desired outputs needed to support policy decision- making. In the EU, the AH surveillance policy is defined as part of the global AH Law which has recently been reviewed.

Recent EU studies have highlighted the need for strengthening the animal diseases EU surveillance programs in order to overcome several key issues e.g. necessity to ensure a comprehensive overview of surveillance in the EU and the need to clarify the purposes/objectives of the surveillance systems.

[1] In health economics, Disability-adjusted life years (DALYs) and Quality-Adjusted life years (QALYs) are well-established, validated, comparable and standardized metrics that reflect a non-monetary benefit. One criticism is that they do not capture the value (different) societies place on lives and how this then shapes our decisions on health.

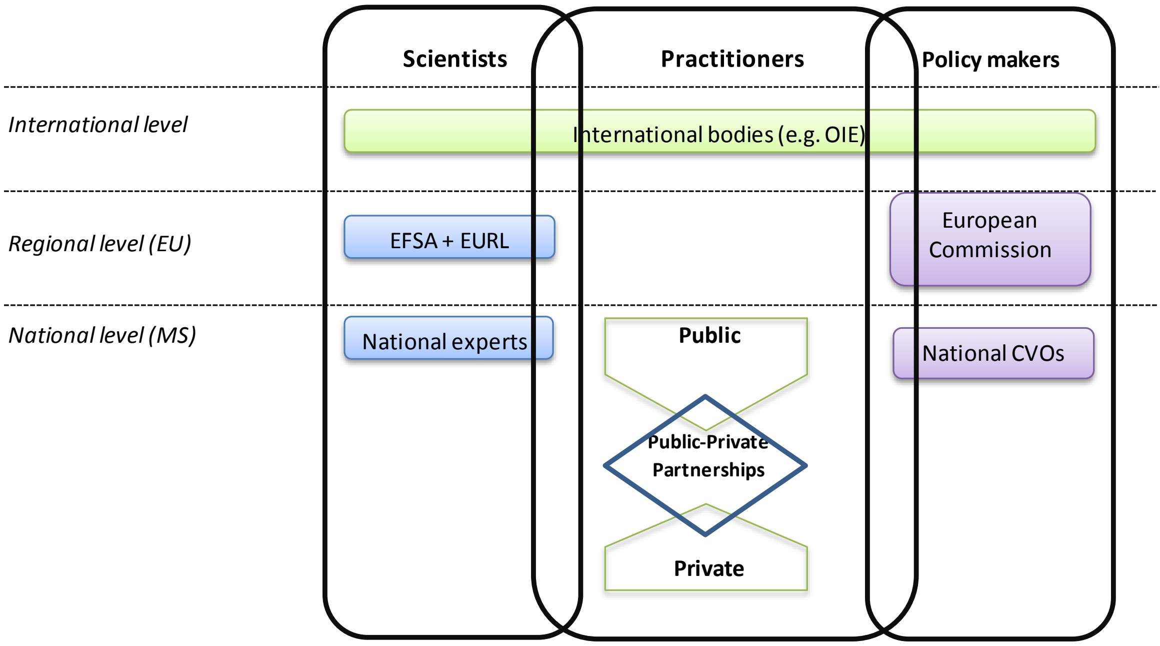
Surveillance activities are also facing new challenges such as climate change and further globalization, and sustainability of the surveillance programs (European Commission, 2007).

Priority setting, affordability, sustainability, social acceptance and communication remain issues that policy-makers have to consider when drafting and implementing their policy (Box 1: Example of DM needs in France). In the context of the new EU Animal Health policy, in the spirit of the "prevention is better than motto", the EU policy-maker (i.e. the European Commission) wishes to promote the early identification of problems before they emerge while being ready to manage outbreaks and crises is a major objective (European Commission, 2007). This approach leads on to reinforcing of biosecurity measures in all areas in which animals are found (farms, markets, border posts, transport vehicles, etc.).

Regulating animal surveillance as part of the animal health policy at EU level allows for coordinated and in the long run (hopefully) cheaper action on EU priorities, making it more effective and less expensive than actions by individual MS. For example, border controls for differing national lists of animal diseases would be highly inefficient and ineffective, given the free movement of commodities on the single EU internal market after import. Therefore diseases of EU significance need to be addressed jointly. An action in one MS may however result in dissemination to others. Third country trade partners might also implement restrictions on imports from the EU as a whole if an outbreak in one of the MS is not properly controlled. The membership of the EU (not only of the individual MS) in International organisations (e.g. WTO) further implies the EU's responsibility to maintain an adequate legal framework compatible with international animal health standards. **The specific added value of the EU's co-financing of mandatory EU surveillance programs is that it provides incentives to MS to put in place surveillance actions which are in the long-term interest of the EU as a whole. Large-scale actions by MS could be difficult without EU support in view of the large costs incurred by the individual MS for the benefit of the EU, even if the overall cost/benefit for the Union as a whole is positive.**

While agreeing with this new approach, Member State Competent Authorities are facing severe issues to finance these activities. **During the last decade, all national administrations have seen their budgets severely reduced resulting in the need to assess cost-effectiveness of their surveillance actions. These assessments are not that easy as the data available to quantify investments in surveillance and monitoring activities are very limited at both national and international levels.** Secondly, surveillance activities overlap with multiple other public interventions and private management of the supply chains, both in terms of personal and resources. **As a consequence it is very difficult to define and allocate partial cost to a surveillance budget and very few attempts have been made to achieve this. MS would value any tool that would theoretically help them to perform a technical and budgetary optimisation of their resources as a key instrument for the definition of their national policy for animal surveillance as part of the animal health policy (Box 1: Example of DM needs in France).**

Figure 1 – Mapping ad typology of stakeholders involved in animal health surveillance



Scientists: Those using or researching new techniques and tools for surveillance*
Practitionners: Those responsible for implementing surveillance programs and wishing to be updated on the latest opportunities and development*
Policy makers: Those responsible for setting national, regional legislative frameworks and international standards*
(*):source: www.animalhealthsurveillance.org. Consulted on 5/11/2013

Box 1. Limits of economic evaluation of animal health management in terms of decision making: example of results from a rapid appraisal in France (adapted from Chaudron M, 2010).

- Importance of political issues at stakes which are unlinked to scientific, technical or economic based evidence:
- Limited use of economic evaluation in decision making process, in part due to the lack of tools.
- Lack of adapted and comprehensive economic evaluation tool which will account for :
  - The Multiplicity of actors and decision making logics which result in a large diversity of behaviours,
  - The Complexity of the links between the actors,
  - The lack of base line situation knowledge,
  - The specific impact of one policy versus the others
  - The differences between programs implemented under the same policy
  - The multiple objectives for one action

These data link back to the following questions which will be addressed within the RISKSUR project:

- What is the logic behind decision making in animal health management?
- How to build up a practical and comprehensive tool to address decision makers'needs?
- How to account for multiple objectives situation?

1.5 Concluding remarks on the context for development of an evaluation support tool

Setting up AH surveillance programmes is becoming increasingly challenging as globalisation, changing livestock trade patterns and climate change impact on animal disease risks and create increasingly complex patterns of disease transmission and spread. In addition, public authorities and decision makers, who are responsible for planning and implementing surveillance programmes, are facing budget cuts which make it increasingly difficult to maintain and promote efficient and effective animal health surveillance. Consequently, they are in need of evidence about how resources can be used in the best possible way to establish fit-for-purpose and effective surveillance systems. Evaluation of surveillance systems is necessary in order to achieve this goal, and ideally this should be done by providing a simple and practical guide for decision makers.

