Deliverable D-WP3.2 – Combined report for all subtasks within the development of output-based metrics

OHEJP JIP MATRIX – WP3

Responsible Partner: 21-APHA
Contributing partners: 34-PIWET

GENERAL INFORMATION

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

DOCUMENT MANAGEMENT

Title OHEJP deliverable	D-WP3.2 – Combined report for all subtasks within the development of output-based metrics
WP and task	WP3-T4 – Development of evaluation strategies for surveillance systems based on output-based metrics
Leader	21-APHA
Other contributors	34-PIWET
Due month of the deliverable	M58
Actual submission month	M58
Type *R: Document, report* *DEC: Websites, patent filings, videos, etc.; OTHER*	R: report Save date: 1-Dec-22
Dissemination level *PU: Public (default)* *CO: confidential, only for members of the consortium (including the Commission Services).*	PU *See updated Grant Agreement*
Dissemination *Author's suggestion to inform the following possible interested parties.*	OHEJP WP 1 ☐ OHEJP WP 2 ☒ OHEJP WP 3 ☐ OHEJP WP 4 ☒ OHEJP WP 5 ☒ OHEJP WP 6 ☒ OHEJP WP 7 ☐ Project Management Team ☒ Communication Team ☒ Scientific Steering Board ☐ National Stakeholders/Program Owners Committee ☐ EFSA ☒ ECDC ☒ EEA ☐ EMA ☐ FAO ☒ WHO ☒ OIE ☒ Other international stakeholder(s): as per OHJEP dissemination plans Social Media: n.a. Other recipient(s): n.a.

Contents

Introduction

The concept of One Health (OH) promotes the decompartmentalisation of human, animal, and environmental health for more efficient and sustainable governance of complex health issues (Bordier et al., 2020). The MATRIX[1] project, forms part of the OH European Joint Programme (OHEJP), a partnership of 44 food, veterinary and medical laboratories and institutes across Europe and the Med-Vet-Net Association. MATRIX aims to build on existing resources within OH Surveillance by creating synergies along the whole surveillance pathway including the animal health, human health and food safety sectors. Key to the effectiveness of any surveillance system is the selection of appropriate methodologies to achieve the objectives of the system. The objective of the MATRIX project WP3[2] is to develop guidelines for the design, implementation, and evaluation of official controls within the food sector using output-based standards (OBS) by looking at the practicalities of implementing output-based surveillance within a OH context.

An OBS is defined by what the surveillance system must achieve in measurable units e.g. surveillance sensitivity or confidence in probability of freedom if a disease is absent, or case finding capacity if a disease is present (Cameron 2012). They therefore allow for variation in how surveillance is conducted, influenced by a variety of country/region specific factors including disease prevalence, performance of the tests used and risk factors of infection. Such factors can dictate risk-based sampling, the choice of alternative sampling and testing methods or the combination of different surveillance components, provided these are agreed between the relevant stakeholders. Using OBS can simplify the comparison of outputs from different surveillance programmes relevant to a region/country's specific situation which can often be a difficult task (Meletis et al., 2022).

Output based standards may be useful in the OH context where surveillance for animal pathogens can act as risk indicators for human health so there may be situations where specific industry focused efforts are most effective and efficient with regard to OBS. For instance, many pathogen control programmes are based on the principle of bottom-up eradication i.e. controlling disease at the animal level to curtail disease at the public health level. However, this may not always be possible if there is no correlation between the presence of the pathogen in the animal sector and food safety/human health.

This deliverable combines the remaining three tasks of the work package:
WP3-T1: Inventory of previous work and current practice
WP3-T2: Identification of operational partners and stakeholders
WP3-T4: Development of evaluation strategies for OBS surveillance systems

The inventory of previous work (WP3-T1) was achieved by carrying out a literature search to gain an understanding of how OBS have been used in surveillance within the animal health, human health, and food safety sectors. The remaining two tasks of stakeholder identification (WP3-T2) and evaluation strategies (WP3-T4) were conducted separately and are here incorporated into guidelines for conducting OBS surveillance. In line with the flexibility of output-based surveillance, these guidelines are non-prescriptive and are based on the three tasks documenting factors which should be considered either prior to designing an OBS system or when evaluating an existing one.

[1] MATRIX is a project of the One Health European Joint Programme (OHEJP), a partnership of 44 food, veterinary and medical laboratories and institutes across Europe and the Med-Vet-Net Association. MATRIX connects existing cross-sectorial One Health programmes in European countries. Today, 19 partner institutes representing the animal health, public health and food safety sectors from 12 countries continue a collaboration that started early in 2020 and will end in December 2022. More information can be found here.

[2] This document is a product of Work Package 3, Tasks 1, 2 and 4 in the MATRIX project. Contributors to these tasks were the OHEJP partners 21-APHA and 34-PIWET. Contributors to other tasks of WP 3 were OHEJP partners 33-NVI, 41-SVA, UCPH, 31-WBVR and 28-IZSAM.

Inventory of previous work and current practice

An input-based standards surveillance programme prescribes which activities must be performed e.g. which animal populations, tests, frequencies, etc. Conversely, OBS prescribe what the surveillance should achieve, for example, the level of confidence the surveillance gives us that disease can be detected at a specific design prevalence. Output-based standards are now becoming internationally accepted as they provide a flexible methodology allowing for risk-based sampling, alternative test methods, combination of different surveillance components and an accumulation of evidence over time. An OBS requires thorough knowledge of the pathogen for which surveillance is aimed at, the testing regimes in place and the 'at risk' populations enabling sampling to be focussed on particularly high-risk batches or populations. This reduces the economic burden of sampling requirements by eliminating testing of batches or populations that are deemed to be low risk whilst increasing the likelihood of detecting disease if it is present.

As OBS surveillance is a relatively new concept, we carried out a literature review to understand how the approach has been used in the three OH sectors of animal health, public health and food safety. The objective was to see whether this type of surveillance was being used with a OH approach in mind i.e. across more than one sector and what the current practice was in terms of how OBS surveillance was being utilised and what objectives it was achieving.

1. Methods

A literature search was conducted in January 2022 in both Scopus (www.scopus.com) and PubMed (www.ncbi.nlm.nih.gov/pubmed) using the search string "output *or* risk *and* based *and* surveillance *or* freedom" in the "title, keyword or abstract". No date range was specified to capture as many articles as possible. Articles were removed using the exclusion criteria as follows: duplicate articles; not in English; not concerning output-based standards (Figure 1). It is acknowledged that no set of search terms will be able to capture all relevant articles and that there will inevitably be some that would not have been captured here. Nevertheless, the search terms used were considered optimal as they were found to give the most comprehensive results despite many articles being rejected.

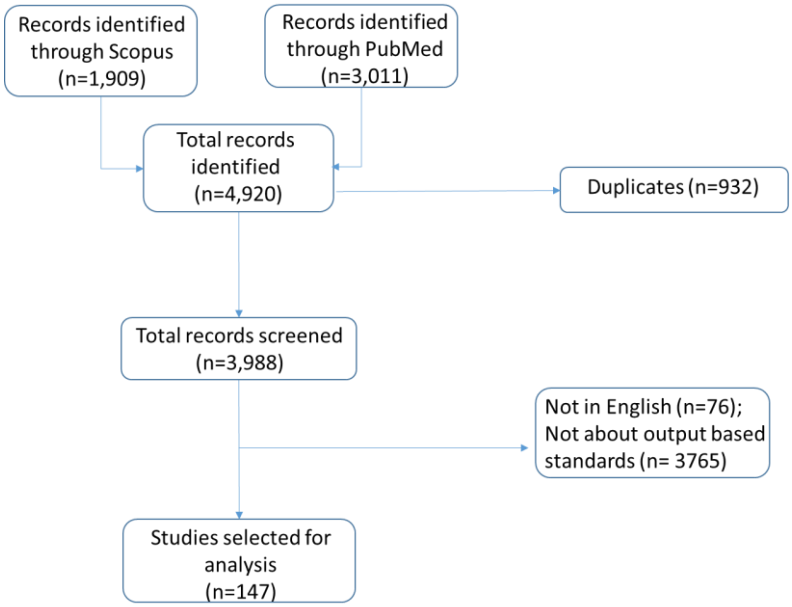


Figure 1, Flow diagram of the decision process and exclusion criteria

2. Results

After reviewing the titles and abstracts, 147 articles meeting the inclusion criteria were selected for the final analysis (See Appendix 1 for full details). The number and scope of the selected articles over time can be seen in Figure 2. There has been a gradual increase in the number of published articles relating to OBS surveillance since 2001 when the first articles began to be published. The majority of papers meeting the inclusion criteria related to the use of OBS in animal health surveillance i.e. the surveillance streams involved testing on farm in the live animal or fallen stock. These articles could then be further split into articles which related specifically to animal health diseases, those that related to zoonotic diseases and those that related to zoonotic diseases for which surveillance was conducted both in the live animal on farm and in the food sector such as slaughterhouse or bulk tank milk (BTM) testing. Animal health diseases (not zoonotic) for which surveillance streams included slaughterhouse or BTM testing were not included in this category. Three articles dealt with OBS surveillance for plant health (Coulston et al., 2008; Mastin et al., 2020; EFSA 2020) whilst six articles were about public health (Ranta et al., 2001; Watkins et al., 2009; Cameron et al., 2020; Epstein et al., 2020; Foddai et al., 2020; Ferri et al., 2021) although three of those related to how OBS surveillance systems used in veterinary health could be applied to help manage the Covid-19 pandemic (Foddai et al., 2020; Foddai et al., 2020; Ferri et al., 2021).

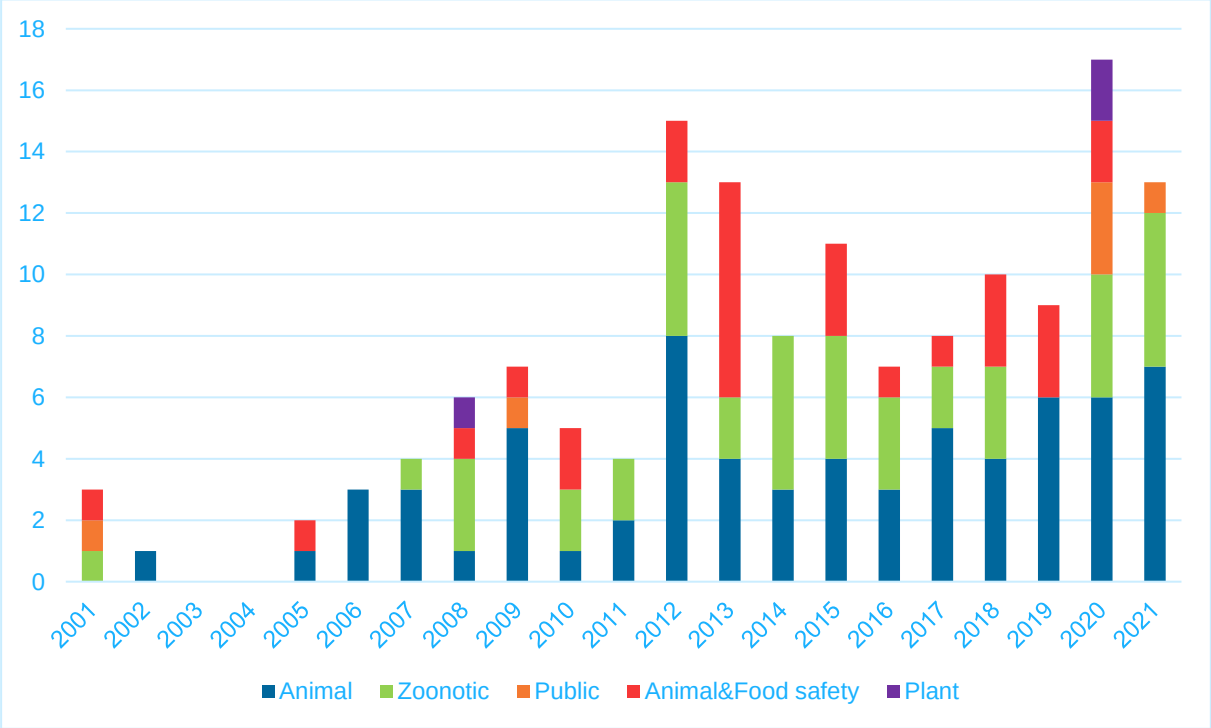


Figure 2, Number of articles meeting the inclusion criteria by year and by category of sector

Articles on the use of output-based surveillance in animal health were split into those relating to specific diseases and categorised as to whether these diseases were animal health only or zoonotic. Zoonotic diseases were then further split into whether surveillance streams contributing to OBS were in the food safety sector or not (Figure 3). A large proportion of articles relating solely to animal health were reviews of how OBS can be used to demonstrate freedom from disease or be used as part of a control or eradication programme for specific diseases, for example, those which are not regulated at European Union level (Hodnik et al., 2020; Koleci et al., 2021; Rapaliute e al., 2021). There were eleven zoonotic diseases covered by the review, but substantially more articles were written about bovine tuberculosis

(bTB) than any other disease (Figure 3). Articles concerning six zoonotic diseases referred to surveillance streams in the food sector namely, egg production (avian influenza), slaughterhouse surveillance (bTB, *Trichinella, Taenia*) and BTM testing (*Brucella, Mycobacterium avium* ssp. paratuberculosis). Animal health diseases (not zoonotic) for which surveillance streams include slaughterhouse or BTM testing were not included in this category. Taking bTB as an example, the majority of the articles were concerned with evaluating the system sensitivity especially with regard to the sensitivity of the screening test used and identification of high-risk herds. Ensuring that high risk populations are identified, and the relevant number of samples are taken with reference to the sensitivity of the tests used can assist in demonstrating confidence in freedom from disease, case identification for disease control and cost efficiency of surveillance strategies.

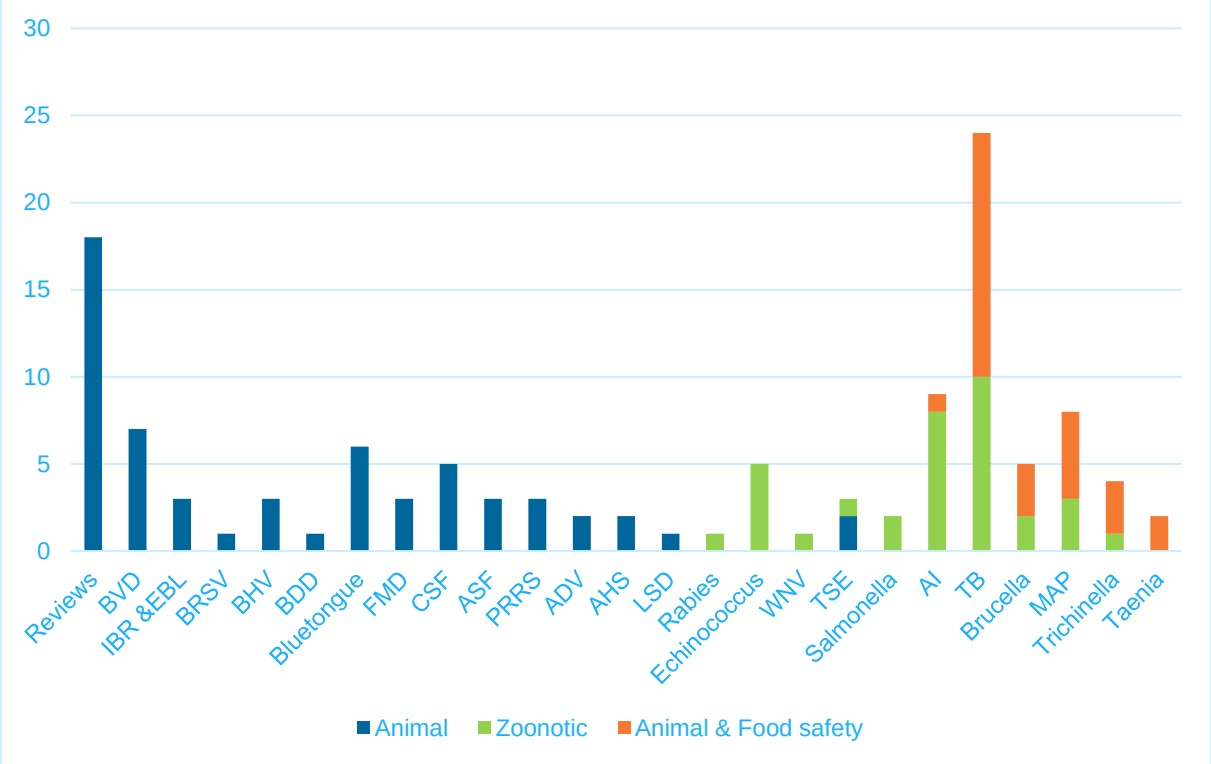


Figure 3, Number of articles meeting the inclusion criteria by pathogen and by category of surveillance (key: BVD= bovine viral diarrhoea, IBR= infectious bovine rhinotracheitis, EBL= enzootic bovine leucosis, BRSV= bovine respiratory syncytial virus, BHV= bovine herpes virus-1, BDD= bovine digital dermatitis, FMD= foot and mouth disease, CSF= classical swine fever, ASF= African swine fever, PRRS= porcine reproductive and respiratory syndrome, ADV= Aujeszky's disease, AHS= African horse sickness, LSD= lumpy skin disease, WNV= West Nile virus, TSE= transmissible spongiform encephalopathy, AI= avian influenza, TB= bovine tuberculosis, MAP= Mycobacterium avium ssp. paratuberculosis)

3. Discussion

From a OH perspective, OBS surveillance appears to be predominantly used in the animal health sector. There is some cross-over with the food safety sector where surveillance for some pathogens in this sector can be one of the components of the surveillance programme. The early detection of disease outbreaks in animal populations can be of public health value. For example, it has been shown that National Control Programmes for *Salmonella* in poultry have correlated with a decline in cases of human salmonellosis in several EU countries (Korsgaard et al., 2009; Tzani et al., 2021). There are a number of critical points where surveillance for zoonotic pathogens can be carried out e.g. on farm, at the slaughterhouse, at the supermarket etc. but utilizing an OBS at the animal surveillance level for some

pathogens may have benefits for human health from a OH approach. Only six articles regarding human/public health met the inclusion criteria. This could be because surveillance for most pathogens in this sector is predominantly outbreak based and is only instigated as part of a tracing operation to identify source and spread of disease should an outbreak occur (see Matrix D-WP2.1 for mapping of the surveillance chain and cross-sectorial linkages: https://zenodo.org/record/6406150#.Y1_s39fP3ct).

Scenario tree modelling was a common approach used in many of the articles included. This methodology presents an objective quantitative analysis of multiple complex data sources to support claims of freedom from disease (Martin et al., 2007). Using a single value for design prevalence implies that all groups in the population have the same average probability of being infected. Scenario-tree modelling effectively divides the population into multiple different risk groups to reflect the group-level probability of infection allowing for situations where there may be clustering of disease. This approach was used in many articles to demonstrate how to increase the efficiency of surveillance, providing equivalent surveillance sensitivity at lower cost (Cameron 2012).

In addition to scenario tree modelling a Bayesian framework was described when conducting OBS surveillance in some articles. This allows the inclusion of prior distributions for surveillance parameters and can be used to estimate true prevalence (Viljugrein et al., 2018). A Bayesian approach is useful if a design prevalence (a prerequisite of scenario tree modelling) is not provided by legislation or when infection has been absent for many years (van Roon et al., 2019). A Bayesian approach to OBS disease surveillance can also be used in place of weighting of populations for a risk-based approach as it allows any uncertainty to be incorporated. A risk-based approach can improve a surveillance system by both increasing the probability of detection and reducing costs (Heisey et al., 2014; Salvador et al., 2018; van Roon et al., 2019).

In conclusion, the articles which met the inclusion criteria in this review were predominantly concerning the use of OBS in animal disease surveillance. Where zoonotic diseases were described surveillance streams were in both the animal health and food safety sectors. Articles relating to solely public/human health pathogens were limited. No articles were retrieved which described an OBS surveillance programme covering all three sectors i.e. from a complete OH perspective.

Guidelines for the design, implementation, and evaluation of official controls within the food sector using output-based standards (OBS).

There are various reasons why OBS surveillance may be appropriate in a One Health (OH) context. Most importantly, it can be adapted to variable disease profiles across legislative regions. This makes it scalable to multi-country, country-level and local implementations.

Here, we outline a guided methodology for how an OBS surveillance system can be designed, implemented and evaluated. It aims to explain some of the details of this surveillance format, helps provide evidence-based decision-making on the best ways of applying it, and showcases analyses to direct improvements to disease surveillance. This guidance builds on WP3 deliverable report on output-based surveillance system selection methodology (https://zenodo.org/record/6984562#.Y1_sLtfP3cs).

This guidance is aimed at those who are considering using OBS as a solution to a surveillance need, whether they are looking to implement a system from scratch, replace a conventional surveillance system, or evaluate potential improvements to an existing OBS system. Because of the broad audience of this guidance, not all sections will be applicable. Equally, while a loose chronology exists throughout the guidance, sections can be completed out of order. See Figure 4 for more details.

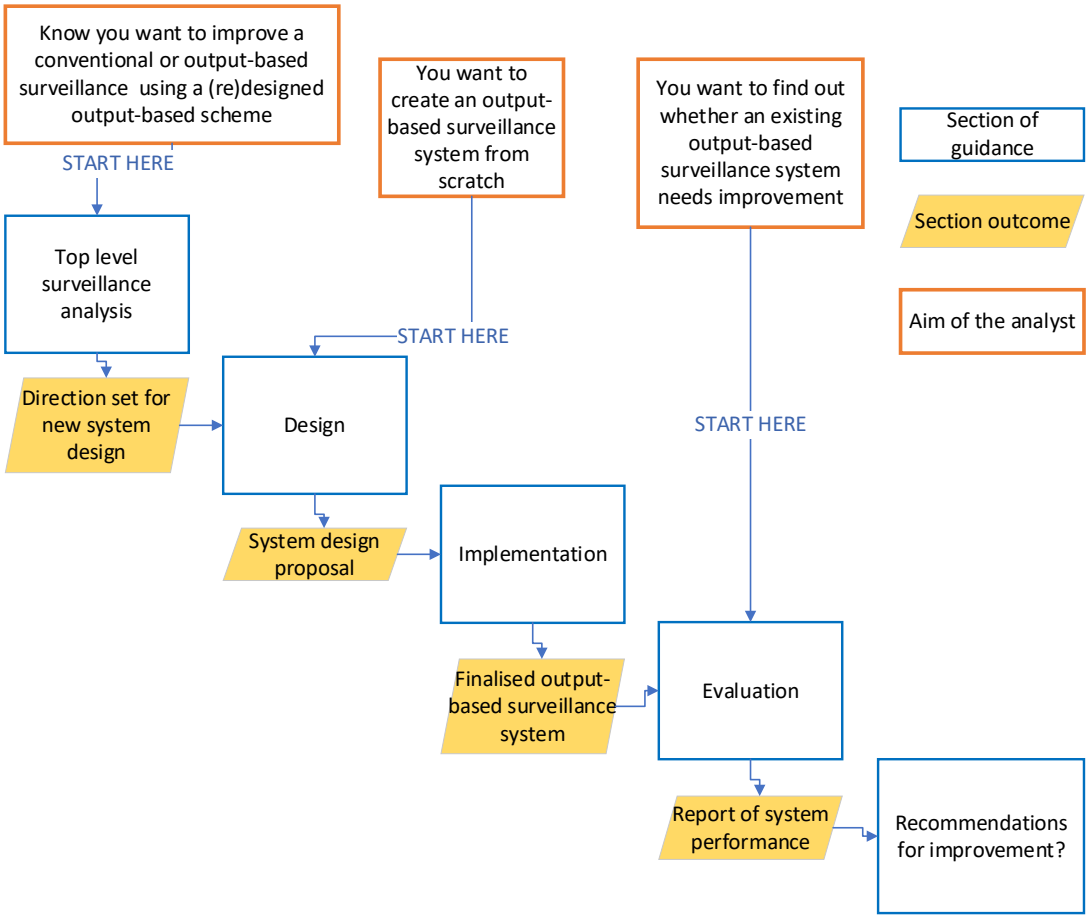


Figure 4, Showing the recommended route an analyst should take through this guidance if they either know they want to improve an existing surveillance system, want to design and implement an output-based surveillance system from scratch, or want to assess the performance of an existing OBS system.

4. Top level surveillance evaluation

If there is already surveillance in place for the target pathogen and population, then a top-level assessment of this surveillance system can establish where it needs strengthening. This functions as a pre-planning stage in the design of any replacement system, framing subsequent design, implementation, and evaluation within the context of the strengths and weaknesses of the existing system. In other words, bringing known issues with the current surveillance to the surface, so that these can be addressed.

Guidance has already been produced for assessing conventional surveillance systems in tools such as SERVAL (Drewe et al., 2015), RISKSUR (https://www.fp7-risksur.eu/), EpiTools (https://epitools.ausvet.com.au/) and OH-EpiCap (WP 4.2 MATRIX - One Health EJP: https://zenodo.org/record/7006654#.Y1_sP9fP3cs and https://onehealthejp.eu/wp-content/uploads/2022/11/OHEJP-MATRIX_OH-EpiCap-flyer.pdf). These tools can also be employed in this pre-planning stage. A common framework for One Health surveillance has also been described by Matrix WP2 here: https://zenodo.org/record/7064374#.Y1_sONfP3cs.

The top-level evaluation described here scores a subset of important attributes from these tools on a fixed scale using expert opinion. These attributes, and their definitions, are shown in **Error! Reference source not found.**

Table 1: Description of the attributes evaluated during a top-level analysis of surveillance.

Attribute	**Definition**
Stakeholder satisfaction	Level of acceptance from stakeholders
Adaptability to sudden scale up	How intuitively surveillance can be upscaled in response to a surge in cases or sudden political interest
One Health interoperability	How well it works with One Health partners
Synergy with other systems	How well it works with other surveillance systems, and whether synergies are being used to their fullest
Management	How fairly roles in the system are distributed, defined, and supported
Appropriate reach	Reaches as much of the target population as it needs to
Population variability captured	Sensitive to differences in disease presentation or population behaviour across the sample area
Accuracy of results	Reflects a prevalence that is close to the true prevalence in the population
Precision of results	How narrow the confidence interval for the results obtained is
Surveillance sensitivity	The probability that the surveillance system will detect an outbreak (sometimes represented by the detection fraction)
Reliability (false positives/negatives)	The probability that the results seen are correct
Economic efficiency	Perceived value for money
Simplicity	How easily the logic of the system can be understood by stakeholders
Data Quality	How often practices within the system (sampling, testing etc.) are quality-controlled
Compatibility	How well the surveillance fits with other organisational practices
Timeliness	How quickly results are collected and reported, for example, surveillance may be conducted seasonally
Repeatable	How easily could this surveillance be applied elsewhere

Attribute	Definition
Grounded by political will/support	How adequately it fulfils the political priorities of surveillance
Relevant and informative	Whether it provides the information you want to know about a disease

Each attribute can be scored as high, medium or low, with scores of 3, 2 or 1 respectively. These scores are defined as follows:

- High: the attribute consistently delivers on its function and sometimes exceeds it. Unlikely to be improved upon further.
- Medium: the surveillance attribute currently performs as intended but has at least one area where it could be improved
- Low: the surveillance attribute needs to be improved for the long-term success of the surveillance system.

