SERVAL: A generic framework for the evaluation of animal health surveillance

Incorporating a worked example from a case study of pre-movement testing for bovine tuberculosis in England and Wales

Developed by staff at the Royal Veterinary College, the Animal Health and Veterinary Laboratories Agency, and the Scottish Agricultural College, UK

Report compiled by Julian Drewe[1] (RVC), Katharina Stärk[2] (RVC), and Toby Floyd[3] (AHVLA)
1. Email: jdrewe@rvc.ac.uk. 2. Email: kstaerk@rvc.ac.uk. 3. Email: Tobias.Floyd@ahvla.gsi.gov.uk.

For more detailed information on the development and use of the SERVAL framework, please see:
Drewe, J.A., Hoinville, L.J., Cook, A.J.C., Floyd, T., Gunn, G. and Stärk, K.D.C. (2013) SERVAL: A new framework for the evaluation of animal health surveillance. *Transboundary and Emerging Diseases* DOI: 10.1111/tbed.12063 (Click here to read the abstract online)

Contents

A. Introduction

Animal health surveillance programmes tend to evolve but are rarely evaluated to ensure that they provide valuable information in an efficient manner. The evaluations that are conducted are often unstructured and therefore incomplete. To address this gap, we have developed SERVAL, a SuRveillance EVALuation framework, which is novel and aims to be generic and therefore suitable to evaluate any animal health surveillance system.

The SERVAL framework has been designed to assist in the comprehensive evaluation of single surveillance components (activities) or entire surveillance programmes. It is intended to be flexible so it may be used to evaluate surveillance for any health condition in any species (or group of species). The inclusion of socio-economic criteria ensures that economic evaluation is an integral part of this framework. It is anticipated that SERVAL's main use will be in the evaluation of existing surveillance systems but we hope it will also prove useful in the design stage of new surveillance programmes. For surveillance systems already in operation, the evaluation should focus on how they are actually implemented which may differ from how they were designed to be implemented.

SERVAL was developed following a technical workshop of international experts followed by a consultation process involving providers and users of surveillance and evaluation data. In order to explore the practicality

and utility of the SERVAL framework, the framework was applied to three case studies – encompassing a range of diseases, species, surveillance aims and evaluation objectives. This document presents SERVAL illustrated by one case study example: Pre-movement testing of cattle for bovine tuberculosis in England and Wales.

The framework is organised in five sections each containing detailed guidance notes and examples to assist the evaluator(s):

1. Define the scope of the evaluation
2. Characterise the surveillance system to be evaluated
3. Design the evaluation
4. Conduct the evaluation
5. Reporting and communication

Anyone familiar with epidemiological concepts and with a reasonable knowledge of the disease under surveillance should be able to use this framework to conduct an evaluation. The evaluation process is likely to require input from people who work on the surveillance system being evaluated. The output of using the framework is expected to be a written evaluation report that includes details of each of the sections listed below under standardised headings. The contents of each section can be tailored to the needs of each surveillance system and so will vary. The evaluation report should be circulated to affected parties including both those implementing the surveillance activities and those using the outputs.

B. The SERVAL framework illustrated with an evaluation case study

Framework section	*Guidance notes*	*Case study: Pre-movement testing for Tuberculosis in cattle*
1. Define the scope of the evaluation	Complete sections a – f below.	–
a. State the evaluation objective(s)	Choose from the following list of six evaluation objectives. This list aims to cover all possible evaluation objectives but excludes higher level strategic decisions, for example "to determine if a surveillance system, or a component of it, should be stopped", which would be based on the output of the evaluation. One or more of the objectives listed here could inform such a decision. For example, "to determine if a surveillance system should be stopped" might be answered by any of objectives i, ii, iii, or iv, and the most relevant one(s) should be chosen.	–
	i. To ascertain if a surveillance system is meeting its objectives. If this is the objective of the evaluation then the objective of the surveillance should be stated	Yes, this is one of the objectives of this evaluation: to ascertain if the pre-movement bovine tuberculosis (bTB) testing of cattle surveillance system is meeting its objectives to detect cases of bTB in animals being moved between farms.
	ii. To ascertain if a foreign surveillance system is reliable enough to accept imports from that country, or if a domestic surveillance system is good enough to support the export of animals or their products.	Not an objective of this evaluation.
	iii. To ascertain if a surveillance system is providing value for money to the funder.	To ascertain if the bTB surveillance system provides value for money to Defra.
	iv. To determine how much benefit (monetary or otherwise) a surveillance system provides to its user groups.	Not an objective of this evaluation.
	v. To identify the strengths and deficiencies of a surveillance system.	Yes, this is one of the objectives of this evaluation.
	vi. To identify potential measures that could improve the performance, efficiency and productivity of a surveillance system.	Not an objective of this evaluation.
b. Formulate the evaluation question	Phrase the evaluation objective as a specific question in a format that the evaluation can seek to address. Where an evaluation is seeking to determine whether a surveillance system meets its objectives the surveillance objective should be clearly stated within the evaluation question.	Does pre-movement bTB testing (PrMT) of cattle from herds in annual and biennial testing interval areas reduce the risk of the spread of bTB in England and Wales?
c. Indicate the motivation for the evaluation	State what prompted the evaluation to be undertaken.	This evaluation is being conducted to explore the practicality and utility of the draft evaluation framework.
d. Define the	Identify the people involved in the evaluation. In some cases a single	This case-study was identified by the SE4302 project team and agreed by the project board, including Jane

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
organisation of the evaluation	body may be responsible for requesting, commissioning and funding the evaluation. In other cases, the body who requests the evaluation may be different to the body who commissions it who in turn may be different to the body that funds it. - Who requested the evaluation? - Who commissioned the evaluation? - Who is funding the evaluation? - Who will lead the evaluation? - Who will contribute to the evaluation? - What other personnel support and administration will be required? - Who will be responsible for communication and reporting? - Who will benefit from the evaluation outputs? Indicate how the engagement of each of these people will be secured.	Gibbens (Customer), John Kinnaird, Katharina Stärk, George Gunn and Alex Cook. This case-study is being conducted as part of project SE4302; a research project funded by Defra and conducted in collaboration between AHVLA, RVC and SAC. The case-study is being led by T Floyd (AHVLA Project Leader) with support from Elizabeth Ely, Kate Harris and Andy Mitchell, as experts in bTB surveillance. Project SE4302 and the evaluation framework under development will be the beneficiary of this case-study. The case-studies will be used to reflect upon the feasibility and utility of and to identify potential improvements to the evaluation framework. These case-studies may also be presented alongside the framework as worked examples of how the framework can/should be applied.
e. Identify the time and resources available for the evaluation	Indicate the staff, funds and time available for the evaluation. Identify the evaluation timeframe, including the start date, delivery date and any interim deadlines.	Thirty-three days or ~£16,500 is available from the SE4302 project budget for this evaluation case-study. The evaluation will be conducted between the 26th of January and the 31st of March 2012, with the draft report on the evaluation case-studies due for submission on the 31st of March 2012 (SE4302 milestone 6).
f. State what will be done with the evaluation outputs	This should be linked with the evaluation objective(s) (Section 1a) and indicate the purposes to which the evaluation results could be put. Thought should be given to how the findings of the evaluation will be reported. The evaluation outputs should be reported in an appropriate format to all relevant parties (see Section 5e).	The main purpose of this evaluation is to exercise the draft evaluation framework. The experience and outputs of this evaluation will be considered and used to produce a set of recommendations for improving the draft framework and included in a report to the project customer (project milestone 4). The case-study may also be presented with the evaluation framework as a worked example or guidance to prospective users of the framework.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
2. Characterise the surveillance system to be evaluated	Complete sections a – g below.	–
a. State the name of the surveillance system to be evaluated.	If just a component of the surveillance system is to be evaluated, name that component, but also indicate which other components are in the surveillance system and indicate how the components relate to each other.	Pre-movement bTB testing in cattle (PrMT) will be evaluated. PrMT is one activity in the approach to mitigation of bTB in GB: wherein the SICCT is applied to cattle moving from higher-risk herds. PrMT is a risk mitigation strategy (to prevent the spread of bTB) rather than purely a surveillance activity and so has objectives other than the generation of surveillance data; but this component still provides information on the health status of animals and herds and the distinction between surveillance and control can sometimes be arbitrary. The surveillance portfolio for bTB in Britain is varied and complex. Other surveillance and mitigation activities focussed on bTB in GB include: Routine and Whole Herd Skin Testing (RHT) and Slaughterhouse inspection for gross pathology. **Note: Requirements for PrMT were amended subsequent to conducting this evaluation (**http://www.defra.gov.uk/animal-diseases/files/tb-infonote-1202-pre-movement.pdf **) and so some of the information provided on PrMT will be out of date.**
b. State the context of the health condition under surveillance	Provide the following information: - Name the health condition under surveillance. This could be a disease, infection, condition or event. Most are likely to be infectious but this framework is designed to apply equally to non-infectious conditions. - State the causal organism or factor (if known). - Indicate if it is zoonotic. - Indicate the context of the health condition under surveillance. Possibilities include: **i.** New (or emerging) health conditions: not previously recognised; not currently recorded as present; may result from the evolution or change in an existing disease agent causing a change of strain, host range, vector, or increase in pathogenicity; or may be the occurrence of any other previously undefined condition. **ii.** Re-emerging health conditions: previously defined conditions that were either absent and have recently re-appeared, or were present at a low level in the population in a defined geographical area and are markedly increasing in prevalence. **iii.** Endemic health conditions: known to be constantly present in the population of interest. **iv.** Exotic health conditions: previously known conditions that cross political boundaries to occur in a country or region in which they are not currently recorded as present.	Bovine tuberculosis in cattle caused by infection with Mycobacterium bovis. Bovine tuberculosis is a notifiable disease and is endemic in large parts of England and Wales. It can affect other mammalian species, including other farmed animals (occasional), companion animals (occasional) and wildlife (badgers in particular are an important reservoir for the disease). It is also a zoonosis, although the risk of contracting the disease in UK is very low as a result of milk pasteurisation and other controls. Bovine tuberculosis is transmitted primarily through exposure to respiratory excretions from infected animals during close contact. The disease is chronic, with clinical signs rarely observed in cattle in GB. Infection is usually first evident in cranial lymph nodes, followed by gross lesions in organs such as the lungs, which may be detected at post-mortem. Consequently, detection of the disease is dependent upon a programme of routine skin testing with the single intradermal cervical comparative tuberculin (SICCT) test and carcase inspection in the slaughterhouse. Detection of bTB has significant economic implications for the farm business. If a reactor is found through pre-movement testing, the herd will be placed under movement restrictions: cattle may only be moved off the premises under license. The Animal Health and Veterinary Laboratories Agency (AHVLA) will arrange for any reactor animals in the herd to be valued and slaughtered. In some cases, non-reactors may be removed as direct contacts. The rest of the herd will undergo a series of tests until AHVLA are satisfied that the herd is free from bTB and restrictions can be lifted (http://animalhealth.defra.gov.uk/about/publications/advice-guidance/documents/7_What_further_testing_.pdf). For dairy herds the farmer must ensure that milk from animals identified as reactors does not enter the human food chain; until movement restrictions are lifted the herd owner must not sell unpasteurised milk to consumers or for use in the manufacturing of unpasteurised milk products, the herd owner must inform their milk purchaser as soon as reactors are identified in the herd.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
c. Summarise the current situation	Briefly summarise the current problem with the health condition under surveillance. - Why is it considered to be a problem? - Briefly indicate the level of current knowledge of the condition. - Identify the Policy objective of the surveillance programme. Examples are given below. The policy objective describes how surveillance information is used by policy makers to inform decisions about how best to support a healthy and sustainable food and farming industry in order to protect the livelihood of producers, other value chain stakeholders and public health and to contribute to national economic development. The specific decisions that surveillance information can assist policy makers with include (but may not be limited to): **i.** Management of outbreaks - whether additional control measures are required to limit the spread of an emerging or exotic disease outbreak. **ii.** Informing trade - whether to permit import or support export of animals or animal products based on the evidence about the prevalence and distribution of disease in the population and the risk of disease spread through the commodity being traded. **iii.** Prioritisation - how to prioritise surveillance and control measures for different health events based on their level of occurrence and impact on animal health and welfare, public health, trade and the wider economy. **iv.** Informing control - Whether the current control measures for particular diseases are effective or should be changed.	Bovine tuberculosis is a serious animal health problem in England and Wales and is the single largest component of animal health related expenditure (>£91million of public money for England in 2010-11). The incidence of herd bTB breakdowns has been increasing over the past 25 years and bTB is now well established in areas of southwest England and Wales; where over 22% of cattle herds were subject to restrictions at some point during 2010. Whilst large areas of north, east and south-east of England experience few breakdowns, there is evidence that the endemic area is spreading north and eastwards. The overall aim of surveillance for bTB in England and Wales is to inform the development of animal health policy on disease control: to protect public health and public confidence in dairy and beef products; to maintain a productive and sustainable cattle industry; to reduce the cost of disease to farmers and taxpayers; to meet domestic and international legal requirements; and to promote animal health and welfare. The more specific policy objective of PrMT is to reduce the risk of spreading bTB (by infected animals) between herds and to areas currently free of infection.
d. Identify the surveillance objective(s)	Choose from the following list of six surveillance objectives. This list aims to cover all possible surveillance objectives but excludes higher level aims, for example, “safeguarding public health” or “maintaining animal welfare” or “prioritisation of threats and resources” which would be decisions to be made at a higher level based on the output of the evaluation. These surveillance objectives are also distinct from higher level aims and policy objectives which are informed by information provided by surveillance activities. Such higher-level aims should be matched to one or more of the objectives listed here. In the example of “prioritisation of threats and resources”, any of objectives i, iii, v, or vi might apply.	–
	i. Monitor the prevalence of infection. While usually aimed at endemic infections, this is also applicable	No, this is not an objective of this surveillance system.