Economic evaluation of health management programs implies to compare different policies to identify a solution that maximizes public welfare or at least generates a net benefit for society. Indeed technical achievements are only a part of decision-making in the policy-formulation process. Economic evaluation of animal health surveillance is technically challenging because of its complex link to intervention, the technical capacity and data requirements needed to assess epidemiological and economic performance, and the lack of standardised metrics and practical tools for economic appraisals. These technical challenges stand in stark contrast with decision-makers' demand for simple and practical guidance for performing systematic evaluation of existing or planned animal health surveillance systems which in turn should result in development of more effective policies.

The evaluation thus must provide a robust scientific foundation for the decision making process while still presenting the findings in a user-friendly and applied format. There is a need for an evaluation tool which will provide guidance on how to select appropriate evaluation question (according to the context of surveillance); information on the methodology to apply and feasibility of the evaluation within a given timeframe and available resources.

2 RISKSUR Evaluation Support Tool (EVA)

Following the needs and gaps identified in the evaluation reviews (D1.2), the RISKSUR project aims at developing an integrated evaluation support (EVA) tool for the economic and epidemiologic evaluation of animal health surveillance systems. The EVA tool builds on existing evaluation framework and methods/tools and aims to provide standardization in the evaluation process to allow for international recognition but without undermining the need for flexibility to account for context and evaluation specificities. The tool should be practical and easy to use and provide step by step guidance on the evaluation process along with a comprehensive toolkit. At this stage of the RISKSUR project, the basic structure and a description of the logic of EVA Tool is available (conceptual model). The list of activities undertaken in the first year of the project to develop this conceptual model is presented in box 2. A road map has been outlined for the next steps to work on the details of the framework, populate it with the required information, and apply it to case studies (Deliverables 5.18; 5.19; 5.20). Both the generic structure of EVA and the roadmap are presented in the following Sections.

Box. 2. Activities undertaken within Year 1 of RISKSUR project to develop the EVA tool conceptual model (all the reports are available on RISKSUR skydrive, under WP1 and WP5 activities)

- Literature reviews (D 1.2)
- Monthly project meetings to discuss task plan, task allocation, theoretical and practical challenges linked to the economic evaluation of animal health surveillance
- Discussions with specialists who have conducted evaluations of surveillance and/or developed frameworks for the evaluation of surveillance
- Internal reports from group discussion and task development
- Workshop on evaluation challenges organized in Montpellier, June 28; to validate the model and initiate case study applications

2.1 Objectives and Principles

The objective of the EVA tool is to provide a comprehensive guidance to decision makers and their technical advisers to plan and conduct evaluations of animal health surveillance systems. The EVA tool will provide DMs with simple step by step guidance and options on what, why and how to measure in order to evaluate animal health surveillance systems (Figure 2).

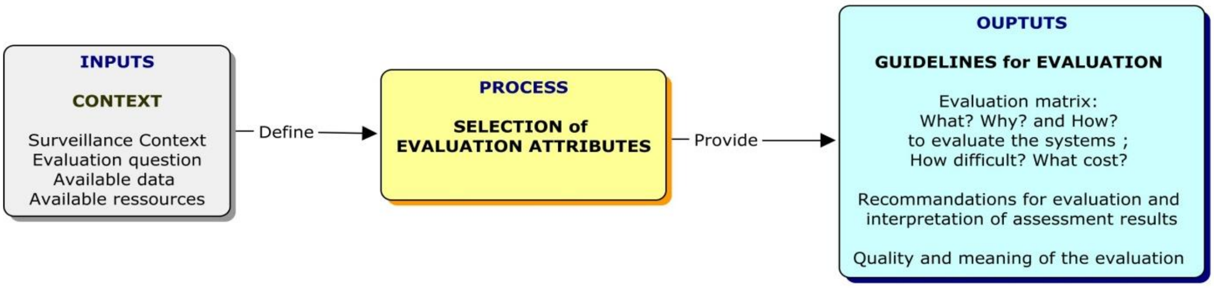


Figure 2. EVA Tool general principle

The components of the tool are to:

1. Select the evaluation question
2. Describe the context including epidemiological situation, surveillance objectives.
3. Select evaluation attributes[2] according to evaluation question and context. (cf. Annex 1 list of evaluation attributes)
4. Selection of methods appropriate for the assessment of the selected attributes
5. Guidance on application of methods
6. Guidance on interpretation of results

In order to allow sufficient flexibility in the evaluation, the selection process will not remove any of the attributes but rather provide an attribute classification based on the following criteria: relevance to address the evaluation question; easiness to be measured; quality of the measure; meaning of the measurement.

The classification process will be based on a combination of scientific evidence and expert opinion to generate information on[3]:

- ranking of performance attributes according to a set of surveillance context and evaluation combinations (cf. prioritisation process, RISKSUR D1.3)
- Structural and functional attributes linked to the performance attributes (cf. linking process, RISKSUR D1.3)
- ranking of methods and tools to measure the attributes (cf. measurement methods and tools, RISKSUR D1.3)

2 Definition of evaluation attributes. Deliverable 1.3

3 These approaches will be describe in details in Deliverable 1.3 Current surveillance characteristics and evaluation criteria

2.2 Development challenges

Expert discussions during the first year of the RISKSUR project have highlighted the need to set clear limits on the scope of the tool and additional challenges to be considered in its development:

Assessment versus evaluation: the objective of an assessment is to provide technical results (either qualitative and/or quantitative) whereas an evaluation implies a judgment on the total value – both technical and economic - surveillance system and recommendations for improvement. A full evaluation would therefore need to consider the broader national context and other health programs. The aim of the tool is to guide the assessment and evaluation of surveillance by end-users who will generate scientific evidence for DMs about SS performance. Further, the tool provides guidance about the interpretation of outputs which allows decision-makers to draw conclusions based on the findings obtained. *Generic tool, comparative analysis and gold standard issues:* the tool needs to be flexible enough to be adapted to each context (socio-economic, epidemiological) and to allow the choice of different evaluation questions (and therefore level/extent of the evaluation. The objective of the tool will be to identify the best effectiveness measure for a specific evaluation context but not to set up at this stage a gold standard for these effectiveness measures. However the issues and challenges linked to the definition of gold standards for animal health surveillance systems will be discussed within the scope of this project.

Development of a new framework: The objective of this work is to promote and facilitate evaluation of the surveillance systems within the EU countries. This project develops a new integrated framework based on the existing ones but adapted to decision maker needs. From the early stages of development the type and needs of decision makers have been taken into account to ensure transferability and sustainability of the tool.

2.3 Outcomes

This first phase of the project has led to the definition of 1) the specific aims of the work to develop the evaluation framework within RISKSUR project including definition of target groups; and 2) the Objectives and Roadmap for the evaluation tool development.

OUTCOME 1: Aims of the development of an Evaluation Framework within RISKSUR project

1. To provide guidance on how to perform an evaluation of the added value of new surveillance design (e.g. considering the interest and benefit of risk-based, taking into account economic/socio-economic and environmental issues) for surveillance systems designed to achieve different purposes (e.g. early detection, freedom from disease, prevalence estimation or detection of cases to facilitate control).

2. To develop a stand-alone integrated economic and epidemiologic evaluation tool for Member States to apply when conducting evaluations of surveillance. The end users of the tool will be the people performing the evaluations, which may be contracted researchers, technical advisors in government, or in some instances even decision-makers themselves. However the tool will not provide standards for evaluation which need to be set up by EU or member states themselves according to their own context and expectations.