The overall score profile for each attribute can indicate areas for improvement in the subsequent design phases.

Example – *Hypothetical top-level analysis of E. multilocularis surveillance system in Poland*

Background:
The *E. multilocularis* surveillance in Poland is an output-based scheme that has been in place for several years across its various municipalities. Top-level surveillance evaluation was conducted to help determine the areas of this system that are performing well and areas for improvement.

Process:
We conducted a hypothetical evaluation of *E. multilocularis* surveillance in Poland using an unvalidated scoring of each of the attributes in Table 1. Scores are based on dummy data intended to represent how this analysis could be done on a known surveillance system. As such, this scoring table is not applicable to the real Polish conditions at the time of analysis.

Results:
The assigned scores for these attributes are presented in Table 2. The sum of all scores was 44. In the dummy data used, high scores were set for 7 attributes, 11 for medium with only 1 low score. This would indicate that in the opinion of the analysts, the surveillance system is performing well, with areas for minor improvements (in, for example, reach, reliability and compatibility) and space for larger improvements in economic efficiency. This analysis provides an impression of the surveillance system and is more indicative than conclusive of where a system could be improved in future designs. Further analysis (see Evaluation section 7) would need to be conducted before drawing concrete conclusions.

Table 2: Output scores for surveillance attributes of the E. multilocularis surveillance system in Poland. High = 3, medium = 2, low = 1. Unknown or N/A is not scored

Attribute	High	Medium	Low	Unknown/Not applicable
Stakeholder satisfaction		2		
Accuracy of results	3			
One Health interoperability	3			
Synergy with other systems	3			
Economic efficiency			1	
Appropriate reach		2		

Reliability (false positives/negatives)		2		
Compatibility		2		
Grounded by political will/support	3			
Precision of results		2		
Repeatable	3			
Population variability captured		2		
Data Quality	3			
Surveillance sensitivity		2		
Simplicity		2		
Well managed	3			
Timeliness		2		
Relevant and informative		2		
Adaptability to sudden scale up		2		
Total	21	22	1	
Overall	44			
Out of	57			

5. Design of an output-based standards surveillance system

The design stage documents the important aspects of a proposed surveillance system. If a proposed OBS system is designed to replace an existing system, then the design stage will help establish what aspects from the previous scheme to retain and which to adapt. If an OBS system is being produced from scratch, then the design stage set out in this guidance will provide a framework for doing so.

Primarily, design is about information gathering, decision-making, and objective setting. In this guidance we set out methodologies to define:

- the objectives of the system
- the key stakeholders to involve
- the pathogen(s) being tested for
- the populations(s) being tested
- the tests that are used
- the distribution of sampling and number of samples taken
- the cost of sampling
- the desired data outputs

5.1. System objectives

5.1.1. Setting the system objectives

What: This section helps establish the objectives of the system i.e. what the surveillance system hopes to achieve from a top-level perspective. The objectives could be to fill a regulatory requirement, to be part of a national strategy, or to assist with disease control at the local level. The objective of an OBS

system could be to demonstrate freedom from disease, or to show the pathogen prevalence in a population with a certain level of confidence. For an OBS system the important attributes which should be considered when setting the objectives are:

- Design Prevalence: This is a fixed prevalence used to determine the hypothesis that disease is present in a population of interest (Stevenson and Sergeant, 2022).
- Confidence levels: This is the level of certainty, generally expressed as a percentage, that the result is correct. That is, if the surveillance process were repeated, the result would be correct X% of the time, where X is the confidence level.
- Surveillance streams: these are made up of a specific population (with associated risk level), where the surveillance occurs e.g. on farm, slaughterhouse etc. and what tests are used. There could be several streams available which can contribute to achieving OBS.
- Probability of introduction: Likelihood of the disease in question being introduced to at least the number of units (e.g. animals) that would be infected given the design prevalence.

Why: The objectives are a thread that runs through all the elements of the system. Documenting the objectives can help ensure they are appropriate to the current context and can serve as a reference for shaping the system during later design phases.

How: One method of compiling a complete list of objectives is to use a hierarchy of objectives (Rahmatian, 1985). This process facilitates the expansion of objectives from their ultimate, top-level goal, down to their practical, ground-level implementation. You can navigate up or down the hierarchy by asking how or why. The 'how?' of an objective should link to an objective in a lower tier and 'why?' with one in a higher tier. This is illustrated in Figure 5.

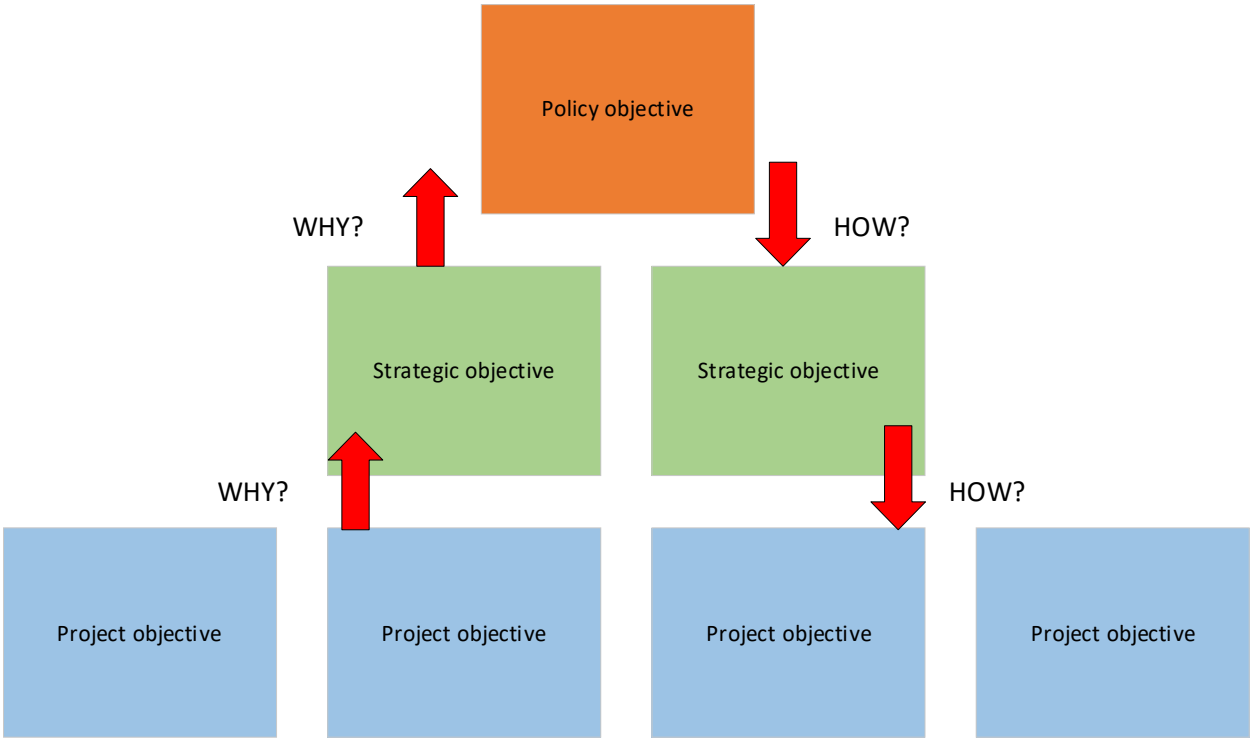


Figure 5, showing the hierarchy of objectives. The policy objective provides the reason why the output-based surveillance is needed, strategic objectives define the strategies in place to meet those policy objectives, while the project objectives provide the practical mechanisms by which these strategies are carried out.

The further down the hierarchy, the more the objectives overlap with the methodologies. Hence, you should think of project objectives as a set of practical constraints and drivers. See the *E. multilocularis* example below for more details.

The objectives can be defined primarily through communication with the prospective system stakeholders (see section 5.2). Once a hierarchy of objectives has been developed, these can again be validated by these stakeholders for accuracy.

Example – *E. multilocularis* surveillance in Great Britain (GB)

Background:
Countries where *E. multilocularis* is not endemic must demonstrate freedom from disease by upholding surveillance in accordance with an output-based scheme prescribed by the European Commission (European Commission, 2017). European Union (EU) member states must demonstrate a prevalence of not more than 1% with a confidence level of at least 95% to be considered free from disease. Although GB has left the EU, this surveillance is still mandated by retained legislation.

Process:
We used the information from this legislation, along with direct input from the system stakeholders, to build a hierarchy of objectives for the GB surveillance system (Figure 6).

Results:

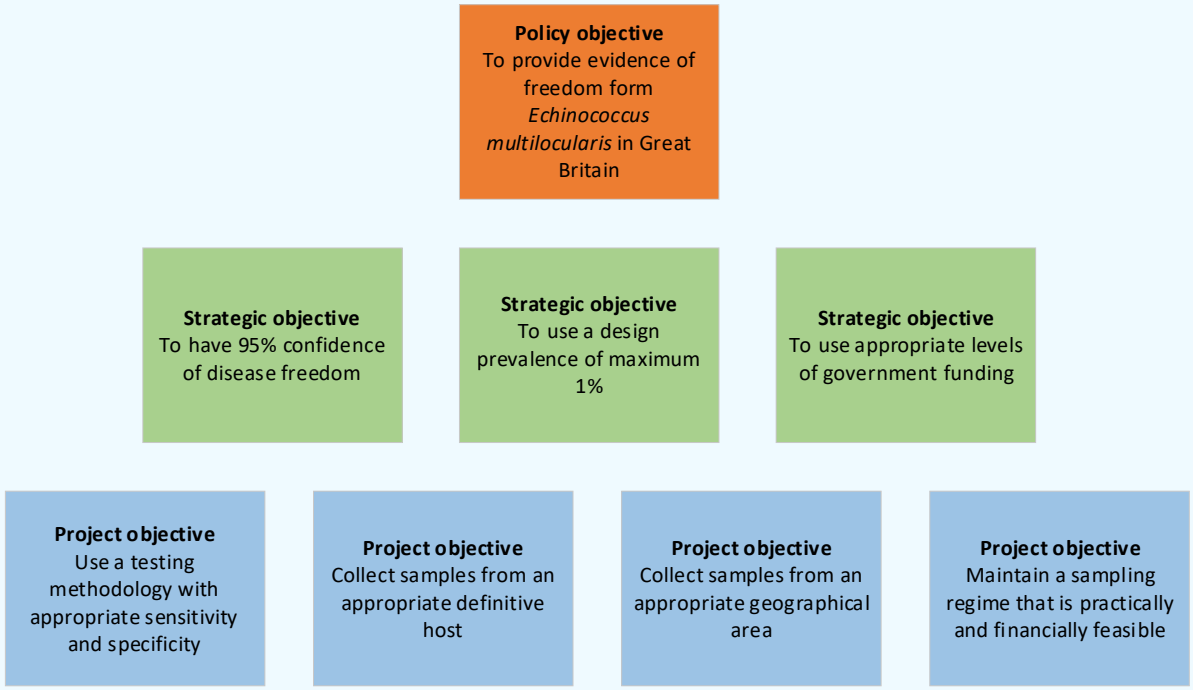


Figure 6, Showing the objective hierarchy of E. multilocularis surveillance in Great Britain.

5.2. Stakeholders

5.2.1. Identifying stakeholders of the system

What: Stakeholders are "any parties who are affected by or who can affect the surveillance system" (Friedman and Miles, 2006).

Why: Knowing the different stakeholders within a system is an essential part of any change management process (Hayes, 2022). Stakeholders, have oversight of the surveillance system and are a useful resource for design choices and as a source of opinions and knowledge to optimise the surveillance system design.

How: The list of stakeholders should be brainstormed based on the available information about the pathogen and the objectives of the system (see section 5.1). This could be internal institutional knowledge of any existing surveillance systems for the pathogen, literature research, or information gathered from your existing professional network. For example, colleagues you have worked and collaborated with on previous projects. Generally, stakeholders of surveillance systems are divided between those who are involved in conducting surveillance and those who the surveillance works to protect (Mazet et al., 2014). These make up three distinct groups. First, governance stakeholders with the influence to set the required output of the surveillance system. These could be those who set regulatory standards, or who uphold quality standards. For example, a national or multinational regulatory authority like the European Food Safety Authority (EFSA). Second, delivery stakeholders who are actively involved in the delivery of the required outputs, either in the collection of samples, laboratory analysis and results reporting, or in the auxiliary planning and strategy roles that enable the surveillance to run smoothly. Finally, beneficiaries who directly or indirectly benefit from the system running well, and whose wellbeing would be directly or indirectly affected by a change to the surveillance system. The general public, for example, are beneficiaries of surveillance systems involving zoonotic pathogens.

Once a list of stakeholders has been established, a strategy for engagement should be devised. Find out who in your organization has had contact with your proposed stakeholders. If your organisation is currently external to the surveillance system and no contact has been previously made with the stakeholders, research into the current surveillance system or stakeholder institutions can provide a person or department to contact. Once contact with at least one stakeholder has been established, these may then be used to establish contact with other stakeholders in the system.

When a full list of stakeholders has been established and contact has been made, they can be used for further information gathering. A structured interview with a pre-planned series of questions is recommended. It is recommended that you read other sections of this guidance to direct the specific input you need from them.

Example: *E. multilocularis* surveillance in Great Britain (GB)

Background

Surveillance stakeholders are defined as "any parties who are affected by or who can affect the surveillance system" (Friedman and Miles, 2006). With this definition, we sought to identify these stakeholders for the current *E. multilocularis* surveillance system in GB.

Process:

Potential stakeholders were identified through a brainstorming session and compiled into a preliminary list. We then used information from direct contact with one of our stakeholders: the Animal and Plant Health Agency (APHA) parasitology lead, to confirm a wider stakeholder list, and to engage several other stakeholders. Finally, we categorised the list into each of the three stakeholder groups.

Results:

The final list of stakeholders was as follows:

Governance:

- The World Organisation for Animal Health (WOAH); who record the disease status of *E. multilocularis* following the compilation of GB results.
- The GB Department for Food, Environment, and Rural Affairs (DEFRA); who compile the results.
- Devolved administrations, who provide oversight of surveillance in Wales and Scotland.

- Local councils, who play a role in maintaining good education on the disease and responding to cases.
- The European Free Trade Association (EFTA); who advise on the measures which should be in place to control *E. multilocularis* given a change in GB's status.

Delivery:
- APHA, who maintain the surveillance system, collecting samples and running analysis. Including:
  - The national reference laboratory (NRL) for *Echinococcus*
  - APHA wildlife management team
  - APHA wildlife risk modelling team.
- Veterinary practitioners, who respond to cases in dogs and hold a stake in maintaining their good health.
- United Kingdom Health Security Agency (UKHSA); who respond to and detect human cases. They would collaborate with APHA in responding to a positive case in humans or positive sample in foxes or dogs.
- Hunters and gamekeepers, who kill foxes and provide carcasses from across the country for testing.

Beneficiaries:
- The Wildlife Trust, who support the welfare and environmental influences of surveillance on fox populations and the general ecology. They have a voice in ensuring surveillance does not severely, or unnecessarily, impact the wellbeing of foxes.
- Fera Science, a wildlife science advice organisation who receive samples from foxes and other wildlife for rodenticide survey, and who could benefit from collection of foxes for this surveillance.
- Science Advice for Scottish Agriculture (SASA), who also receive samples from foxes and other wildlife for rodenticide survey, and who could benefit from collection of foxes for this surveillance.
- Pet owners, who hold a stake in making sure their pets remain healthy, and who are the most likely to become infected given transfer from foxes to pets, due to contact with their pets and infected intermediate hosts.
- Media outlets, who have an interest in distributing information on the quality of surveillance and in the event of case detection.
- The general public: good surveillance ensures that any incursion of *E. multilocularis* reaches as few members of the public as possible.

5.3. Surveillance Parameters

It is recommended that a method such as scenario tree modelling approach is used to model the process of disease detection via each surveillance system component. The tree should include all factors affecting the probability of infection or detection of a surveillance unit for a given design prevalence and sample size. Bayesian approaches can also be used, in particular, where no design prevalence is given or when incorporating prior distributions for surveillance parameters.

5.3.1. The pathogen of interest

What: The target pathogen and its epidemiological features should be known before designing your surveillance system.

Why: This heavily impacts any downstream practical decisions on how the system will function, including the choice of definitive host, and sampling method.

How: Structured interviews with the appropriate stakeholders (see section 5.2 for more information on stakeholders) may provide knowledge about the target pathogen, along with literature research. This information can then be compiled into a succinct pathogen profile. Any relevant information can be

added to this profile, but it should aim to be a complete overview covering all one-health aspects. If the pathogen is zoonotic, particularly if it is a foodborne pathogen, this should be flagged at this stage.

Example – *E. multilocularis* surveillance in GB

Background

E. multilocularis poses a threat to humans and animals across GB, so maintaining freedom from this pathogen is a major policy priority.

Process

A literature review was conducted using the snowball sampling methodology to assemble the relevant information on the pathogen (Lecy and Beatty, 2012).

Results

E. multilocularis is a tapeworm whose larval stages cause disease in several domestic and wild species as well as alveolar echinococcosis in humans. Generally, animals do not show any clinical signs of infection, but can sometimes present with:

- Bowel pain
- Fluid accumulation in the abdomen
- Weight loss
- Jaundice
- Along with clinical signs that resemble tumours and sometimes alveolar cysts in the liver, brain, or other areas of the body (Centers for Disease Control and Prevention, 2019).

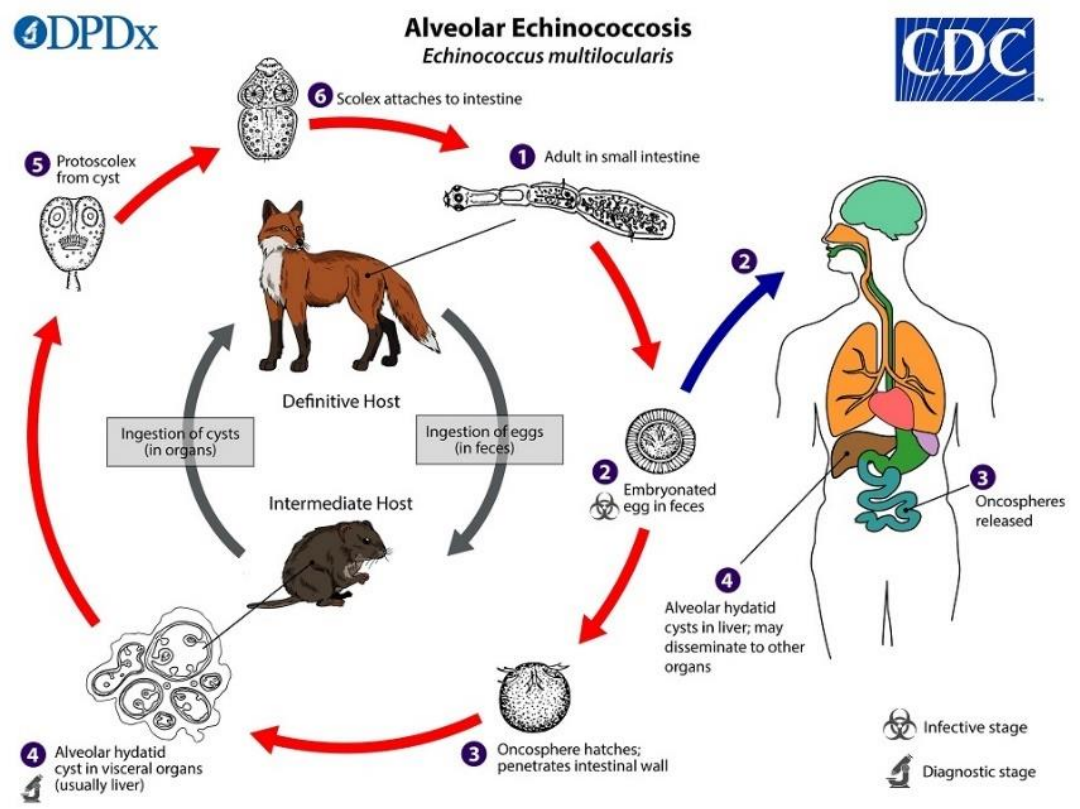


Figure 7, Life cycle of Echinococcus multilocularis (Centers for Disease Control and Prevention, 2019)

In this life cycle (Figure 7), foxes (particularly red foxes), raccoon dogs and other canids can all take the place of the definitive host, while rodents represent the intermediate host (World Health Organization, 2021). In contrast, humans are considered an aberrant host, and while infection can cause alveolar echinococcosis, humans are dead end hosts (DEFRA and APHA, 2019).

5.3.2. The population of interest

What: As with the target pathogen, the target population is a key determinant in your system design. This is usually the population that is considered to be most at risk and therefore the one in which you are most likely to detect a positive case.

Why: The choice of population has implications on almost all areas of the workflow, including sampling type and method, and geographical area(s) sampled. Generally, the population or populations being tested are the one(s) most at risk of being infected. Choosing an appropriate population to sample gives greater confidence that results correlate to the true prevalence of the pathogen.

How: Engaging the system's stakeholders through structured interviews (see section 5.2 for more information on stakeholders) is a good way to determine the population of interest, as well as provide further information about why that population or species has been chosen. This can also be combined with literature reviews to further enhance that knowledge. The population(s) or sources to be tested will vary with the pathogen, and the nature of the pathogen will dictate the most relevant population or source to test. For example, if the surveillance is concerned with a foodborne pathogen, the most appropriate sample source or surveillance stream might be sampling at slaughter, testing of bulk milk, or testing animal products at retail or before import.

Example – *E. multilocularis* surveillance in GB

Background:
There are a number of viable hosts for *E. multilocularis* that are present in GB including red foxes, domestic dogs and some rodent species. Each species represented different challenges in terms of workflow and sampling procedures.

Process:
To determine the species sampled for *E. multilocularis* testing in GB, we initially conducted literature searches looking at viable hosts of *E. multilocularis* in GB. Subsequently, we engaged delivery stakeholders from APHA, Fera Science and SASA through structured interviews.

Results:
Although *E. multilocularis* has various domestic and wild hosts, the red fox is the most relevant wild host in GB. Microtine rodents such as *Microtus arvalis* (the common vole) and rodents in the Arvicolinae subfamily can also be infected with *E. multilocularis* (World Organisation for Animal Health (WOAH), 2022), however, these rodent species represent intermediate hosts.

By sampling from individual foxes rather than taking environmental samples or sampling from intermediate hosts, the results can be identified at animal level, ensuring the species and approximate location ID is known. Additionally, the samples are taken from foxes that are killed by hunters and gamekeepers. This can serve to reduce costs by negating the need for bespoke sample collection systems while strengthening engagement with hunters and gamekeepers as stakeholders.

Hence, in GB the red fox is the most appropriate host choice for the surveillance scheme.

5.3.3. Sampling methods and distribution

What: The method of sampling and geographical distribution of samples taken is decided here.

Why: Knowing how the target population is distributed and how it will be sampled is an essential practical detail that will help inform the type of test used, and how the final design proposal will be implemented.

How: Samples may be taken using a risk-based framework or by taking randomly from the entire population. The extent of risk-based selection should be stated. Convenience sampling is not

recommended for OBS surveillance as it would be unlikely to support representative sampling of the host population. Delivery stakeholders can provide the contextual knowledge to inform the type of sampling that is most appropriate and feasible (see section 5.2 for more information on stakeholders). Additional external information sources such as population surveys could provide further information to support the chosen sampling type. The specifics of sampling link closely to the testing method chosen (section 5.3.4) as the number of samples required will vary based on the sensitivity of the test used. In order to confirm the number of samples required, and to validate confidence in the test results, a sample size calculator such as EpiTools can be used (Sergeant, 2018a). Regardless of the sampling method chosen, research should include all sample sources that are relevant to the probability of introduction of the pathogen. For farmed or kept animals, this will likely include multiple surveillance streams such as slaughter animals, imports and movements. In contrast, for wild animals, relevant surveillance streams may include trapped or hunted animals, resident populations, and transient or migratory populations, particularly where they cross borders.

Example - *E. multilocularis* surveillance in Great Britain (GB)

Background:
The target of this sampling, the red fox (*Vulpes vulpes*), has a varied distribution across GB (example distribution map in section 5.3.2).

Process:
We gathered information about the number of carcasses to be sampled through structured interviews with the APHA wildlife team. Because of the variable nature of both the distribution and number of samples taken per year, the wildlife team carry out spatial modelling annually to ensure that the samples being tested represent a random sampling distribution across the population area.

Results:
In this system, an excess of samples are taken and stored. Spatial modelling is then used to clarify which samples to test to generate an optimally random sampling distribution.

As part of this section, EpiTools (https://epitools.ausvet.com.au/) was used to confirm the required sample size to fulfil the requirements of the surveillance system. This links closely with sections 5.3.4 and 6.2 as the sample size calculation required additional information about the test sensitivity and population size. Using the egg flotation test as the initial design choice, the required sample size to detect disease at a 1% prevalence with 95% confidence given a random sampling distribution was 383.

The input used with the EpiTools calculator was:
- Population: 357,000 (Mathews et al., 2018)
- Design prevalence: 1%
- Test sensitivity: 0.78
- Test specificity: 1

Required sample size:	383
Cut-point number of positives:	0
Type I error:	0.0498
Type II error:	0
Population-level sensitivity:	0.9502
Population-level specificity:	1
Interpretation:	If a random sample of **383** units is taken from a population of **357000** and **0** or fewer reactors are found, the probability that the population is diseased at a prevalence of **0.01** is **0.0498.**
Method:	Simple binomial (large population)

Figure 8, EpiTools output for calculation of required sample size based on population size, test sensitivity and design prevalence of the system

5.3.4. Testing methods

What: Here the method of testing is decided which will optimally detect the target pathogen in the target population. The selected method must also be approved by all stakeholders involved.