Framework section	*Guidance notes*	*Case study: Pre-movement testing for Tuberculosis in cattle*
	to new and re-emerging infections and may form part of an assessment of the impact of control programmes on infection incidence.	
	ii. Case finding of infected animals: detection of as many cases as possible of a known infection to facilitate control. The emphasis here is on finding those individuals who are infected in order to intervene in some way such as culling or vaccination. This will usually apply to an endemic disease.	Yes, case-finding of infected animals is an objective of this surveillance system. The primary surveillance objective of PrMT is to detect as many cases of bTB as possible; this is coupled directly to the outcome of mitigating the risk of spread of bTB through the movement of infected cattle between holdings.
	iii. Early detection of new or re-emerging infection. Early detection could be defined as detection of infection before an outbreak becomes uncontrollable: this timeframe will vary by health condition and should be estimated. If this objective is chosen, a statement should be included to define how early the system aims to detect infection.	No, this is not an objective of this surveillance system.
	iv. Demonstrate freedom from infection. If this objective is chosen, a statement should be included to define the prevalence and associated confidence level which are considered to indicate disease freedom. These concepts are presented with examples in: Dufour B, et al. Proposed criteria to determine whether a territory is free of a given animal disease. Veterinary Research 32: 545-563.	No, this is not an objective of this surveillance system.
	v. Identify changes in the population at risk. Here, risk factors rather than an infectious agent are the target for surveillance. This might lead to Identification of new population groups at risk and in need of targeted prevention measures.	No, this is not an objective of this surveillance system.
	vi. Improve epidemiological understanding of a disease. Generating knowledge about a disease, for example academic research or hypothesis generation. It is anticipated that this objective will usually relate to a new health condition.	No, this is not an objective of this surveillance system.
e. Specify the target population for the surveillance system	This is the animal population which the surveillance system was designed to cover. Quantify it as precisely as possible including species, breed, age, sex, production type and geographical location. It may be helpful to indicate if the target population is vertically or horizontally integrated. Vertically integrated means a single producer raises animals from birth through to death (e.g. for fish this would include the hatchery, smoltery and marine pens) and therefore the one producer is a single epidemiological unit. Horizontally integrated means several producers each farm a different life stage (and	The unit of interest operates at both the animal and herd level, as the result of animal-level testing will affect the status of the herd. The target population is all cattle 42 days of age or over moving from holdings in 1 and 2 yearly testing intervals (TIs) in England and Wales. Scotland is recognised as Officially TB Free (OTF) and does not undertake PrMT. The number of herds classified as TI 1-2 varies annually (e.g. from 36911, or 43% of GB herds, in 2009 to 44186, or 53%, in 2010) and is increasing; although there are factors other than the burden and distribution of bTB that affect trends in the classification of cattle herds (e.g. changes in policy).

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
	therefore each is a separate epidemiological unit).	Exemptions from this requirement for PrMT include: - cattle that would be subject to 3 or 4 yearly testing if not for reasons of public health (e.g. producers of raw milk) - those moving off the premises within 30 days of arrival - cattle at artificial insemination centres - moving to exempt or approved finishing units, markets or approved bTB collection centres - moving direct to slaughter - moving to agricultural shows - moving under specific veterinary license (e.g. for treatment) - cattle moving within a shared occupancy agreement group or between premises sharing rights of the same common - cattle that have been tested within the past 60 days for other reasons (e.g. RHT).
f. Describe the structure of the surveillance system	Give details of how the surveillance system works by detailing the components present in each of these four categories: i. Data collection (inputs); ii. Data management (processes); iii. Data analysis (outputs); iv. Data dissemination (outcomes). Consider presenting this information in flow-chart format. During this process it may help to think about the characteristics of the surveillance system in these three areas: i. Agent (infectious/non-infectious, incubation time, life cycle); ii. Host/herd (susceptibility, contacts); iii. Sampling (test quality, sample size, sample frequency).	Cows (42 days and older) moving from herds with 1- or 2-yearly testing intervals are required to have been tested negative to bTB within 60 days prior to movement. These cows might be tested at a routine or whole herd test, provided this falls within the required time window, otherwise the herd owner must arrange for the PrMT to be performed by an Official veterinary surgeon (OVS). The OVS performs SICCTs on the eligible cattle on the farm of origin. The test is read 72 hours after the injection of tuberculin. Results of the SICCT for each animal are recorded and a copy is kept by the herd owner. The OVS reports the PrMT, the number of cattle tested and their classification (positive, inconclusive or negative reactor) to the AHVLA and these data are recorded in the central VetNet database. The results of PrMT must be reported to AHVLA within 1 day for positive and inconclusive reactors or 5 days for PrMT where all cattle tested were non-reactors. More animal level data (e.g. animal ID and the skin measurements) are recorded on the practises own TBMaster database or, occasionally, on the central VeBus database. From September 2011, a new data system (SAM) will collate all testing information in a single database and also allow the data input directly by the OVS. Descriptive analyses of bTB surveillance data collated in VetNet are conducted and published on the Defra external website on a monthly basis. More in-depth epidemiological analyses of bTB surveillance data are conducted on a 6-monthly basis. These surveillance reports are published on the Defra secured intranet, with a summary form published on the Defra external website; the Welsh Government publishes the full TB surveillance report for Wales on its external website.
	Data collection: Use of appropriate data sources and collection methods and the existence of a case definition and data collection protocol. Consider each of the following: - Who provides the data? - Who collects the data?	SICCT performed on animals at the origin farm by the Official Veterinary Surgeon (OVS). The tester returns to assess the result of SICCT for each cow 72 hours following the injection of tuberculin. A standardised method and interpretation of the test are prescribed but compliance with the protocol is not monitored.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
	- Where are data collected? - How are data collected? - How are data recorded (e.g., on paper or electronically)? - What type of data are being dealt with (e.g., active/passive, threat-specific/syndromic)? - Is there a data collection protocol? - How are staff trained to collect data? - What is the case definition?	
	Data management: Use and documentation of systems for processing information, including data processing protocols and data verification procedures. Consider each of the following: - How are data managed? - What data security measures are in place? - How are data stored? - How are this documented? - Are quality assurance procedures followed? - Are there data processing protocols? - Describe the data verification procedures.	The data management system is undergoing significant changes currently. Prior to 2011 it consisted of several databases (VetNet, VeBus and numerous practice-based TBMaster databases) which are to be replaced by a single central database (SAM). Results (including animal ID, test results and interpretation) of on-farm testing are entered into two central databases (VetNet and VeBus) by AHVLA staff from paper copies of on-farm testing results submitted by OVSs or AH veterinary officers. For many veterinary practices animal-level data is recorded and stored in the practice's TBMaster databases and this data is not routinely collated centrally. Surveillance data stored in VetNet and VeBus are periodically assessed for validity and accuracy. Compliance is monitored by AHVLA Gloucester through comparison of bTB testing data (SAM/VetNet) and cattle movement data (CTS). A proportion (5%) of apparently non-compliant movements is investigated further – most of which are subsequently found to be compliant – and a very small proportion of these are referred to local authorities for follow-up.
	Data analysis: Methods used for the analysis and interpretation of data. Consider each of the following: - How are data analysed and interpreted? - Are performance indicators used and if so, which ones and how are they calculated?	Basic descriptive statistical analyses of bTB surveillance data collated in SAM/VetNet are conducted on a monthly basis: examining trends in the incidence of herd breakdowns and the prevalence of herds under restrictions. In addition, more in-depth epidemiological analyses and interpretation of trends are conducted by AHVLA epidemiologists every 6 months.
	Data dissemination: Methods used for information exchange between people involved at all levels of the surveillance system. Consider each of the following: - Which methods are used to exchange information between people involved in the surveillance system (providers, analysers and users of surveillance data)? These might include: case reporting cards, emails, letters, phone calls, interim reports of surveillance data, websites for disseminating information, and feedback given to the data providers.	Results are available to the farmer/owner at the time of testing. Results inform the permission to move the animals and will also affect the OTF status of the herd: Detection of reactor or inconclusive reactor cattle requires further testing and investigation and finding a reactor results in suspension or withdrawal of herd OTF status. A weekly download of data from VetNet, cleaned and combined with data from other sources (e.g. CTS) is used in a web-based application to visualise and interrogate bTB surveillance data (SPIDA). SPIDA is hosted on the Defra GSI and so can be accessed by anyone in the Defra family. Monthly and annual statistical summaries of bTB surveillance data are published on the Defra website (http://www.defra.gov.uk/statistics/foodfarm/landuselivestock/cattletb/); although PrMT data are not

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
	- How frequently are data or reports disseminated? - To date, what actions (if any) have been taken as a result of the surveillance activity? These might include: details of mitigation measures imposed; decreased incidence of diseases; use of surveillance data for policy and programme decisions; and appropriateness of outbreak response.	included specifically in these monthly statistical summaries. The GB 6-monthly and annual bTB surveillance reports, providing a more in-depth epidemiological analysis and interpretation of the data are produced, and a summary version of them is available within the Defra Website GSI; the annual surveillance report for Wales is published in full on the Welsh Assembly Government website (http://wales.gov.uk/topics/environmentcountryside/ahw/disease/bovinetuberculosis/researchandevidence/?lang=en). Extracts of data from VetNet and VeBus may also be used for epidemiological studies.
g. Identify the study design of the surveillance system	Outline the study design and indicate how the sampling frame and testing protocol are decided. Describe the general structure of the surveillance system including: - origin of data (whether active, passive or enhanced passive) - disease focus (whether hazard-specific or general) - study design (e.g. case reports, survey or continuous collection - sample size calculation and sampling strategy including whether a risk-based strategy is used	Study design: Pre-movement testing is carried out on a continuous basis, as required for the movement patterns of cattle in England and Wales. Approximately 43,944 PrMT tests were performed on 20,543 unique holdings in 2010. The requirement for testing in the English and Welsh PrMT system is based around the bTB testing interval of the herd of origin. Cattle holdings are designated their TI by their location and herd history with regard to bTB incidence; the SAM/VetNet data system is used to identify holdings, their classification by TI and bTB status. Cattle moved from herds in annual or two-yearly routine bTB testing areas must have been skin-tested with a negative test result within the 60 days preceding the movement date.
h. Identify and engage the surveillance system users	Identify the people involved in the surveillance system that is being evaluated: - Who pays for the surveillance? - Who provides the surveillance data? - Who analyses the surveillance data? - Who uses the resulting information? - Who benefits from any action resulting from the surveillance? - Who pays for disease mitigation - Who (if anyone) might lose out if disease is reported (e.g. it might be thought that famers' reputations may be tarnished if they declare disease in their herd)? Identify how these people will be engaged in the evaluation process. Note that the people identified in this section may be different from the people identified in section 1d where the focus was on the people involved in the evaluation itself.	PrMT is funded in part by the government (i.e. the cost of tuberculin, equipment for testing and training of OVSs) and in part by the farmer (i.e. the costs of OVS time to conduct the test, additional labour to handle the animals etc). The farmer does have the option of using government-funded tests (e.g. RHT) to serve as PrMTs if the timing of these tests is appropriate. Surveillance data is provided by the OV and inputted directly to SAM/VetNet via a web portal. Surveillance data collated in SAM/VetNet is available to the AHVLA and is used for many purposes: from visualising distribution of breakdown herds and interrogating specific results (SPIDA) to preparation and publication of bTB statistics on a quarterly and annual basis (see 2 f). The direct beneficiary of the data provided by PrMT is the farmer, who will then be granted to move the (negative) tested cattle. The herd of destination also benefits from the additional assurance that the cattle are a low risk of introducing bTB. As the data contributes to the body of data available from all bTB surveillance activities, AHVLA and Defra also benefit from PrMT: information on the burden and distribution of bTB is used to inform the development of and implementation of bTB policy. Defra pays compensation (market value) for reactor cattle slaughtered but this may not necessarily cover the true value of the animal or the total costs of replacement. The costs of other control measures (e.g. restriction on movement or sale of milk) will be borne by the farmer. These people will not be engaged in this case-study evaluation but under other circumstances they might be engaged through a series of meetings throughout the evaluation process: to plan the evaluation and discuss the results or in connection to specific questions or attributes (e.g. participation).
i. Outline the organisational structure	Indicate who leads and manages the surveillance system being evaluated and briefly describe their roles. Identify whether there are appropriate steering and scientific committees and describe their roles and responsibilities.	The Bovine TB Eradication Advisory Group for England (TBEAG) has been set up to advise on the development and implementation of the strategy for eradicating bovine TB. It is an expert group with responsibility for advising the Animal Health and Welfare Board for England (AHWBE) and Defra ministers. It is a subgroup of AHWBE, which has overall responsibility for TB strategy.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
		The membership of the group includes representatives from the farming industry, the veterinary profession, Defra and the AHVLA. The TBEAG meets several times per year (roughly every 2 months) to review and discuss bTB surveillance data, strategy and control measures. There are also several Working Groups focussing on specific aspects of bTB mitigation.
3. Design the evaluation	Consult with experts in the relevant disease, species or epidemiology of the condition under surveillance, for assistance with selecting and assessing relevant attributes. Guidance for selection of attributes is presented below. It is not the job of the evaluator(s) to set thresholds / targets / success criteria for attributes (this is the job of higher-level decision makers).	–
a. Select the attributes to be assessed	A master list of 22 attributes and their definitions appears in the Appendix. It is not necessary to assess all of these attributes in any single evaluation. The Attribute Selection Matrix (next page) provides a guide to assist the attribute selection process. Attributes have been classified as primary, secondary or tertiary attributes dependent on the surveillance objective. To enable a balanced evaluation, it is suggested that the aim should be to assess all primary attributes listed for that objective. Secondary attributes should be assessed in addition to primary attributes if data and resources allow, but are not essential to the evaluation process. The attribute classifications presented here should be considered as a guide rather than being prescriptive. It may be varied and the exact choice of which attributes to assess is left to the evaluator. The choice of attributes may be influenced by the purpose of the evaluation, the disease type, and the surveillance objective(s). Note that four attributes (benefit, communication, cost, sensitivity) are classified as primary attributes under every surveillance objective and so should be assessed as part of every surveillance evaluation.	For the purposes of this case study, 12 attributes were assessed. Their importance was prioritised into primary, secondary and tertiary attributes as detailed below.