3. To apply the tool as a pilot trial to evaluate the efficiency of different surveillance designs (e.g. risk-based versus conventional sampling surveillance) according to the different contexts of the EU countries (based on surveillance system mapping/typology)

OUTCOME 2. Scope and Road map for the EVA Tool development

The basic structure of the EVA tool (conceptual model):

- Builds on existing evaluation frameworks to develop comprehensive guidelines for practical evaluation of animal health systems
- Is flexible to allow for many different context and evaluation questions
- Is simple and practical
- Targets end users' needs

In the next steps, this basic structure will be developed in more detail, populated with data and information, and applied to a selection of case studies while aiming to maintain these key features.

The EVA tool presents a "basic recipe" approach to the evaluation of surveillance by guiding the user through various elements that are relevant to the evaluation including selection of the evaluation question, description of the context, selection of attributes of interest and relevant for the context described, selection of appropriate methods to perform the assessment,

Explicitly excluded functionalities/features:

- Perform any analysis of surveillance data to provide assessment or evaluation results, i.e. there will not be any data entry features with related automatic calculations of for example costs or benefits of surveillance. The reason for this is that the large heterogeneity of surveillance, contexts and evaluation questions require rather specific data collection for a wide range of combinations which is not within the scope of this project. This process will be supported by the development of approaches within the RISKSUR project to perform epidemiological (WP2-4) and economic (WP5) evaluation of surveillance systems and/or components and these will be applied to selected case studies.

To address the challenges presented above it was agreed that the EVA tool will be applied and further developed in different phases, starting with simple evaluation questions (e.g. least-cost analysis, cost-effectiveness comparative analysis) and build up onto more complex issues (e.g. cost-benefit) within the time frame of the project.

3 EVA Tool Conceptual Model

The Eva tool conceptual model describes the links between the three parts: INPUTS, PROCESS and OUTPUTS (Figure 3). It guides the user to define and enter the **input parameters** (e.g. surveillance context, evaluation question, data available) (cf. INPUTS section), the tool will process the INPUTS to display relevant attributes and facilitate the selection of attributes for the evaluation (cf. EVA PROCESS section) and provide a **comprehensive evaluation methodology including selection of methods for their assessment adapted to the context of the evaluation** (cf. OUTPUTS section) (Figure 3). The latter will include a description of the advantages/limits, data requirements and resource requirements for the assessment along with recommendations on how to perform the epidemiologic and economic assessment and guidelines for interpretation of results.

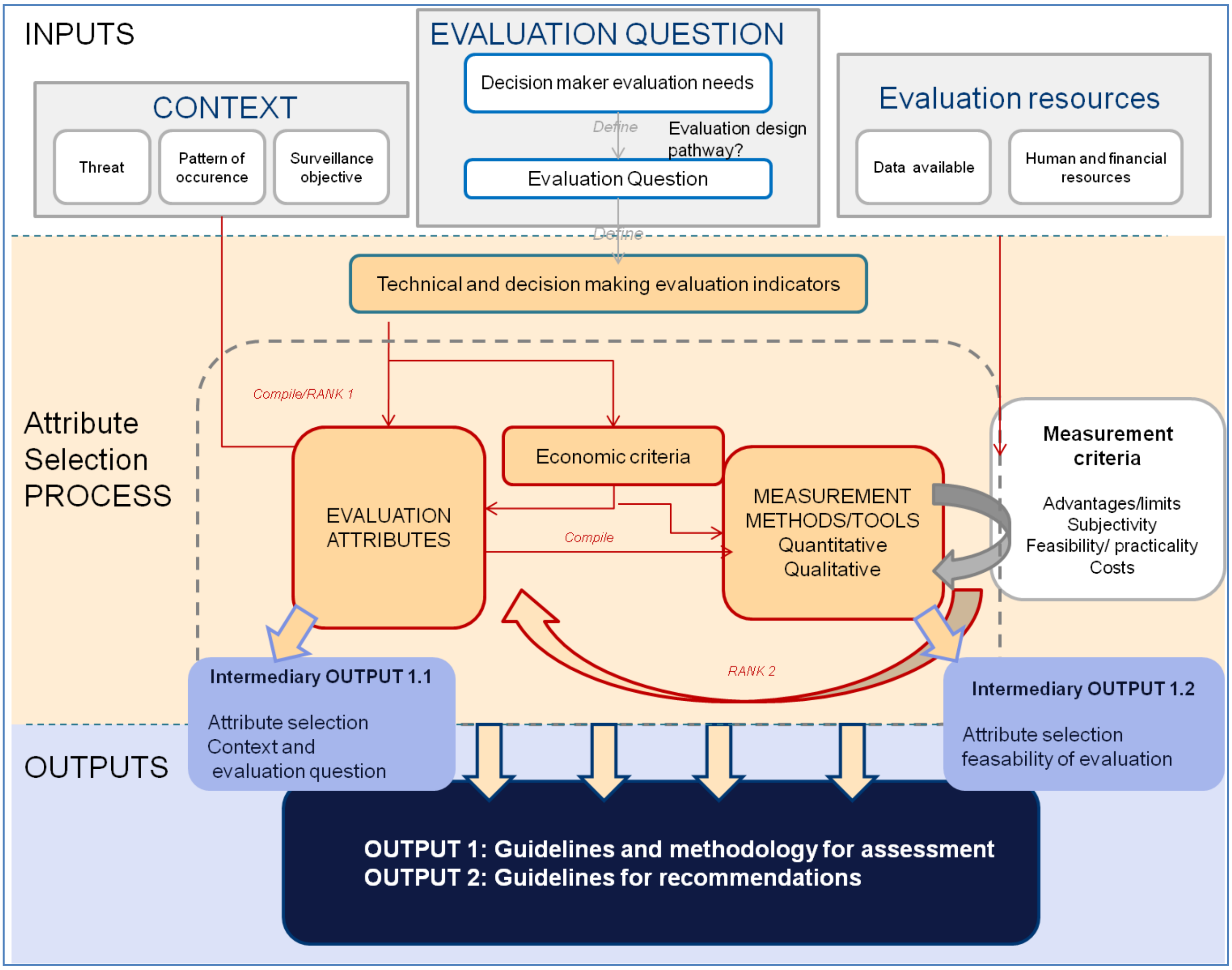


Figure 3. Process and functionality model for the EVA support tool

3.1 EVA Tool Development

The conceptual model was developed by the consortium experts within the first year of the project (cf. Box 3). Some of the data used to parameterise the tool were also defined and validated.

Part 1: INPUTS. The list of the type of inputs relevant for the evaluation was validated during the meeting through the application of three working case studies covering three different surveillance purposes (early detection; freedom from disease; case detection to facilitate disease control) (cf. List of validated inputs, Annex 2). Two different approaches were suggested regarding the definition of the EVALUATION QUESTION: 1) to develop a pre-defined list of evaluation questions to choose according to DMs' needs and 2) to develop an evaluation question design pathway to allow for more guidance in the definition of the evaluation question by the user (cf. Evaluation question Annex 3).

Part 2: PROCESS: The internal process of the tool will consist of scenario tree algorithms and/or a matrix which will allow for selection (and/or ranking) of: 1) evaluation indicators and 2) a set of evaluation attributes relevant to address a particular evaluation question under a specific context.