Why: When designing a surveillance system, whether output-based or otherwise, a testing method must be chosen that fulfils the objectives of the system.

How: The sample collection medium (for example, faeces or tissue) and test used can be decided together; one will determine the other. Through structured interviews with the stakeholders (see section 5.2) and literature review, the testing options can be collated. From there, the most appropriate method can be chosen, considering the budget and resources available, the sensitivity and specificity of the testing method, and the population available for testing. The sensitivity of different testing methods can be assessed when associated with different surveillance streams. The primary concerns of the testing methods are:

- Sensitivity: Probability that a positive test result is true.
- Specificity: Probability that a negative test result is true.

Example - *Echinococcus multilocularis* surveillance in Great Britain

Background:
Multiple testing methods are available for detection of *E. multilocularis*. Each has associated advantages and drawbacks, so compiling a list of the available options allows for easier evaluation in subsequent steps.

Process:
Through engagement with stakeholders and by reviewing the available literature, we compiled a list of available test types along with their corresponding sensitivity values.

Results:
The compiled list of available tests for detection of *E. multilocularis* is:

- Zinc egg flotation and PCR (APHA parasitology lead, 2022)
- Sedimentation and counting technique (SCT) (APHA parasitology lead, 2022)
- Real time PCR (multiple methods available depending on the target, primers and probes) (World Organisation for Animal Health (WOAH), 2022)
- PCR with direct DNA extraction from faeces (World Organisation for Animal Health (WOAH), 2022)
- Coproantigen ELISAs (World Organisation for Animal Health (WOAH), 2022)

Speaking with the parasitology team at APHA, the egg flotation and SCT were both tests that the staff had experience in. The real time PCR is not currently a WOAH recommended test for large scale population surveillance. However, the literature suggests a higher sensitivity than many of the other techniques, so this was retained as an option for further evaluation. From this list, the options taken forwards were:

*Table 3: Testing methods for E. multilocularis. *Test sensitivity as recommended by EFSA. +Test sensitivity from the average of the range found in literature.*

Parameter	Test		
	Egg flotation	**SCT**	**qPCR**
Species sampled	Fox	Fox	Fox

Test sensitivity	0.78*	0.78*	0.89+
Test specificity (assumed)	1	1	1

For the purposes of this selection, the test specificity is set at 1, as any non-negative results were assumed to undergo further confirmatory testing.

5.3.5. Test costing

What: This section discusses how to calculate the annual and/or per sample cost of testing.

Why: Understanding the costs of testing helps determine whether surveillance is achievable within the budgetary constraints of your system. It also allows for assessment of which surveillance streams give the best value for money balanced against the test sensitivity. Later, in the evaluation section, this can also be used to calculate the cost-effectiveness of the chosen surveillance option.

How: Cost of testing can be broken down into:

- Consumables and reagents: This will cover any routine consumables costs such as reagents, personal protective equipment (PPE), laboratory, or field consumables.
- Staff: This will cover all costs relating to staff e.g. cost of staff time for sampling, testing, training and travel.
- Equipment: This covers the cost of all equipment used in the system. This may, for example, include the cost of purchasing and maintaining laboratory equipment.
- Other operational costs: This section covers all other costs not accounted for, such as sample transport.

Structured interviews with delivery stakeholders may be able to provide cost data (see section 5.2 for more information on stakeholders). Interviewing delivery stakeholders related to different parts of the system will capture a broader range of cost data. If further information is needed, for example, for the unit costs of consumables, an average price per item can be sought from internal cost sheets in testing laboratories, if available, or from the price lists of online retailers.

Example - *E. multilocularis* surveillance in Great Britain

Background:
The egg flotation test is the method routinely used at APHA for the detection of *E. multilocularis* eggs in fox faeces. Calculating the cost of the existing surveillance testing method can provide a baseline value to compare against either financial targets, or alternate testing methods during the evaluation of the system (section 7).

Process:
The standard operating procedures (SOP) for the egg flotation method was used to create a list of required consumables, reagents and equipment. The costs were then calculated using information available from supplier websites or, where data were unavailable, are hypothetical values. Hypothetical cost rates were also used for staff time.

Results:

Table 4: Hypothetical data showing the cost breakdown per test of the egg flotation test, and the data sources associated with these costs.

Parameter	Unit	
Test		Egg flotation

Species sampled	-	Fox
Test sensitivity	-	0.78
Test specificity	-	1
Consumables and reagents	Per test	€56.88
Staff time (testing)	Per test	€9.26
Operational costs (excluding testing)	Annual cost	€269,162.88
Equipment	Annual cost	€894.15
Tests required at 1% prevalence	No. of tests	383
Cost of testing at 1% prevalence	€	€270,123.17

5.3.6. Data reporting

What: This section documents what data will be reported, how it is presented, and who it will be reported to.

Why: Data reporting ensures the system is reporting the right information to the right stakeholders. It should summarise the objectives, the activities conducted in the surveillance system, and the results to give full transparency of all important aspects of the system.

How: The kind of data to report will depend on the specifics of the surveillance program. However, a system can broadly expect to report:

- The timeframe of reporting (for example, whether annual or monthly)
- The sampling strategy
- Testing method used, with sensitivity data
- Target population, potentially with justification for the choice
- Number of samples taken in the sampling period
- Methodology for results analysis
- Results of the testing (i.e. number of positive/negative samples detected from the total population sampled)

Commonly, these data are provided in scientific reports to the governance stakeholders. These reports are often then further compiled into annual reports produced by governing bodies which can be made available to the wider public. An example of surveillance reporting guidelines in animal health can be seen in the AHSURED guidelines (Comin et al., 2018).

Example - *E. multilocularis* surveillance in Great Britain (GB)

Background:

The full data reporting for GB can be found in the annual reports produced by EFSA prior to 2021 (European Food Safety Authority, 2021b). More recent surveillance results are not yet published, but in future will be presented by DEFRA rather than EFSA. In the absence of these, the submission and reporting of past data to/by EFSA shall be explored in this example.

Process:
We gathered information about the data reported to EFSA initially via a review of published EFSA reports on annual *E. multilocularis* surveillance (European Food Safety Authority, 2021b). From there we engaged with stakeholders at APHA to discuss the data provided to EFSA from their perspective.

Results:

From the 2019/2020 sampling year, GB reported results for 464 samples taken between March 2019 and January 2020 (European Food Safety Authority, 2021a) (Figure 9) with the sample distribution shown in **Error! Reference source not found.** and an estimated red fox population density across the country in **Error! Reference source not found.** (Croft et al., 2017).

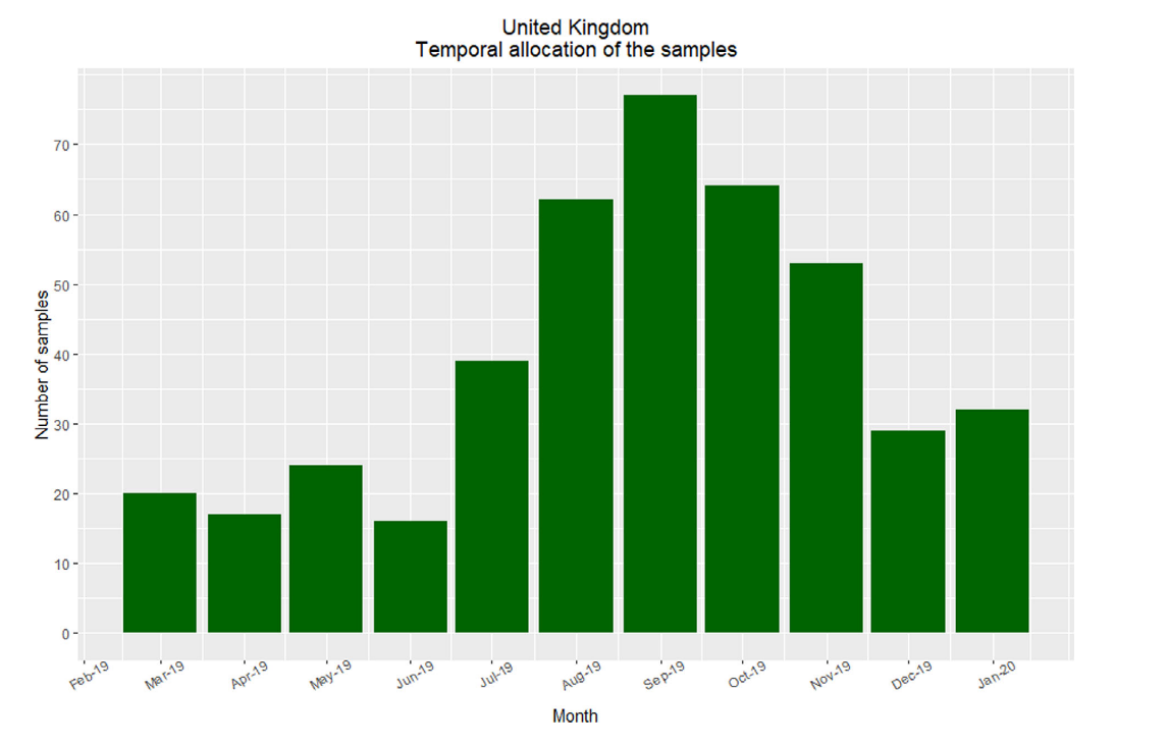


Figure 9, Temporal distribution of red fox faecal samples taken in UK in the sampling year 2019/2020

These samples were tested using an egg flotation test with Cest1-Cest2 NAD1 PCR (European Food Safety Authority, 2021a) with an overview of the methodology provided in the report (European Food Safety Authority, 2021b).

The sampling strategy used is random sampling, with the sample size calculated by the RIBESS tool (European Food Safety Authority, 2012) and based on the test sensitivity and estimated population size for detection at 1% prevalence with a 95% confidence interval.

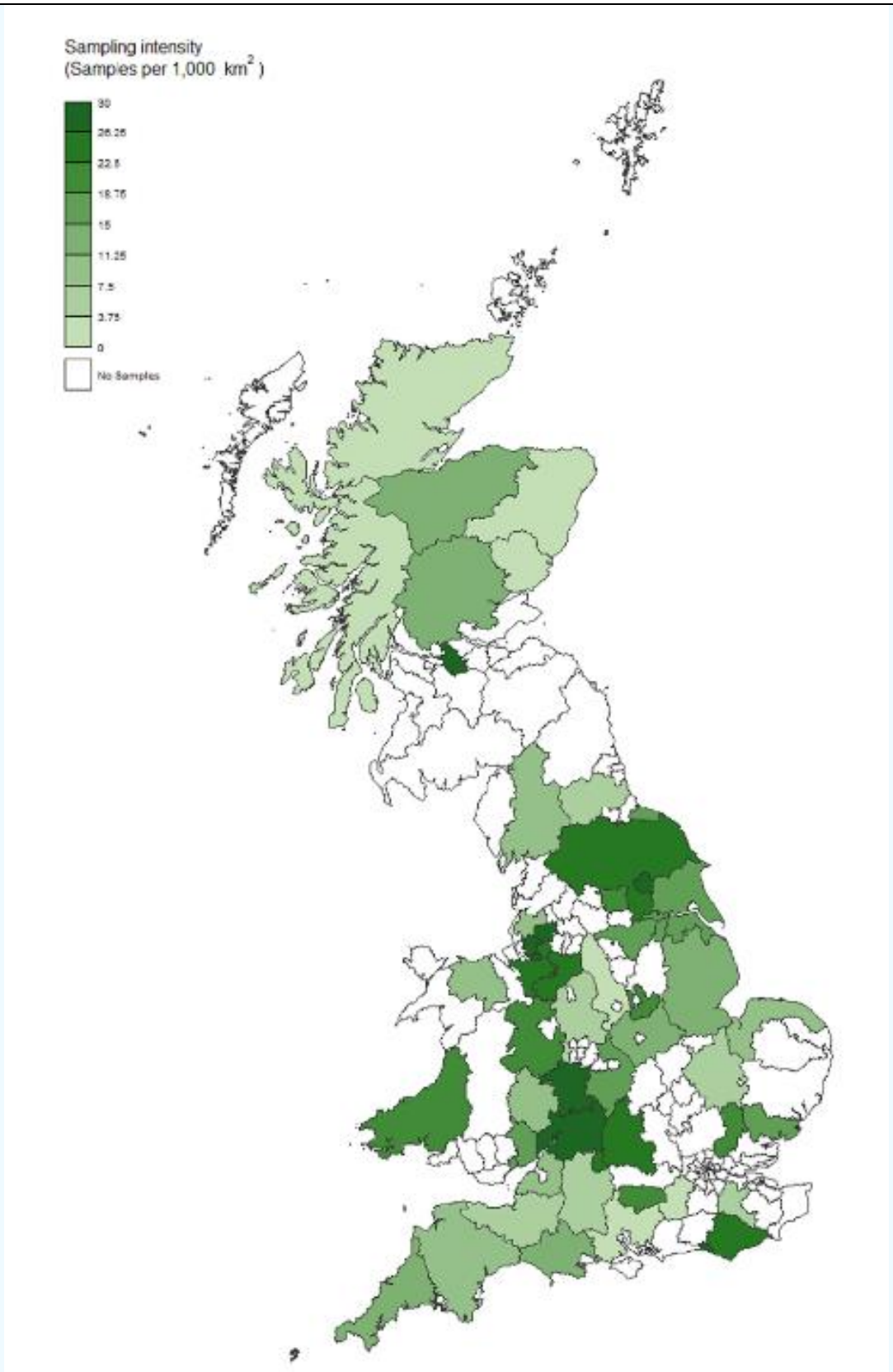


Figure 10, Sampling intensity (samples per 1,000km^2) from sampling period 2019/2020

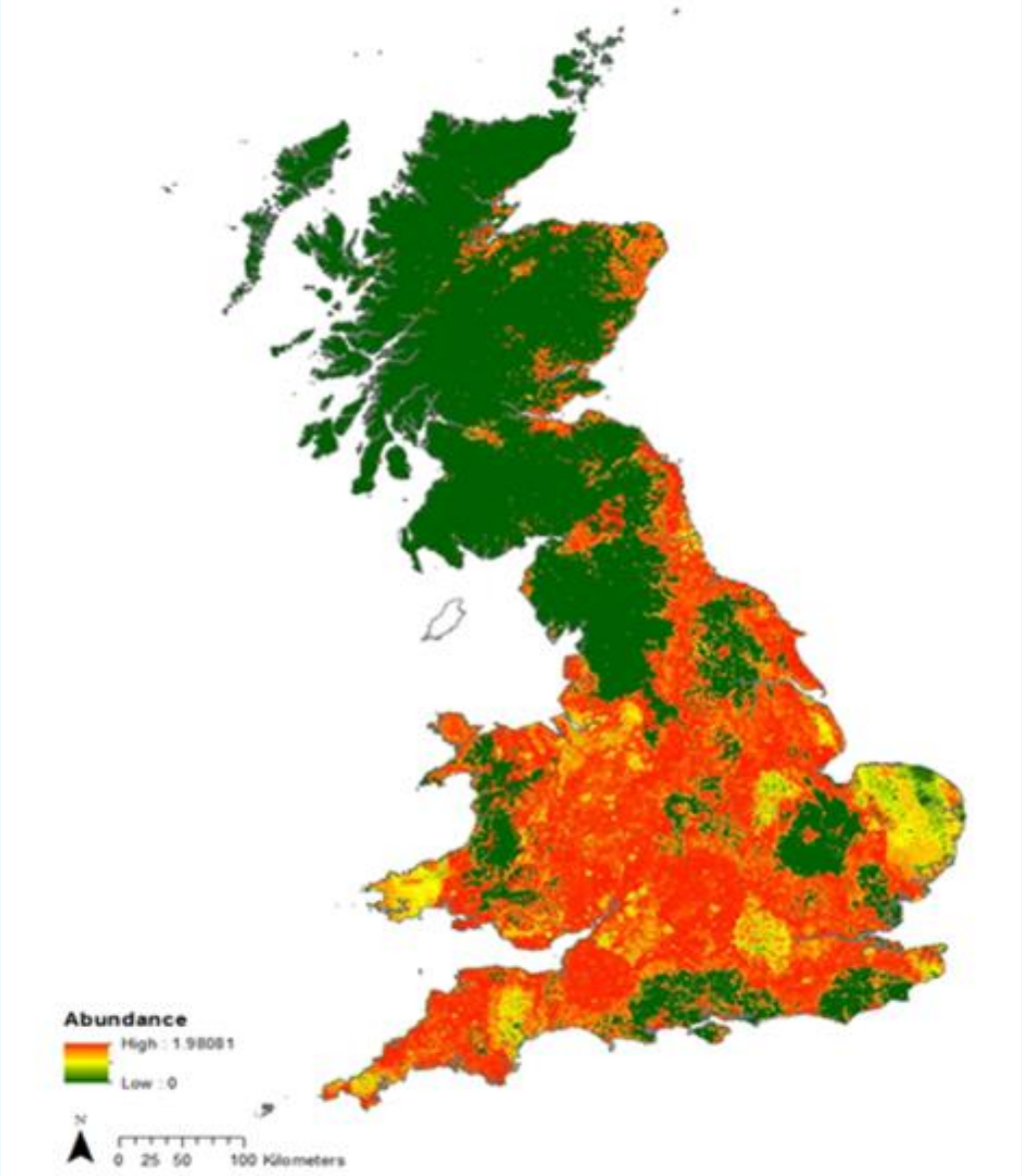


Figure 11, Estimate of fox abundance in GB using NBN presence data and published density data (Croft et al., 2017)

From there, EFSA would evaluate the information and data provided and summarise whether it fulfilled the legal requirements of the legislation and proves freedom from disease.

6. Implementation of an output-based standards surveillance system

There are several important exercises that can be done to aid system implementation. First, it is important to outline how you hope the system will function in an easily communicable way. This will build a common understanding amongst the stakeholders of your design proposal. Stakeholders can then provide feedback on your proposed system and suggest improvements to make it more practically or economically viable. Once the proposed system has been agreed, a strategy can be devised for maintaining the continued quality of the system through test validation and accreditation.

6.1. System mapping

What: System mapping provides a flow diagram showing all processes from the point of sample collection to the reporting of results.

Why: Visualising the entire system holistically in this way is good for documenting the chronology of the surveillance system. It makes the function of the system easily communicable and forms an essential first step for any system evaluation.

How: The simplest method for system mapping is constructing a flow diagram with direct input from your stakeholders (COHESIVE, 2022). This should describe the chronological steps from sample acquisition to results analysis. Most of the system structure will already have been determined in the design process. However, any remaining aspects of the system that are unclear should be highlighted in this flow diagram and clarified by the stakeholders. The map should outline which stakeholders will be involved at each stage in the process.

The system structure map can also be used to represent any synergistic systems linked to the surveillance. For example, if the same samples could be used for other purposes. This helps document the interfaces of the surveillance system with other useful activities and highlights opportunities to make sampling more mutually practical and beneficial.

Example - *Echinococcus multilocularis* surveillance in Great Britain
Background: The surveillance system for *E. multilocularis* has multiple stakeholders each contributing to, and benefitting from, various stages. In order to better understand the flow of information through the system, as well as the system chronology and potential areas for improvement, it is important to visualise the system holistically. **Process**: Contact with the APHA parasitology team lead and APHA wildlife team enabled the production and validation of a system structure map. This shows the sequence of events from carcass collection to results reporting. **Results**:

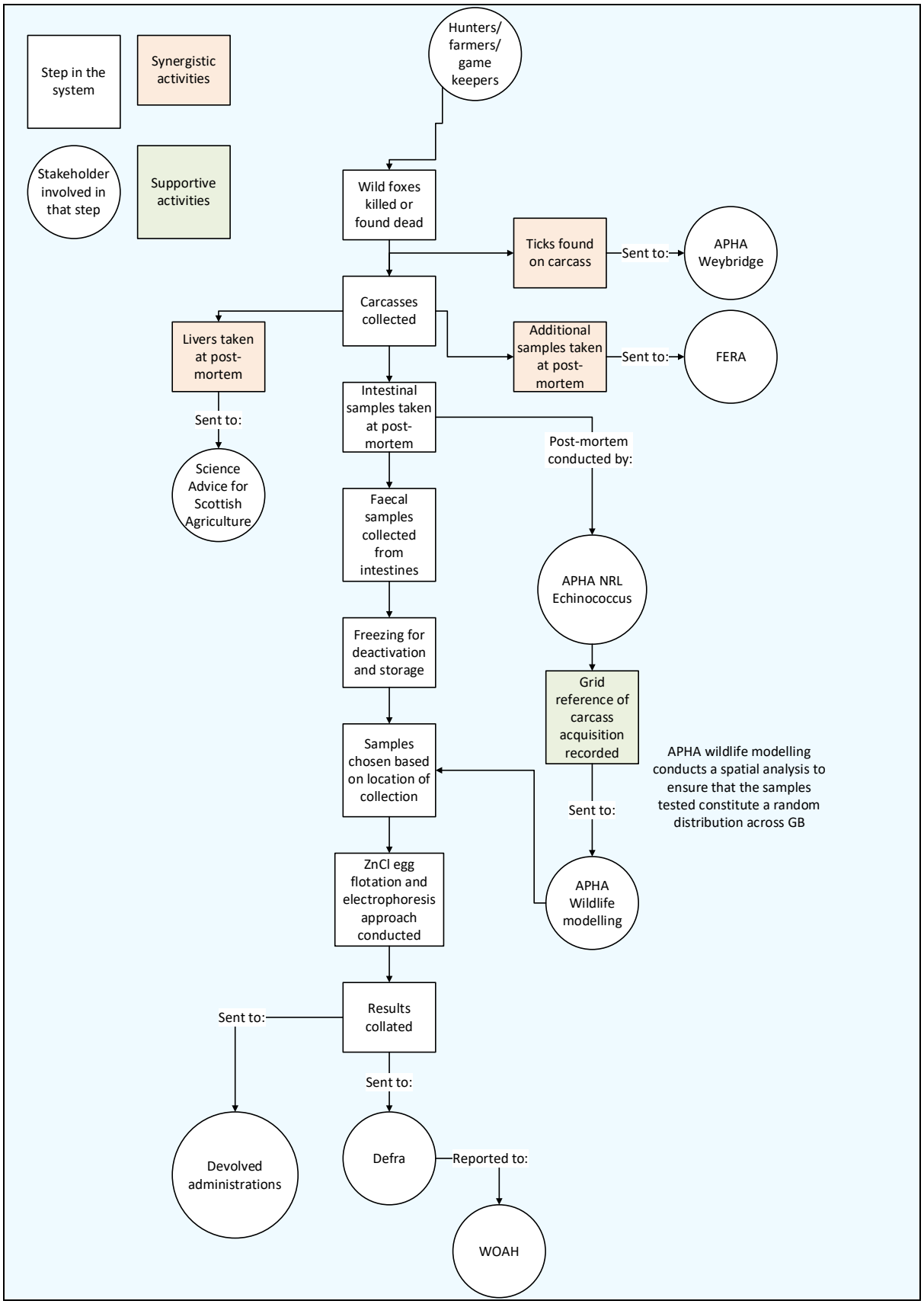

Step in the system
Synergistic activities
Stakeholder involved in that step
Supportive activities
Hunters/ farmers/ game keepers
Wild foxes killed or found dead
Ticks found on carcass
Sent to:
APHA Weybridge
Carcasses collected
Livers taken at post-mortem
Sent to:
Science Advice for Scottish Agriculture
Additional samples taken at post-mortem
Sent to:
FERA
Intestinal samples taken at post-mortem
Post-mortem conducted by:
Faecal samples collected from intestines
APHA NRL Echinococcus
Freezing for deactivation and storage
Grid reference of carcass acquisition recorded
Samples chosen based on location of collection
Sent to:
APHA wildlife modelling conducts a spatial analysis to ensure that the samples tested constitute a random distribution across GB
ZnCl egg flotation and electrophoresis approach conducted
APHA Wildlife modelling
Results collated
Sent to:
Sent to:
Devolved administrations
Defra
Reported to:
WOAH

Figure 12, showing the system structure and chronology from carcass collection to result reporting. Rectangles represent steps in the system while circles represent stakeholders involved in relevant steps (ZnCl=zinc chloride).

6.2. Project management planning

What: Effective project management is required to implement your proposed surveillance design in practice.

Why: Having a thorough and well-planned strategy for implementing surveillance is essential for success in the desired timescale. Without a well thought out and documented plan, the approach to implementation can become uncoordinated and inefficient, which can ultimately lead to delays in implementation or even to a resultant system that is not fit for purpose.

How: There are a wide array of project management tools available, and each organisation likely has a preferred method. Project management for the implementation of large complex systems requires a full guidance document on its own. However, there are several best practices and rules of thumb, which can be taken from other scientific fields.

For this, we can draw inspiration from systems engineering practices. Systems engineering is centered around the delivery of complex engineering projects and has a good track record of use across a range of science and technology-focused projects (Leal, 2020, Emes* et al., 2012). Three core concepts of systems engineering which are directly applicable to the implementation of output-based surveillance systems are: project left-shift, continuous integration, and detailed documentation.

Project left-shift (Figure 13) focusses on shifting the project resource, funding, and input to the start of a project rather than the end of it (Emes et al., 2014). The concept derives from the idea that while the budget across a project is usually fixed, the value of that budget steadily drops due to inflation as the months and years of its duration pass by. This means that early investments in a project are worth more than later ones. Left shift is also linked to the idea of path dependency. Making a bad decision early on in a project and later working to resolve the problems that the decision caused is wasteful, stressful, and often unnecessary. By investing more time and effort into early planning and pre-planning stages of your implementation, you make significant savings in the long term and will be more likely to finish a project by the deadline.