Framework section	***Guidance notes***							***Case study: Pre-movement testing for Tuberculosis in cattle***
	ATTRIBUTE SELECTION MATRIX							Classifications chosen for this case study are indicated below.
Key to classification: ***1 Primary attributes: aim to assess all of these for the chosen surveillance objective.*** ***2 Secondary attributes: assess these attributes if data and resources allow.*** ***3 Tertiary attributes: assess these attributes only if considered important.***	Attribute	Surveillance objective: 1. Monitor prevalence	2. Case finding	3. Early detection	4.Demonstrate freedom	5. Identify changes	6. Improve understanding	
	Benefit	1	1	1	1	1	1	Primary attribute - considered essential
	Bias	1	3	2	1	3	2	Attribute not assessed
	Communication	1	1	1	1	1	1	Primary attribute - considered essential
	Cost	1	1	1	1	1	1	Primary attribute - considered essential
	Coverage	3	1	1	3	3	3	Primary attribute - considered essential
	Data analysis	2	3	3	1	2	2	Attribute not assessed
	Data collection	2	2	2	2	2	2	Secondary attribute - considered important
	Data completeness	2	2	1	1	2	2	Secondary attribute - considered important
	Data management	3	3	3	2	3	3	Tertiary attribute – considered useful
	Flexibility	3	3	1	3	3	3	Attribute not assessed
	Historical data	1	3	1	1	1	1	Attribute not assessed
	Impact	1	2	1	1	2	2	Secondary attribute - considered important
	Lab management	2	2	2	2	3	2	Attribute not assessed
	Multiple utility	2	2	2	2	2	2	Attribute not assessed
	Participation	2	1	1	1	2	2	Primary attribute - considered essential
	Precision	2	3	3	3	2	3	Attribute not assessed
	Repeatability	2	3	3	3	2	3	Attribute not assessed
	Representativeness	1	2	1	1	1	2	Attribute not assessed
	Sensitivity	1	1	1	1	1	1	Primary attribute - considered essential
	Specificity	1	1	1	1	2	1	Primary attribute - considered essential
	Stability/sustainability	1	2	1	2	2	3	Attribute not assessed
	Timeliness	1	1	1	1	1	2	Primary attribute - considered essential

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
b. Clarify which type of economic evaluation will be conducted	An economic evaluation should be an integral part of any surveillance evaluation. Options include: - cost-effectiveness analysis - cost-benefit analysis - qualitative - semi-quantitative - quantitative Information is provided by these attributes: benefit, cost, impact.	We did not have the time and resources available to conduct economic analyses of PrMT in this case study. However, Annex 1 of the 2010 review of PrMT: (http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/pre-movement-testing-review.pdf) and before this, the regulatory impact assessment: (http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/prmt-regulatory.pdf) have conducted cost-benefit analyses of PrMT.
c. Identify which methods and tools will be used for data collection and analysis	Guidance for this part of the framework can be found in the Appendix at the end of this document (page 25 onwards).	Most attributes will be qualitatively assessed given the time and budget constraints faced by the case-study. Where time and data permit, quantitative approaches will be incorporated.
4. Conduct the evaluation	Complete sections a – f below.	–
a. Assess the surveillance objective(s)	Are the surveillance objectives identified in section 2d clearly defined and relevant to disease situation?	Yes, the surveillance objectives identified in section 2d are clearly defined and relevant to disease situation.
b. Collect data	Guidance for this part of the framework will be provided by the outputs of Task 2.3 (additional evaluation tools).	Data were collected using the sources listed in section 3c.
c. Analyse the data	Guidance for this part of the framework will be provided by the outputs of Task 2.3 (additional evaluation tools).	Data were analysed using the techniques summarised alongside each attribute's assessed value in the next section (section 4d).
d. Assess the chosen attributes	Present a summary measure for each attribute which contributes information on the surveillance system's performance. See guidance information for each attribute given in Section 3a.	**Key to traffic-light coding of attributes:** ● Excellent or very good. ● Good, though room for improvement. ● Poor: in need of attention. **Primary attributes for bTB surveillance** 1. Benefit ● Direct benefits of PrMT include: • Reduced risk of introduction of bTB to the destination herd via infected animals • Earlier detection of disease in the origin herd • Reduced spread of bTB via infected cattle; particularly into lower risk areas but also between herds within endemic regions Less tangible and indirect benefits of PrMT include: • Changes to the patterns and timing of cattle movements • Information on the burden and distribution of bTB in the British cattle population • Reduced risk to public health from bTB

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
		The direct benefits of PrMT could be quantified as losses avoided by earlier detection of infected cattle. The average cost of a herd breakdown has been estimated at £30,000, £10,000 of which accrues to the individual farmer (http://www.defra.gov.uk/animal-diseases/a-z/bovine-tb/). It is expected that the benefit for farms in 4 yearly TIs would be greater than those tested annually, as a greater period of time might elapse before introduced cattle are retested. The requirement for PrMT may influence farmers' choices regarding the timing of herd tests, the timing of cattle movements or even whether or not to move a cow; these changes being both benefits and opportunity costs. Due in part to sanitary measures like the pasteurisation of milk, the risk of bTB to public health in Britain is believed to be negligible; and so the benefit of PrMT to safeguarding public health is likely to be similarly small. It is difficult to quantify the economic benefit of the additional information on the burden and distribution of bTB generated by PrMT – and also the additional assurance provided to those purchasing cattle – but these benefits cannot be overlooked. Arguably the benefits of PrMT accrue mostly to the cattle industry and in particular to the destination herds. However, it is difficult to dissect the benefits of PrMT from other aspects of the surveillance and control of bTB in Britain. For further consideration of the costs and benefits of PrMT, see Annex 1 of Defra's 2010 review of PrMT: http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/pre-movement-testing-review.pdf. 2. Communication Communication should be relatively good between herd owners and veterinary practitioners since they will maintain a relationship through other services. The results of testing are available at the time of reading and a copy of the test results are kept by the farmer. Defra provide online and telephone advice services to herd owners regarding all aspects of bTB testing, including what to if a reactor is found in the herd. If a reactor is found in the herd, prompt communication between herd owner and animal health officers is required in valuing the cattle prior to slaughter. Bovine tuberculosis surveillance data are analysed and reported monthly and statistics are published on Defra's website. This includes statistics on the number of reactors, new herd breakdowns and compliance but excludes data from the PrMT. Published statistics on bTB are often reported upon in the farming industry press and online etc, but there are no products of surveillance data by AHVLA with interpretation deliberately tailored to the needs and interests of the farming industry. The 6-monthly and annual bTB surveillance reports – with a more in-depth epidemiological analysis and interpretation of trends – are produced by AHVLA for Defra. These reports cover all TB surveillance data, including PrMT. The annual report is very large and cumbersome because of the volume of bTB surveillance data produced and the various information demands of the customer. Communication of the results of surveillance for bTB generally could be improved within Defra/AHVLA and with the cattle industry through better engagement between AHVLA and Defra, industry stakeholders, industry commentators and opinion leaders.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
		3. Cost ● The average cost of PrMT to the farmer was observed by Bennett (2009) to be £8.85 per animal or £253 per test (batch), half of which is veterinary costs and the other half the cost of labour in handling animals. But this estimate would vary by herd and the number of animals tested per batch. From a fitted curve of the farm-level cost of PrMT produced by Bennett (2009), the average batch of 13 cows would cost the farmer £156 (or £12 per cow). For dedicated PrMT tests, Defra cover the cost of tuberculin: around 90p per animal or £12 per batch of 13. Defra/AHVLA are also responsible for the cost of monitoring compliance and enforcement of the regulation and the costs of managing and reporting on the data generated. These costs have not been quantified here. Defra are also responsible for the cost of training and evaluating bTB-testers and providing the testing equipment. These costs will be divided among other bTB surveillance activities and have not been estimated here. It is important to consider the cost of surveillance in the context of the costs of disease mitigation where infection has been detected. These will include follow-up, contiguous and tracing testing; slaughter and compensation of reactor cattle and the cost of restriction placed on the breakdown herd. The average cost of a herd breakdown has been estimated at £30,000, £10,000 of which accrues to the individual farmer (http://www.defra.gov.uk/animal-diseases/a-z/bovine-tb/). So >90% of the costs of surveillance accrue to the (origin) farmer, whereas the most part of the costs of mitigation fall to Defra/AHVLA. For further consideration of the costs and benefits of PrMT, see Annex 1 of Defra's 2010 review of PrMT: http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/pre-movement-testing-review.pdf. 4. Coverage ● In 2010 53% (n=44186) of British cattle herds registered on VetNet were categorised as 1 or 2 yearly TI, most of which would be eligible for PrMT. Analysis of data from VetNet and CTS estimates compliance with the PrMT requirement appears to be around 90% of eligible movements. Further scrutiny of 5% of apparently non-compliant movements finds that most of these are in fact complaint: giving an estimated compliance of >97%. The average batch size for PrMT is 13 cows, although the distribution of testing batch size is positively skewed with around 67% of testing batches containing 15 or more cattle. The mean herd size of cattle herds in TI1-2 areas is around 100 cows. So, whilst coverage of the GB cattle population and within herd is relatively low and biased, coverage of the target population (as measured by compliance) is good. 5. Participation ● Stakeholders include: farmer, PVS (OV), AHVLA and Defra, +/- farming industry bodies Participation of herd owners is mandatory for pre-movement testing, apart from where exemptions apply, although the herd owner must be able to provide evidence of this (e.g. test chart, cattle passport, approval license etc. For details see: http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/premove-booklet.pdf).

Framework section	*Guidance notes*	*Case study: Pre-movement testing for Tuberculosis in cattle*
		Movement of cattle from a holding may be considered a breach of statutory requirements if you cannot provide evidence of exemption and could result in a reduction being applied to their Single Payment. Further, herd owners may face prosecution if they do not comply with pre-movement testing. AHVLA Field services Gloucester monitor compliance with the PrMT requirement by comparing CTS movement data with VetNet testing data; a small proportion of apparently non-compliant movements (~5%) are investigated further and most of these are subsequently found to be compliant. Greater than 97% of eligible cattle movements are compliant with PrMT. There were a total of 59 suspected breaches of the requirement reported to local authorities up until 2010. 6. Sensitivity ● Whilst diagnostic test sensitivity is the probability that a herd will test positive if tested, the surveillance sensitivity also depends on the probability that the herd is tested. From SE3238 (meta-analysis of validation of diagnostic tests for bTB) the median test sensitivity of SICCT (standard interpretation) is 0.64 (95% CI: 0.48-0.78, model B). With 100% of moving animals tested, the batch-level sensitivity will be 0.64, assuming one cow is infected per batch. If more than one cow is infected per batch the batch-level sensitivity will be higher: e.g. if two cows were truly infected in a PrMT batch of 13, the probability of detecting at least one of these cows would be 87%. In terms of the sensitivity of PrMT to detect infection in the herd of origin (a secondary objective in PrMT), as only part of the herd is tested the probability of identifying infection in the herd will be lower. Herd-level sensitivity will depend upon the number of truly infected cattle in the herd (design prevalence), the herd size and PrMT batch size and the probability of selecting an infected cow for movement. Assuming an average herd size of 100 with 3 infected cows (based on the average number of reactors identified at the disclosing test in herd breakdowns) a PrMT batch size of 13 and no association between the risk of infection and the risk of movement; the probability of detecting reactors in an infected herd would be 23%. The estimates of batch and herd-level sensitivity should be treated as guides and we would expect a fair degree of variation around these estimates. The relatively low sensitivity to detect disease in the herd of origin might be accepted as this is not the primary purpose of PrMT. Whilst more sensitive diagnostic tests exist, SICCT is deemed fit for purpose for the application to PrMT. 7. Specificity ● Specificity of SICCT at animal and herd level as above. From SE3238 (meta-analysis of validation of diagnostic tests for bTB) the median specificity of SICCT (standard interpretation) is 1 (95%CI 0.99-1, model A) SICCT is a highly specific test it seems. So even under 'worst case' conditions (Sp=0.99) the likelihood of getting a false positive cow in a batch is very low: the risk only really becomes significant at the level of whole herd tests (e.g. >100 animals). 8. Timeliness ● The SICCT takes 72 hours to complete. The results of testing are available at the time of reading and a copy of the test results are kept by the farmer. If any reactors are found, restrictions will be put in place immediately.