Evaluation indicators: A matrix will define relevant evaluation indicators (e.g. effectiveness measures and/or economic criteria) to be measured according to each combination of evaluation question and context parameters. A template of the matrix has been developed and will be further completed based on available scientific evidence and expert opinion (cf. Evaluation matrix template 1). The list of indicators to be considered to address RISKSUR evaluation questions on cost-effectiveness and other evaluation questions will be developed and validated in the second year of the project through the case study applications.

Attribute selection: This selection will be based on both scientific evidence and expert opinion. The attributes will be selected according to 1) their importance to the specific context and evaluation question and 2) how easy the attributes can be measured, their quality and meaning of this measurement. The list of attributes and economic criteria relevant for the evaluation process has been validated[4]. The different approaches to be developed for the selection and prioritisation of the attributes have been selected and described (e.g. expert opinion, linking of attributes by network analysis). These methods are described in details in a Deliverable 1.3 report. A review of the advantages and limits of the methods and tools available to measure the attributes has been performed and will be further completed and validated trough case study application to provide information on the feasibility, quality and meaning of the evaluation (RISKSUR D1.3).

Part 3: OUTPUTS. The tool will provide two main outputs and two intermediary outputs. ***Intermediary output 1.1*** a ranking/prioritisation of the evaluation attributes according to the context and evaluation question (Table 1: Evaluation matrix 1); ***Intermediary output 1.2*** the feasibility of the evaluation (Table 2: Evaluation matrix 2). The main output (**OUTPUT 1)** will consist in a combination of the two intermediary outputs to provide a comprehensive assessment approach (What to assess

[4] adapted from Hoinville et al. previous work (ref)

and Why?) along with its feasibility, advantages, limits and cost (How to do the assessment?). The tool will also provide information on the quality of the assessment to be performed according to the data available and a list of data to be collected to improve the assessment result (How good is the assessment?) (Table 3).

The second output **(OUTPUT 2)** will provide guidance on the meaning of the assessment outputs and which type and extent of recommendations could be drawn to address the evaluation question.

The tool will be developed in 4 phases (Table 4): phases 1 and 3 will provide a tool for cost-effectiveness evaluation addressing directly the objectives of RISKSUR (this tool will be used to evaluate RISKSUR case studies); phases 2 and 4 will provide additional elements to further develop the tool to address more complex evaluation questions and the cost-benefit evaluation.

Table 4. EVA Tool research and development phases

Phase	*Tool*	*Objectives*	*Practical application*	*Time frame*
1	Pilot assessment tool 1	To cover the following evaluation questions addressing RIKSUR objectives: - Least cost-analysis - Comparing 2 surveillance modalities	Conceptual model	Jan- Oct 2013
			Process for attribute selection	Jan-Mar2014
			Validation of the tool using specific case studies	Jan-Mar2014
			Software development (see section 4.2)	Jan-Mar 2014
2	Pilot assessment tool 2	Add in more complex and additional evaluation questions such as cost-effectiveness, cost-BENEFIT evaluation	Process for attribute selection, meaningful effectiveness measure	Mar-June 14
			Validation of the tool using specific case studies	June-Dec 14
			Inclusion into EVA tool software development	Jan-Mar 15
3	Pilot evaluation concept tool	conceptual model to provide guidance for data interpretation and evaluation recommendations	Development of conceptual model using case studies	June-Dec14
4	Pilot evaluation tool	development of tool algorithm to provide guidance for data interpretation and evaluation recommendations	Validation of the tool using case studies	Jan-March 15

Table 1. TEMPLATE FOR THE EVALUATION MATRIX 1 (intermediary OUPTUT 1.1)

EVALUATION QUESTIONS	Evaluation Indicators	SURVEILLANCE CONTEXT AND SURVEILLANCE PURPOSE								
		Early detection			*Freedom from disease*			*Endemic disease*		
		Early detection to implement rapid control	*Other purpose?*	*Other purpose?*	To prove disease eradication	To confirm freedom from exotic disease	*Other objective?*	To detect cases for disease control (keeping under certain prevalence threshold)	To measure disease prevalence o evaluate control measure efficacy	To measure disease prevalence to rank diseases
Is my surveillance system effective? *(Question 4. i.e. does it meet its objectives?)*	*Effectiveness measure*	*Meaningful effectiveness measure*								
	Economic criteria	*None*								
What would be the cost-effectiveness of different surveillance component? *(Questions 4 and 9)*	*Effectiveness measure*	*Meaningful effectiveness measure* ….								
	Economic criteria	*Cost effectiveness*								
Which surveillance modalities is the most efficient? *(i.e. Question 7 or 8. Identify the surveillance modality that achieve the objectives at least cost or generate most benefits)*	*Effectiveness measure*	*Meaningful effectiveness measure*								
	Economic criteria	*Cost Benefit*								

Table 2. TEMPLATE FOR THE EVALUATION MATRIX 2 (intermediary OUTPUT 1.2)

Evaluation Indicators	Evaluation Attributes	Measurement Methods and tool						
		Name	Description	Applicability/ Practical Tools	Data requirement	Outputs	Advantages	Limits
Effectiveness measure	Sensitivity (Se)	Capture/recapture	Statistical analysis of minimum 2 independent list of surveillance data	Models available Open software ("R")	2 independent sources of surveillance data or 3 sources	Se Precision (Real Pe) Infection rate	Straightforward to compare 2 sampling strategies	Only applicable for endemic diseases or need for simulated data sets.
		Scenario-tree modeling	…	…	…	…	…	…
		Bayesian modeling	…	…	…	…	…	…
Economic criteria	Cost	Cost analysis (OASIS module)	…	…	…	…	…	…
		Cost analysis (CDC module)	…	…	…	…	…	…

Table 3. TEMPLATE for EVA Tool OUPUT 1: What, Why and How to perform an assessment of surveillance systems to address an evaluation question X under a context Y.

WHAT TO ASSESS ?			WHY ?	HOW ?					
Evaluation Indicators	**List of attributes/criteria**	**Assessment Characteristics**	**Importance/ context**	**Method**	**Easiness criteria**	**Data required**	**Feasibility**	**Timing and Cost**	**Quality of the assessment**
Effectiveness measure	*Attribute 1* *Attribute 2* *Attribute 3* …	*Direct* *Indirect* *Proxy for attribute 1*	*Critical* *High* *Medium* *Low* *Not relevant*	*Method A* *Method B* *Method C* ….	Easy Difficult	Description of the data	YES/NO	Estimated time and cost	*To be defined*
	…..	…	…		…				
Economic criteria	*Criteria 1* *Criteria 2* *Criteria 3* ….	*Direct* *Indirect* *Proxy for X*	*Critical* *High* *Medium* *Low* *Not relevant*	*Method a* *Method b* *Method c* ….	Easy Difficult	Description of the data	YES/NO	Estimated time and cost	*To be defined*

Assessment characteristics: this section defines if the evaluation attribute allows for a direct measure of the effectiveness indicator; has an indirect effect on the effectiveness indicator or is linked other attributes which could be a measurement proxy.

Importance/context: this section provides the outputs of the prioritisation/ranking of attributes according to the context and evaluation question. The information could be qualitative (from "critical" to "not relevant") or quantitative (attributes ranked from 1 to X).

Method: this section provides information on the best method available to assess the attributes according to its simplicity of implementation (column "Easiness criteria") and the availability of the data ("Data required"). If the data are not available the method will be listed at the lowest and information on the data required to implement it provided in the column "Data required"

Simplicity criteria: this section provides information on the degree of difficulty linked to the application of the measurement method. This degree of difficulty will be based on expert opinion and will take into consideration the level of expertise required and the practical tools available to implement the method.