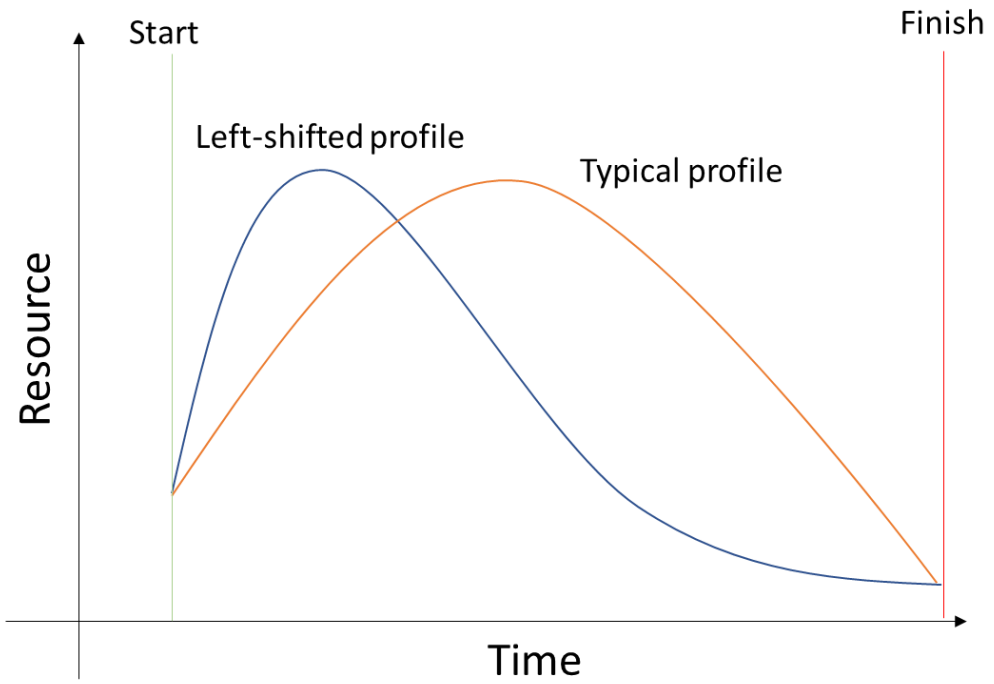


Figure 13, showing a left-shifted project profile compared to a typical project profile. In a left-shifted project, far more resource is invested early in the project, allowing for efficiency savings later.

Continuous integration is the commitment to appropriate levels of validation and verification with the stakeholders of the system throughout the implementation process. The project is first partitioned, with its core components identified and designed and then implemented. As each component of the system is implemented, they must integrate with one another, and must perform the same function in practice as intended in their design. An integrated project can be represented by a V-model, with the partitioning stages of the project represented on the left side of the V, and integration on the right (Weilkiens, 2011). In the context of implementing an output-based surveillance system, a V-model could resemble Figure 14. In this, each stage of integration is validated and verified against the original design, first of that specific component, and then of the system as a whole. Doing so puts a check and balance on the overall output at each stage of implementation, making sure it delivers on the original design proposal.

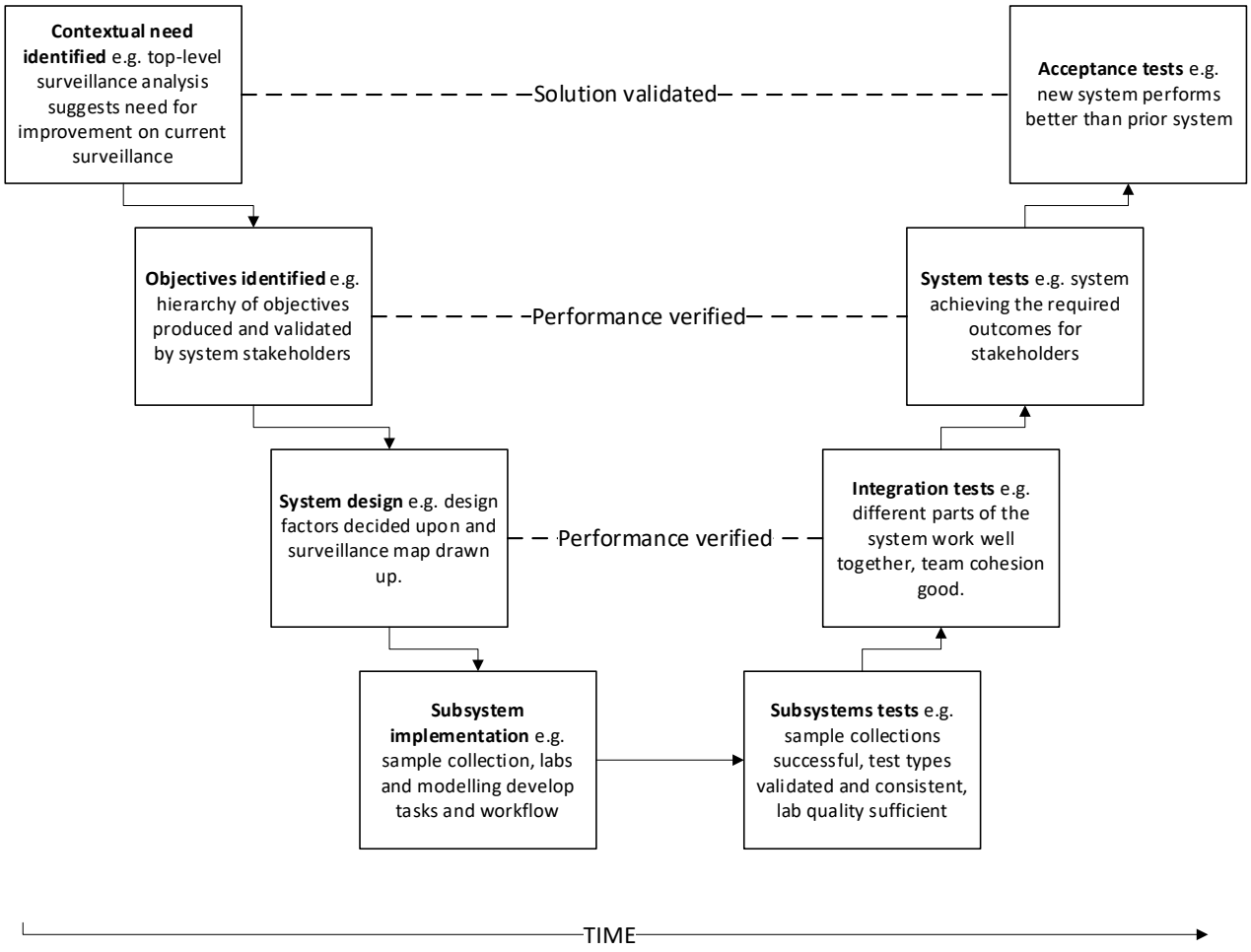


Figure 14, showing an example V-model for the implementation of output-based surveillance. Adapted from Weilkins, 2011.

Project documentation is key to maintaining a common purpose and shared vision across your implementation team. There are three core documents that should be maintained in any project. First, the project risk register, discussed in more detail in section 6.3. Secondly, the project management plan. This includes all the documents produced during the design process and any scheduling plans you have made prior to implementation. For example, your V-model (see Figure 14) and/or Gantt charts (see Wilson (2003)) which outline the list of project tasks and a visual timeline for when these tasks are expected to be completed. These make up your baseline: essentially the intended structure of your surveillance system, and your way of getting there (Lester, 2006). Any changes to this baseline must undergo a formal change management process and should be communicated with all the relevant stakeholders. Part of change management is documenting the intended changes and the impacts of these changes on the project in a change log (Davison, 2016). This log is then integrated into the project management plan. Finally, a work breakdown structure should be created. This contains a long-term, medium-term and short-term breakdown of what each individual stakeholder (where applicable) and team member will be doing over the course of a project to implement the surveillance system (Devi and Reddy, 2012).

More detailed guidance should be sought before implementing a complex output-based surveillance system. However, keeping these three concepts in mind will help minimise ambiguity throughout the implementation of your new system.

6.3. Assessing implementation risks

What: This step assesses the organisational and practical issues that may emerge between designing and implementing a surveillance system.

Why: Even if a system *looks* able to achieve its objectives, it may still lack essential practical details. Assessing the practicality of the system minimises the chance of unforeseen difficulties when implementing it.

How: Assessing practicality requires operational risk analysis. Operational risks, as opposed to disease risks, are uncertain events which may have a positive or negative impact on project objectives (Raanan and Kenett, 2011). The operational risk analysis process consists of risk identification, risk assessment, and risk management.

Risk identification can be done using interviews and workshops with the delivery stakeholders. The stakeholders should understand all the details of your proposed system design. Ideally, if workshops are done in groups, members of these groups should be operating in a similar part of the surveillance system. With these individuals and groups, encourage them to list all potential events that could impact the success of the current system and document these in a risk register (Table 5). The risk register should include a unique risk number, either single number system or letter-number system to denote the stage the risk may occur. For example, S1 for the first risks in the sampling stage, T1 for the first risk in the testing stage, and A1 for the first risk in the analysis stage.

Table 5: Showing an example risk register. Each potential risk should be documented with input from stakeholders to determine their probability of occurring and their impact. The actions taken in response to these risks should be decided upon and approved by the stakeholders.

Risk no.	Risk description	Probability of risk occurring	Impact of risk	Actions taken	Comments
S1	E.g. Collectors don't sample from wide enough area	E.g. Low/Medium/High	E.g. Low/Medium/High	E.g. Avoid/Accept	E.g. steps taken to avoid risk or reason for accepting risk
S2					
T1					
T2					
A1					

Risk analysis assesses the probability of each of these risks occurring and the impact if these risks occur. Probability and impact can be measured as either Low, Medium, or High. Any risk above low in either category warrants consideration. Any risk that is above low in both categories demands actions be taken (Figure 15).

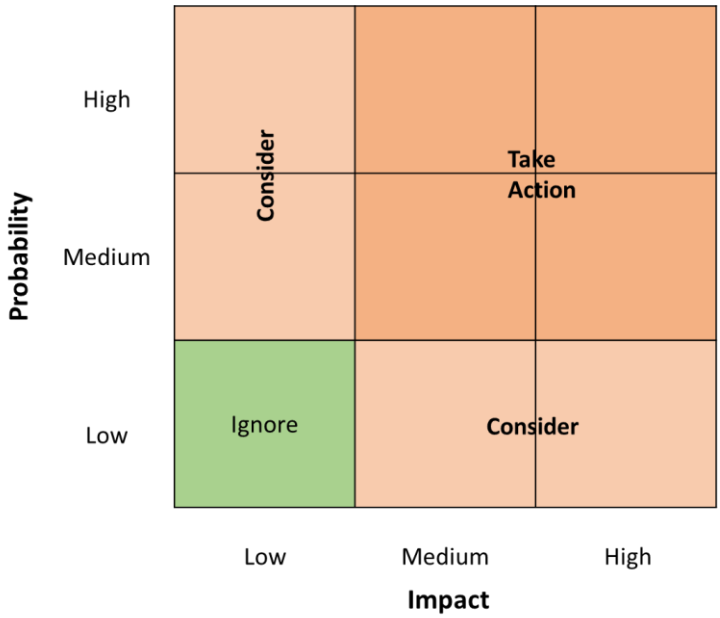


Figure 15, showing the matrix of operational risks. Any risks considered of above low probability or consequence should be considered. Action should be taken for any risks deemed of above low risk in both categories. Adapted from Ramler and Felderer (2016).

Risk management in this context is the action taken to reduce the probability or impact of these risks. Within output-based surveillance systems, the options available are avoidance or acceptance. Avoidance of a risk demands that the system is adapted to prevent a risk from occurring or to reduce the probability of a risk occurring. This will likely require aspects of the system design to be reconsidered. All changes made to avoid a risk must be updated in the prior design and implementation stages where relevant and recorded in a change log within your project management plan (see section 6.2). Acceptance is usually decided upon if no reasonable avoidance mechanisms can be implemented. An accepted risk will not require further action.

Stakeholders must agree with the outcomes of risk analysis. Specifically, they must agree to any changes to the system design made in response to the risk assessment, and any accepted risks. Once all stakeholders are satisfied, the final design can be approved.

6.4. Test validation and test assurance

What: Gather documentation relating to validation, assurance, or audit of the test/process in question.

Why: With a surveillance system approved for implementation, consideration should be made over its long-term maintenance and viability. Finalising a methodology for regular test validation provides assurance to all stakeholder groups that the system will continue to deliver on its objectives after it is implemented for both the delivery stakeholders carrying out the testing and those receiving the results.

How: Generally, this information will be available from delivery stakeholders carrying out the testing. But depending on the organisation, this could also be provided by the relevant quality assurance department or test validation team. The type of surveillance will impact what kind of validation, assurance or audits are relevant. However, in general, questions to ask relating to validation and assurance can include:

- Is the testing process validated?
- Is this internal validation accredited by an external body e.g., United Kingdom Accreditation Service (UKAS)?

- How is test performance measured and how often does this occur. For example, is the testing process and SOP accredited, and does this involve frequent quality standard checks and an annual test report?
- Are ring trials conducted e.g. by European reference labs; WOAH lab

7. Evaluation of an output-based standards surveillance system

This section provides a range of evaluation exercises for existing OBS schemes. These exercises are designed to inform decision-making on potential improvements to the system. Within surveillance objectives, the historical trends and background section aims to document the political and disease context that the surveillance system exists in, important when evaluating or re-evaluating the objectives of the system, while the flexibility analysis determines its resilience to changing political landscapes and disease profiles. The stakeholders section assesses the human aspects of a surveillance system; engaging with those involved in it. Assessments of result accuracy and financial viability are explored in the surveillance parameters section.

7.1. System objectives

7.1.1. Historical trends and background

What: The first step to evaluation is to understand the context of the existing surveillance, including trends in the number of samples collected, the disease prevalence detected, and the dynamics of the disease in question. These will inform whether the original objectives of the system are still relevant to the current disease situation.

Why: Understanding historical trends can be used to inform the success of the surveillance streams relative to other surveillance options. For example, if a new a test has been tried somewhere else in the world, how does it compare with the test currently implemented in your context?

How: The most efficient method of understanding the historical trends in a surveillance system is through a combination of literature review and stakeholder engagement. Through either of these approaches, seek to answer the following questions:

- Has the level of detection changed since the first implementation of the surveillance system? Has prevalence of the pathogen been demonstrated to have increased/decreased or changed in its geographical distribution?
- Has new evidence come to light on the dynamics of the pathogen under surveillance? For example, have new competent hosts been found?
- Have new tests been developed for the same pathogen and host as the original surveillance system? And do these new tests promise improved sensitivity and/or specificity than the current implemented option; do they offer other advantages such as cost efficiency?
- Have any aspects of the surveillance system been recognised to be operating particularly well? For example, have other groups taken inspiration from the current system and implemented the same methods elsewhere?
- Are the surveillance streams which make up the surveillance system still relevant with regard to the likelihood of detecting the pathogen if present?
- Have any issues or doubts about aspects of the surveillance system been raised? Are any of these corroborated by data?
- Has the political or legislative context of surveillance changed? Has the target pathogen or population become higher or lower priority to governing bodies? Is the need for surveillance brought in to question by these changes?

7.1.2. Robustness and flexibility analysis

NOTE: It is advised to create a system structure map for your system as in section 6.1 before conducting this evaluation

What: This evaluation aims to determine how changes to the system would affect the various stakeholders and affect the system's ability to deliver on its core objectives. It includes consideration of impacts such as the perceived engagement of stakeholders, the financial or research synergies that could be built, or may be lost by changes to the system, and the practical stresses that may be put on the system by potential changes. Flexibility is defined by the capacity of a system to change structurally, while maintaining robustness: its ability to withstand changes without compromising core functions – the delivery of its objectives with confidence in its results. This evaluation does not necessarily need to be done with particular changes in mind, but rather as an exercise to identify the surveillance system's flexibility to change.

Why: No surveillance system is built to last forever. They are designed for a specific purpose in response to the situation at the time of their development. It is therefore expected that, even if no alternative surveillance systems exist now to replace it, every system will undergo changes at some point to match the situation of the day. A good output-based surveillance system therefore needs to be adaptive to technological, practical, or political changes that influence them so that they can continue delivering value for their stakeholders. Compared to traditional input-based surveillance systems, OBS are generally more flexible, so taking an OBS approach can make it easier to build this adaptability into your system.

How: Determining the flexibility of the system requires systems thinking. Following prior analyses, the surveillance system components should be mapped to show their interrelationships (Section 6.1). In addition, the roles of each stakeholder in the system should have been determined (Section 5.2.1). The goal of this evaluation is to identify how each system component influences each other system component, and how these in turn positively or negatively influence the stakeholders.

We recommend using causal loop diagrams to illustrate these links. To produce these diagrams, the first step is to identify which system components directly affect each stakeholder. For example, carcass collection stakeholders will be directly impacted by the sampling requirements of the test chosen, as they may need to collect more carcasses to fulfil this need. Hence, there is a link between the test chosen, the sampling requirements, the carcasses required for collection, and therefore the carcass collection stakeholders.

Similarly, the confidence level in the probability of freedom is inherently influenced by the number of samples tested. This is influenced by the sensitivity of the test, which is influenced by the type of test used. These links should be traced back and drawn to show these causal relationships. When demonstrating these links, it is essential to show whether the relationship is positive or negative. For example, higher test sensitivity has a negative effect on the number of tests required: a higher sensitivity results in fewer samples taken. The number of tests required positively influences the number of samples taken: more tests required means more samples will be needed to be tested.

While tracing back, it may become clear how other stakeholders might be affected. More tests might increase sampling costs, for example, which might affect delivery stakeholders working on a tight budget. The causal loop diagram helps to identify how each stakeholder might be affected by a change to one or other of the system components. From this, it is possible to predict stakeholder advocacy for given changes.

Finally, engage stakeholders to determine their tolerance to change. This is how you ultimately determine the flexibility of the system. If stakeholders operate under fixed constraints these should be identified and documented. For example, delivery stakeholders may be working within a budgetary range. If they can agree to an increase in sampling rate, what is their cut-off sample number? Governance stakeholders may have some tolerance in the design prevalence or testing confidence they expect to see from a surveillance system. What is this tolerance and to what extent could the system adapt before those tolerances are exceeded?

Example - *E. multilocularis* surveillance in Great Britain (GB)
Background: Although the system for *E. multilocularis* surveillance has been in place for a number of years, it is important to ensure that the system can be flexible in response to changes. **Process**: Through the system map produced in section 6.1 and engagement with stakeholder at various points of the design and evaluation process, we were able to identify key system nodes as well as system components associated with each stakeholder. **Results:** In determining the flexibility of the GB system, we first identified the key parts of the system that could be subject to change and addressed which stakeholders would be affected by those changes. The results of this exercise are shown in **Error! Reference source not found.**

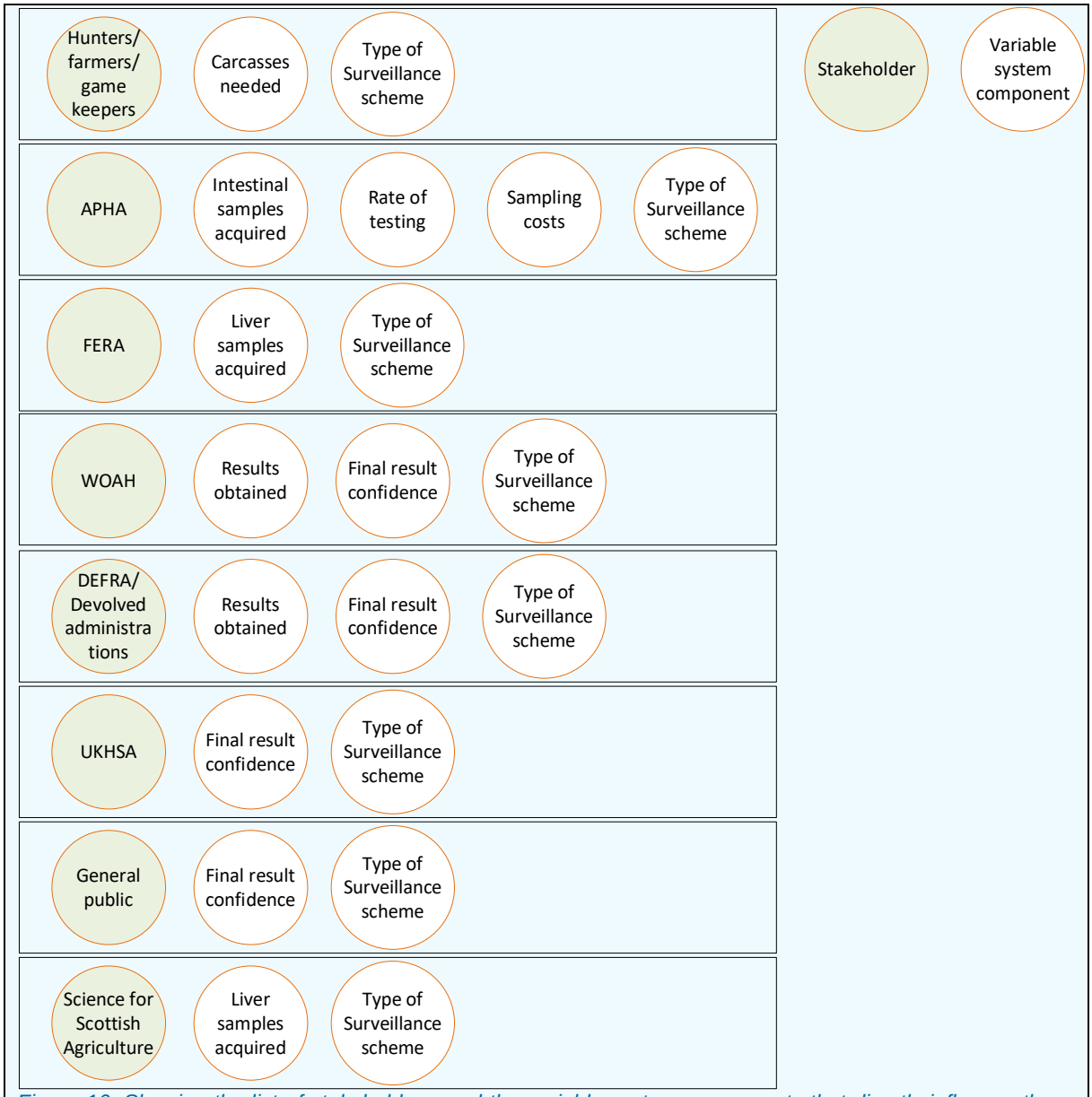


Figure 16, Showing the list of stakeholders, and the variable system components that directly influence those stakeholders.

From the results of this initial exercise, and from the process diagram generated in section 6.1, we were able to create a causal loop diagram showing the interrelationship between the system components and stakeholders. This was designed to show the positive or negative influences changes each of the system components would have both on other system components and on the stakeholders.

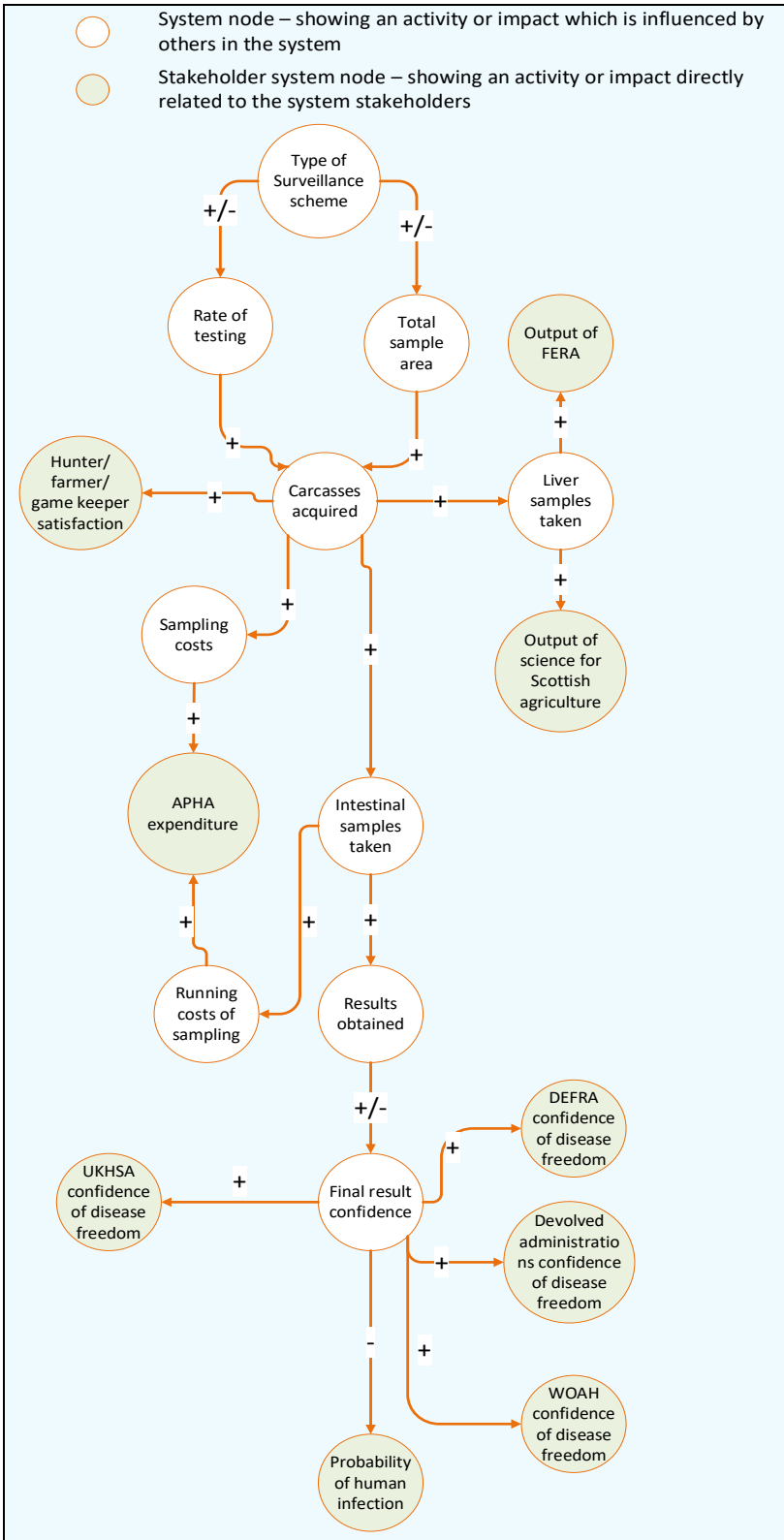


Figure 16, Example causal loop diagram illustrating the causal relationships between stakeholder and system variables in the GB E. multilocularis surveillance system

By representing all aspects of the system in this way, it was possible to isolate five key interrelationships between the system nodes:

1. The chosen surveillance scheme will affect how many carcasses are collected, and where they are collected from (for example, if collected according to risk-based sampling rather than random sampling). This has ripple effects on every other part of the system.

2. A higher sample requirement would mean more time and money spent collecting those samples. It would also demand more from farmers, hunters and gamekeepers to provide carcasses for analysis. This could strengthen or damage relationships with these stakeholders, depending on their appetite for collaboration, and thereby increase or decrease their satisfaction with the system and their willingness to supply samples (APHA carcass collection coordinator, 2022).

3. More carcasses collected means more of all sample types are available for commercial collaborators. Fera Science and SASA who use livers and other parts of the foxes for research purposes may glean increased research opportunities from a higher sampling rate.