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
		Prior to the implementation of the web portal, test cards needed to be submitted to the local AHVLA office within 1 day where reactors were identified or 5 days where all cattle tested negative. The SPIDA application for interrogating and visualising bTB surveillance data receives weekly downloads of data from SAM/VetNet. So overall, the collection and management of PrMT surveillance data has been reasonably timely and this is hoped to be improved by the full implementation of the new data management system (SAM). **Secondary attributes for bTB surveillance** 1. Data collection ● The SICCT is a well-established diagnostic assay with a standardised method. However, training of OVSs in the SICCT method and interpretation of the results is brief – with a great reliance on 'on-the-job' learning. Initial assessment of competency in the practical application of a test is limited in most cases to a single testing session and there is no repeat assessment; nor are there mechanisms to provide quality assurance in the delivery of the test on an ongoing basis. A review of bTB skin testing in England and Wales in 2006 observed that deviations from the established test protocol was common among both private veterinary testers and State Veterinary Service staff; but it was not felt that these deviations would undermine the accuracy of the test (http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/dnv-report.pdf). Nonetheless, the want of repeat assessment or assurance mechanisms introduces a degree of uncertainty to the test accuracy and repeatability. 2. Data completeness and correctness [no colour code allocated]. Completeness of data in the VetNet database is expected to be high at both the record and field level. It is understood that the data downloaded from VetNet to the SPIDA application each week undergoes a certain amount of cleaning but we did not have enough information in hand on completeness and accuracy of the data to properly assess this attribute. 3. Impact ● PrMT appears to be well-received by the farming industry and the level of compliance with the requirement is very high. Given also that sensitivity of the test method is considered fit for purpose, it may be concluded that PrMT meets its stated objective of reducing the spread of bTB via movement of infected cattle. Analysis of bTB surveillance data demonstrates a potential impact of the advent of PrMT on the number of newly infected parishes, particularly in 4 yearly TI areas; but this data is complex and there are many other potential influences. The advent of PrMT appears to have an impact of farmer behaviour with regard to cattle movements and the timing of routine herd tests: including that more cows are included in routine herd tests than previously, thereby increasing the probability of detecting infection. The impact of PrMT in the first 3 years of implementation has been reviewed by the bTB Eradication Group (the forerunner to the TBEAG): http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/pre-movement-testing-review.pdf

Framework section	***Guidance notes***	***Case study: Pre-movement testing for Tuberculosis in cattle***
		<u>Tertiary attribute considered potentially important for bTB surveillance</u> 1. Data management ● There are several separate databases comprising the data system for collating surveillance data from PrMT (and other components of bTB surveillance). VetNet and VeBus are central databases collecting herd level and animal level data respectively. Many private veterinary practices also keep their own database on individual animals (bTB Master): this is a standardised database but is not networked. Finally, the CTS database records data on the movements and location of cattle in GB. VetNet is the primary database of interest in terms of identifying herd bTB status and breakdowns. The structure of VetNet is reasonably good, with definition of data fields and primary keys; although validation constraints and internal consistency might be a concern as errors in the data do appear: including duplicate CPH numbers (the primary key) and instances where the number of cattle tested is greater than the number of cattle in the herd. A data dictionary exists for VetNet but a summary overview of the dataset, including an entity relationship diagram was not apparent. Information around the structure of the data system was gleaned from users of the data. The data system is not covered by ISO9000 but the data protection implications are clear and defined and quality control checks are periodically conducted. A major drawback of this data system as of 2010 is the lack of integration of the separate databases. The VeBus database contains surveillance data from bTB testing conducted by ex-AH personnel but not necessarily data generated by OVSs in practice. The PVS (bTB Master) databases are not networked and surveillance data is not automatically collated centrally. This data could be accessed but would require retrieving data from each OV practice in the country individually. Improving the integration of the data system and so the availability of surveillance data for interrogation and analysis was identified as an objective in the development of the new database SAM; but the SAM data system has suffered serious setbacks in implementation and the many potential advantages of the new system have yet to be realised.
e. Perform an economic analysis	Use the information collected for assessed attributes, particularly the cost, impact and benefit attributes but also those measuring other aspects of surveillance effectiveness e.g. timeliness, to perform the cost-benefit (or, if appropriate, a cost-effectiveness) analysis identified in section 3b.	We did not have the time and resources available to conduct economic analyses of PrMT in this case study. However, Annex 1 of the 2010 review of PrMT (http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/pre-movement-testing-review.pdf) and before this, the regulatory impact assessment (http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/tb/documents/prmt-regulatory.pdf) have conducted cost-benefit analyses of PrMT. Estimates of the net annual benefit of PrMT in the 2010 review varied between £9.32m (2006) and £12.63m (2008).
f. Synthesise the results and use them to make suggestions for improvements to the surveillance system	Draw together the results of the individual attribute assessments and the economic analysis to reach conclusions about the evaluation question(s) listed in section 1b. Identify evidence-based suggestions for possible improvements to the surveillance system. Possible ideas for enhancing the collection of surveillance data include: - use of portable technology (e.g., collecting data using digital devices	The question that this evaluation set out to answer was: Does the pre movement testing for bovine TB effectively detect cases to reduce the risk of disease spread? Through the high level of compliance and use of an adequately sensitive diagnostic test, PrMT meets its primary objective of reducing the spread of bTB via movement of infected cattle and it may be stated that this component of bTB mitigation is fit for purpose. Further, economic analyses conducted separately demonstrate a positive net benefit. PrMT achieves a high level of compliance – and so coverage of the target population – provides timely

Framework section	*Guidance notes*	*Case study: Pre-movement testing for Tuberculosis in cattle*
	rather than paper forms); - risk-based requirement or sampling; - review of sampling strategies including the sample size, pooling of samples, and integration of data from different sources. The value of surveillance might be improved by changing the methods used to analyse or disseminate information	dissemination and use of surveillance data and employs a highly specific diagnostic test. The main areas that might be considered for improvement are quality assurance of test application, management of surveillance data and the interpretation and communication of surveillance data. <u>Quality assurance of test application</u> Currently the training and assessment of those implementing the SICCT on-farm is limited and there are no mechanisms in place to assure test-protocols are being adhered to. This creates uncertainty about the true (field) accuracy of the results of diagnostic testing. Studies on the field accuracy of SICCT and adherence to test protocol in the UK would be welcome. Mechanisms for quality assurance of application of SICCT (e.g. spot-inspections) or periodic reassessments of bTB-testers (e.g. every 5 years) could be considered. It is proposed that consideration be given to such measures by the bTB Eradication Group. <u>Data System Management</u> The main problem with the management of bTB surveillance data up to 2010 was the lack of integration of the various databases, in particular the lack of centralised collation of data stored in TBMaster databases in private veterinary practices around the country. This will be addressed by the implementation of SAM data system. Currently, failures and delays in the implementation of SAM are having a serious impact on the collation, analysis, interpretation and communication of bTB surveillance data and, more importantly, delivery of bTB surveillance and control measures. However, implementation of the SAM data system was deliberately not addressed in this case-study. <u>Interpretation and communication of surveillance data</u> Currently the products of analyses of bTB surveillance data are either relatively unsupported by interpretation (monthly statistics) or overly large and cumbersome (annual report). Communication of the results of surveillance for bTB generally could be improved within Defra/AHVLA and with the cattle industry through better engagement between AHVLA and Defra, industry stakeholders, industry commentators and opinion leaders. The information requirements of these stakeholders could be better defined through consultation. As the bTB Eradication Advisory Group includes representatives from policy, AHVLA and the farming industry, one of the regular meetings of the group would be an excellent forum for reviewing the outputs of the PrMT and other bTB surveillance data.

Framework section	*Guidance notes*	*Case study: Pre-movement testing for Tuberculosis in cattle*
5. Reporting and communication	Complete sections a – g below.	–
a. Identify the target audience(s) for the outputs from the evaluation	The primary audience is often (but not always) the evaluation funder. Secondary audiences are other users of the output and people involved in the surveillance system under evaluation. Ensure that all relevant people identified previously in sections 1d and 2g are included.	Primary audience are the SE4302 project team and Defra customer. Secondary audience might include those involved in planning and implementing PrMT and other components of bTB surveillance but this evaluation should be discussed within the AHVLA TB epidemiology group before further dissemination.
b. Consider which communication medium is most appropriate	More than one may be necessary in order to reach all affected parties.	This document will be sufficient for the case study report but in the case of a commissioned evaluation we would expect a report more along the lines of the following: - Executive summary including the key messages and recommendations of the evaluation (2 pages). - Main body of report based on this framework document providing additional detail and justification for the findings (20-30 pages). - References to published supporting information, including data sources used to complete sections 2 and 4 of this framework. - Further information provided in appendices, such as methods, results and discussion of economic analyses and other quantitative analyses of the surveillance system as appropriate.
c. State any uncertainties in the results of the evaluation and recommend further work	State the level of uncertainty associated with the results summarised in section 4f and any caveats in their interpretation. Make recommendations for any further work required to complete the evaluation of this system	Economic analyses (4e) could not be conducted because of time and resource constraints, but these have been conducted for PrMT over the first 3 years of implementation. The assessment of the Impact of PrMT on the spread of bTB by movement of infected cattle was approached from a theoretic point of view. In reality PrMT (and other components of bTB surveillance) produces a great deal data which could be used to assess the impact of PrMT on the spread of bTB both within high-risk areas (TI1-2) and to lower risk areas (TI4). Analyses of data from the first 3 years of PrMT were conducted in the 2010 review of PrMT but could not be repeated for this evaluation. The Data Completeness and Accuracy attribute could not be adequately assessed because of difficulty accessing resource to examine the databases. The bTB data system is currently undergoing significant changes: with SAM replacing VetNet, VeBus and the practice-based TBMaster systems. Difficulties in implementing the new system have caused significant problems in bTB surveillance and control and SAM remains a sensitive issue. Consequently we deliberately chose to assess the Data Management of the system in use in 2010. The main problem with Data Collection in PrMT and all other components of bTB surveillance using SICCT is the uncertainty surrounding tester compliance with the established testing method: Further work to establish the level of compliance with the agreed and standardised method would be beneficial.
d. Identify strengths and weaknesses of the	Indicate the main strengths and weaknesses of the surveillance	The main strengths of PrMT include the high level of participation and compliance; timeliness of testing results and communication between the farmer and OVS; and the high specificity of the chosen diagnostic

surveillance system	system.	test. The main limitations of PrMT include poor integration of the various surveillance databases; the lack of quality assurance surrounding the application of the SICCT on-farm; and difficulties in communicating the surveillance data generated. It is important to note that these limitations apply to bTB surveillance in its broader context.
e. Make recommendations	Make recommendations which indicate how the suggestions for improving the surveillance system identified in section 4f could be practically implemented and any recommendations for further evaluation provided in section 5c. Make it as easy as possible for the evaluation outputs to lead to actions to influence decisions and policy. Clearly communicate how the question(s) asked by the commissioners was dealt with (translated) in the evaluation process.	1. Consideration should be given to either the periodic reassessment of bTB-testers and their application of the SICCT and/or systematized quality assurance mechanisms assessing the adherence of bTB-testers to the SICCT method 2. Consideration should be given to the value to the target audiences of the current outputs of surveillance data (i.e. the monthly bTB statistics and the annual surveillance reports). Key stakeholders should be identified and engaged with a review of the content and format of these outputs in order to define better the information needs of Defra and the farming industry with regard to bTB surveillance. 3. An estimation of the actual (rather than theoretical) sensitivity of current testing using available data would be useful. For example, to examine the evidence for movements resulting in breakdowns due to failure to detect cases (suboptimal sensitivity) or due to herds that animals were moved from having cases that were likely to have been infected when the test was carried out. Concept notes for these recommendations could be produced and presented for consideration by the TB Eradication Advisory Group for England, Defra and the Welsh Government prior to more detailed investigation. Note that the implementation of the AHVLA SAM database has not been specifically addressed in this evaluation.
f. Indicate ways for follow-up by the funder	This might include a recommendation on when next to repeat the evaluation.	Not applicable for this case-study; but this might involve a repeat assessment of those attributes specifically affected by implementation of the recommendations.
g. Measure what effect the evaluation output had	Assess how fully the outputs outlined in Section 1f were achieved. This may need to be done 6-12 months after the end of the evaluation.	Not applicable.