Feasibility/Timing and Cost: these two sections provide information on the feasibility, timing/cost and on the quality of the evaluation proposed according to the type of data and resources available.

Quality of the assessment: This section will provide information on the level of quality of the evaluation if only one or a combination of attributes is being measured. This information will be based on the robustness of the method used for the assessment and could be either quantitative or qualitative according to the method used to assess the attributes.

4 Next steps

4.1 Case study application

The case studies used for the development of the conceptual model of the EVA tool have been selected to cover the three surveillance purposes considered in the RISKSUR project workpackages 2, 3 and 4; along with additional case studies from challenging environments (limited number and quality of the data) to test the model under different situations.

- WP2 (early detection): early detection of HPAI in poultry in UK;
- WP3 (freedom from disease): freedom from CSF in pigs in Germany and
- WP4 (endemic diseases): measuring salmonella prevalence in pigs in Sweden; measuring salmonella prevalence in pigs in Vietnam; measuring HPAI prevalence in poultry in Vietnam .

Additional case studies will be also considered for further development of the tool such as early detection of vector born diseases; early detection of ASF and CSF in Corsica.

Matrix 1 and 2 will be further developed by working those case studies but expending to other evaluation questions and by also working on additional case studies covering additional surveillance specific objectives.

Each case study exercise will provide:

1 The meaningful effectiveness measure. To test the new approach of development of a unique measure of effectiveness (which integrates all the evaluation attributes of the surveillance performances) (cf. RISKSUR D1.3 report).
2 The list of attributes relevant to each context/EVA question combination under the case study situation (attribute prioritisation)
3 The links between attributes under the case study situation (attribute linking)
4 The advantages and limits of the measurement methods and tools to be used under the case study situation
5 The feasibility and quality of the evaluation under the case study situation
6 Guidance on result interpretation and evaluation recommendations

A sensitivity analysis of the outputs of the case studies will be performed to assess which parameters have most influence on the evaluation framework and the need for further validation of the matrix through expert opinion elicitation for example.

The tool will be submitted to the project scientific and advisory boards to be tested and officially validated by the end users. The end users will also provide input on the best method to define the evaluation question, whether to guide them through this process or to simply provide a list of evaluation question options for them to select.

4.2 EVA tool software development

A user friendly computed version of the EVA tool will be developed from year two of the project. This will be conducted in close collaboration with a “user group” (advisory group or technical staff nominated by the advisory group) as well as with the programming partners in RISKSUR (Trace Tracker).

The different steps for the EVA tool software development will be considered as followed:

1. Planning and feasibility: meetings will be organised between the “user group”, WP5 and WP6 partners to assess the needs, the requirement of the tool and the software development feasibility. Scenario pathway will be defined for each step of the EVA tool conceptual model (represented by the arrows in Figure 2) to build up a conceptual algorithm for development of the software. This part will allow to select the most adapted development option and to plan the next steps.
2. Product design: this step includes overall design of the product, design of database and design of data structure. And will be performed by WP6 in close collaboration with WP5.
3. Coding
4. Testing: the pilot software will be tested by WP5 partners and in a second step by the “user group”
5. Installation and maintenance will be taken over by TraceTracker.

4.3 Provisional calendar

Activities YEAR 2	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24
Tool development: (D5.18)												
Evaluation Questions												
Attribute priorisation/Linking												
Quality of evaluation												
Interpretation and evaluation												
Deliverable 5.18 Report												
Case study application												
EVA tool application using case studies												
Deliverable 5.20 (Case studies selected and described) Report												
Protocols for evaluation data collection (D 5.19)												
Software development												
Algorithm development												
Planning meeting												
WP6 coding												
EVA tool workshop												

5 General Conclusion

Year one of the project has allowed to develop a conceptual model for a decision support tool to guide the decision makers and their technical adviser in the process of economic evaluation of their animal health surveillance systems (EVA tool). Extensive expert discussion and meetings have allowed listing and addressing all the challenges linked to the evaluation activities to be performed within the project and to the development of the EVA tool. This work is the result of a close collaboration between WP1 (Evaluation theme transversal activity) and WP5 (Evaluation). Evaluation activities within the second year of the project will be centred on further development of the tool through practical implementation of real case studies.

6 References

Antoine-Moussiaux N, Binot A, Chaudron M, Goutard F, Roger F, Peyre M (2011). Socio-economic evaluation of animal health surveillance: principles and prospects. CIRAD-AFD Report, 41pp. Online access : http://revasia.cirad.fr/publications/rapports

Antoine-Moussiaux N, Delabouglise A, Vu DT, Nguyen VK, Goutard F, Roger F, Peyre M (2012) Social impact of HPAI surveillance network : Participatory evaluation at the community level in Vietnam 13th ISVEE conference, Maastricht, Netherlands, August 20-24

Babo Martins, S., Rushton, J. (in press), Cost-effectiveness analysis – adding value to animal health, welfare and production assessment, Revue scientifique et technique (World Animal Health Organisation).

Chaudron M (2010). *French only*. Rentabilité de la surveillance épidémiologique des maladies animales dans les pays en développement : Contribution de l'évaluation économique et perspectives de recherche (Revue bibliographique et entretiens d'acteurs). Rapport de stage Master CEAV-SPV. AFD-CIRAD

Dehove, A., Commault, J., Petitclerc, M., Teissier, M., & Mace, J. 2012: Economic analysis and costing of animal health: a literature review of methods and importance, *Rev.Sci.Tech.* 31, 605-604.

Drewe, J. A., L. J. Hoinville, et al. (2012). Evaluation of animal and public health surveillance systems: a systematic review. Epidemiology and Infection 140(4): 575-590.

European Commission, 2007. A new Animal Health Strategy for the European Union (2007-2013) where "Prevention is better than cure". Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the regions, COM 539 (2007) final

Granger, L. M. (2011). Global surveillance: suggestions for a strategic approach. FAO animal production and health proceedings, ISSN 1810-0732 ; 14.

Häsler B., Howe K. S., and Stärk K. D. C. (2011). "Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for economic analysis.," BMC health services research, vol. 11 (1), p. 225.

Hoinville L.J., Alban L., Drewe J.A., Gibbens J.C., Gustafson L., Häsler B., Saegerman C., Salman M., Stärk K.D.C. (2013) Proposed terms and concepts for describing and evaluating animal-health surveillance systems, Preventive Veterinary Medicine, Volume 112, Issues 1–2, Pages 1-12, http://dx.doi.org/10.1016/j.prevetmed.2013.06.006.

Hoinville L, (2011) Animal Health Surveillance Terminology Final Report from Pre-ICAHS Workshop December 2011 (version 1.0)

Howe, K.S., Häsler, B., Stärk, K.D.C. (2013), *Economic principles for resource allocation decisions at national level to mitigate the effects of disease in farm animal populations,* Epidemiology and Infection 141:91–101. doi: 10.1017/S095026881200060X.

Jann W, Wegrich K. (2007). 4 Theories of the Policy Cycle. Handbook of public policy analysis, untag-smd.ac.id

Jebara KB. (2004). **Surveillance, detection and response: managing emerging diseases at national and international levels.** Rev Sci Tech. 23(2):709-15.

Mclnerney, J.P., Howe, K.S., Schepers, J.A. (1992), *A framework for the economic analysis of disease in farm livestock*, Preventive Veterinary Medicine 13:137–154.