4. A higher sampling rate, or improvement in the geographical spread of collected samples will increase the overall confidence in the surveillance system. It will increase the probability that cases in wildlife will be detected before the disease becomes established in the wild population. This will reduce the number of human cases, and therefore provide a higher benefit to society at large.

5. A change in the costs of maintaining the system, and the downstream effects on the benefit to stakeholders, will affect the benefit-cost ratio of the surveillance system. A higher benefit-cost ratio means the surveillance system generates greater value for money.

7.2. Stakeholders

7.2.1. Stakeholder engagement

What: This evaluation is concerned with understanding the stakeholders and their current engagement in the system. In this case, evaluation focusses on determining and depicting the level of interest and influence current stakeholders have in the system.

Why: Stakeholders have diverse views and roles. Thus, to understand them, it is a useful exercise to categorise them. Doing so can, for example, identify the most influential stakeholders in the system, or those who hold the largest stake in it achieving its objectives. You can then determine whether their position in the system is still appropriate.

How: A modified Mendelow matrix is an effective way to categorise stakeholders. This is a two-dimensional matrix plotting the interest stakeholders have in the system on the x axis and the influence they have on it on the y axis (**Error! Reference source not found.**(Mendelow, 1981)). It provides information about which stakeholders are the most engaged, and which are most influential.

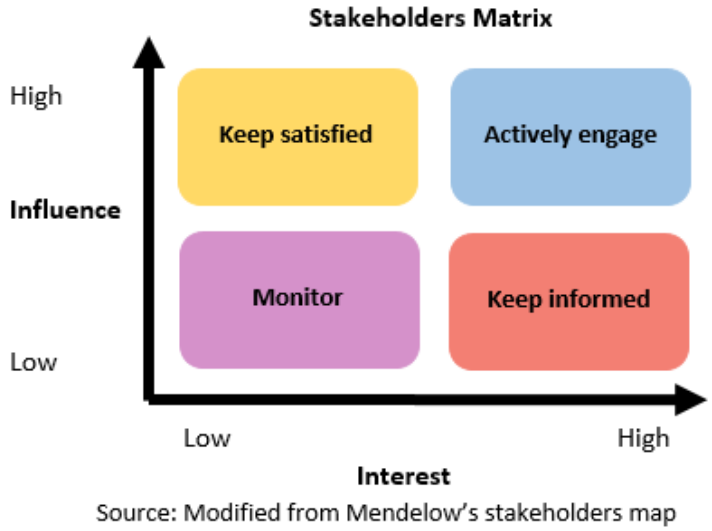


Figure 17, Mendelow's matrix, showing stakeholder interest and influence (COHESIVE, 2022)

Structured interviews should be used to determine the level of influence and interest in the system. Direct questions are a good starting point, for example 'what is your perceived level of influence on the system?'. However, stakeholders might find these difficult to answer descriptively without a frame of reference. It can be useful to follow up these questions with more descriptive questioning.

A question which asks the stakeholders how they might implement change to a system could return more tangible insights into the barriers stakeholders face when trying to implement change. A stakeholder with high influence will likely have a strong idea of how to enact change to the system. They may have been directly involved in making prior changes to the system. Stakeholders with lower influence may not know how they would change a system or may reference other stakeholders in describing how they would do so.

The level of interest in the system is concerned with how stakeholders would be affected by changes to the system. When ascertaining the interest of stakeholders, questions that explore hypothetical scenarios may yield richer results. Asking, for example, how a stakeholder might be affected by increasing or decreasing the sample numbers taken, or by changing the objectives of the system. If their answers indicate they would need to take immediate action as a result of these changes, this illustrates a high level of interest in the system. The challenge comes for beneficiaries of output-based surveillance systems, such as the general public, who may not be aware of the implications of changes to it on their own health and wellbeing. A judgment can be made in these cases based on the prior information compiled.

Another tool could be survey-based questions rating interest and influence on a quantitative scale, for example from 1-10. It is important to accompany any survey questions with clear definitions of each attribute.

With interviews and surveys, every effort should be made to contact as many stakeholders as possible from across the system. Where this is not possible, a proxy can be used to show the influence and interest these stakeholders have. The profile of these missing stakeholders can be built by asking other stakeholders about them. Some will have worked closely with those stakeholders or may know how they operate by virtue of working in the same system. If you take this approach, it is important to get input about missing stakeholders from as many other stakeholders as possible.

Once the bulk of information has been compiled on each stakeholder, they can be placed on the Mendelow matrix. A completed matrix of all stakeholders should then be verified by the stakeholders themselves.

Finally, you should evaluate whether the position of the stakeholders on the matrix is still appropriate, particularly with regard to the influence they have on the system. This can be done by asking stakeholders whether they think they should have more or less influence on the system. Stakeholders can be represented on the Mendelow matrix with arrows to denote whether they think they need more or less influence on the system. In addition to showing the effectiveness and appropriateness of stakeholder engagement in the OBS, it also provides an indication of stakeholder satisfaction.

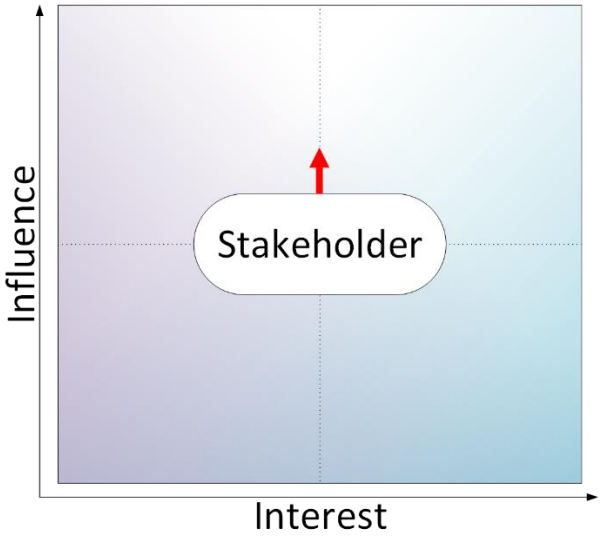


Figure 18, showing a hypothetical Mendelow matrix, with a stakeholder placed between two continuous scales of interest and influence. A stakeholder wanting more influence than they currently have on the system is represented with a red arrow.

Example – *E. multilocularis* surveillance in Great Britain (GB)

Background:
In section 5.2 we identified the stakeholders in the current system. For these stakeholders, we sought to map individual influence and interest to better understand both their role in the system, gauge the effectiveness and appropriateness of current engagement practices, and assess stakeholder satisfaction.

Process:
We reached out to as many stakeholders as possible, either via email or through arranging interviews. From the information we gained through these communications, we then mapped each stakeholder onto a modified Mendelow matrix and sought feedback on this from the stakeholders involved.

Results:
Interviewing key stakeholders
We organised hour-long interviews with each of the following stakeholders:

- APHA Parasitology discipline lead and laboratory coordinator for *E. multilocularis* surveillance in GB.
- Carcass collection coordinator for *E. multilocularis* surveillance in GB.
- APHA discipline lead for wildlife epidemiology and modelling, leading *E. multilocularis* sample selection, and risk modelling.
- SASA research coordinator rodenticide sampling in wildlife

- Fera Science research coordinator rodenticide sampling in wildlife

In addition, we engaged the following stakeholders via email:

- UKHSA Emerging Infectious Zoonoses Team
- DEFRA

We attempted to engage, but could not reach:

- WOAH

We discussed the following topics with each stakeholder:

- The role of the stakeholder within the system
- The perceived roles of other stakeholders in the system
- Their perceived understanding of how the surveillance system practically functioned to deliver outputs
- Their perceived influence on the system
- Their satisfaction with the system, particularly with regards to the level of influence they had on it.

For stakeholders that could not be contacted directly, attributes were estimated from the expert knowledge of the other stakeholders; from their past interactions with these stakeholders and their experience working within the system.

Mapping stakeholders
With the information compiled in the interviews, it was possible to map each stakeholder on a Mendelow matrix.

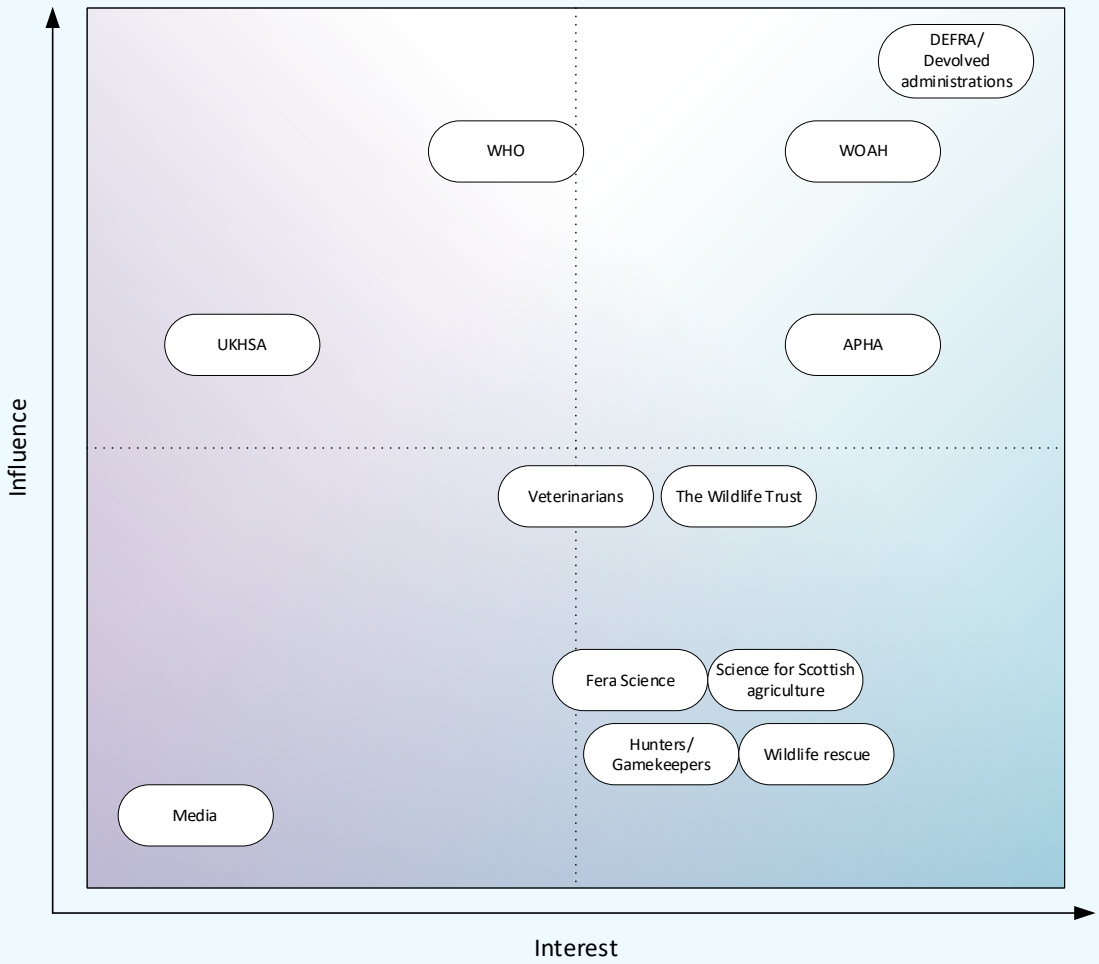


Figure 19, Stakeholders involved in GB E. multilocularis surveillance mapped to a Mendelow matrix, sorted by level of influence and interest in the surveillance system.

In the future, DEFRA will receive the annual reports of the surveillance, therefore, they have both high interest and high influence on the matrix. APHA and WOAH are also in this quarter of the matrix as they are responsible for carrying out the surveillance, producing the annual reports to prove disease freedom.

With the current GB situation for *E. multilocularis,* the UKHSA is in the low interest, high influence quarter of the Matrix. However, this would likely change to high interest, high influence, if there were changes to the status of *E. multilocularis* in GB.

When asked, satisfaction was very high: no stakeholder felt they needed more or less influence on the system.

7.3. Surveillance parameters

7.3.1. Minimum sample size evaluation for prescribed design prevalence

What: This evaluation calculates the minimum sample size to allow for disease detection to a set design prevalence.

Why: The minimum sample size should be calculated before conducting monitoring studies. This calculation is relevant for monitoring the disease in the population. If the sample size is too big it will require additional financial cost. If the sample size is too small, it can lead to incorrect study results.

How: Scientific publications, international and governmental statistical data, hunting associations or other professional organisational data, expert opinions, and grey literature can all provide relevant population size data and information about test sensitivity. Furthermore, the sensitivity of the test can also be determined via validation studies and in the case of a commercial test, via the test manufacturer. This information can then be used to calculate the minimum sample size needed for surveillance using the online EpiTools calculator - "Sample size for demonstration of freedom (detection of disease) in a finite population" (available online at https://epitools.ausvet.com.au/freedomfinitepop) (Sergeant, ESG, 2018).

This tool can calculate the sample size needed to achieve the required probability of detecting disease (herd-sensitivity) at the defined design prevalence for a finite population, assuming diagnostic assay with known sensitivity and 100% specificity. These calculations use an approximation of the hypergeometric distribution (MacDiarmid, 1988, Sergeant, 2018b). According to MacDiarmid (1988) the probability (β) that there are no test-positive animals in the sample tested can be calculated as:

$$\beta = \left(1 - \frac{n\,SE}{N}\right)^{pN}$$

where:

- p = true prevalence of infection
- SE = sensitivity of the test
- N = herd size
- n = sample size

The required parameters (inputs) for the calculator are:

- Population size
- Test sensitivity
- Desired herd-sensitivity
- Design (target) prevalence

The main output of this EpiTools analysis is the number of samples that should be examined to provide the desired herd sensitivity for a specified design prevalence. Some results of such analysis are shown in Table 6, and includes countries where *E. multilocularis* is present, and those where it has not been detected. Calculations concerned *E. multilocularis* in the red fox population in selected European countries. In these calculations, the EpiTools calculator inputs were set as follow:

- red fox population size - defined according to the data from publications and reports (Table 6)
- sensitivity of *E. multilocularis* detection test (sedimentation and counting technique (SCT), intestinal scraping technique (IST) or PCR method)- derived from publications and reports as reported in the table in section 6.4
- desired herd-sensitivity - was set at 0.95
- design (target) prevalence – here was set in accordance to the calculated true prevalence (section 7.3.2).

Furthermore, this EpiTools calculator can generate graphs of the sample sizes needed to achieve the desired herd sensitivity, for a defined test sensitivity and range of population size and design prevalence. See the *E. multilocularis* in Poland and GB example below.

Table 6: Calculation of the number of samples required to detect E. multilocularis in the red fox population in selected European countries covering both countries free from E. multilocularis, and where it is present.

Country	Red fox population								Sample size for demonstrating detection of disease						
	References	2009	2010	2011	2012	2013	2014	2022	2009	2010	2011	2012	2013	2014	2022
Poland	[1]				193402	210332	198679					19	19	19	
Latvia	[2]	35000	34800						9	9					
Denmark	[3]						31100							405	
Hungary	[4]							78000							60
Romania	[5]; [6]		53292							63					
Finland	[7]; [8]						150000							384	
Ireland	[7]; [9]; [10]						150000							339	
Great Britain	[7]; [11]						240000							353	
Norway	[12]; [7]		70000		70000	70000	151000			475				476	

[1] - The Forest Data Bank (2022); [2] -Kirjušina et al. (2015); [3] - Danish Centre For Environment And Energy (2022); [4] - European Health and Digital Executive Agency European Food Safety Authority (2015)(HaDEA); [5] - Şuteu et al. (2014) [6] - Romanian National Institute of Statistics (2008); [7] - Assessment of *E. multilocularis* surveillance reports submitted in 2015 in the context of Commission Regulation (EU) No 1152/2011; [8] - Kauhala (2007); [9] - Hayden and Harrington (2000); [10] - Marnell et al. (2009) [11] - DEFRA and APHA (2019); [12] - Sviland et al. (2014).

Example – Calculation of the minimum sample size required to demonstrate disease freedom/disease detection using the EpiTools calculator: *E. multilocularis* infection of red foxes in Great Britain and Poland

Background:

Calculation of the minimum sample size required to detect a pathogen at a prescribed design prevalence is important to both ensure that sample size remains appropriate for the test being used, particularly in the event of a change or update to the testing procedure, and to minimise the costs of sampling and testing. Using a sample size that is too small for the sensitivity of test being used or population size, can result in an underpowered analysis which does not provide sufficient evidence for disease freedom. Conversely, using a sample size which is too large results in unnecessary additional costs. This is particularly important in the context of the system objectives, as many systems will have a financial or budgetary objective to ensure the testing is financially feasible, and to provide good value for money.

Process:

We searched relevant literature sources and engaged with stakeholders to determine the test sensitivities and population size in order to calculate the minimum sample size required. We then used EpiTools to calculate the minimum sample size to demonstrate pathogen detection at the required design prevalence.

Results:

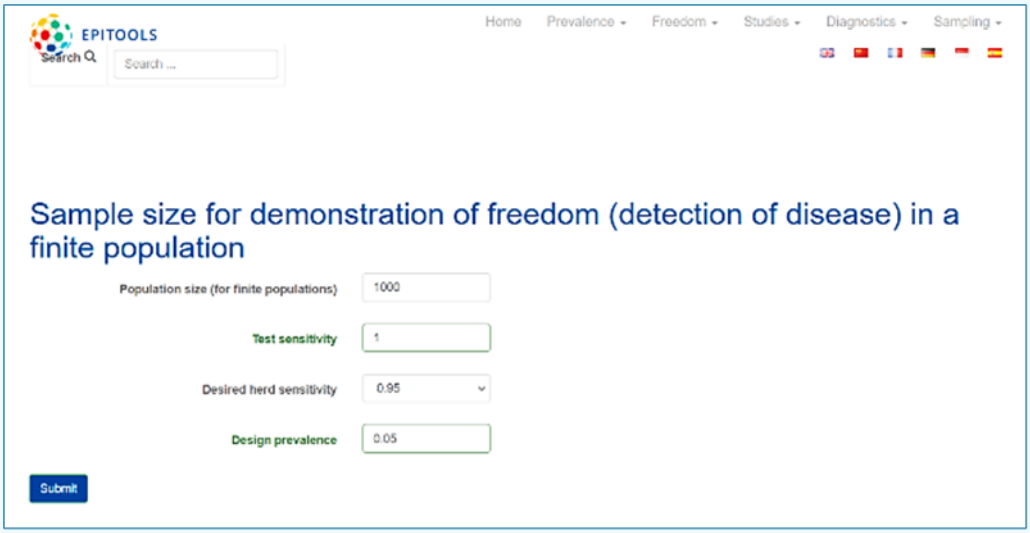


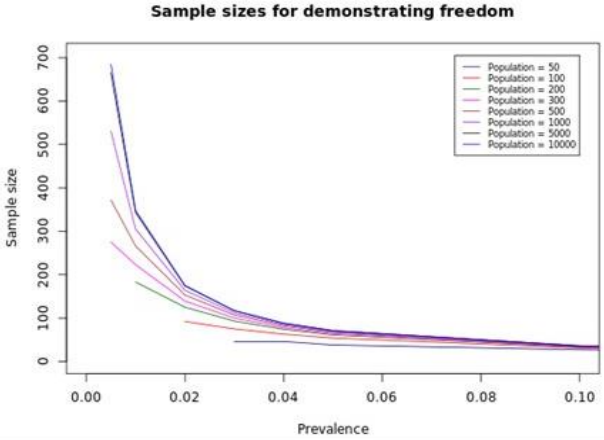
Figure 20, Input parameters of the EpiTools calculator - "Sample size for demonstration of freedom (detection of disease) in a finite population"

The EpiTools calculator was used to calculate the minimum sample size needed to detect *E. multilocularis* infection in red foxes in Great Britain (A) and Poland (B) at the required design prevalence. The input data was set according to those listed in *Figure 22*.

The plots below show the sample sizes required to provide the specified probability of detection of *E. multilocularis* in red foxes in both countries. These plots were generated by the EpiTools calculator to show predictions for different prevalence levels and population sizes for a specified test sensitivity.

A

Inputs	
Population size	240000
Test sensitivity	0.85
Confidence level	0.95
Design prevalence	0.01



B

Inputs	
Population size	198679
Test sensitivity	0.885
Confidence level	0.95
Design prevalence	0.1864

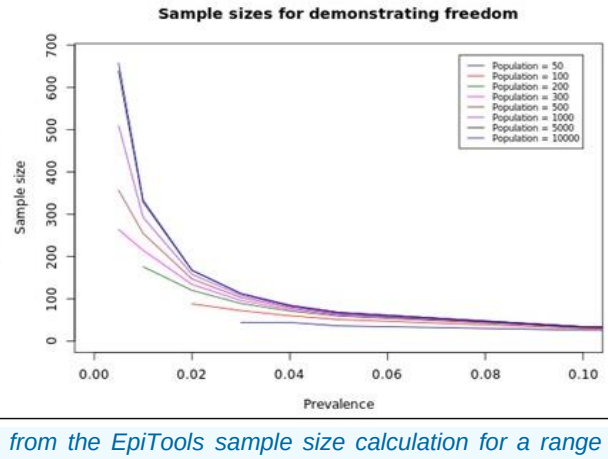


Figure 21 showing the inputs and outputs from the EpiTools sample size calculation for a range of design prevalences in Great Britain (A) and Poland (B)

7.3.2. True prevalence evaluation

What: This section estimates the true prevalence to correct previously calculated prevalence of disease (apparent prevalence).

Why: Most diagnostic tests have imperfect sensitivity and specificity. Calculation of true prevalence (the proportion of a population that is actually infected) considers the sensitivity and specificity of the applied test. Calculating the true prevalence can determine whether the choice of design prevalence for the system is still appropriate. This is more accurate than calculations of apparent prevalence (the proportion of the population that tests positive for the disease) which are reported in the majority of epidemiological studies/reports and do not include these parameters.

How: Sources which can be used to provide the data needed for the calculation of true prevalence are the following: scientific publications, international and governmental reports, expert opinions, and grey literature. Having identified and collated the relevant data, the true prevalence may be calculated.

A useful tool for calculating true prevalence is the EpiTools calculator - "Estimated true prevalence and predictive values from survey testing" (https://epitools.ausvet.com.au/trueprevalence). This tool calculates the true prevalence, as well as positive and negative predictive values, and likelihood ratios based on testing results using an assay of known sensitivity and specificity (Sergeant, 2018b). The inputs required to perform computations by the EpiTools calculator are as follows:

- sample size
- number of positive samples
- test sensitivity
- test specificity
- confidence level
- type of confidence interval for apparent prevalence
- type of confidence interval for true prevalence

To determine the true prevalence (TP) from these data, EpiTools applies the Rogan-Gladen estimator which is as follows:

$$TP = \frac{[AP + (SP - 1)]}{[SP + (SE - 1)]}$$

where:

- AP = apparent prevalence
- SP = specificity
- SE = sensitivity

Example – calculation of true prevalence using EpiTools: *E. multilocularis* in red foxes in Great Britain and Poland

Background:
As the sensitivity and specificity varies between diagnostic tests, there can be discrepancies between the apparent prevalence and true prevalence of a disease. This can help confirm whether the design prevalence of the OBS is still appropriate. Including true prevalence in the calculations can strengthen the evidence presented for this purpose and acts as an additional check to ensure that the surveillance system is fulfilling its policy objective.

Process:
The EpiTools calculator "Estimated true prevalence and predictive values from survey testing" was used to determine the true prevalence of *E. multilocularis* in red foxes in selected countries of the EU (*Table 7*). Furthermore, an assessment of the suitability of this tool for calculating true prevalence in sub-national administrative units was conducted on the example of the regions of Poland (Table 9). To perform these calculations, the input for the calculator was as follows:

- Number of examined samples obtained from red foxes (intestines or faeces samples) and number positive samples - set according to data from publications and reports as indicated in *Table 7* and *Table 9*.
- Sensitivity and specificity of the method (SCT, IST or PCR method) - determined in accordance with the study results listed in the *Table 7*
- Confidence level - was set at 0.95
- Type of confidence interval for apparent prevalence - Wilson CI was used
- Type of confidence interval for true prevalence – Blaker was used

Results:

The EpiTools online calculator enables the generation of plots comparing the confidence intervals and predictive values calculated for true prevalence and apparent prevalence. Plots for both prevalence values of *E. multilocularis* in red foxes in GB and Poland were generated using the above inputs, as well as plots for the three following regions of Poland: Opolskie, Śląskie, and Podkarpackie.

These three regions of Poland were selected to demonstrate the usefulness of the EpiTools calculator for estimating true prevalence for areas where high (Podkarpackie), medium (Śląskie) and low (Opolskie) numbers of disease cases were found. All plots are shown in the Figures below.

Table 7: Calculation of the true prevalence of E. multilocularis in red foxes in selected European countries.

Country	Apparent prevalence calculation					True prevalence calculation				
	Survey references	Number of tested samples	Number of postive results	Method	Apparent prevalence (%)	Sensitivity and specificity references	Method sensitivity	Method specificity	True prevalence (%)	95% CI
Poland	[1]	1546	255	SCT	16.5	[12]	0.885	1	18.64	16.64-20.82
Latvia	[2]	45	16	SCT	35.6	[12]	0.885	1	40.18	26.24-56.68
France	[3]	3307	562	SCT	17	[12]	0.885	1	19.2	17.8-20.69
Germany (northern)	[4]	3094	523	SCT	16.9	[12]	0.885	1	19.1	17.65-20.64
Denmark	[5]	546	4	SCT	0.73	[12]	0.885	1	0.83	0.32-2.11
Hungary	[6]	100	5	SCT	5	[12]	0.885	1	5.65	2.43-12.63
Romania	[7]	561	27	IST/SCT	4.8	[13]	0.78	1	6.17	4.27-8.86
Belgium	[8]	990	243	IST	24.55	[13]	0.78	1	31.47	28.16-35.03
Slovakia	[9]	660	49	IST/SCT	7.4	[13]	0.78	1	9.52	7.26-12.41
Estonia	[10]	17	5	SCT	29.4	[12]	0.885	1	33.23	15.01-60.04
Finland	[11]	265	0	PCR	0	[11]	0.78	1	0	0-1.83
Ireland	[11]	331	0	SCT	0	[12]	0.885	1	0	0-1.3
Great Britain	[11]	434	0	PCR	0	[11]	0.85	1	0	0-1.03
Norway	[11]	523	0	PCR	0	[11]	0.63	1	0	0-1.16

[1] - Karamon et al. (2014); [2] - Bagrade et al. (2008), [3] - Combes et al. (2012), [4] - Berke et al. (2008); [5] - Enemark et al. (2013); [6] - Sréter et al. (2004); [7] - Sikó et al. (2011); [8] - Hanosset et al. (2008); [9] - Bagrade et al. (2008), 2001; [10] - Moks et al. (2005); [11] - Assessment of *E. multilocularis* surveillance reports submitted in 2015 in the context of Commission Regulation (EU) No 1152/2011; [12] -Otero-Abad et al. (2017); [13] - Hofer et al. (2000)

Table 8: Calculation of the true prevalence of E. multilocularis in red foxes in individual provinces of Poland. Apparent prevalence was obtained from Karamon et al. (2014).