C. Discussion

The aim of this project was to develop and trial a generic evaluation framework for animal health surveillance systems. The result is the SERVAL framework which is intended primarily for use in GB, but its flexible and generic nature mean it should be also applicable to other EU members and potentially worldwide use.

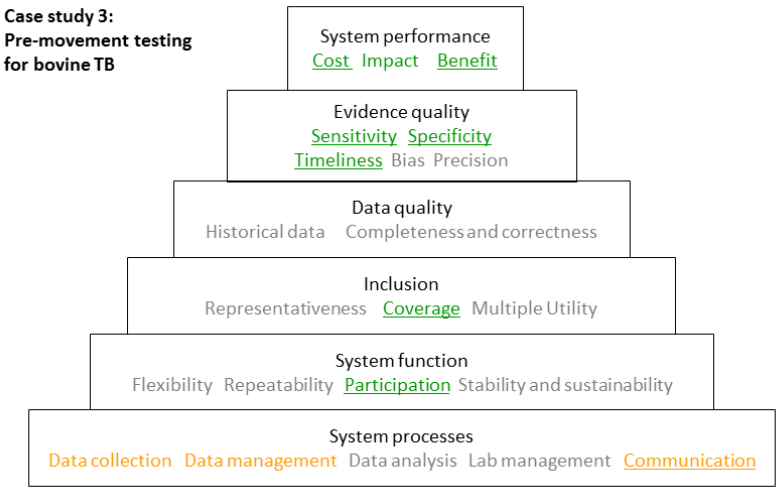
We used data and information that was already available when undertaking case studies. We did not set out to identify performance indicators and then proceed to collect indicator data, a process which would allow a more detailed evaluation. The objective of the case-studies was to provide a proof of concept approach which shows that the framework appears robust, complete and user-friendly. The evaluations provide a comprehensive evaluation of the selected systems using available data. Recommendations for obtaining additional information to complete more detailed evaluations of specific attributes are provided: for example, developing quality assurance mechanisms for the application of the SICCT. However, the evaluation carried out did not result in recommendations for substantial changes to PrMT.

A major difference between SERVAL and other existing frameworks is the high number of attributes (up to 22) from which evaluators are encouraged to select a shortlist of those most appropriate to each evaluation. Guidance is offered in the choice of attributes in the Attribute Selection Matrix (section 3a of SERVAL) but no prescription is made. This allows the evaluation to be easily tailored to any evaluation objective. This flexibility in choice of attributes was highlighted as a very positive development in a consultation process involving people with experience of being restricted to assessing the same 10 or so attributes by previous evaluation tools. Whilst the optimal number of attributes for assessment is likely to vary depending on the objectives of each evaluation, between five and 10 attributes per evaluation are likely to be required to provide a complete evaluation. However, assessing this number appears to be rare: most published evaluations assess three or fewer attributes, although these often take a detailed and quantitative approach to the evaluation of the selected attributes. The SERVAL framework is designed to provide a comprehensive, systematic assessment by evaluating a variety of relevant attributes while avoiding the problems that may occur by defining too many attributes which could detract from the evaluation's goal by making it a huge task to gather data and making interpretation difficult. The SERVAL attribute selection matrix has been designed to assist evaluators in choosing and prioritising attributes without unduly restricting choice for those situations where it may be appropriate to assess more or fewer attributes. Conducting these case studies prompted discussions about which attributes should be considered as primary attributes for different surveillance systems, further use of the framework will help to clarify any remaining issues although it is likely that some subjectivity in the choice attributes will remain confirming the requirement for a flexible framework.

A common suggestion during the consultation process when SERVAL was undergoing development was for guidance on how much time an evaluation using SERVAL would be expected to take. Our case studies indicate that a complete qualitative evaluation may be accomplished in approximately 6-8 person-days, although the exact amount of time is likely to vary depending on the system being evaluated, availability of expertise and information, and the depth of evaluation required. If data need to be collected for specific indicators, or if broader interviewing of stakeholders involved in a surveillance programme is conducted, considerably more time will be required. A longer period would also be needed in order to conduct a more detailed evaluation involving rigorous quantitative approaches. Crucially, SERVAL allows the evaluation to be scaled according to needs and resources so that even a small scale evaluation will provide value. The SERVAL framework has been designed to be thorough without being unnecessarily long and so evaluations are straightforward and efficient, aided by the prompting notes. During the consultation no-one identified any aspects of the framework as redundant, although most conceded that some parts might be more important than others according to the situation in which it is applied.

Additionally, the case study evaluations provide an indication of approximately how much text is expected in response to each question in the framework. Responses are expected to be detailed but brief. A report of a single surveillance evaluation would therefore be expected to be about 30 pages long. Figures that summarise the findings of the evaluation – for example colour-coding the attribute assessments (Figure 1: on next page) – are likely to be an effective way of communicating the results.

Figure 1. Summary of attribute assessment outcomes from the pre-movement testing for tuberculosis case study.



Key
Attribute assessment: **Excellent/very good** **Room for improvement** **In need of attention** **Not assessed**
Attribute priority: **Primary** **Secondary or Tertiary**

Definitions of terms and concepts in the framework are not always agreed and hence a glossary of terms and their definitions is included (Appendix). The experts consulted during the development of SERVAL generally understood and agreed with the definitions of terms used in the framework, although they acknowledged their own familiarity with the concepts and attributes presented through their previous experience in planning, coordinating and evaluating animal health surveillance. Thus the Appendix is needed. These definitions have been developed using the output from discussions at a workshop to discuss surveillance terminology prior to the ICAHS conference in May 2011. These discussions are ongoing and the SERVAL and surveillance terminology initiatives will continue to work together to finalise the definitions of evaluation attributes. There was also some discussion during the conduct of the case studies about how individual attributes should be measured and the project team are currently working on the production of guideline to assist with these decisions.

Economic evaluation is an integral part of the SERVAL framework because the costs of obtaining surveillance information should be balanced against the benefits derived. However, this part of the evaluation was difficult to perform in the case studies. Fortunately in the case of PrMT, economic evaluations of the surveillance component had been undertaken as part of a previous review. Examining the outputs of surveillance without considering the resources used (or vice versa) represents only half the process. Further, it is hard to rate the benefit and cost attributes in isolation: What level of benefit is good? What cost is good? A possible solution would be to consider whether the potential benefits of surveillance have been fully realised by the surveillance system under evaluation.

One issue that arose during the case study process was whether assessing 'To ascertain if a surveillance system is providing value for money to the funder' should be considered an objective of all evaluations carried out for Defra. This is because improving the efficiency of surveillance is one of the main objectives of the current surveillance review process and funders are almost always going to be interested in whether the efficiency of their current surveillance could be improved. This does not necessarily mean a quantitative cost-effectiveness analysis is required but it is expected that such evaluations will identify whether there are any alternative strategies that should be considered and investigated further as a means to improve efficiency.

Ideally, the value of a surveillance evaluation should be realised by more than simply those who fund it. Information provided by following the SERVAL framework is likely to be increasingly valuable to the farming industry as they become more involved in commissioning and funding animal health surveillance in GB. The output from an evaluation, such as the case studies presented in this report, could be a useful tool for

communicating to such parties how an animal health surveillance system operates and the information and value for money that it provides. In addition, because SERVAL provides a logical, clear and structured approach, the output could become a source of assurance and credibility for the system examined. Finally, the SERVAL framework could be used to communicate the design and performance of animal health surveillance systems with international bodies such as the OIE and the EU to influence decisions made by these bodies on legislation for animal health surveillance.

In conclusion, undertaking the case studies has shown the SERVAL framework to be a comprehensive and generic framework for the evaluation of animal health surveillance. It is straightforward to use and flexible enough to accommodate a range of surveillance and evaluation objectives. SERVAL is freely available on the internet. We encourage its use by those involved in animal health surveillance and would be delighted to receive feedback on users' experiences.

If you do use SERVAL, please acknowledge the source by citing the website www.rvc.ac.uk/serval. Thank you.

D. Acknowledgments

SERVAL was developed by staff at the Royal Veterinary College, Animal Health and Laboratories Agency, and the Scottish Agricultural College. We thank all those who contributed to the development and testing of the framework through participation in workshops, consultations and review meetings. This project was funded by Defra, project SE4302.

E. Appendix: Definitions of the 22 attributes included in the framework, with examples of reports demonstrating their assessment.

Attribute	Definition	Guidance Notes	References
Benefit	Direct and indirect advantages produced by the surveillance system. Not limited to money, benefits might include any losses avoided due to information provided by surveillance system: financial savings, better use of resources, improved animal production, improved public health, increased understanding about a disease, or increased trade.	The benefits of a surveillance activity should be listed and, where possible, quantified. This information will be valuable to the economic evaluation in section 4 of the framework document. An evaluation of the benefits of a surveillance activity should include : 1. A complete list and characterisation of all the potential benefits of the surveillance activity 2. Where possible, quantify market benefits in financial terms 3. Where possible, quantify non-monetary benefits by alternative methods. For example, Quality-adjusted life years for public health benefits (HM Treasury 2003, Zinsstag et al 2007) or using points-system (Dufour 1999) 4. Consider how the benefits are distributed among stakeholders, including: producers, consumers, the livestock industry or society Points to consider whilst assessing the benefits of surveillance include: - Surveillance and disease control are often integrated: That is to say, surveillance provides information that informs control and so many benefits of surveillance are often realised by control measures. As with costs, it is important to understand the benefits of surveillance in the broader context of disease mitigation. Benefits of surveillance may be considered as disease losses and mitigation costs avoided by detection of disease. So it may be useful to begin by listing all the losses and costs resulting from disease and disease mitigation measures. It may be difficult in some instances to distinguish between the direct benefits of surveillance and those arising from mitigation. - The benefits of surveillance for early detection of disease outbreaks can be quantified as the losses and costs avoided through earlier detection and control - The primary benefit of surveillance providing evidence of disease freedom is access to international markets (for both live animals and animal products). The economic value of international trade can be attributed as a benefit to surveillance. Officially recognised disease-free status often permits the disease-free country/region to maintain border security measures against introduction of the disease (eg restriction on trade and movement of animals or requirement for	Dijkhuizen *et al*. Economic analysis of animal diseases and their control. *Preventive Veterinary Medicine*. 1995. 25:135-149 Dufour B. Technical and economic evaluation method for use in improving infectious animal disease surveillance networks. *Veterinary Research* 1999; 30: 27-37. Häsler *et al*. Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for economic analysis. *BMC Health Services Research* 2011. 11:225 HM Treasury. The green book: Appraisal and evaluation in central government. 2003. http://www.hm-treasury.gov.uk/d/green_book_complete.pdf ILRI. Veterinary epidemiology and economics in Africa: a manual for use in the design and appraisal of animal health policy. 1998 http://www.fao.org/wairdocs/ILRI/x5436E/x5436e00.htm#Contents McInerney *et al*. A framework for the economic analysis of disease in farm livestock. *Preventive Veterinary Medicine* 1992. 13: 137-154 (avoidable costs) Moran D, Fofana A. An economic evaluation of the control of three notifiable fish diseases in the United Kingdom. *Preventive Veterinary Medicine*

Attribute	Definition	Guidance Notes	References
		pre-export testing) – thus mitigation of risk of incursion is also a benefit of surveillance for freedom from disease.. - Surveillance for case-detection and monitoring prevalence of endemic disease provides information for the improved control and management of disease; including prioritisation of diseases and allocation of resources. - Improved public health is an obvious advantage to surveillance for zoonoses. Increased consumer confidence is another – although consumer confidence may also be of significance to other high-profile, non-zoonotic diseases. - Consider potential indirect or secondary benefits of surveillance; externalities or spill-over of benefit to other livestock sectors or industries. It may be helpful to consider potential benefits both upstream (eg animal feed producers) and downstream (eg food processors) of the production system. Examining the value chain will aid in this.	2007; 80: 193-208. Morris S*, et al.* The costs and effectiveness of surveillance of communicable disease: A case study of HIV and AIDS in England and Wales. *Journal of Public Health Medicine* 1996; 18: 415-422. Rushton 1999 (costs and benefits often difficult to quantify) rev sci tech 18:315 Zinsstag *et al.* Human benefits of animal interventions for zoonosis control. *Em Inf Dis* 2007. 13(4): 527-531
Bias	The extent to which a prevalence estimate produced by the surveillance system deviates from the true prevalence value. Usually (if not always) refers to endemic diseases. Bias is reduced as representativeness is increased.	Assessing the bias of a system is most relevant to surveillance of endemic diseases where the objective is to monitor the prevalence of a disease. Bias may lead to erroneous conclusions about the burden or distribution of disease in the population. For some surveillance activities – such as risk-based surveillance aimed at detecting cases to facilitate control – surveillance may be intentionally biased toward sub-groups of the population at higher risk of disease. So the context and objective of surveillance will determine whether bias is acceptable or not. Either way, an evaluation of bias should include: 1. An assessment of whether any prevalence estimates produced are likely to be biased based on an assessment of the potential sources of bias in a surveillance system and its outputs 2. Where possible bias should be quantified and the outputs adjusted accordingly. Bias in epidemiology may be categorised into misclassification bias and selection bias: - Misclassification bias concerns the sensitivity and specificity of the case-definition	Del Rio Vilas VJ, Pfeiffer DU. The evaluation of bias in scrapie surveillance: A review. *Veterinary Journal* 2010. 185:259-264. Del Rio Vilas VJ, Böhning D. Application of one-list capture-recapture models to scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 2008. 85: 253-266 Doohoo I, Martin W, Stryhn H (eds). Veterinary epidemiologic research. 2003. AVC inc. Guasticchi, G., *et al* 2009. Syndromic surveillance: sensitivity and positive predictive value of the case definitions. Epidemiol. Infect. 137, 662-671. Hendrikx P*, et al.* Development of performance indicators for the