Murray, C.J., Evans, D.B., Acharya A., Baltussen, R.M.P.M. (2000), *Development of WHO guidelines on generalized cost-effectiveness analysis*, Health Economics 9:235–51.

RISKSUR D1.2. (September 2013) Evaluation methods of surveillance systems and current practices.

RISKSUR D1.3. (December 2013) Current surveillance characteristics and evaluation criteria.

Rushton J., (2009) Economics of Animal Health and Production. CABI Publishing, 352 pp.

Shahab, S. (2009). Finding value in the evaluation of public health syndromic surveillance systems from a policy perspective. K. u. Search Canada, Alberta Health Services: 24.

Stärk KDC, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, Davies P. (2006). Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches. BMC Health Services Research, 6:20. Online access: http://www.biomedcentral.com/1472-6963/6/20

Taylor L H, Latham S M, and Woolhouse M E (2001). Risk factors for human disease emergence. Philos Trans R Soc Lond B Biol Sci. 356(1411): 983–989.

7 ANNEXES

7.1 Glossary

A glossary of the following terms will be completed and adapted to the context of health surveillance within the framework of RISKSUR project

Assessment	To determine, estimate or judge the value of. An assessment provides technical results (either qualitative and/or quantitative) which may or may not be linked to a judgment on the validity/quality of those results
Impact assessment	A process aimed at structuring and supporting the development of policies. It identifies and assesses the problem at stake and the objectives pursued. It identifies the main options for achieving the objective and analyses their likely impacts in the economic, environmental and social fields. It outlines advantages and disadvantages of each option and examines possible synergies and trade-offs
Effectiveness	is the capability of producing a desired result. When something is deemed effective, it means it has an intended or expected outcome,
Efficacy	is the capacity to produce an effect. In medicine, efficacy indicates the capacity for beneficial change (or therapeutic effect) of a given intervention. When talking in terms of efficacy vs. effectiveness, effectiveness relates to how well a treatment works in the practice of medicine, as opposed to efficacy, which measures how well treatment works in clinical trials or laboratory studies
Efficiency	The extent to which a resource is used for the intended purpose. Efficiency describes the extent to which time, effort or cost is well used for the intended task or purpose. It is often used with the specific purpose of relaying the capability of a specific application of effort to produce a specific outcome effectively with a minimum amount or quantity of waste, expense, or unnecessary effort.
Efficiency assessment of a program	Cost-benefit or cost-effectiveness analysis assesses the efficiency of a program. Evaluators outline the benefits and cost of the program for comparison.
Evaluation	is a systematic determination of a subject's merit, worth and significance, using criteria governed by a set of standards. It can assist an organization, program, project or any other intervention or initiative to assess any aim, realisable concept/proposal, or any alternative, to help in decision-making; or to ascertain the degree of achievement or value in regard to the aim and objectives and results of any such action that has been completed The primary purpose of evaluation, in addition to gaining insight into prior or existing initiatives, is to enable reflection and assist in the identification of future change. Evaluation implies a judgment on the total value – both technical and economic.

Program evaluation	a systematic method for collecting, analyzing, and using information to answer questions about projects, policies and programs, particularly about their effectiveness and efficiency. In both the public and private sectors, stakeholders often want to know whether the programs they are funding, implementing, voting for, receiving or objecting to are producing the intended effect. While *program evaluation* first focuses around this definition, important considerations often include how much the program costs per participant, how the program could be improved, whether the program is worthwhile, whether there are better alternatives, if there are *unintended* outcomes, and whether the program goals are appropriate and useful. Program evaluation may be conducted at several stages during a program's lifetime. Each of these stages raises different questions to be answered by the evaluator, and correspondingly different evaluation approaches are needed. Rossi, Lipsey and Freeman (2004) suggest the following kinds of assessment, which may be appropriate at these different stages: • Assessment of the need for the program • Assessment of program design and logic/theory • Assessment of how the program is being implemented (i.e., is it being implemented according to plan? Are the program's processes maximizing possible outcomes?) • Assessment of the program's outcome or impact (i.e., what it has actually achieved) • Assessment of the program's cost and efficiency
Impact evaluation (assessing effectiveness)	The impact evaluation determines the causal effects of the program. This involves trying to measure if the program has achieved its intended outcomes. Assesses the changes that can be attributed to a particular intervention, such as a project, program or policy, both the intended ones, as well as ideally the unintended ones. In contrast to outcome monitoring, which examines whether targets have been achieved, impact evaluation is structured to answer the question: how would outcomes such as participants' well-being have changed if the intervention had not been undertaken? This involves counterfactual analysis, that is, "a comparison between what actually happened and what would have happened in the absence of the intervention." Impact evaluations seek to answer cause-and-effect questions. In other words, they look for the changes in outcome that are directly attributable to a program
Policy analysis	Determining which of various alternative policies will most achieve a given set of goals in light of the relations between the policies and the goals". However, policy analysis can be divided into two major fields. Analysis **of** policy is analytical and descriptive—i.e., it attempts to explain policies and their development. Analysis **for** policy is prescriptive—i.e., it is involved with formulating policies and proposals (e.g., to improve social welfare). The area of interest and the purpose of analysis determines what type of analysis is conducted. A combination of policy analysis together with program evaluation would be defined as Policy studies

7.2 Annex 1: Provisional list of evaluation attributes

Evaluation attributes and economic criteria that could be used in the RISKSUR evaluation algorithm

Hoinville, L *et al.* 2013

Sections 1-4 of this document provide definitions of all the evaluation attributes identified from various sources that are thought to be most relevant to the evaluation of animal health surveillance systems. Section 5 lists the economic efficiency criteria that can be used to choose between different surveillance options. These lists of attributes and criteria were selected from the complete list provided in Appendix 1.

1. Organisational attributes

1.1. Attributes aimed at evaluating the management processes

Organisation and management	An assessment of organisational structures of the surveillance including whether the objectives are relevant and clearly defined and the existence of steering and technical committees whose members are representative of the surveillance stakeholders. The members of these committees should have appropriate expertise, clearly defined roles and responsibilities and should hold minuted meetings regularly to oversee the function of the system.
Training provision	Provision of adequate initial training and an ongoing program of training for those implementing the surveillance system, particularly those collecting the data
Performance indicators and evaluation	Whether performance indicators are routinely used to monitor system performance and whether periodic external evaluations are used to assess the system outputs in relation to its objectives
Resource availability	An assessment of the financial and human resources available for implementing the surveillance activity including the expertise and capability of personnel

1.2. Attributes aimed at evaluating the technical processes

Data collection	The use of appropriate data sources and collection methods including automation of data collection where appropriate and the existence of a case definition and data collection protocol including an appropriate sampling strategy
Sampling strategy	The use of appropriate sampling strategies including the use of risk-based approaches and pooled sampling where appropriate. This could include a risk-based requirement calculations or risk-based sampling. The basis of the risks used in the design of the risk-based sampling strategy should be assessed.
Data storage and management	Appropriate use and documentation of data management systems for processing information, including data processing protocols, and effective use of data verification procedures and data storage and back-up procedures
Internal communication	An assessment of the methods used and ease of information exchange between all those involved in providing, managing, analysing and disseminating information for the surveillance system . The methods used to provide feedback to data providers and to increase their awareness about hazards and surveillance activities should also be assessed.
External communication and dissemination	An assessment of the data and information provided to those outside the surveillance system including the timeliness and types of output produced. The efforts made to disseminate these outputs including the use of web-based systems should also be assessed.