Provinces	No. of examined foxes	Apparent prevalence (%)	Wilson 95% CI of apparent prevalence	No. of positive foxes	True prevalence (%)	Blaker 95% CI of true prevalence
Dolnośląskie	102	2	0.54-6.87	2	2.22	0.61-7.76

Kujawsko-Pomorskie	103	3.9	1.52-9.56	4	4.39	1.72-10.81
Łódzkie	104	5.8	2.67-12.02	6	6.52	3.02-13.58
Małopolskie	98	28.6	20.57-38.19	28	32.28	23.24-43.15
Mazowieckie	120	30.8	23.27-39.58	37	34.84	26.3-44.73
Lubuskie	107	4.7	2.01-10.48	5	5.28	2.27-11.84
Opolskie	100	0	0-3.7	0	0	0-4.18
Podkarpackie	106	47.2	37.93-56.6	50	53.3	42.86-63.96
Podlaskie	100	34	25.46-43.72	34	38.42	28.77-49.4
Pomorskie	100	3	1.03-8.45	3	3.39	1.16-9.55
Śląskie	102	11.8	6.86-19.45	12	13.29	7.75-21.97
Świętokrzyskie	97	17.5	11.24-26.29	17	19.8	12.7-29.7
Warmińsko-Mazurskie	98	50	40.29-59.71	49	56.5	45.52-67.47
Wielkopolskie	119	2.5	0.86-7.15	3	2.85	0.97-8.08
Zachodniopomorskie	90	5.6	2.4-12.35	5	6.28	2.71-13.96
Total	1,546	16.5	14.73-18.43	255	18.64	16.64-20.82

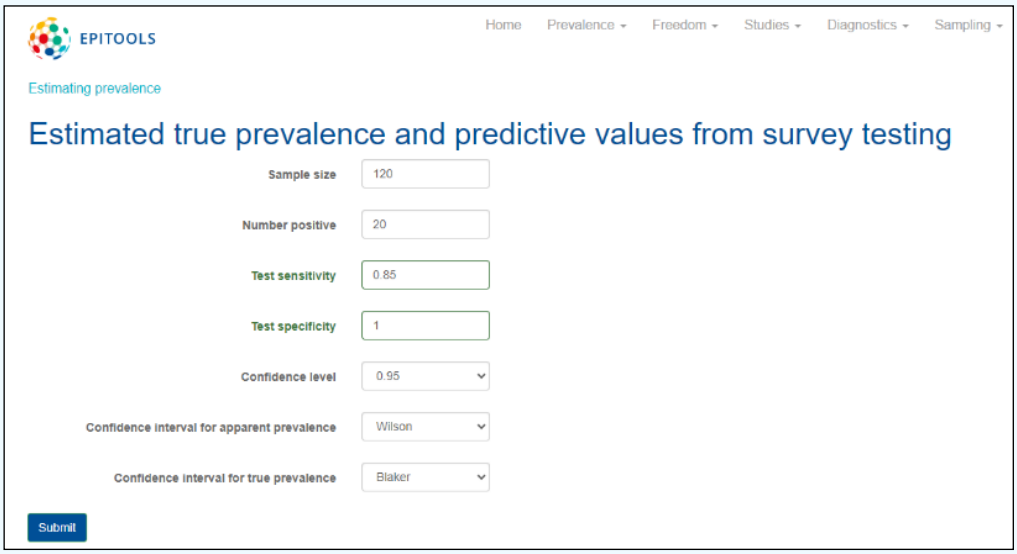


Figure 22, Data input for calculating the true prevalence of E.multilocularis using the EpiTools calculator

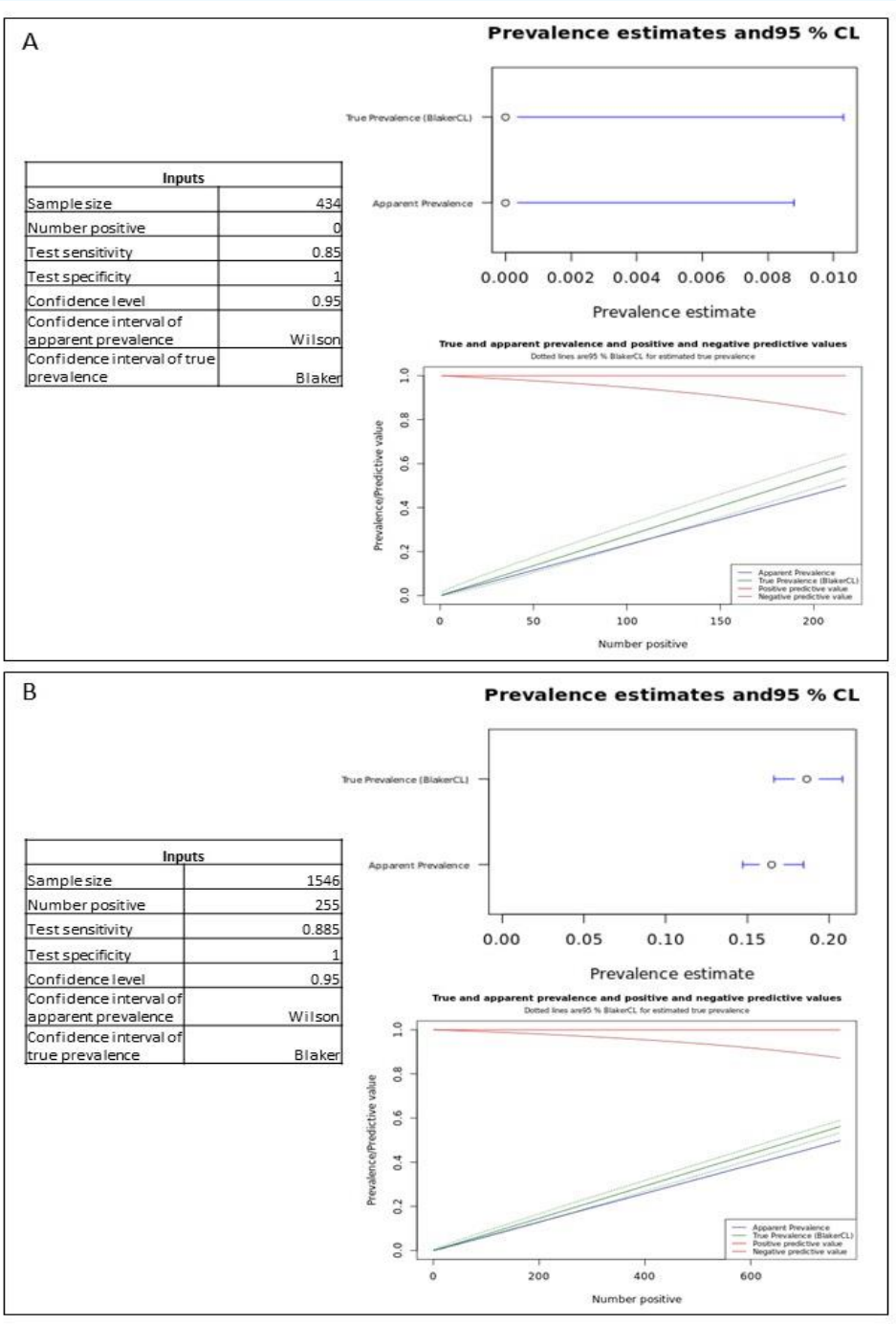


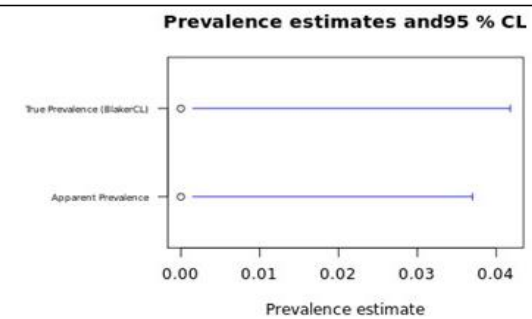
Figure 23, Comparison of confidence intervals and predictive values calculated for true prevalence and apparent prevalence of E. multilocularis in red foxes in the Great Britain (A) and Poland (B) using input data from Table 3.

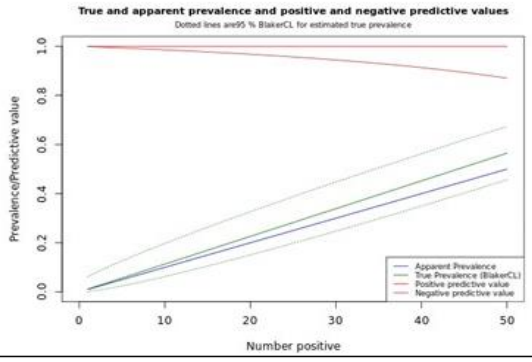
Se - method sensitivity; Sp - method specificity; CL - confidence level

Example – calculation of true prevalence using EpiTools: *E. multilocularis* in Great Britain and Poland

C

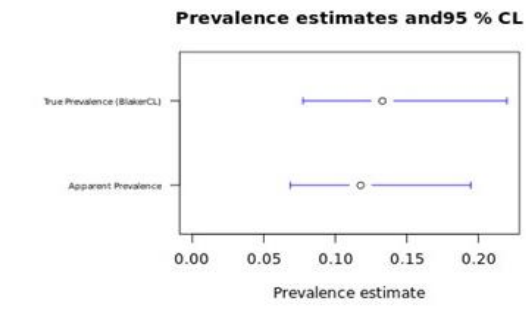
Inputs	
Sample size	100
Number positive	0
Test sensitivity	0.885
Test specificity	1
Confidence level	0.95
Confidence interval of apparent prevalence	Wilson
Confidence interval of true prevalence	Blaker





D

Inputs	
Sample size	102
Number positive	12
Test sensitivity	0.885
Test specificity	1
Confidence level	0.95
Confidence interval of apparent prevalence	Wilson
Confidence interval of true prevalence	Blaker



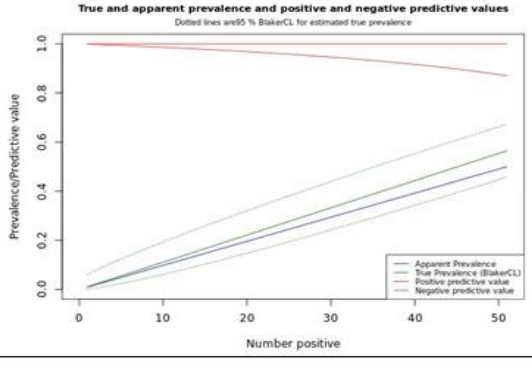
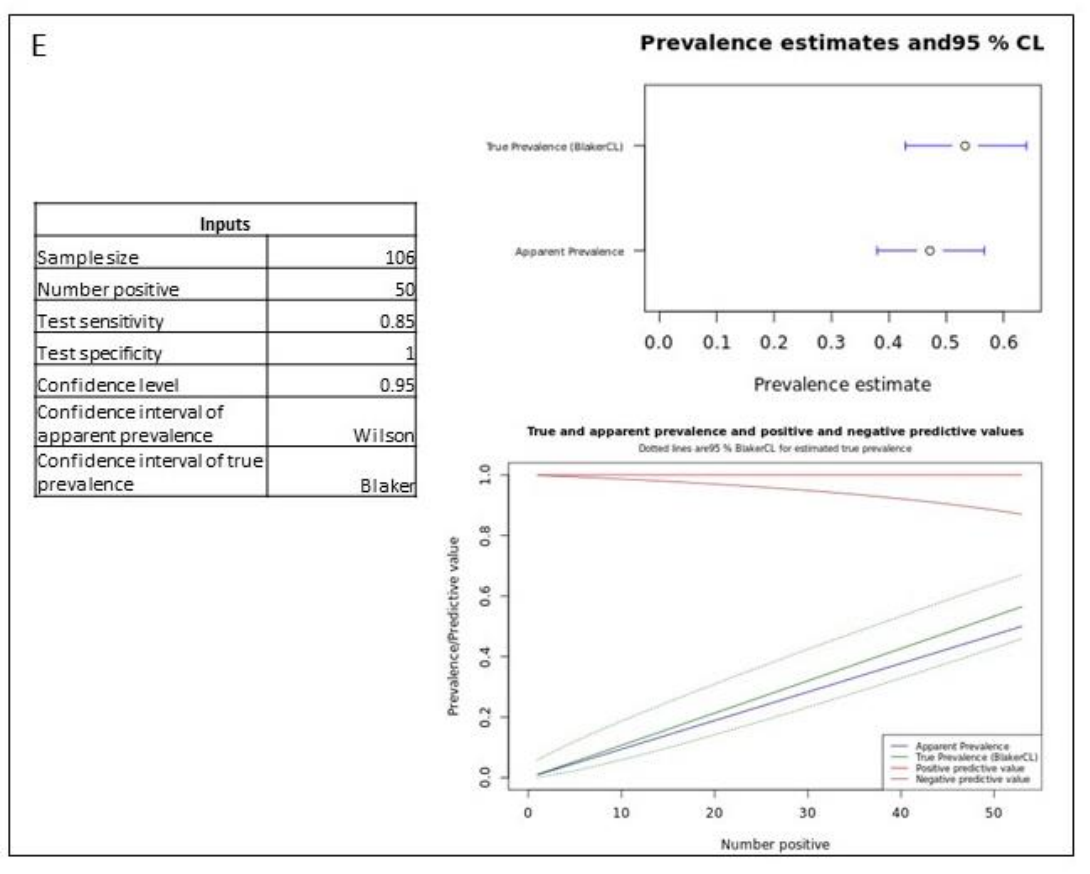


Figure 24, Confidence intervals and predictive values calculated for true prevalence and apparent prevalence of E. multilocularis in red foxes in the following selected regions of Poland using input data from Table: Opolskie (C), Śląskie (D) ,and Podkarpackie (E).

Se - method sensitivity; Sp - method specificity; CL - confidence level

Example – calculation of true prevalence using EpiTools: *E. multilocularis* in Great Britain and Poland



Inputs	
Sample size	106
Number positive	50
Test sensitivity	0.85
Test specificity	1
Confidence level	0.95
Confidence interval of apparent prevalence	Wilson
Confidence interval of true prevalence	Blaker

Se - method sensitivity; Sp - method specificity; CL - confidence level

7.3.3. Cost-effectiveness analysis

What: This evaluates whether the surveillance system is cost effective and provides good value for money.

Why: It is important that the testing process and the overall cost of the wider surveillance scheme is as cost effective as possible to make best use of the budget available. This likely also affects stakeholder satisfaction with the system and may affect the long-term sustainability of the system.

How: Cost effectiveness analyses (CEA) measure the input cost required for the system to produce a given output. Unlike some other economic analysis approaches, the 'effectiveness' component of a CEA can be defined by the analyst. Some studies use the quality adjusted life year, or disability adjusted life year, as metrics for effectiveness (Vallejo-Torres et al., 2018, Benedictus et al., 2009, Pitter et al., 2018). These are measurements of wellbeing within the human population, calculated from the perceived impacts of a human becoming infected with a disease. In output-based surveillance, the output is already defined at the operational level (to detect a pathogen at a stated design prevalence with a stated

confidence). CEA can easily be applied here to measure the cost input required to meet these outputs. This can then be compared directly to alternative approaches.

Gathering data on the cost inputs into a system first requires an inventory of all materials and reagents used, of average staff time required, and of any transport and sample collection costs. Materials and reagents required can be found using laboratory standard operating procedures. The price of each cost component may be attainable through contact with stakeholders working within the system. Alternatively, these may be found on supplier websites. Staff time should ideally be derived through contact with the staff themselves, preferably staff who have a holistic view of the system from sample acquisition to result reporting.

When collecting data on alternative test types, which are not yet in use, it may be useful to use proxies. Proxies can be similar tests already conducted for other diseases, and hence already have internal cost lists in the organisation. Data on alternative tests may also be found on supplier websites. Every test type will be different so it's important to try and avoid biases wherever possible. For example, if you are calculating costs over a year and a piece of key equipment needs maintenance every four years, then this cost needs to be considered fairly: it should not be ignored but should also not be considered in full for a single year of testing. A fair solution would be to divide this cost by the years between maintenance activities to make it a normalised annual cost output.

Data for each testing type must be calculated per test and multiplied by the required sample size (see section 5.3) based on the sensitivity of each test. This can be calculated using, for example, the EpiTools online resource. Doing so allows for direct comparison between the cost-effectiveness of each test type.

Example - *E. multilocularis* surveillance in Great Britain

Background:
Previously in test costing (section5.3.5) we used hypothetical data as an example of the cost of the egg flotation test for *E. multilocularis* surveillance. An objective for this surveillance is to ensure that the system uses a method that is practically and financially feasible. To evaluate the system against the financial aspect of this objective, we could either evaluate one method against (for example) a budget breakdown, or financial targets. Alternatively, multiple methods of testing, or other parts of the system, can be evaluated against each other to determine which is the most appropriate choice. For *E. multilocularis* we produced a cost-effectiveness analysis (CEA) comparing the hypothetical costs of multiple testing methods; the egg flotation test, and two alternate methods identified in the sampling methods section (Section 5.3.3).

Process:
To evaluate the cost effectiveness of the testing, the SOPs for these tests were used to create the consumables, reagents and equipment lists. Hypothetical costs of each of these components were then derived. Costs for two alternative methods of testing identified in section 5.3.4 were also produced based on protocols found through literature searches, and the three methods compared in a cost effectiveness analysis.

Results:

Egg flotation test

Table 9: Hypothetical cost breakdown per test of the annual cost of the egg flotation test, and the data sources associated with these costs.

Parameter	Value	Reference or supplier
Consumables, reagents, and equipment		
Pipettes	€123.89	Hypothetical cost data
Balance	€57.92	
Micro-centrifuge	€269.10	
Microscope	€443.25	
Staff		
Lab time	€22.75	(Eckert, 2003)
Additional costs (including sample collection and post-mortem examination)	€228.38	Hypothetical cost data
Other		
Test sensitivity	78%	(APHA parasitology lead, 2022)

Comparing cost-effectiveness of alternative test types
Comparing the cost-effectiveness of alternative test types requires available alternatives to be found and their individual cost-effectiveness calculated. These can then be normalized based on the sensitivities of each proposed test and compared directly with one another.

Available tests
Cataloguing the tests available was done through discussions with the stakeholders and through literature research. The annual EFSA report on *E. multilocular*is surveillance in Europe was an essential resource, summarizing how each country in Europe was conducting their tests, describing a range of alternative options (European Food Safety Authority and Zancanaro, 2021).

Costs for each test
The SOP of each test were used to determine what materials, reagents and instruments were required for that testing methodology. The costs of each of these reagents was determined using data available from potential supplier websites, or hypothetical cost data along with information in available literature. Similarly, hypothetical values were also generated for staff time, sample transport and post-mortems. All cost values were then added together to provide the annual costs of maintaining a surveillance system using each test type, including the costs for sample collection, post-mortem, testing, and epidemiological services linked to the system.

Normalised costings for each testing methodology
The total costs for a year applying each testing methodology were converted into a mean cost per test. The required testing output was used to calculate the required sample size to be taken from the population. The system was assumed to adopt a simple random testing design. The number of samples to be taken was found using EpiTools, an online sample size calculator developed by AUSVET (AUSVET, 2022). Since positive results were assumed to be followed up and confirmed, the specificity of all tests was set to 1.

The minimum number of tests required to detect a 1% prevalence with 95% confidence with the sensitivities specified by these tests was then multiplied by the cost per test to provide the overall cost of each testing methodology.

The sedimentation and counting technique (SCT), which is recognised as the gold standard for *E. multilocularis* testing, involves taking segments from the intestine of infected hosts and washing these with a saline solution. This yields a sediment after several periods of vigorous shaking and supernatant removal. These sediments are then observed under a microscope to count any eggs in the sample (European Food Safety Authority, 2021b).

The qPCR test, which is not yet widely used has been recommended as an alternative to the SCT test. This involves lysis and extraction of *E. multilocularis* DNA from faecal samples, followed by magnetic separation and amplification using real-time PCR (Maksimov et al., 2019).

For each test, a range of parameters were assembled, to allow for accurate comparison of their cost-effectiveness. These included the sensitivities of these tests, the consumables and reagents required for each along with the costs for staff time and equipment maintenance.

Sedimentation and Counting Technique (SCT) test

APHA conducts the SCT as part of the external quality assurance and proficiency testing schemes provided by the European Union Reference Laboratory for Parasites (EURLP) for the detection of *Echinococcus* sp. worms in intestinal mucosa. The instructions and procedure provided by the EURLP for this testing was used to broadly determine the consumables, reagents and equipment required for this test (European Union Reference Laboratory for Parasites, 2022). Prices per test were generated using hypothetical data.

Table 10: Hypothetical data showing the cost breakdown per test of the SCT, and the data sources associated with these costs.

Parameter	Value	Reference or supplier
Consumables, reagents, and equipment maintenance		
Freezer safe sample tubes	€0.42	(QIAGEN technologies, 2012)
Petri dish	€0.22	Hypothetical data
Scalpel	€0.98	
Forceps	€0.53	
Rectangular plastic dishes	€0.88	
NaCl >99.5% purity	€0.72	
Pipettes	€123.89	
Microscope maintenance	€443.25	
Balance PM	€57.92	
Staff		
Lab time	€22.75	(Eckert, 2003)
Additional costs (including sample collection and post-mortem examination)	€228.38	Hypothetical data
Other		
Test sensitivity	78%	(European Food Safety Authority, 2021b)

The staff time spent processing samples, 'lab time', was calculated using an average sample throughput of 15 samples per day based on information from literature (Eckert, 2003). The additional time costs including sample collection and post-mortems ('non-lab time') were assumed to be the same for all methods, and therefore are set at a blanket cost per sample(hypothetical data).

Real-time PCR

The real time PCR method used in this evaluation is the QIAamp Fast DNA Stool Mini Kit (QT) combined with a TaqMan PCR, the method for which has been previously described in literature (Maksimov et al., 2019, Knapp et al., 2014). A combination of this literature, and in-house SOPS were used to populate a list of consumables, reagents, and equipment (Central Unit for Sequencing and PCR (CUSP), 2022) which were then assigned hypothetical costs.

Table 11: Hypothetical data showing the cost breakdown per test of the real time qPCR test, and the data sources associated with these costs.

Parameter	Value	Reference or supplier
Consumables, reagents, and equipment		
QIAamp DNA Stool kit	€6.06	(QIAGEN technologies, 2012)
TaqMan Gene expression MM	€1.78	SLS
Primers	€3.61	Merk

Probe	€0.18	Merk
PCR plates	€0.0793	Agilent
Plate seals	€0.104	Agilent
Safelocks	€0.104	SLS
Pipette tips	€0.117	Starlab
Disposable spatulas	€0.04	SLS
Freezer safe sample tubes	€0.42	SLS
Service contract	€7,160.40	Hypothetical data
Replacement cartridges	€11,700.00	
Software	Ad hoc – less than annual cost	
Staff		
Lab time	€10.32	Hypothetical data
Additional costs (including sample collection and post-mortem examination)	€228.38	Hypothetical data
Other		
Test sensitivity	89%	(Maksimov et al., 2019)

Comparative cost-effectiveness

The costs of achieving the target output of the surveillance system were compared between each testing methodology. For annualised costs, such as sample collection and post-mortem, the per test cost was calculated based on the number of samples collected in GB for the sampling year 2021-2022: 800 (APHA parasitology lead, 2022). This was multiplied by the number of tests required, determined using the sensitivities of the tests and the EpiTools calculator.

Table 12: Showing the cost-effectiveness of three different testing methodologies for E. multilocularis at detecting a 1% prevalence detection with 95% confidence (hypothetical data).

Parameter	Unit	Test		
		Egg flotation	SCT	qPCR
Species sampled	-	Fox	Fox	Fox
Throughput	-	batch of 20 every 12h	10-20 per day (Average 15)	12-30 min per sample (Average 21)
Test sensitivity	-	0.78	0.78	0.89
Test specificity	-	1	1	1
Consumables and reagents	Per test	€56.88	€3.74	€12.48
Staff time (testing)	Per test	€9.26	€17.57	€10.32
Operational costs (excluding testing)	Annual cost (800 tests)	€291,593.12	€291,593.12	€291,593.12
Equipment	Annual cost	€894.15	€625.05	€18,860.40
Tests required at 1% prevalence	No. of tests	383	383	336
Cost of testing at 1% prevalence	€	€165,823.53	€150,408.31	€148,989.54

The test sensitivity of 0.78 for the zinc egg flotation (EF) and SCT methods is the value recommended for use by EFSA for this type of testing, whereas test sensitivity for the qPCR method is the average of those sourced from literature. From these data the qPCR is the most sensitive of the testing methods, reflected in the lower number of tests required per year to detect 1% prevalence with 95% confidence.

For this hypothetical scenario, the SCT is the most economical when it comes to consumables and reagents, costing an estimated €3.74 per test compared to the €12.48 and €56.88 required for the PCR and EF respectively. This is also true for the estimated annual cost of equipment and maintenance, with the SCT requiring an estimated €625.05 per year compared to €894.15 for the EF and €18.860.40 for the PCR equipment. This difference is mainly due to the comparatively large maintenance cost for real time PCR equipment. Where these outputs differ, however, is the cost of staff time associated with each test. We estimated the cost-per-test of both the EF and PCR at between €9-11 whereas due to the time intensive nature of the SCT, the per cost test was determined to be €17.57 based on staff processing an average of 15 samples per day (Eckert, 2003).

The additional operational costs were calculated by taking the total hypothetical cost for the entire current testing system over the sampling year and subtracting the consumables, reagents, and staff time for processing samples using EF in the laboratory. The remaining value, therefore, is expected to cover all other costs including transport, sample collection, post-mortem, registration, and administrative tasks. These costs could differ slightly for different testing types, but since all types require collection, post-mortem, and administration, we estimate these would remain relatively constant between them. Therefore, the same baseline value has been used for all testing types in this model. This was then scaled to the number of tests being performed. Variability from cost differences between sampling years was thought to be consistent across testing types.