Attribute	Definition	Guidance Notes	References
		(often intimate to the diagnostic protocol) - Selection bias results when the probability of being sampled is associated with the probability of disease (and therefore the probability of other factors associated with disease). In this regard the bias, sensitivity, specificity, coverage and representativeness of a surveillance system are related concepts. Bias in the prevalence estimates obtained using surveillance data may result from poor sensitivity, specificity, coverage or representativeness in the system. Some potential sources of bias to consider include: - Sensitivity/specificity of the diagnostic method - Under-reporting in passive surveillance activities - The sample source population: For example, sampling at abattoirs may lead to an under-estimate of the prevalence of many diseases as these animals are from a healthy (and younger) sub-population, whereas sampling fallen stock may lead to an over-estimate of burden - Selection bias may also be introduced in terms of geography, production type, herd/flock size, species or age category of the animal Bias in the surveillance output can be examined and quantified by several methods: - Simple comparison of multiple surveillance data sources examining the same population. In some instances a separate survey might be designed and implemented to specifically examine potential biases (eg a postal survey to explore under-reporting in passive surveillance activities) - More sophisticated statistical or simulation methods can be applied to existing data o Capture-Recapture (CRC) methods have been applied to make inferences about the unobserved cases and so the completeness of surveillance data (Del Rio Vilas and Böhning 2008, Guasticchi et al, 2009,). o Morignat et al (2006) used simulation models to explore potential biases (identified *a priori*) in scrapie surveillance data. o Mathematical models might also be used to simulate the spread of disease and generation of surveillance data to explore potential bias in	bovine clinical salmonellosis surveillance network in France. *Journal of Veterinary Medicine Series B-Infectious Diseases and Veterinary Public Health* 2005; 52: 465-475. Morignat *et al*. Estimates of the prevalence of spongiform encephalopathies in sheep and goats in France in 2002. *Veterinary Record,* 2006. 158:683-687 Wells, S.J., Ebel, E.D., Williams, M.S., Scott, A.E., Wagner, B.A., Marshall, K.L., 2009. Use of epidemiologic information in targeted surveillance for population inference. Preventive Veterinary Medicine 89, 43-50. Williams, M.S., Ebel, E.D., Wells, S.J., 2009. Population inferences from targeted sampling with uncertain epidemiologic information. Preventive Veterinary Medicine 89, 25-33.

Attribute	Definition	Guidance Notes	References
		the system. If bias can be identified and measured, then it should be possible to adjust the prevalence estimate. - Correcting an estimate of prevalence for incomplete test accuracy is easily achieved with knowledge of the sensitivity and specificity of the diagnostic protocol (see equation 5.16 in Doohoo *et al* 2003) Prevalence estimates can also be adjusted for selection bias where the strength of association (ie relative risk or odds ratio) and the distribution of the risk factors in the background population are known (Wells et al 2009, Williams et al 2009). - Morignat *et* al (2006) uses simulation methods to both explore and partially correct of bias in surveillance data If bias is deemed to be significant and unacceptable and cannot be satisfactorily corrected for during analysis and interpretation of the data, one might consider reviewing the design and implementation of the surveillance activity.	
Communication	An assessment of the methods and ease of information exchange between people involved at all levels of the surveillance system (providers, analysers and users of surveillance data) including an assessment of the information provided, timeliness, types of outputs and a description of the efforts made to disseminate the	Communication concerns the dissemination of information and provision of feedback into the system. Communication in a surveillance system is often related to various other attributes, including participation, timeliness and impact. An assessment of communication should include: - A list of the outputs that are generated from the surveillance data; Who are these intended for and do they meet all information needs of the target audience? - An assessment of who has access to the surveillance outputs; Are all stakeholders represented? - An assessment of whether the surveillance outputs are produced sufficiently frequently. Do they contain up-to-date data of sufficient quality? Are the data presented with sufficient discussion of its meaning, limitations and biases from an epidemiological perspective? - A list of other feedback provided to those contributing to the surveillance system e.g. data quality checks Qualitative or semi-quantitative social science approaches are likely to be of value in	Buehler JW*, et al.* Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recomm Rep* 2004; 53: 1-11. Riera-Montes and Velicko 2011. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2 Roberts M*, et al.* Implementing and evaluating a practice-based surveillance program for equine infectious disease in North Carolina. In: Proceedings of the International Society of Veterinary Epidemiology and Economics, 6-11 August 2006,

Attribute	Definition	Guidance Notes	References
	surveillance information.	assessing this attribute. Consultation with the key stakeholder groups of the surveillance system will be useful, including: - Providers of surveillance data (eg farmers, veterinarians, laboratory staff etc) - Those analysing and interpreting the surveillance data (ie generating information and knowledge from the data and disseminating it) - Users of surveillance data, including the direct customer (funder) but also other beneficiaries of the information as appropriate(eg government, the farming industry or academia)	Cairns, Australia.
Cost	List and quantify each of the resources required to operate the surveillance system. For example: time, trained personnel, finance, standards and guidelines, communication facilities, forms for surveillance, computers, other equipment.	Cost and the breakdown of cost is important – also who pays – this can be considered in relation to benefits or other attributes in the economic analyses An assessment of costs should include: 1. A listing of the areas of expenditure to be quantified 2. Estimates of the cost of each 3. Consideration of the distribution of costs among stakeholders, including: producers, consumers, the livestock industry or society This list can be built upon the characterisation of the surveillance activity in section 2 of the evaluation framework. All areas of the planning and implementation of surveillance should be considered: - Planning and design of the surveillance activity - Operational management - Sample collection and handling - Laboratory testing or other diagnostic services - Data collection, management and analysis - Interpretation, reporting and dissemination of surveillance information It may be helpful to distinguish between fixed and variable costs: - Fixed costs vary only in the long term and are incurred regardless of the level of surveillance (e.g. costs of planning, salaries of permanent staff, laboratory	Moran D, Fofana A. An economic evaluation of the control of three notifiable fish diseases in the United Kingdom. *Preventive Veterinary Medicine* 2007; 80: 193-208. Morris S*, et al.* The costs and effectiveness of surveillance of communicable disease: A case study of HIV and AIDS in England and Wales. *Journal of Public Health Medicine* 1996; 18: 415-422. Phillips VL*, et al.* Evaluation of program performance and expenditures in a report of performance measures via a case study of two Florida tuberculosis programs. *Eval Program Plann* 2010; 33: 373-378.

Attribute	Definition	Guidance Notes	References
		facilities etc) - Variable costs vary in the short term and with the level of surveillance (e.g. sample collection costs, test reagents etc) Consider also how the costs of surveillance are divided among the stakeholders (e.g. what is paid for directly by the producer, by industry levy or public funds). Considering the distribution of surveillance costs and benefits is an important part of economic analyses. Effort should be made to distinguish the costs of surveillance from costs of disease control measures but, where surveillance and control are closely integrated, this may be hard to do. Regardless, it can be useful to consider the costs of surveillance in the broader context of the costs of mitigation and the costs of disease.	
Coverage	Proportion of the population of interest that is included in the surveillance activity.	The Coverage of a surveillance system is often related to the Representativeness, Bias and Sensitivity. A high coverage is particularly important to surveillance for the early detection of exotic or new (emerging) diseases. An assessment of coverage should include 1. At the very least, the sampled and target populations should be characterised and compared qualitatively 2. Where sufficient data on the target population exists, simple calculations of the proportion coverage can be made (eg 75% of the national herd and 45% of cattle holdings are sampled annually). 3. Where sufficient information on the background population is lacking, more sophisticated statistical techniques might be employed (eg Capture-Recapture analysis) Some considerations when assessing the coverage of a surveillance activity are the target population and the unit of interest: - Coverage should be measured against the population of interest (the target population) as defined in section 2 of the framework. This may not include all animals or holdings in a country that are susceptible (eg post-import testing of cattle say is focussed on a sub-population of holdings that receive livestock from overseas and not all holdings keeping cattle). At this point, it may be worth considering whether the target population has been adequately defined (ie	Brooker S*, et al.* The use of schools for malaria surveillance and programme evaluation in Africa. *Malaria Journal* 2009; 8: 231. Del Rio Vilas VJ, Pfeiffer DU. The evaluation of bias in scrapie surveillance: A review. *Veterinary Journal* 2010. 185:259-264. Del Rio Vilas VJ, Böhning D. Application of one-list capture-recapture models to scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 2008. Del Rio Vilas *et al*. A case of capture-recapture methodology using scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 2005. 67: 303-317 Walker N*, et al.* Epidemiological analysis of the quality of HIV sero-surveillance in the world: how well do we track the epidemic? *AIDS* 2001; 15: 1545-1554.

Attribute	Definition	Guidance Notes	References
		whether the exclusion of certain animals or holdings is merited) - The unit of interest – in which the level of coverage is measured – is often the unit of interest of surveillance (eg animal or holding). If insufficient data exists for this, or alternative perspectives are desired, coverage might be assessed at other aggregate levels (eg geographical areas) or relevant intermediate steps in the surveillance pathway (eg the proportion of veterinary practices submitting diagnostic samples, private laboratories submitting data or participating abattoirs or markets). - In certain contexts it may be worth establishing a timeframe of reference (eg annual coverage). The choice of timeframe should reflect the epidemiology of the disease.	
Data analysis	Appropriate methods used for analysis and interpretation of data.	Surveillance systems that perform well in this attribute will use analytical methods that are appropriate to the data and the information needs of users of the data whilst exploiting the data to its fullest extent. In this regard there is a relationship between this attribute and those of Data collection, Data management, Communication and Impact. An evaluation of data analysis should include: 1. The identification of the analysis methods applied to surveillance data: o No analysis o Basic descriptive statistics o Examination of trends o More sophisticated statistical approaches (eg time series analyses, spatial analyses) 2. An assessment of whether the limitations of data have been understood and accounted for in statistical analyses? 3. An indication as to whether the body of data available being fully exploited or could further use of data be made? It may help to review demands for information made by users of the surveillance data in the past, to determine whether their needs were met by the methods applied.	

Attribute	Definition	Guidance Notes	References
Data collection	The use of appropriate data sources and collection methods and the existence of a case definition and a data collection protocol.	A surveillance system that scores well on this attribute will have a clear and comprehensive case definition; make use of appropriate diagnostic tests; have a written protocol that describes collection of data (and samples); and the limitations of the collection methods will be clearly defined and understood. There is a relationship between this attribute and those of Data completeness, Data management and Laboratory management. Questions to consider when assessing data collection include: 1. Is there a written case definition for this surveillance system that is clearly defined and complete with specified inclusion and exclusion criteria? 2. Does the case definition include relevant details of the case signalment, clinical and pathological signs and epidemiological information as appropriate? 3. Does the case-definition include laboratory diagnosis? a. If applicable, are the chosen diagnostic methods appropriate to the case definition, including in terms of diagnostic samples being collected and the expected pathophysiology of disease? b. Have the sensitivity and specificity of the tests been assessed? 4. Is there a written sample and data collection protocol and are there appropriate assurance mechanisms to ensure the protocols are followed? 5. Are there data collected that are not used in analysis or interpretation (redundancy)? 6. Are there information needs for which data are not currently collected and feasibly could be? It may help to review demands for information made by users of the surveillance data in the past, to determine whether their needs were met by the data available.	
Data completeness and correctness	Proportion of data intended to be collected that were collected and stored	Completeness of surveillance data is relatively simple to measure and should be considered at two levels: fields and records. Most commonly Data completeness is measured as the proportion of records with missing	Harpaz R*, et al.* Lessons learned from establishing and evaluating indicators of the quality of measles surveillance in the United States, 1996-1998.