Laboratory testing and analyses	Whether testing is carried out using appropriate methods, including an assessment of diagnostic test sensitivity and specificity, with quality assurance scheme and timely and accurate delivery of results.
Data analysis	Whether appropriate methods are used for the analysis and interpretation of data at an appropriate frequency
Quality assurance	Whether the laboratory or other surveillance processes are quality assured or accredited

2. Functional attributes

2.1. Attributes aimed at evaluating the system function

Stability and sustainability	The ability to function without failure (reliability), to be operational when needed (availability) and the robustness and ability of system to be ongoing in the long term (sustainability).
Acceptability and engagement	Willingness of persons and organisations to participate in the surveillance system, the degree to which each of these users is involved in the surveillance. Could include an assessment of stakeholder awareness of the system and their understanding of it. Could also assess their beliefs about the benefits or adverse consequences of their participation in the system including the provision of compensation for the consequence of disease detection.
Simplicity	Refers to the surveillance system structure, ease of operation and flow of data through the system.
Flexibility	The ability to adapt to changing information needs or operating conditions with little additional time, personnel or allocated funds. The extent to which the system can accommodate collection of information about new health-hazards or additional/alternative types of data; changes in case definitions or technology; and variations in funding sources or reporting methods should be assessed.
Portability	Evaluating the possible use of the system in other circumstances or at a different location
Interoperability	Compatibility with and ability to integrate data from other sources and surveillance components

2.2. Attributes aimed at evaluating the quality of the data collected

Data completeness and correctness	The proportion of data that was intended to be collected that actually was and the proportion of data entries that correctly reflect the true value of the data collected
Historical data	Quality and accessibility of archived data

3. Attributes related to surveillance effectiveness

3.1. Attributes aimed at evaluating inclusion

Coverage	The proportion of the population of interest (target population) that is included in the surveillance activity.
Representativeness	The extent to which the features of the population of interest are reflected by the population included in the surveillance activity, these features may include herd size, production type, age, sex or geographical location or time of sampling (important for some systems e.g. for vector borne disease)
Multiple utility	Whether the system captures information about more than one hazard

3.2. Attributes aimed at evaluating the quality of the evidence provided

False alarm rate (inverse of specificity)	Proportion of negative events (e.g. non-outbreak periods) incorrectly classified as events (outbreaks). This is the inverse of the specificity but is more easily understood than specificity.
Bias	The extent to which a prevalence estimate produced by the surveillance system deviates from the true prevalence value. Bias is reduced as representativeness is increased
Precision	How closely defined a numerical estimate is. A precise estimate has a narrow confidence interval. Precision is influenced by prevalence, sample size and surveillance approach used.
Timeliness	Timeliness can be defined in various ways • This is usually defined as the time between any two defined steps in a surveillance system, the time points chosen are likely to vary depending on the purpose of the surveillance activity. • For planning purposes timeliness can also be defined as whether surveillance detects changes in time for risk mitigation measures to reduce the likelihood of further spread The precise definition of timeliness chosen should be stated as part of the evaluation process. Some suggested definitions for the RISKSUR project are; **For early detection** Measured using time - Time between introduction of infection and detection of outbreak Measured using case numbers - Number of animals/farms infected when outbreak detected **For demonstrating freedom** Measured using time - Time between introduction of infection and detection of presence by surveillance system Measured using case numbers – Number of animals/farms infected when infection detected **For case detection to facilitate control** Measured using time - Time between infection of animal (or farm) and their detection Measured using case numbers – Number of other animals / farms infected before case detected **For detecting a change in prevalence** Measured using time - Time between increase in prevalence and detection of increase Measured using case numbers - Number of additional animals/farms infected when prevalence increase is identified.
Sensitivity	Sensitivity of a surveillance system can be considered on three levels. • **Surveillance sensitivity (case detection)** refers to the proportion of individual animals or herds in the population of interest that have the health-related condition of interest that the surveillance system is able to detect • **Surveillance sensitivity (outbreak detection)** refers to the probability that the surveillance system will detect a significant increase (outbreak) of disease. This may be an increase in the level of a disease that is not currently present in the population or the occurrence of any cases of disease that is not currently present. **Surveillance sensitivity (presence)** – refers to the probability that disease will be detected if present at a certain level (prevalence) in the population.
PPV	Probability that health event is present given that health event is detected
NPV	The probability that no health event is present given that no health event is detected
Repeatability	How consistently the surveillance component performance can be maintained over time.
Robustness	The ability of the surveillance system to produce acceptable outcomes over a range of assumptions about uncertainty by maximising the reliability of an adequate outcome. Robustness can be assessed using info-gap models.

4. Attributes assessing surveillance value

4.1. Attributes aimed at assessing value

Cost	The evaluation should list and quantify each of the resources required to operate the surveillance system and identify who provides this resource. These resources could include: time and personnel (labour), services (e.g. laboratory tests, postage), travel, consumables, and equipment.
Technical impact	This indicates the changes that have been based on the results of the surveillance providing a measure of the usefulness of the surveillance system in relation to its aims. This should include details of actions taken as a result of the information provided by the surveillance system e.g. changes in protocols or behaviour and changes in mitigation measures and particularly changes in disease occurrence
Benefit	The benefit of surveillance quantifies the monetary and non-monetary positive direct and indirect consequences produced by the surveillance system and assesses whether users are satisfied that their requirements have been met. This includes financial savings, better use of resources and any losses avoided due to the existence of the system and the information it provides. These avoided losses may include the avoidance of • Animal production losses • Human mortality and morbidity • Decrease in consumer confidence • Threatened livelihoods • Harmed ecosystems • Utility loss Often, the benefit of surveillance estimated as losses avoided can only be realised by implementing an intervention. Hence, it is necessary to also assess the effect of the intervention and look at surveillance, intervention and loss avoidance as a three-variable relationship. Further benefits of surveillance include maintained or increased trade, improved ability to react in case of an outbreak of disease, maintaining a structured network of professionals able to react appropriately against a (future) threat, maintaining a critical level of infrastructure for disease control, increased understanding about a disease, and improved ability to react in case of an outbreak of disease.

5. Economic efficiency criteria

Optimal economic efficiency	The net benefit to society shall be maximised. Achieved where the marginal costs of least-cost combinations of surveillance and intervention resources equal the marginal benefits of mitigation (=loss avoidance).
Economic acceptability	Ensuring that the benefits (=loss avoidance) generated by a mitigation policy at least cover the costs for surveillance and intervention.
Least-cost choice	Ensuring that a technical target for disease mitigation (e.g. time to detection) is achieved at minimum cost without quantifying the benefit.