Overall, in this model the qPCR is shown to be the most cost-effective testing method due to its lower number of tests required per year and consequently the lowest overall cost.

7.3.4. Propose improvements to the system (if applicable)

Each evaluation from the previous section will have developed an understanding of how well the surveillance system currently functions. This may have highlighted areas where the surveillance system needs improvement. Improvements do not necessarily mean increases in testing output, but rather changes to the system that make it more effective at achieving its objectives at the time of evaluation.

Examples of potential improvements include:

- Changes to test type to increase cost-effectiveness or accuracy of surveillance
- Changes to design prevalence to detect a higher or lower pathogen prevalence with greater confidence
- Changes to sample number to better reflect the chosen design prevalence
- Changes to the objectives of the system to reflect the current legislative requirements
- Changes to levels of stakeholder involvement to improve organizational efficiency in the system and boost stakeholder satisfaction

Any proposed improvements to the system constitute a change to the design proposal of the surveillance system. Hence, it may be necessary to go through the stages of design (section 5) and implementation (section 6) to ensure improvements are properly considered from all angles by the relevant stakeholders.

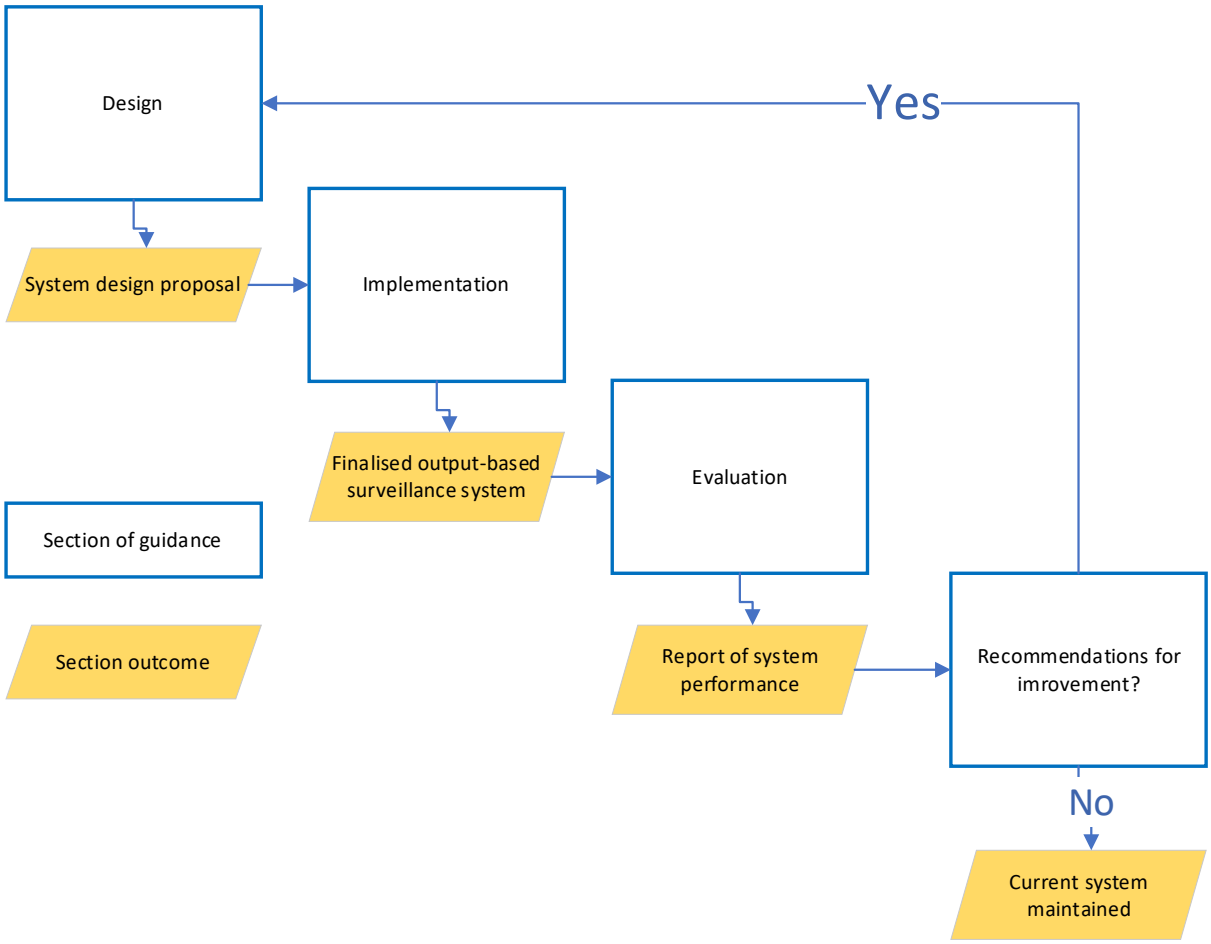


Figure 25, showing the outcome of each stage of this guidance. If any improvements can be made as a result of the evaluation of the system, then a revised system design which includes any improvements must be considered and implemented.

8. Discussion

Output based standards can allow for variation in surveillance activities to achieve the same objective and may be useful in the OH context where surveillance for animal pathogens can act as risk indicators for human health. The flexibility which using OBS allows, however, also requires more transparency to assist stakeholders, trade partners, decision-makers and risk assessors in interpreting the validity of the surveillance outcomes (Comin et al., 2019).

Key to the effectiveness of any surveillance system is the selection of appropriate methodologies to achieve the objectives of the system, whether it be increasing efficiency, introducing mitigations, or changing the surveillance components. A robust methodology which can evaluate the different aspects of surveillance and demonstrate confidence in the OBS is essential for stakeholder engagement. Similarly, stakeholder engagement is essential in order to produce a good evaluation of a surveillance system, therefore it is key to engage as many of the stakeholders as possible during this task.

In the design section of this guidance, we show that clear objectives are key to creating a robust OBS framework. We emphasise the importance of identifying all the stakeholders acting within your system and demonstrate how stakeholder engagement can guide the design of successful surveillance systems with their expertise and knowledge. Decisions on target pathogen, high risk populations, and testing method feed into the choice of sampling method and required sample size. For OBS, risk-based or

random sampling are generally most appropriate, since convenience-based methods often do not provide a representative sample of the target population. However, a convenience sampling approach can be made representative through spatial modelling techniques, as in the GB *E. multilocularis* example, if modelling and high sample availability can provide a random sampling distribution within the collected sample set. Adjacent to this, the population size and test sensitivity will directly influence the required sample size, which can be determined using, for example, the EpiTools calculator (Sergeant, 2018a). Determining the projected or actual test costs during the design phase feeds into both the cost-effectiveness objectives, but also into the test method used. Finally, as part of the design process, we show how overall test costs can be calculated with direct input from stakeholders, to ensure fulfilment of financial objectives.

Implementation planning is key in ensuring that a system design is applied logically and with due consideration of the real-world context of your system. This guidance showcases three useful implementation exercises to help bring your system design to life. We show how systems mapping can be used to visualise the steps and stakeholders involved in surveillance. Doing so helps to communicate the intended system design to all relevant stakeholders from an early stage. This is particularly important with regards to an OBS system which may be a novel approach to an existing surveillance programme. This guidance outlines the process of partitioning and integration (important in breaking down and building up large, complex, systems), highlights the importance of left shift to project efficiency, and suggests some key documents that should be kept for effective change management, work allocation, and scheduling. Finally, the guidance provides a method for recording and managing operational risks during the implementation process, which is a key stage for obtaining stakeholder satisfaction with the implementation plan along with identifying barriers to implementation early in the planning process.

To make OBS surveillance sustainable, regular performance evaluation is highly recommended. The evaluation section described in this guidance explores the evaluation of surveillance through five distinct lenses. Through analysis of historical trends and background, we can establish the relevance of the system to the contemporary disease and legislative context. Then, by applying technical evaluation tools such as EpiTools, we can measure the accuracy of our prevalence estimations and chosen sample sizes using current parameters if they are found to have changed. This provides an indication of whether individual surveillance streams could be upscaled or downscaled to meet the required output of the system. Along with a technical performance assessment, this guidance provides advice on how to evaluate the human factors within your system through stakeholder evaluation. Budgetary viewpoints are considered in the cost-effectiveness analysis put forward in this guidance, including consideration of alternative testing options. Finally, incorporating the results gathered from prior evaluations, the flexibility of the system to adaptation and change is analysed. Therefore, in completing the full evaluation, the technical, human, economic, and practical elements of the system can be visualised in the wider context of the current disease situation. This enables improvements to be made to this system with a holistic evidence-base supporting them.

EpiTools is used throughout the guidance in the examples for *E. multilocularis* in Poland and GB but it is acknowledged that other epidemiological calculators exist. For example, the calculator by Iowa state university provides sample size calculators and probability of detection calculators (Iowa state university, 2022). We chose EpiTools for the examples because of its broad range of available analysis applications, including sample size estimations using both hypergeometric and binomial approaches and true prevalence estimations using Bayesian and pooled computational approaches. This range of analyses makes it applicable to OBS with large or small population sizes, and with a broad design prevalence range. In addition, the tool is free and has had usage across several published articles, making it readily accessible to analysts from a range of backgrounds (Villarta Jr and Asaad, 2014, Laurin et al., 2021, Charan and Kantharia, 2013).

The inclusion of practical resources like EpiTools in this guide is indicative of the hands-on approach we wanted to take with it. Because this guidance is designed for OBS surveillance systems only, the guidance it provides can be more tailored than other surveillance evaluation tools such as SERVAL and RISKSUR EVA, which are generic to all forms of surveillance. The guidance is tied directly to worked examples that highlight immediate practical recommendations rather than top-level areas for improvement.

This guide provides a range of relevant activities for the design, implementation, and evaluation of OBS surveillance. However, depending on the context of its users there may be gaps that require additional research. This is expected given the broad scope of OBS in different situations, and as such this guidance should be taken alongside other training and literature from other sources. The evaluations also have some limitations. Any changes to an existing surveillance system or set up of a new system will inevitably incur implementation costs. Hence, in addition to the recommendations derived from the evaluation activities, it would also be beneficial to consider the costs of training staff, designing a new workflow, and purchasing new equipment before making any formal decisions to change the system design. In summary, the guidance outlined here provides a range of beneficial activities that can support the design, implementation and evaluation of OBS systems which can provide a valuable framework to facilitate the increasing interest in OBS surveillance.

9. References

APHA 2022. Laboratory costs breakdown: materials and reagents. *In:* AGENCY, A. A. P. H. (ed.).

APHA PARASITOLOGY LEAD. 2022. *RE: Interview with APHA parasitology lead.*

AUSVET. 2022. *Sample size to achieve specified population level (or herd, flock, cluster, etc) sensitivity* [Online]. AUSVET. Available: https://epitools.ausvet.com.au/freedomss [Accessed 23/05/2022].

BAGRADE, G., ŠNÁBEL, V., ROMIG, T., OZOLIŅŠ, J., HÜTTNER, M., MITERPÁKOVÁ, M., ŠEVCOVÁ, D. & DUBINSKÝ, P. 2008. Echinococcus multilocularis is a frequent parasite of red foxes (Vulpes vulpes) in Latvia. *Helminthologia,* 45**,** 157-161.

BENEDICTUS, A., HOGEVEEN, H. & BERENDS, B. 2009. The price of the precautionary principle: Cost-effectiveness of BSE intervention strategies in the Netherlands. *Preventive Veterinary Medicine,* 89**,** 212-222.

BERKE, O., ROMIG, T. & VON KEYSERLINGK, M. 2008. Emergence of Echinococcus multilocularis among red foxes in northern Germany, 1991–2005. *Veterinary parasitology,* 155**,** 319-322.

CENTERS FOR DISEASE CONTROL AND PREVENTION. 2019. *Parasites - Echinococcosis Biology,* [Online]. Available: https://www.cdc.gov/parasites/echinococcosis/biology.html [Accessed].

CENTRAL UNIT FOR SEQUENCING AND PCR (CUSP). 2022. *RE: RE:qPCR reagents costs.*

CHARAN, J. & KANTHARIA, N. 2013. How to calculate sample size in animal studies? *Journal of Pharmacology and Pharmacotherapeutics,* 4**,** 303-306.

COHESIVE. 2022. *One Health: Setting up a risk analysis system for zoonoses* [Online]. Available: https://www.ohras.eu/page/home [Accessed 17/05/2022].

COMBES, B., COMTE, S., RATON, V., RAOUL, F., BOUÉ, F., UMHANG, G., FAVIER, S., DUNOYER, C., WORONOFF, N. & GIRAUDOUX, P. 2012. Westward spread of Echinococcus multilocularis in foxes, France, 2005-2010. *Emerg Infect Dis,* 18**,** 2059-62.

CROFT, S., CHAUVENET, A. L. & SMITH, G. C. 2017. A systematic approach to estimate the distribution and total abundance of British mammals. *PloS one,* 12**,** e0176339.

DANISH CENTRE FOR ENVIRONMENT AND ENERGY. 2022. Available: https://dce.au.dk/en/ [Accessed].

DAVISON, N. 2016. The change control process. *In:* RESOURCES, U. P. M. (ed.).

DEFRA & APHA. 2019. *Echinococcus multilocularis: how to spot and report the disease* [Online]. Available: https://www.gov.uk/guidance/echinococcus-multilocularis-how-to-spot-and-report-the-disease [Accessed 30/09/2022 2022].

DEVI, T. R. & REDDY, V. S. 2012. Work breakdown structure of the project. *Int J Eng Res Appl,* 2**,** 683-686.

DREWE, J., HOINVILLE, L., COOK, A., FLOYD, T., GUNN, G. & STÄRK, K. 2015. SERVAL: a new framework for the evaluation of animal health surveillance. *Transboundary and emerging diseases,* 62**,** 33-45.
ECKERT, J. 2003. Predictive values and quality control of techniques for the diagnosis of Echinococcus multilocularis in definitive hosts. *Acta Tropica,* 85**,** 157-163.
EMES, M., SMITH, A. & JAMES, A. 2014. 'Left-shift'vs The Time Value of Money.
EMES*, M. R., SMITH, A., JAMES, A. M., WHYNDHAM, M. W., LEAL, R. & JACKSON, S. C. 8.1. 2 Principles of Systems Engineering Management: Reflections from 45 years of spacecraft technology research and development at the Mullard Space Science Laboratory. INCOSE International Symposium, 2012. Wiley Online Library, 1069-1084.
ENEMARK, H. L., AL-SABI, M. N., KNAPP, J., STAAHL, M. & CHRÍEL, M. 2013. Detection of a high-endemic focus of Echinococcus multilocularis in red foxes in southern Denmark, January 2013. *Euro Surveill,* 18**,** 20420.
EUROPEAN FOOD SAFETY AUTHORITY 2012. A framework to substantiate absence of disease: the risk based estimate of system sensitivity tool (RiBESS) using data collated according to the EFSA Standard Sample Description - An example on Echinococcus multilocularis. *EFSA Supporting Publications,* 9**,** 366E.
EUROPEAN FOOD SAFETY AUTHORITY 2015. Assessment of Echinococcus multilocularis surveillance reports submitted in 2015 in the context of Commission Regulation (EU) No 1152/2011. *EFSA Journal,* 13**,** 4310.
EUROPEAN FOOD SAFETY AUTHORITY 2021a. Annual assessment of Echinococcus multilocularis surveillance reports submitted in 2020 in the context of Commission Delegated Regulation (EU) 2018/772. *EFSA Journal,* 19**,** e06382.
EUROPEAN FOOD SAFETY AUTHORITY 2021b. Annual assessment of Echinococcus multilocularis surveillance reports submitted in 2021 in the context of Commission Delegated Regulation (EU) 2018/772. *EFSA Journal,* 19**,** e06945.
EUROPEAN FOOD SAFETY AUTHORITY & ZANCANARO, G. 2021. Annual assessment of Echinococcus multilocularis surveillance reports submitted in 2021 in the context of Commission Delegated Regulation (EU) 2018/772. *EFSA Journal,* 19**,** e06945.
EUROPEAN UNION REFERENCE LABORATORY FOR PARASITES. 2022. *Proficiency Testing* [Online]. Available: https://www.iss.it/en/web/iss-en/eurlp-proficiency-testing [Accessed].
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.
HAYDEN, T. J. & HARRINGTON, R. 2000. *Exploring Irish mammals*, Town House.
HAYES, J. 2022. *The theory and practice of change management*, Bloomsbury Publishing.
HOFER, S., GLOOR, S., MÜLLER, U., MATHIS, A., HEGGLIN, D. & DEPLAZES, P. 2000. High prevalence of Echinococcus multilocularis in urban red foxes (Vulpes vulpes) and voles (Arvicola terrestris) in the city of Zürich, Switzerland. *Parasitology,* 120 (Pt 2)**,** 135-42.
IOWA STATE UNIVERSITY. 2022. *Sample size calculators* [Online]. Available: https://fieldepi.research.cvm.iastate.edu/calc/ [Accessed 26/10/2022].
KARAMON, J., KOCHANOWSKI, M., SROKA, J., CENCEK, T., RÓŻYCKI, M., CHMURZYŃSKA, E. & BILSKA-ZAJĄC, E. 2014. The prevalence of Echinococcus multilocularis in red foxes in Poland—current results (2009–2013). *Parasitology research,* 113**,** 317-322.
KAUHALA, K. 2007. Paljonko Suomessa on pienpetoja?
KIRJUŠINA, M., DEKSNE, G., MARUCCI, G., BAKASEJEVS, E., JAHUNDOVIČA, I., DAUKŠTE, A., ZDANKOVSKA, A., BĒRZIŅA, Z., ESĪTE, Z. & BELLA, A. 2015. A 38-year study on Trichinella spp. in wild boar (Sus scrofa) of Latvia shows a stable incidence with an increased parasite biomass in the last decade. *Parasites & vectors,* 8**,** 1-8.
KNAPP, J., MILLON, L., MOUZON, L., UMHANG, G., RAOUL, F., ALI, Z. S., COMBES, B., COMTE, S., GBAGUIDI-HAORE, H., GRENOUILLET, F. & GIRAUDOUX, P. 2014. Real time PCR to detect the environmental faecal contamination by Echinococcus multilocularis from red fox stools. *Veterinary Parasitology,* 201**,** 40-47.
LAURIN, E., BRADSHAW, J., HAWLEY, L., GARDNER, I. A., GARVER, K., JOHNSON, S. C. & THAKUR, K. K. 2021. Importance of sample size for estimating prevalence: a case example of infectious hematopoietic necrosis viral RNA detection in mixed-stock Fraser River sockeye salmon (Oncorhynchus nerka), British Columbia, Canada. *Canadian Journal of Fisheries and Aquatic Sciences,* 78**,** 589-598.

LEAL, R. 2020. The foundations of systems engineering. *In:* ENGINEERING, U. C. F. S. (ed.). London.
LECY, J. D. & BEATTY, K. E. 2012. Representative literature reviews using constrained snowball sampling and citation network analysis. *Available at SSRN 1992601*.
LESTER, A. 2006. *Project management, planning and control: managing engineering, construction and manufacturing projects to PMI, APM and BSI standards*, Elsevier.
MACDIARMID, S. 1988. Future options for brucellosis surveillance in New Zealand beef herds. *New Zealand Veterinary Journal,* 36**,** 39-42.
MAKSIMOV, P., ISAKSSON, M., SCHARES, G., ROMIG, T. & CONRATHS, F. 2019. Validation of PCR-based protocols for the detection of Echinococcus multilocularis DNA in the final host using the Intestinal Scraping Technique as a reference. *Food and Waterborne Parasitology,* 15**,** e00044.
MARNELL, F., KINGSTON, N. & LOONEY, D. 2009. Ireland Red List No. 3: Terrestrial Mammals. *National Parks and Wildlife Service, Department of the Environment, Heritage and Local Government, Dublin, Ireland*.
MATHEWS, F., KUBASIEWICZ, L., GURNELL, J., HARROWER, C., MCDONALD, R. A. & SHORE, R. 2018. *A review of the population and conservation status of British mammals: technical summary*, Natural England.
MAZET, J., UHART, M. & KEYYU, J. 2014. Stakeholders in one health. *Rev Sci Tech,* 33**,** 443-452.
MENDELOW, A. L. 1981. Environmental Scanning--The Impact of the Stakeholder Concept.
MOKS, E., SAARMA, U. & VALDMANN, H. 2005. Echinococcus multilocularis in Estonia. *Emerg Infect Dis,* 11**,** 1973-4.
OTERO-ABAD, B., ARMUA-FERNANDEZ, M. T., DEPLAZES, P., TORGERSON, P. R. & HARTNACK, S. 2017. Latent class models for Echinococcus multilocularis diagnosis in foxes in Switzerland in the absence of a gold standard. *Parasites & vectors,* 10**,** 1-14.
PITTER, J., VOKÓ, Z., JÓZWIAK, Á. & BERKICS, A. 2018. Campylobacter control measures in indoor broiler chicken: critical re-assessment of cost-utility and putative barriers to implementation. *Epidemiology & Infection,* 146**,** 1433-1444.
QIAGEN TECHNOLOGIES. 2012. *QIAamp DNA Stool Handbook,* [Online]. Available: https://www.qiagen.com/us/resources/download.aspx?id=e5f66d36-1f1c-4e89-828e-257088bdad53&lang=en [Accessed 01/06/2022].
RAANAN, Y. & KENETT, R. S. 2011. *Operational Risk Management: a practical approach to intelligent data analysis*, John Wiley & Sons.
RAHMATIAN, S. 1985. The hierarchy of objectives: Toward an integrating construct in systems science. *Systems Research,* 2**,** 237-245.
RAMLER, R. & FELDERER, M. Requirements for integrating defect prediction and risk-based testing. 2016 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 2016. IEEE, 359-362.
ROMANIAN NATIONAL INSTITUTE OF STATISTICS. 2008. Available: http://www.insse.ro/cms/files/statistici/comunicate/com_anuale/fond%20cinegetic/Fondcinegetic_2008.pdf [Accessed].
SERGEANT, E. 2018a. *Epitools Epidemiological Calculators* [Online]. Ausvet. Available: http://epitools.ausvet.com.au [Accessed 2022].
SERGEANT, E. 2018b. *Epitools epidemiological calculators* [Online]. Available: https://epitools.ausvet.com.au/ [Accessed].
SIKÓ, S. B., DEPLAZES, P., CEICA, C., TIVADAR, C. S., BOGOLIN, I., POPESCU, S. & COZMA, V. 2011. Echinococcus multilocularis in south-eastern Europe (Romania). *Parasitol Res,* 108**,** 1093-7.
SRÉTER, T., SZÉLL, Z., SRÉTER-LANCZ, Z. & VARGA, I. 2004. Echinococcus multilocularis in Northern Hungary. *Emerg Infect Dis,* 10**,** 1344-6.
STEVENSON, M. & SERGEANT, E. 2022. *Design and Analysis of Disease Surveillance Programs Using epiR* [Online]. Available: https://cran.r-project.org/web/packages/epiR/vignettes/epiR_surveillance.html [Accessed 24/10/22 2022].
ŞUTEU, O., MIHALCA, A. D., PAŞTIU, A. I., GYÖRKE, A., MATEI, I. A., IONICĂ, A., BALEA, A., OLTEAN, M., D'AMICO, G. & SIKÓ, S. B. 2014. Red foxes (Vulpes vulpes) in Romania are carriers of Toxoplasma gondii but not Neospora caninum. *Journal of Wildlife Diseases,* 50**,** 713-716.