Attribute	Definition	Guidance Notes	References
	accurately.	or invalid data in the data fields – where data fields are variables containing demographic, clinical, pathologic or epidemiological information recorded for each sample. Key data fields (eg animal id, holding of origin, diagnostic result etc) should be identified and the proportion of completeness measured. Measurement of the proportion of records or observations that have been collated in the data system may also be considered. This will require comparison with an alternative source of data (eg the sample frame or paper records of sampling and laboratory test results). Poor data completeness may indicate problems in the Data collection, Data Management or Communication and engagement attributes.	*Journal of Infectious Diseases* 2004; 189: S196-S203. Miller M*, et al.* Evaluation of Australia's National Notifiable Disease Surveillance System. *Commun Dis Intell* 2004; 28: 311-323. Pipino et al 2002. Data quality assessment. Communications of the ACM 45(4): 211. http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCACMApr02.pdf Riera-Montes and Velicko 2011. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2 Rumisha SF*, et al.* Monitoring and evaluation of integrated disease surveillance and response in selected districts in Tanzania. *Tanzan Health Res Bull* 2007; 9: 1-11.
Data Management	Appropriate structure and documentation of data management systems for processing information, including data processing protocols, and effective use of data verification procedures.	Data management is a broad area concerning the collation, storage and maintenance of data, including but not limited to matters of data quality, accessibility, usefulness and security. Assessing this attribute will require an intimate understanding of the data systems employed by the surveillance activity. More detailed guidelines on assessing data management are provided in the references. An assessment of this attribute should include: 1. Consideration of whether the database structure has been correctly designed: - Has each field of data been tightly defined to ensure correctness, conciseness and consistency across records? - Have primary keys, uniquely identifying each record, been assigned? - Has the database been normalised, to ensure data is stored in the most	Mosley 2008. DAMA DMBOK Functional guide version 3. http://www.dama.org National Office for Stats guidelines. http://www.nationalarchives.gov.uk/information-management/projects-and-work/implementation-guides.htm Pipino et al 2002. Data quality assessment. Communications of the ACM 45(4): 211. http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCACMApr02.pdf Woolhouse et al 2011. Guide to good

Attribute	Definition	Guidance Notes	References
		parsimonious, transparent and useable way? - Have validation constraints, preventing the input of invalid data, and internal cross-consistency checks been applied? - Is the data stored in a way that allows the required interrogation and analysis? 2. Consideration of whether documentation of the data is sufficient to facilitate interpretation and understanding of the data: - Is there a document providing a summary overview of the data and collection methods and explaining any idiosyncrasies relevant to the analysis and interpretation of the data? - Is there a data dictionary that clearly defines each field? - Is there an entity relationship diagram that explains how the data relate? 3. Consideration of whether there are adequate protocols for managing data quality and security: - Is the data management system covered by a data quality standard (eg ISO9000, Good Clinical Practice or Good Laboratory Practice)? - Are Data Protection implications defined and is the Information Asset Owner identified? - Are periodic data quality control checks implemented? - Are records management issues clearly defined, including policy on the retention of data?	practice for quantitative veterinary epidemiology. Chapter 1. http://www.qve-goodpracticeguide.org.uk/guide
Flexibility	Ability to adapt to changing information needs or operating conditions with little additional time, personnel or allocated funds.	Flexible systems can accommodate new health-related events, changes in case definitions or technology, and variations in funding or reporting sources (CDC 2001). This attribute is determined more by the planning and management of the surveillance system than the operation of the system. Simpler or more generic systems are likely to be more flexible. An evaluation of the flexibility of the system may be made by considering how the surveillance system has responded to changes in the past. Potential changes or events to	Aavitsland P*, et al.* Anonymous reporting of HIV infection: An evaluation of the HIV/AIDS surveillance system in Norway 1983-2000. *European Journal of Epidemiology* 2001; 17: 479-489. CDC Updated guidelines for evaluating

Attribute	Definition	Guidance Notes	References
		consider include: - Changes in the information needs of the users of surveillance - Changes in relevant national or international legislation or guidelines - Changes in the demography of the target population - Changes in the epidemiology of disease (including outbreaks) or the emergence of new disease threats - Changes or improvements to the methods of surveillance, including adoption of new technologies (eg development of new diagnostic methods) - Changes to behaviour or influences on behaviour of key actors and agents in the system (eg changes to reporting behaviour or the costs of diagnostic services) An assessment of how likely it is that such changes may occur in the future and whether the surveillance system would be able to respond to these changes should also be made. Assessment of this attribute will be aided by consultation with key stakeholders of the system.	public health surveillance systems. *Morbidity and Mortality Weekly Report* vol 50. 2001 Colby et al. Evaluation of two systems for managing emergency poultry diseases in intensive poultry production regions. *International journal of Poultry Science* 2003. 2(3):234-241 Jefferson H*, et al.* Evaluation of a syndromic surveillance for the early detection of outbreaks among military personnel in a tropical country. *Journal of Public Health* 2008; 30: 375-383. Riera-Montes and Velicko 2011. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2
Historical data	Quality and accessibility of archived data.	Maintaining historical data is more important to surveillance activities designed to provide evidence for freedom from disease or for monitoring trends in prevalence of endemic disease. Historical data can also be valuable to epidemiological research. This attribute is related to those of Data management and Repeatability. Questions to consider include: - How many years of data are stored? - How complete and reliable are the data? - Are the data stored in a way that allows the required interrogation and analysis? - Is there a summary overview of the data and collection methods explaining key idiosyncrasies of the data and changes to the data or collection methods over time? - What use is currently made of historical surveillance data?	Gazarian M*, et al.* Evaluation of a national surveillance unit. *Archives of Disease in Childhood* 1999; 80: 21-27.

Attribute	Definition	Guidance Notes	References
Impact	A measure of the usefulness of the surveillance system. Should include details of actions taken as a result of the information provided by the surveillance system, e.g., changes in protocols or behaviour.	The Impact (called 'Usefulness' in CDC 2001) of a surveillance system is related to the Benefit derived from the system where assessment should consider specific examples or events where information generated by the surveillance system has influenced disease mitigation efforts. In this regard it will be useful to measure Impact retrospectively, through consultation with relevant stakeholders of the system. As with Benefits, the Impact of surveillance in some cases may be realised through its relationship with disease control measures (Haesler *et al* 2011). An assessment of impact should consider: - How do the objectives of the surveillance system reflect the stated needs of policy and the industry it serves? - How are outputs generated from the surveillance data used? Who are they intended for and how well received are they? - What questions have been asked of the surveillance data previously and have these information needs been met? - How has information generated by the surveillance system influenced the development of national or international disease control policy (eg changes to requirements for surveillance or control of disease)? - How has information generated by the surveillance system contributed to the prioritisation of disease threats within the industry, country or globally? - How has the surveillance system contributed to mitigation of endemic disease or earlier detection and control of exotic disease outbreaks?	Carrieri MP, *et al.* Evaluation of the SIMI system, an experimental computerised network for the surveillance of communicable diseases in Italy. *European J of Epidemiology* 2000; 16: 941-947. Hesterberg U, *et al.* Evaluation of the sensitivity of the British brucellosis surveillance system using stochastic scenario tree modelling. In: Proceedings of the 12th meeting of the International Society of Veterinary Epidemiology and Economics, 10-14 August 2009, Durban, South Africa, 2009. Häsler *et al.* Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for economic analysis. *BMC Health Services Research* 2011. 11:225
Laboratory management	Testing carried out using appropriate methods with quality assurance scheme and timely and accurate production of results.	Diagnostic laboratories should aim to produce reliable, accurate, unbiased results within a suitable time frame and at acceptable cost. With the emphasis on a quality service and value for money a laboratory should have quality control procedures for monitoring the validity of tests undertaken. Questions to consider in assessing this attribute include: - Does the laboratory implement a structured and systematic quality management system?	International Accreditation Forum. http://www.iaf.nu/ International Laboratory Accreditation Cooperation. http://www.ilac.org/ International Organisation for Standardisation. http://www.iso.org/

Attribute	Definition	Guidance Notes	References
		- Does the system include internal quality control processes (eg checking that test kits and reagents are performing within specifications, ensuring regular use of internal controls and certified reference materials)? - Does the laboratory participate in inter-laboratory comparison or proficiency testing? - Is the laboratory accredited to international standards of operation (ie ISO 9001 and ISO 17025)?	
Multiple utility	The ability of a surveillance system to capture information on several diseases or health conditions: a measure of how generic the system is.	Multiple Utility in a system should always be considered when examining the cost-effectiveness of a system. Firstly one should assess the *realised* multiple utility of the system but it will also be of benefit to assess the *potential* multiple utility – an outcome of assessing potential multiple utility might be recommendations on how to add value to the system currently implemented. An assessment of multiple utility should consider: - What additional information is or could be gathered during sample collection (eg on animal health or husbandry and demographics)? - What other types of samples are or could be collected at the time of sampling? - What other diseases are or could be tested for with the samples collected? - How long are samples stored following testing and could they be used for other purposes (including other research purposes)? For a surveillance system to offer value to other diseases or information needs, the objectives and processes of the system should be aligned to other systems. So it may be expected that more simple systems are likely to have more potential for multiple utility. For example, a simple random survey of holdings, repeated annually and with good coverage and representativeness could be useful for various diseases; whereas a risk-based design aimed at a specific threat may be of limited value for other diseases with differing epidemiology.	Izadi M*, et al.* A Bayesian network model for analysis of detection performance in surveillance systems. *AMIA Annu Symp Proc* 2009; 2009: 276-280. Malecki KC*, et al.* Effective Environmental Public Health Surveillance Programs: A Framework for Identifying and Evaluating Data Resources and Indicators. *J Publ Health Man Prac* 2008; 14: 543-551.
Participation	A description of the extent to which people in each of the user groups identified in	Participation (defined as Acceptability in Buehler *et al* 2004) examines the involvement or engagement of stakeholders in the planning, design and implementation of the surveillance activity. The efficacy of any surveillance system that is greatly dependent on voluntary participation or human behaviour (eg passive surveillance activities) will be	Buehler JW*, et al.* Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the

Attribute	Definition	Guidance Notes	References
	section 2(g) get involved in the surveillance process.	vulnerable to problems with engagement. An assessment of participation should include the identification of the factors likely to increase or prevent stakeholder participation and an assessment of the likely impact of these factors on levels of participation Qualitative or semi-quantitative social science approaches are likely to be of value in assessing participation. Consultation with all those involved in generating, analysing, reporting and using surveillance data will be valuable. Factors that may influence participation include: - What communication pathways exist internal to the surveillance system (eg between those collecting or providing data and those analysing and reporting the data)? Are these pathways formalised in any fashion? - Does information and feedback flow freely between those implementing surveillance and those using surveillance data? - How are each of the key stakeholders represented in the planning, design and implementation stages of the surveillance activity? - What are the incentives (e.g. compensation payments) or barriers (e.g. consequences of reporting) for participation	CDC Working Group. *MMWR Recomm Rep* 2004; 53: 1-11. Riera-Montes and Velicko 2011. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2
Precision	How closely defined a numerical estimate obtained from the study population is. A precise estimate has a narrow confidence interval. Precision is influenced by sample size, the chosen confidence level and data completeness and correctness.	Precision in surveillance activities designed to monitor prevalence is a measure of the degree of certainty around the point estimate of prevalence or incidence (ie the confidence interval or standard error). NB A related concept in surveillance designed to provide evidence for freedom from disease is the measure of confidence in disease freedom derived from the Sensitivity of the surveillance system. The precision of point estimates in epidemiological studies is dependant upon disease prevalence, sample size and the approach to sample selection (ie the design effect, Doohoo *et al* 2003). Precision of a surveillance activity will determine the how sensitive the surveillance system is to changes in prevalence. The desired level of precision will be set by the epidemiology of disease, surveillance objectives and the optimal allocation of resources.	Doohoo I, Martin W, Stryhn H (eds). Veterinary epidemiologic research. 2003. AVC inc.

Attribute	Definition	Guidance Notes	References
Repeatability	How consistently the study results can be reproduced over time.	Repeatability is a concept often applied to validating diagnostic tests and is related to precision. In terms of a surveillance system, Repeatability is also related to the attributes of Historical data and Stability and sustainability. A surveillance activity that performs well in this attribute produces data that can be easily compared across years and where changes to the data and data collection methods over time are clearly defined and understood. One might consider changes to legislation; changes to diagnostic methods, including improvements of adoption of new technology; changes to surveillance design; or influences on disease reporting behaviour in passive surveillance activities. - How have these impacted on the comparability of surveillance data over the time period of interest? - Have these influences been identified and examined and can they be accommodated in interpretation of the surveillance data?	Walker N*, et al.* Epidemiological analysis of the quality of HIV sero-surveillance in the world: how well do we track the epidemic? *AIDS* 2001; 15: 1545-1554.
Representativeness	Extent to which features of the population of interest are reflected in the surveillance data that are collected. Features may include: herd size; herd type (e.g., breeding, fattening, milk, meat); age; sex; location. A surveillance system that is representative accurately describes the distribution of infection in the population by place	The Representativeness of a surveillance system is related to the attributes of Coverage and Bias; it is a comparison of the sample and target populations with regard to a number of key features or risk factors. As such, the first step will be to identify and characterise key characteristics of the target population upon which to measure representativeness. These characteristics might be risk factors for the disease threat – knowledge of the associations between these characteristics, selection in the sample population and disease will inform the understanding of bias. Examples of relevant features include: - Livestock sector or production type - Herd/flock size - Age, sex or purpose of animal - Geographic location The second consideration of assessing representativeness is whether there is sufficient and accurate data on the identified features in both the target and sample populations.	Del Rio Vilas VJ, Pfeiffer DU. The evaluation of bias in scrapie surveillance: A review. *Veterinary Journal* 2010. 185:259-264. Lynn T*, et al.* An evaluation of scrapie surveillance in the United States. *Preventive Veterinary Medicine* 2007; 81: 70-79. Macarthur C, Pless IB. Evaluation of the quality of an injury surveillance system. *American Journal of Epidemiology* 1999; 149: 586-592. Van Benthem BHB, van Vliet JA. Reflections on an evaluation of the Dutch Infectious diseases Surveillance Information System. *Euro Surveill* 2008; 13.