7.3 Annex 2: EVA Tool Inputs

Context	Evaluation design pathway	Resources
1. What is the disease, threat or hazard of interest 2. What is the target population (species, sector and geographical area covered) and period covered 3. What is the pattern of disease occurrence in the target population, options include: - Endemic - Sporadic - Exotic - Re- emerging, Emerging	Why is this surveillance being carried out, what is the policy purpose for this surveillance activity, options include • Management of outbreaks – facilitate implementation of risk mitigation measures to limit the extent of an outbreak • Inform control - facilitate risk mitigation for a disease that is currently present to reduce economic impact • Inform control - facilitate risk mitigation for a disease that is currently present to reduce human exposure • Inform trade – provide evidence to restrict import or support export • Prioritize hazards – provide evidence to inform the requirement for risk mitigation What action will be taken based on the results from this surveillance activity What is the purpose of the evaluation, options include • assess whether the existing surveillance is meeting objective • identify the best surveillance approach to meet the objective • identify how to improve the existing surveillance to better meet its objective What are the national or international legal requirements for surveillance for this threat What economic method will be used, options include • Cost-effectiveness • Cost-benefit • Cost-minimization • More options to be added….	Which data is available for the assessment of evaluation attributes What resources are available to implement the evaluation

7.4 Annex 3. Provisional list of evaluation questions

Evaluation question	Description	Examples
1. *Ascertain if one or more surveillance component(s) or system(s) is/are capable of meeting a technical objective or target*	A surveillance objective and related target have been defined and the evaluator wants to know whether the surveillance component(s) or system(s) under evaluation is/are capable of achieving the objective and target. This is a binary question with a yes/no answer. It does not provide information about how the surveillance can be improved if it is not able to meet its objective and target. It can be applied to current surveillance and potential surveillance options.	*Example 1:* Surveillance objective: Demonstrate freedom from disease X in population z. Target: Demonstrate with 95% confidence that prevalence of disease X in population z is below 0.02. *Example 2:* Surveillance objective: Detect disease Y early Target: Detect disease Y in population j within 48 hours of incursion. *The evaluation question for both is whether the surveillance meets the defined target (yes/no question)?*
2. *Assess the costs of surveillance component(s) or system(s) (out of two or more) that achieve(s) a defined objective and rank them according to costs to identify the least-cost option(s)*	A surveillance objective and related target have been defined and the evaluator has at least two surveillance options that are capable of achieving the objective and target. The evaluator is interested in comparing the costs of these options which all achieve the target (binary outcome as in no. 1) and identifying the least-cost option(s). This requires calculating the surveillance costs. It can be applied to current surveillance and potential surveillance options.	*Example 1:* A range of surveillance systems with different characteristics to demonstrate freedom disease X in population z. No 1 uses blood sampling of cattle and antigen testing, No 2 uses bulk milk sampling and antibody testing, No 3 is a combination of both. They all fulfil the target. The evaluation question is what the costs are of the three options and which one is the cheapest.
3. *Assess the effectiveness of one or more surveillance component(s) or system(s) in relation to a surveillance objective and rank the options accordingly*	A surveillance objective and related target have been defined and the evaluator wants to know the effectiveness of the surveillance component(s) or system(s) under evaluation to achieve the objective and target. Contrary to no. 1 the outcome is not of the binary yes/no form, but provides a quantitative measure of the effectiveness in relation to the objective and target. It can be applied to current surveillance and potential surveillance options.	*Example 1:* Surveillance objective: To detect cases of infection with pathogen Y to facilitate interventions in an eradication programme. Target: Detect as many holdings that have animals infected with Y as possible in a defined area. The effectiveness measure in the case could be the sensitivity of surveillance. *Example 2:* Surveillance objective: To detect disease M early in population s. Target: Detect the disease as early as possible after incursion. The effectiveness in measure in this case could be the timeliness, e.g. number of days from introduction until detection; number of days from introduction until outbreak response; number of infected animals or holdings detected within the first week of the outbreak, etc.

Evaluation question	Description	Examples
4. *Assess if there is/ are (a) surveillance component(s) or system(s) that achieve a higher effectiveness than another one at the same cost*	A surveillance objective and related target have been defined and the evaluator wants to know if it is possible to achieve a higher effectiveness of the surveillance component(s) or system(s) under evaluation to achieve the objective and target at the same cost. Like in no. 3, the effectiveness in relation to the objective and target needs to be quantified and compared to the cost. This requires calculating the surveillance costs. It can be applied to current surveillance and potential surveillance options.	Same examples as for no. 4, but with the added question whether a higher effectiveness can be achieved while keeping the costs constant.
5. *Ascertain if a surveillance component or system generates a net benefit in monetary terms for society, industry, animal holder*	A surveillance objective and related target have been defined and the evaluator wants to know if the cost of the surveillance expressed in monetary units is lower or higher than the resulting benefit expressed in monetary units. Or in other words, the evaluator wants to know whether the programme generates a net benefit (or a net loss). This requires calculating the surveillance costs, the intervention costs, and the loss avoidance from having in place surveillance and intervention. Intervention is relevant here because disease cannot be avoided by having in place surveillance alone. The loss avoidance refers to difference in disease losses with the surveillance in question and the disease losses in the counterfactual (e.g. a situation without surveillance). It can be applied to current surveillance and potential surveillance options. It can be estimated for as many surveillance component(s) or system(s) as desired by the evaluator.	*Example:* Surveillance objective: To detect disease M early in population s. Target: Detect the disease as early as possible after incursion. The evaluator may want to know whether it is worthwhile to make an investment for a new surveillance system. Hence, it is necessary to think about the consequences of not detecting disease early, i.e. what would an outbreak look like without surveillance and what would it be with surveillance. In both cases, there will be a certain (but supposedly different) number of animals or holdings affected, the associated production losses, and resource use for outbreak control (and wider reaching consequences depending on the disease). All of these can be valued and the difference calculated.

Evaluation question	Description	Examples
6. *Ascertain if a surveillance component or system generates a net benefit in non-monetary terms for society, industry, animal holder*	A surveillance objective and related target have been defined and the evaluator wants to know if the cost of the surveillance is lower or higher than the resulting benefit expressed in non-monetary units. Or in other words, the evaluator wants to know whether the programme generates a net benefit (or a net loss). Contrary to no. 5, the loss avoidance is not estimated in monetary units, but non-monetary proxies may be used instead. Such an approach is recommendable where no or few market prices exist (e.g. consumer confidence, ecosystem health, animal welfare), where there may be ethical problems when attributing a money value (e.g. value of a human life) or where monetary benefits cannot be calculated to resource or data constraints and a proxy is to be used instead. Similar to a cost-benefit analysis as described under no. 5, the benefit needs to be estimated and compared to the costs of surveillance (and in some cases intervention). To determine whether an investment is worthwhile, the decision-maker and/or evaluator need to determine a cost-effectiveness threshold that reflects the value of the effectiveness outcome. It can be applied to current surveillance and potential surveillance options. It can be estimated for as many surveillance component(s) or system(s) as desired by the evaluator.	*Example 1:* Surveillance objective: To detect cases of infection with pathogen Y to facilitate interventions to eliminate the pathogen in wildlife species q. Target: Identify all areas in a country where infection in species q is present. Following detection, an intervention is implemented. The benefit identified from elimination is the conservation of the species. Hence, the survival of species q is the non-monetary benefit. *Example 2:* Exactly the same scenario as in no. 5, but instead of attributing a value to the consequences, a measure of timeliness (e.g. time until detection) is used as a proxy. It is important that the decision-maker has an
7. *Identify the surveillance system (out of two or more) that generates the biggest net benefit in monetary terms and/or non-monetary terms for society, industry, animal holder*	This is the same as in 5 and 6, but with a clear focus on identifying the surveillance component(s) or system(s) that create the biggest net value for society, industry, or animal holders	-

Evaluation question	Description	Examples
8. *Identify how surveillance attributes could be improved*	If a surveillance is not meeting its objective and/or target, if it not effective enough, and/or does not produce a net benefit in monetary or non-monetary terms, the evaluation focuses on assessing the strengths and deficiencies of the surveillance attributes that impact on performance and highlighting areas for improvement.	-

References

Dehove, A., Commault, J., Petitclerc, M., Teissier, M., & Mace, J. 2012: Economic analysis and costing of animal health: a literature review of methods and importance, *Rev.Sci.Tech.* 31, 605-604.