SVILAND, S., JOHANSEN, T., KLEVAR, S., VALHEIM, M. & JONSSON, M. 2014. Surveillance programmes for terrestrial and aquatic animals in Norway. *Annual Report*.
THE FOREST DATA BANK. 2022. Available: https://www.bdl.lasy.gov.pl/portal/en [Accessed].
VALLEJO-TORRES, L., GARCÍA-LORENZO, B. & SERRANO-AGUILAR, P. 2018. Estimating a cost-effectiveness threshold for the Spanish NHS. *Health economics,* 27**,** 746-761.
VILLARTA JR, R. L. & ASAAD, A. S. 2014. Sample Size Determination in an Epidemiologic Study using the EpiTools Web-Based Calculator. *Acta Medica Philippina,* 48.
WEILKIENS, T. 2011. Chapter 1: the V model. *Systems engineering with SysML/UML: modeling, analysis, design.* Elsevier.
WILSON, J. M. 2003. Gantt charts: A centenary appreciation. *European Journal of Operational Research,* 149**,** 430-437.
WORLD HEALTH ORGANIZATION. 2021. *Echinococcosis* [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/echinococcosis [Accessed 18/10/22 2022].
WORLD ORGANISATION FOR ANIMAL HEALTH (WOAH) 2022. ECHINOCOCCOSIS (INFECTION WITH ECHINOCOCCUS GRANULOSUS AND WITH E . MULTILOCULARIS). *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals.*

Appendix 1: Articles meeting the inclusion criteria of the literature review

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Ranta J., Hovi T., Arjas E.	Poliovirus surveillance by examining sewage water specimens: Studies on detection probability using simulation models	2001			1			
Doherr M.G., Heim D.., Fatzer R., Cohen C.H., et al.	Targeted screening of high-risk cattle populations for BSE to augment mandatory reporting of clinical suspects	2001	1	1				
Cousins D.V., Roberts J.L.	Australia's campaign to eradicate bovine tuberculosis: The battle for freedom and beyond	2001	1			1		
Hadorn D.C., Rüfenacht J., Hauser R., Stärk K.D.C.	Risk-based design of repeated surveys for the documentation of freedom from non-highly contagious diseases	2002	1					1
De Vos CJ, Saatkamp HW, Huirne RB.	Cost-effectiveness of measures to prevent classical swine fever introduction into The Netherlands	2005	1					
Fischer E.A.J., Van Roermund H.J.W., Hemerik L., Van Asseldonk M.A.P.M., et al.	Evaluation of surveillance strategies for bovine tuberculosis (*Mycobacterium bovis*) using an individual based epidemiological model	2004	1			1		
Corbellini L.G., Schwermer H., Presi P., Thür B., et al.	Analysis of national serological surveys for the documentation of freedom from porcine reproductive and respiratory syndrome in Switzerland	2006	1					
Racloz V., Straver R., Kuhn M., Thur B., et al.	Establishment of an early warning system against Bluetongue virus in Switzerland	2006	1					
Stärk, K.D., Regula, G., Hernandez, J. et al.	Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches.	2006	1					1

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Martin PA, Cameron AR, Greiner M.	Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees	2007	1					1
Knopf L., Schwermer H., Stärk K.D.C.	A stochastic simulation model to determine the sample size of repeated national surveys to document freedom from bovine herpesvirus 1 (BoHV-1) infection	2007	1					
Martin P.A.J., Cameron A.R., Barfod K., Sergeant E.S.G., et al.	Demonstrating freedom from disease using multiple complex data sources. 2: Case study-Classical swine fever in Denmark	2007	1					
de Vos CJ, Saatkamp HW, Ehlers J.	Simulation evaluation of *Salmonella* monitoring in finishing pigs in Lower Saxony, Germany	2007	1	1				
Alban L., Boes J., Kreiner H., Petersen J.V., et al.	Towards a risk-based surveillance for *Trichinella* spp. in Danish pig production	2008	1	1				
Yamamoto T., Tsutsui T., Nishiguchi A., Kobayashi S.	Evaluation of surveillance strategies for bovine brucellosis in Japan using a simulation model	2008	1			1		
Hadorn DC, Haracic SS, Stärk KD.	Comparative assessment of passive surveillance in disease-free and endemic situation: example of *Brucella melitensis* surveillance in Switzerland and in Bosnia and Herzegovina	2008	1	1				
Hadorn DC, Stärk KD.	Evaluation and optimization of surveillance systems for rare and emerging infectious diseases	2008	1					1
Coulston J.W., Koch F.H., Smith W.D., Sapio F.J.	Invasive forest pest surveillance: Survey development and reliability	2008					1	
Martin P.A.J.	Current value of historical and ongoing surveillance for disease freedom: Surveillance for bovine Johne's disease in Western Australia	2008	1	1				

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
More S.J., Cameron A.R., Greiner M., Clifton-Hadley R.S., et al.	Defining output-based standards to achieve and maintain tuberculosis freedom in farmed deer, with reference to member states of the European Union	2009	1			1		
Frössling J., Ågren E.C.C., Eliasson-Selling L., Lewerin S.S.	Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: Scenario-tree modelling of the surveillance system	2009	1					
Watkins R.E., Martin P.A.J., Kelly H., Madin B., et al.	An evaluation of the sensitivity of acute flaccid paralysis surveillance for poliovirus infection in Australia	2009			1			
Schwermer H, Reding I, Hadorn DC.	Risk-based sample size calculation for consecutive surveys to document freedom from animal diseases	2009	1					1
Williams MS, Ebel ED, Wells SJ.	Poisson sampling: a sampling strategy for concurrently establishing freedom from disease and estimating population characteristics	2009	1					1
Nielsen SS, Toft N.	A review of prevalences of paratuberculosis in farmed animals in Europe	2009	1					1
Hadorn DC, Racloz V, Schwermer H, Stärk KD.	Establishing a cost-effective national surveillance system for Bluetongue using scenario tree modelling	2009	1					
Schuppers M.E., Rosenberg G., Graf R., Eidam V., et al.	A study to demonstrate freedom from *Trichinella* spp. in domestic pigs in Switzerland	2010	1			1		
Knight-Jones TJ, Hauser R, Matthes D, Stärk KD.	Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza	2010	1	1				
Alba A, Casal J, Napp S, Martin PA.	Assessment of different surveillance systems for avian influenza in commercial poultry in Catalonia (North-Eastern Spain)	2010	1	1				

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Schuppers M.E., Frey C.F., Gottstein B., Stärk K.D.C., et al.	Comparing the demonstration of freedom from *Trichinella* infection of domestic pigs by traditional and risk-based surveillance	2010	1			1		
Gustafson L., Klotins K., Tomlinson S., Karreman G., et al/	Combining surveillance and expert evidence of viral hemorrhagic septicemia freedom: A decision science approach	2010	1					
Christensen J., Stryhn H., Vallières A., Allaki F.E.	A scenario tree model for the Canadian Notifiable Avian Influenza Surveillance System and its application to estimation of probability of freedom and sample size determination	2011	1	1				
Wahlström H, Isomursu M, Hallgren G, Christensson D, et al.	Combining information from surveys of several species to estimate the probability of freedom from *Echinococcus multilocularis* in Sweden, Finland and mainland Norway	2011	1	1				
Blickenstorfer S., Schwermer H., Engels M., Reist M., et al.	Using scenario tree modelling for targeted herd sampling to substantiate freedom from disease	2011	1					1
EFSA Panel on Animal Health and Welfare	Scientific Opinion on bluetongue monitoring and surveillance	2011	1					
Welby S., Govaerts M., Vanholme L., Hooyberghs J., et al.	Bovine tuberculosis surveillance alternatives in Belgium	2012	1			1		
Cameron A.R.	The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease	2012	1					1
Schuppers M.E., Stegeman J.A., Kramps J.A., Stärk K.D.C.	Implementing a probabilistic definition of freedom from infection to facilitate trade of livestock: Putting theory into praxis for the example of bovine herpes virus-1	2012	1					

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Comin A., Stegeman A., Marangon S., Klinkenberg D.	Evaluating surveillance strategies for the early detection of low pathogenicity avian influenza infections	2012	1	1				
EFSA Panel on Animal Health and Welfare (AHAW)	Scientific Opinion on the use of a gamma interferon test for the diagnosis of bovine tuberculosis	2012	1	1				
European Food Safety Authority	Scientific and technical assistance on *Echinococcus multilocularis* infection in animals	2012	1	1				
Goutard F.L., Paul M., Tavornpanich S., Houisse I.,et al.	Optimizing early detection of avian influenza H5N1 in backyard and free-range poultry production systems in Thailand	2012	1	1				
Reist M., Jemmi T., Stärk K.D.C.	Policy-driven development of cost-effective, risk-based surveillance strategies	2012	1					1
Willeberg P., Nielsen L.R., Salman M.	Designing and evaluating risk-based surveillance systems: Potential unwarranted effects of applying adjusted risk estimates	2012	1					1
Willeberg P.	Animal health surveillance applications: The interaction of science and management	2012	1					1
Caporales V., Giovannini A., Zepeda C.	Surveillance strategies for foot and mouth disease to prove absence of disease and absence of viral circulation	2012	1					
Reber A, Reist M, Schwermer H.	Cost-effectiveness of bulk-tank milk testing for surveys to demonstrate freedom from infectious bovine rhinotracheitis and bovine enzootic leucosis in Switzerland	2012	1					
Rutten N, Gonzales JL, Elbers AR, Velthuis AG.	Cost analysis of various low pathogenic avian influenza surveillance systems in the Dutch egg layer sector	2012	1			1		

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Murphy TM, Wahlström H, Dold C, Keegan JD, et al.	Freedom from *Echinococcus multilocularis*: an Irish perspective	2012	1	1				
Häsler B, Howe KS, Presi P, Stärk KD.	An economic model to evaluate the mitigation programme for bovine viral diarrhoea in Switzerland	2012	1					
More S.J., Sergeant E.S.G., Strain S., Cashman W., et al.	The effect of alternative testing strategies and bio-exclusion practices on Johne's disease risk in test-negative herds	2013	1			1		
European Food Safety Authority	Assessment of *Echinococcus multilocularis* surveillance reports submitted 2013 in the context of Commission Regulation (EU) No 1152/2011	2013	1	1				
Welby S., Méroc E., Faes C., De Clercq K., et al.	Bluetongue surveillance system in Belgium: A stochastic evaluation of its risk-based approach effectiveness	2013	1					
Vanderstichel R, Christensen J, Stryhn H, Hurnik D.	Standards for reporting surveillance information in freedom from infection models by example of *Trichinella* in Canadian market hogs	2013	1			1		
Boklund A., Dahl J., Alban L.	Assessment of confidence in freedom from Aujeszky's disease and classical swine fever in Danish pigs based on serological sampling-Effect of reducing the number of samples	2013	1					
Calvo-Artavia F.F., Alban L., Nielsen L.R.	Evaluation of surveillance for documentation of freedom from bovine tuberculosis	2013	1			1		
Bessell P.R., Orton R., O'Hare A., Mellor D.J., et al.	Developing a framework for risk-based surveillance of tuberculosis in cattle: A case study of its application in Scotland	2013	1			1		
Frössling J., Wahlström H., Ågren E.C.C., Cameron A., et al.	Surveillance system sensitivities and probability of freedom from *Mycobacterium avium* subsp. paratuberculosis infection in Swedish cattle	2013	1			1		

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Kostoulas P, Nielsen SS, Browne WJ, Leontides L.	Sample size estimation to substantiate freedom from disease for clustered binary data with a specific risk profile	2013	1	1			1	
Calvo-Artavia FF, Nielsen LR, Alban L.	Epidemiologic and economic evaluation of risk-based meat inspection for bovine cysticercosis in Danish cattle	2013	1			1		
Eichenberger RM, Lewis F, Gabriël S, Dorny P, et al.	Multi-test analysis and model-based estimation of the prevalence of *Taenia saginata* cysticercus infection in naturally infected dairy cows in the absence of a 'gold standard' reference test.	2013	1			1		
Oidtmann B., Peeler E., Lyngstad T., Brun E., et al.	Risk-based methods for fish and terrestrial animal disease surveillance	2013	1					1
Diserens N., Presi P., Bernet D., Schüpbach-Regula G., et al.	Risk assessment for the design of a risk-based surveillance programme for fish farms in Switzerland (in accordance with Council Directive 2006/88/EC of the European Union)	2013	1					
Whist A.C., Liland K.H., Jonsson M.E., Sæbø S., et al.	Designing a risk-based surveillance program for *Mycobacterium avium* ssp. paratuberculosis in Norwegian dairy herds using multivariate statistical process control analysis	2014	1	1				
Christensen J, El Allaki F, Vallières A.	Adapting a scenario tree model for freedom from disease as surveillance progresses: the Canadian notifiable avian influenza model	2014	1	1				
Norstrom et al	Estimation of the probability of freedom from Bovine virus diarrhoea virus in Norway using scenario tree modelling	2014	1					
Ferrer E., Alfonso P., Ippoliti C., Abeledo M., et al.	Development of an active risk-based surveillance strategy for avian influenza in Cuba	2014	1	1				

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Rivière J., Carabin K., Le Strat Y., Hendrikx P., et al.	Bovine tuberculosis surveillance in cattle and free-ranging wildlife in EU Member States in 2013: A survey-based review	2014	1	1				
European Food Safety Authority	Assessment of *Echinococcus multilocularis* surveillance reports submitted in 2014 in the context of Commission Regulation (EU) No 1152/2011	2014	1	1				
Heisey D.M., Jennelle C.S., Russell R.E., Walsh D.P.	Using auxiliary information to improve wildlife disease surveillance when infected animals are not detected: A Bayesian approach	2014	1					1
Rodríguez-Prieto V, Vicente-Rubiano M, Sánchez-Matamoros A, Rubio-Guerri C, et al.	Systematic review of surveillance systems and methods for early detection of exotic, new and re-emerging diseases in animal populations.	2014	1					1
Martin P.A.J., Langstaff I., Iglesias R.M., East I.J., et al.	Assessing the efficacy of general surveillance for detection of incursions of livestock diseases in Australia	2015	1					
de Vos C.J., van der Goot J.A., van Zijderveld F.G., Swanenburg M., et al.	Risk-based testing of imported animals: A case study for bovine tuberculosis in The Netherlands	2015	1	1				
Craighead L, Gilbert W, Subasinghe D, Häsler B.	Reconciling surveillance systems with limited resources: an evaluation of passive surveillance for rabies in an endemic setting	2015	1	1				
Foddai A., Nielsen L.R., Willeberg P., Alban L.	Comparison of output-based approaches used to substantiate bovine tuberculosis free status in Danish cattle herds	2015	1			1		
Rivière J, Le Strat Y, Dufour B, Hendrikx P.	Sensitivity of Bovine Tuberculosis Surveillance in Wildlife in France: A Scenario Tree Approach	2015	1	1				

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Schärrer S., Widgren S., Schwermer H., Lindberg A., et al.	Evaluation of farm-level parameters derived from animal movements for use in risk-based surveillance programmes of cattle in Switzerland	2015	1					
More S.J., Cameron A.R., Strain S., Cashman W., et al.	Evaluation of testing strategies to identify infected animals at a single round of testing within dairy herds known to be infected with *Mycobacterium avium* ssp. paratuberculosis	2015	1			1		
Marques A.R., Pereira M., Ferreira Neto J.S., Ferreira F.	Design and prospective evaluation of a risk-based surveillance system for shrimp grow-out farms in northeast Brazil	2015	1					
Oseguera Montiel D, Bruce M, Frankena K, Udo H, et al.	Financial analysis of brucellosis control for small-scale goat farming in the Bajío region, Mexico	2015	1	1				
Rossi G., De Leo G.A., Pongolini S., Natalini S., Vet al.	Epidemiological modelling for the assessment of bovine tuberculosis surveillance in the dairy farm network in Emilia-Romagna (Italy)	2015	1			1		
Santman-Berends IM, Mars MH, van Duijn L, van Schaik G.	Evaluation of the epidemiological and economic consequences of control scenarios for bovine viral diarrhea virus in dairy herds	2015	1					
Ojimelukwe AE, Prakarnkamanant A, Rushton J.	Estimating the sensitivity of passive surveillance for HPAI H5N1 in Bayelsa state, Nigeria	2016	1	1				
Sonnenburg J., Schulz K., Blome S., Staubach C.	The challenge of detecting classical swine fever virus circulation in wild boar (*Sus scrofa*): Simulation of sampling options	2016	1					
El Allaki F, Harrington N, Howden K.	Assessing the sensitivity of bovine tuberculosis surveillance in Canada's cattle population, 2009-2013	2016	1			1		

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Alban et al	Comparison of risk-based versus random sampling in the monitoring of antimicrobial residues in Danish finishing pigs	2016	1	1				
Christensen J, Vallières A.	Scenario tree model for animal disease freedom framed in the OIE context using the example of a generic swine model for Aujeszky's disease in commercial swine in Canada	2016	1					
Lyngstad T.M., Hellberg H., Viljugrein H., Bang Jensen B., et al.	Routine clinical inspections in Norwegian marine salmonid sites: A key role in surveillance for freedom from pathogenic viral haemorrhagic septicaemia (VHS)	2016	1					
Miller A.L., Olsson G.E., Sollenberg S., Skarin M., et al.	Support for targeted sampling of red fox (*Vulpes vulpes*) feces in Sweden: A method to improve the probability of finding *Echinococcus multilocularis*	2016	1	1				
VanderWaal K., Enns E.A., Picasso C., Alvarez J., et al.	Optimal surveillance strategies for bovine tuberculosis in a low-prevalence country	2017	1	1				
Arruda A.G., Poljak Z., Knowles D., McLean A.	Development of a stochastic agent-based model to evaluate surveillance strategies for detection of emergent porcine reproductive and respiratory syndrome strains	2017	1					
Rivière J, Le Strat Y, Hendrikx P, Dufour B.	Cost-effectiveness evaluation of bovine tuberculosis surveillance in wildlife in France (Sylvatub system) using scenario trees	2017	1	1				
Schulz K., Peyre M., Staubach C., Schauer B., et al.	Surveillance strategies for Classical Swine Fever in wild boar-a comprehensive evaluation study to ensure powerful surveillance	2017	1					
Bisdorff B., Schauer B., Taylor N., Rodríguez-Prieto V., et al.	Active animal health surveillance in European Union Member States: Gaps and opportunities	2017	1					1

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Hénaux V., Calavas D.	Evaluation of the cost-effectiveness of bovine brucellosis surveillance in a disease-free country using stochastic scenario tree modelling	2017	1	1				
Veldhuis A., Santman-Berends I., Schauer B., Mars J., et al.	Epidemiological performance and subsequent costs of different surveillance strategies to control bovine herpesvirus type 1 in dairy farms	2017	1					
Welby S, van Schaik G, Veldhuis A, Brouwer-Middelesch H, et al.	Effectiveness and Cost Efficiency of Different Surveillance Components for Proving Freedom and Early Detection of Disease: Bluetongue Serotype 8 in Cattle as Case Study for Belgium, France and the Netherlands	2017	1					
Hansen R.K., Nielsen L.H., El Tholth M., Haesler B., et al.	Comparison of alternative meat inspection regimes for pigs from non-controlled housing - Considering the cost of error	2018	1	1				
Rüegg S.R., Welby S., Yassin H., Van der Stede Y., et al.	Optimising cost-effectiveness of freedom from disease surveillance—Bluetongue Virus Serotype 8 as an example	2018	1					
Nugent G., Gormley A.M., Anderson D.P., Crews K.	Roll-back eradication of bovine tuberculosis (TB) From wildlife in New Zealand: Concepts, evolving approaches, and progress	2018	1	1				
Kao S.-Y.Z., VanderWaal K., Enns E.A., Craft M.E., et al.	Modeling cost-effectiveness of risk-based bovine tuberculosis surveillance in Minnesota	2018	1	1				
Ågren E.C.C., Lewerin S.S., Frössling J.	Evaluation of herd-level sampling strategies for control of *Salmonella* in Swedish cattle	2018	1	1				
Salvador L.C.M., Deason M., Enright J., Bessell P.R., Kao R.R.	Risk-based strategies for surveillance of tuberculosis infection in cattle for low-risk areas in England and Scotland	2018	1	1				
Schulman K, Lyytikäinen T.	The effect of genetic susceptibility and targeting of sampling on the sensitivity of the surveillance system and	2018	1					

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
	certainty-of-freedom for classical scrapie in Finland in 2008-2014							
Willeberg PW, McAloon CG, Houtsma E, Higgins I, et al.	The Herd-Level Sensitivity of Abattoir Surveillance for Bovine Tuberculosis: Simulating the Effects of Current and Potentially Modified Meat Inspection Procedures in Irish Cattle	2018	1	1				
Toftaker I, Ågren E, Stokstad M, Nødtvedt A, et al.	Herd level estimation of probability of disease freedom applied on the Norwegian control program for bovine respiratory syncytial virus and bovine coronavirus	2018	1					
Tratalos JA, Barrett DJ, Clegg TA, O'Neill RG, et al.	Sampling Methodology to Maximize the Efficient Use of National Abattoir Surveillance: Using Archived Sera to Substantiate Freedom From Bluetongue Virus Infection in Ireland	2018	1					
Comin A., Grewar J., Schaik G.V., Schwermer H., et al.	Development of Reporting Guidelines for Animal Health Surveillance—AHSURED	2019	1					1
Napp S., Ciaravino G., Pérez de Val B., Casal J.,et al.	Evaluation of the effectiveness of the surveillance system for tuberculosis in cattle in Spain	2019	1	1				
Isoda N, Asano A, Ichijo M, Ohno H, et al.	Assessment of the cost effectiveness of compulsory testing of introduced animals and bulk tank milk testing for bovine viral diarrhea in Japan	2019	1					
Häsler B., Garza M., Bisdorff B., Léger A., et al.	Assessing the Adoption of Recommended Standards, Novel Approaches, and Best Practices for Animal Health Surveillance by Decision Makers in Europe	2019	1					1

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Meyer A., McAloon C.G., Tratalos J.A., More S.J., et al.	Modeling of alternative testing strategies to demonstrate freedom from *Mycobacterium avium* ssp. paratuberculosis infection in test-negative dairy herds in the Republic of Ireland	2019	1	1				
Grewar J.D., Sergeant E.S., Weyer C.T., van Helden L.S., et al.	Establishing post-outbreak freedom from African horse sickness virus in South Africa's surveillance zone	2019	1					
van Roon AM, Santman-Berends IMGA, Graham D, More SJ, et al.	STOC Free: An Innovative Framework to Compare Probability of Freedom From Infection in Heterogeneous Control Programmes	2019	1					1
Yang DA, Johnson WO, Müller KR, Gates MC, et al.	Estimating the herd and cow level prevalence of bovine digital dermatitis on New Zealand dairy farms: A Bayesian superpopulation approach	2019	1					
de la Cruz ML, Pozo P, Grau A, Nacar J, et al.	Assessment of the sensitivity of the bovine tuberculosis eradication program in a high prevalence region of Spain using scenario tree modeling	2019	1	1				
Jamin C, Rivière J.	Assessment of bovine tuberculosis surveillance effectiveness in French wildlife: An operational approach	2020	1	1				
Epstein A., Namuganga J.F., Kamya E.V., Nankabirwa J.I., et al.	Estimating malaria incidence from routine health facility-based surveillance data in Uganda	2020			1			
Costa L., Duarte E.L., Knific T., Hodnik J.J., et al.	Standardizing output-based surveillance to control non-regulated cattle diseases: Aspiring for a single general regulatory framework in the European Union	2020	1					1

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
van Roon A.M., Santman-Berends I.M.G.A., Graham D., More S.J., et al.	A description and qualitative comparison of the elements of heterogeneous bovine viral diarrhea control programs that influence confidence of freedom	2020	1					
Grewar JD, Porphyre T, Sergeant ES, Theresa Weyer C, et al.	Post-outbreak African horse sickness surveillance: A scenario tree evaluation in South Africa's controlled area.	2020	1					
Belsare A., Gompper M., Keller B., Sumners J.,et al.	Size matters: Sample size assessments for chronic wasting disease surveillance using an agent-based modeling framework	2020	1					
Cameron A.R., Meyer A., Faverjon C., Mackenzie C.	Quantification of the sensitivity of early detection surveillance	2020	1					1
Foddai A., Lubroth J., Ellis-Iversen J.	Base protocol for real time active random surveillance of coronavirus disease (COVID-19) – Adapting veterinary methodology to public health	2020			1			
Foddai A., Lindberg A., Lubroth J., Ellis-Iversen J.	Surveillance to improve evidence for community control decisions during the COVID-19 pandemic – Opening the animal epidemic toolbox for Public Health	2020			1			
Rosendal T., Widgren S., Ståhl K., Frössling J.	Modelling spread and surveillance of *Mycobacterium avium* subsp. paratuberculosis in the Swedish cattle trade network	2020	1	1				
Mastin AJ, Gottwald TR, van den Bosch F, Cunniffe NJ, et al.	Optimising risk-based surveillance for early detection of invasive plant pathogens	2020					1	
Jordan A.G., Citer L.R., McAloon C.G., Graham D.A.,et al.	Johne's disease in Irish dairy herds: Considerations for an effective national control programme	2020	1	1				

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Guétin-Poirier V, Crozet G, Gardon S, Dufour B, et al.	Integrating data of veterinarians' practices in assessing the cost effectiveness of three components of the bovine tuberculosis surveillance system by intradermal tuberculin testing in French cattle farms through a scenario-tree approach	2020	1	1				
Guétin-Poirier V, Rivière J, Dufour B.	Cost-effectiveness of two different protocols for animal tracing investigations of bovine tuberculosis outbreaks in France	2020	1	1				
Foddai A., Floyd T., McGiven J., Grace K.,et al.	Evaluation of the English bovine brucellosis surveillance system considering probability of disease introduction and non-random sampling	2020	1	1				
EFSA	General guidelines for statistically sound and risk-based surveys of plant pests	2020					1	1
Calistri P., De Clercq K., Gubbins S., Klement E., et al.	Lumpy skin disease epidemiological report IV: data collection and analysis	2020	1					
Rapaliute E., van Roon A., van Schaik G., Santman-Berends I.,et al.	Existence and Quality of Data on Control Programs for EU Non-regulated Cattle Diseases: Consequences for Estimation and Comparison of the Probability of Freedom From Infection	2021	1	1				
Viljugrein H., Hopp P., Benestad S.L., Våge J., et al.	Risk-based surveillance of chronic wasting disease in semi-domestic reindeer	2021	1					
Desvaux S., Urbaniak C., Petit T., Chaigneau P., et al.	How to Strengthen Wildlife Surveillance to Support Freedom From Disease: Example of ASF Surveillance in France, at the Border With an Infected Area	2021	1					

Authors	Title	Year	Animal	Zoonotic	Public	Food safety	Plant	Review
Nielsen S.S., Alvarez J., Bicout D.J., Calistri P., et al.	ASF Exit Strategy: Providing cumulative evidence of the absence of African swine fever virus circulation in wild boar populations using standard surveillance measures	2021	1					
Garner G, Vosloo W, Tapsuwan S, Bradhurst R, et al.	Comparing surveillance approaches to support regaining free status after a foot-and-mouth disease outbreak	2021	1					
Farber JM et al	Alternative approaches to the risk management of Listeria monocytogenes in low risk foods	2021			1			
Scaramozzino, P., Carvelli, A., Bruni, G. et al.	West Nile and Usutu viruses co-circulation in central Italy: outcomes of the 2018 integrated surveillance	2021	1	1				
Pineda P, Santa C, Deluque A, Peña M, et al.	Evaluation of the sensitivity of the classical swine fever surveillance system in two free zones in Colombia	2021	1					
van Roon et al	Key Learnings During the Development of a Generic Data Collection Tool to Support Assessment of Freedom of Infection in Cattle Herds	2021	1					1
Ferri M., Lloyd-Evans M.	The contribution of veterinary public health to the management of the COVID-19 pandemic from a One Health perspective	2021	1	1				1
Graham D, More SJ, O'Sullivan P, Lane E, et al.	The Irish Programme to Eradicate Bovine Viral Diarrhoea Virus-Organization, Challenges, and Progress	2021	1					
Autio T et al	Overview of Control Programs for Cattle Diseases in Finland	2021	1	1				
Santman-Berends IM, Mars MH, et al	Control and Eradication Programs for Six Cattle Diseases in the Netherlands	2021	1	1				

Appendix 2 – Abbreviations and acronyms

AP	Apparent prevalence
APHA	Animal and Plant Health Agency
BTM	Bulk tank milk
bTB	Bovine tuberculosis
CEA	Cost-effectiveness analysis
CL	Confidence level
DEFRA	Department for Environment, Food and Rural Affairs
ECDC	European Centre for Disease Prevention and Control
EF	Egg flotation technique
EFSA	European Food Safety Authority
EFTA	European Free Trade Association
EURLP	European Union Reference Laboratory for Parasites
IST	Intestinal scraping technique
NRL	National Reference Laboratory
OBS	Output based standard
OH	One Health
OHEJP	One Health European Joint Programme
RiBESS	Risk based estimate of system sensitivity tool
RISKSUR	Risk-based animal health surveillance systems
SASA	Science & Advice for Scottish Agriculture
SCT	Sedimentation and Counting Techniques
SE	Sensitivity
SERVAL	Surveillance evaluation framework
SOP	Standard Operating Procedure
SP	Specificity
UKAS	United Kingdom Accreditation Service
UKHSA	United Kingdom Health Security Agency
WOAH	World Organisation for Animal Health