Attribute	Definition	Guidance Notes	References
	and animal. Bias increases as representativeness reduces.	Where sufficient data exists, representativeness might be explored through: - simple descriptive analyses - statistical analyses (eg cross-tabulation and regression techniques, or Capture-Recapture methods) - spatial visualisation, exploration and analyses with GIS (Geographic Information Systems) tools	
Sensitivity	Sensitivity can be considered on three levels: 1) Surveillance sensitivity (case detection) refers to the proportion of individual animals or herds that have the condition of interest that the surveillance system is designed to detect; 2) Surveillance sensitivity (outbreak detection) refers to the probability that the surveillance system will detect an outbreak of disease (what constitutes an outbreak should be defined); 3) Surveillance sensitivity (presence) refers to the probability that disease will be detected if	Sensitivity is the most commonly assessed attribute of surveillance systems. Combined with timeliness, it is of particular importance to surveillance for early detection of outbreaks. With representativeness it is frequently scrutinised when evaluating surveillance activities intended to provide evidence for disease freedom. When monitoring the prevalence of endemic diseases, poor sensitivity will contribute to bias in the surveillance outputs. Surveillance sensitivity (case detection)can be assessed by - Comparing prevalence estimates from multiple systems or studies (Lynn *et al* 2007) - Considering biases and limitations in the data available and their likely impact on the estimate of sensitivity. - Using statistical methods like capture recapture methods to address the issue of availability of gold standard comparison (del Rio Vilas *et al* 2005, del Rio Vilas and Bohning, 2008) - Bayesian approaches can also be useful to estimate sensitivity (Branscum et al, 2006) in the absence of a reference test or population. Surveillance sensitivity (outbreak detection) can be assessed by - Quantifying the proportion of outbreaks of disease detected by a specific surveillance component - Applying simulation modelling methods (Audigé and Becket 1999, Willeberg *et al* 2011).	Audigé L and Becket S. A quantitative assessment of the validity of animal-health surveys using stochastic modelling. *Preventive Veterinary Medicine* 1999. 38: 259-276 Branscum AJ et al Sample size calculation for disease freedom and prevalence estimation surveys Statistics in Medicine 2006 25 2678-2674 Buckeridge DL Outbreak detection through automated surveillance: a review of the determinants of detection Journal of Biomedical Informatics 2007 40: 370-9 Cannon RM. Demonstrating disease freedom – combining confidence levels. *Preventive Veterinary Medicine*. 2002. 52: 227-249 del Rio Vilas VJ et al A case study of capture-recapture methodology using scrapie surveillance data in Great Britain 2005 PVM 67 303-17 – two list del Rio Vilas VJ and Bohning D (2008)Application of one-list capture-recapture models to scrapie surveillance data in

Attribute	Definition	Guidance Notes	References
	present at a certain level (prevalence) in the population.	- In the public health field the methods used to assess the ability of detection algorithms to detect outbreaks have been reviewed (Buckeridge, 2007, Watkins et al 2006) These include comparing the outbreaks detected by these outbreaks to previously identified outbreaks in recorded data or to simulated outbreaks sumperimposed on surveillance data (Mandl *et al* 2004, Jackson et al 2007). Surveillance sensitivity (presence) can be assessed by - Considering whether the design of the system is likely to achieve the sensitivity specified in the design of the system - Using probabilistic methods or other methods (Martin et al 2007, Hood et al 2009) - Sensitivity (presence) is usually used to assess surveillance for demonstrating freedom but can also assess performance of surveillance for early detection Some considerations when assessing the sensitivity of surveillance include - The probability of selection into the surveillance system must be defined and quantified. This may be a simple random sample of animals from a single homogenous population or a complex pathway of epidemiologic and behavioural factors describing the observation, reporting and subsequent investigation of notifiable disease (ie passive surveillance) - The probability of diagnosis (ie the sensitivity of the diagnostic protocol, including that of laboratory tests) - The choice of design prevalence (ie the expected prevalence of disease that the system is designed to detect) is a key assumption. Setting a very low design prevalence will result in a low estimate of sensitivity, placing unreasonable demand upon resources; whereas setting a high design prevalence will give an inflated estimate of sensitivity, thereby undermining credibility of the result. Choice of the design prevalence should be based upon understanding of the epidemiology of the disease. Sometimes legislation offers guidance on the design prevalence of surveillance for exotic diseases.	Great Britain PVM 85 253-66 Fujii H*, et al.* Evaluation of a sentinel surveillance system for influenza, 1995-2000, Kyoto City, Japan. *Japanese Journal of Infectious Diseases* 2002; 55: 23-26. Hadorn DC and Stärk K. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Veterinary Research* 2008. 39: 57 Hood GM, et al. Alternative methods for computing the sensitivity of complex surveillance systems. *Risk analysis* 2009 29 1686-98 Jackson ML, et al A Simulation study comparing aberration detection algorithms for syndromic surveillance BMC medical informatics and decision making 2007 7:6 Kleinman KP, Abrams AM. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 2008; 27: 4057-4068. Knight-Jones TJD*, et al.* Evaluation of the effectiveness and efficiency of wild bird surveillance for avian influenza. *Vet Research* 2010; 41. Lynn T*, et al.* An evaluation of scrapie surveillance in the United States. *Preventive Veterinary Medicine* 2007; 81: 70-79. Mandl *et al*. Measuring outbreak-detection performance by using controlled feature set simulations. *Morbidity and*

Attribute	Definition	Guidance Notes	References
			Mortality Weekly Report 2004. 53(Supp): 130-143 Martin PAJ *et al*. Demonstrating freedom from disease using multiple complex data sources: A methodology based on scenario trees. *Preventive Veterinary Medicine* 2007; 79:71-97 Watkins RE et al Approaches to the evaluation of outbreak detection methods BMC Public Health 2006 6:263 Willeberg P *et al*. Epidemiological models to support animal disease surveillance activities. *Rev Sci Tech* 2011. 30(2): 603-614
Specificity	Proportion of true non-events correctly classified as such. The inverse of this is the false alarm rate.	Evaluation of the specificity of a surveillance system is especially important for surveillance activities designed to detect outbreaks and cases because it is related to the misdirection of resources: ie expenditure on disease investigation and mitigation measures that are needlessly applied. The specificity of many surveillance activities will be very high or complete (100%), because of the consequences of confirming disease; this is especially true for surveillance for exotic diseases carrying implications for trade. Specificity can be considered at several levels, depending upon the epidemiology of the disease and the objectives and design of the system: - the specificity of pre-diagnostic indicators of disease (eg clinical signs) - the specificity of screening and confirmatory diagnostic tests applied - the rate of false-positive signals raised by detection algorithms applied to surveillance data - the proportion of reports of suspect cases of disease that are subsequently negated (NB this metric actually concerns the Positive Predictive Value of a system; a related concept which has been assessed in some evaluations	Del Rio Vilas *et al*. A case of capture-recapture methodology using scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 2005. 67: 303-317 Kleinman KP, Abrams AM. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 2008; 27: 4057-4068.

Attribute	Definition	Guidance Notes	References
Stability and sustainability	Reliability (function without failure), availability (operational when needed) and sustainability (ability of the system to be ongoing in the long term).	The Stability and Sustainability of a system is possible most pertinent to surveillance intended for early detection of new/emerging or exotic (notifiable) diseases. This attribute can be measured retrospectively by 1. Looking at the incidence of minor and major faults over a defined period of time or 2. Giving a measure of the proportion of time that the system is fully functional Assessment of this attribute will benefit from consultation with those involved in the generation, management and analysis of surveillance data. If performance indicators have been implemented in the surveillance process, historical data from these will give a good insight into the ongoing functioning of the system.	Clothier HJ*, et al.* An evaluation of the Australian Sentinel Practice Research Network (ASPREN) surveillance for influenza-like illness. *Commun Dis Intell* 2005; 29: 231-247. Hendrikx P*, et al.* Development of performance indicators for the bovine clinical salmonellosis surveillance network in France. *Journal of Veterinary Medicine Series B-Infectious Diseases and Veterinary Public Health* 2005; 52: 465-475.
Timeliness	Timeliness is the time between any two defined steps in a surveillance system.	The timeliness of a surveillance system is especially important to surveillance for the early detection of emerging or exotic disease threats – where the intention is to implement control measures as soon as possible. The time points chosen are likely to vary depending on the purpose of the surveillance activity. For outbreak detection this can be defined using various time points including the time between exposure to the infectious agent and the initiation of risk mitigation measures or the time between when disease could have been detected and when it actually was reported. For planning purposes, timeliness can also be defined as whether surveillance detects changes in time for risk mitigation measures to reduce the likelihood of further spread. The precise definition of timeliness chosen should be stated as part of the evaluation process. For surveillance systems designed to detect cases of disease the CDC guidelines (CDC 2001) describe a useful approach to measuring the timeliness of a surveillance system which is focussed on the time taken to process surveillance data. In brief: 1. Map the surveillance process, from sample collection and handling, through the diagnostic process, management and analysis of data and reporting and	Carpenter TE. Evaluation and extension of the cusum technique with an application to Salmonella surveillance. *Journal of Veterinary Diagnostic Investigation* 2002; 14: 211-218. CDC Updated guidelines for evaluating public health surveillance systems. *Morbidity and Mortality Weekly Report* vol 50. 2001 CDC for evaluating public health surveillance systems for early detection of outbreaks 2004 MMWR 53:1-11 Colby *et al*. Evaluation of two systems for managing emergency poultry diseases in intensive poultry productions regions. *Int J Poul Sci*. 2003 2(3): 234-241 Del Rocio Amezcua *et al*. Evaluation of a

Attribute	Definition	Guidance Notes	References
		dissemination of results. The description of the surveillance system developed in section 2 of the framework will aid in this. 2. Identify key time intervals for measurement. The timeliness measure should be aligned with the objectives of the system and in some cases more than one measure may be required to gain sufficient understanding. Some examples of relevant timeliness measures include: a. For passive surveillance activities, the interval between observation of the first clinical signs of disease and laboratory investigation b. For post-import testing for exotic notifiable disease, the interval between entry to the country and the return of a laboratory result c. For ongoing active surveillance of endemic diseases, one measure might be the frequency of analysis and publication of surveillance data 3. Just as the timeliness measure may differ between systems, the criterion for timeliness will differ between disease threats. The timeliness of a system should be assessed with consideration of the epidemiology of the disease of interest (eg by comparison to the generation interval of infectious diseases) 4. When the key time intervals have been identified and fully characterised, data should be collected to measure the timeliness. Where sufficient event data exists, calculations will be relatively simple. In the absence of sufficient valid data, simulation models might be developed and applied. 5. Assessment should also consider factors which influence the timeliness of a system. For example, availability of human resources for the collection of samples of investigation of reported disease, availability of laboratory facilities, adoption of new technology to streamline laboratory investigation, adoption of automated approaches to the collation and management or the analysis and reporting of surveillance data, etc. Examples of studies assessing the timeliness of surveillance for case detection include Jajosky and Groseclose, 2004 and Takahoshi et al 2004 For surveillance systems designed to detect outbreaks, timeliness is often assessed in combination with sensitivity (Kleinman and Abrams 2008). Guidance on the assessment of timeliness for this type of surveillance system is provided by the CDC framework (2004)	veterinary-based syndromic surveillance system implemented for swine. *Canadian Journal of Veterinary Research.* 2010. 74(4):241-251 Jackson ML et al A Simulation study comparing aberration detection algorithms for syndromic surveillance BMC medical informatics and decision making 2007 7:6 Jajosky and Groseclose 2004. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* 4:29 Kleinman KP, Abrams AM. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 2008; 27: 4057-4068. Mandl *et al*. Measuring outbreak-detection performance by using controlled feature set simulations. *Morbidity and Mortality Weekly Report* 2004. 53(Supp): 130-143 Riera-Montes and Velicko 2011. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2 Siegrist et al Bio-ALIRT Biosurveillance Detection Algorithm Evaluation MMWR Suppl 2004 53 152-8 Takahashi T et al Evaluation of a public health Salmonella surveillance system in King County, Washington American Journal

Attribute	Definition	Guidance Notes	References
		which focuses more on the time taken to detect outbreaks. Timeliness for these surveillance systems has been assessed by: - assessing the time to detect naturally occurring outbreaks (Siegrist et al, 2004) - assessing the time to detect simulated outbreaks (Mandl *et al* 2004, Jackson et al 2007) - using a simulation model to predict the time to detect outbreaks (Yamamoto et al, 2008)	of Infection Control 2004 32 7-11 Yamamoto T et al Evaluation of surveillance strategies for bovine brucellosis in Japan using a simulation model PVM 2008 86 57-74