

---

## REVIEW ARTICLE

# Evaluation of animal and public health surveillance systems: a systematic review

---

J. A. DREWE<sup>1\*</sup>, L. J. HOINVILLE<sup>2</sup>, A. J. C. COOK<sup>3</sup>, T. FLOYD<sup>2</sup>  
AND K. D. C. STÄRK<sup>1</sup>

<sup>1</sup> *Centre for Emerging, Endemic and Exotic Diseases, Royal Veterinary College, University of London, UK*

<sup>2</sup> *Centre for Epidemiology and Risk Analysis, Animal Health & Veterinary Laboratories Agency, Addlestone, Surrey, UK*

<sup>3</sup> *Veterinary Surveillance Department, Animal Health & Veterinary Laboratories Agency, Addlestone, Surrey, UK*

*(Accepted 6 October 2011; first published online 11 November 2011)*

### SUMMARY

Disease surveillance programmes ought to be evaluated regularly to ensure they provide valuable information in an efficient manner. Evaluation of human and animal health surveillance programmes around the world is currently not standardized and therefore inconsistent. The aim of this systematic review was to review surveillance system attributes and the methods used for their assessment, together with the strengths and weaknesses of existing frameworks for evaluating surveillance in animal health, public health and allied disciplines. Information from 99 articles describing the evaluation of 101 surveillance systems was examined. A wide range of approaches for assessing 23 different system attributes was identified although most evaluations addressed only one or two attributes and comprehensive evaluations were uncommon. Surveillance objectives were often not stated in the articles reviewed and so the reasons for choosing certain attributes for assessment were not always apparent. This has the potential to introduce misleading results in surveillance evaluation. Due to the wide range of system attributes that may be assessed, methods should be explored which collapse these down into a small number of grouped characteristics by focusing on the relationships between attributes and their links to the objectives of the surveillance system and the evaluation. A generic and comprehensive evaluation framework could then be developed consisting of a limited number of common attributes together with several sets of secondary attributes which could be selected depending on the disease or range of diseases under surveillance and the purpose of the surveillance. Economic evaluation should be an integral part of the surveillance evaluation process. This would provide a significant benefit to decision-makers who often need to make choices based on limited or diminishing resources.

**Key words:** Animal pathogens, public health, surveillance, veterinary epidemiology.

### INTRODUCTION

Disease surveillance in both animal and public health fields involves the ongoing systematic collection, analysis, interpretation and timely communication of

health-related data [1]. The purpose of surveillance activities may include monitoring of endemic diseases and the impact of control measures or the identification of (re-)emerging and exotic diseases that may have a significant impact upon public health, animal health, welfare and international trade [2]. Animal health surveillance includes animal conditions which may pose a threat to human health – either directly or

\* Author for correspondence: Dr J. A. Drewe, Centre for Emerging, Endemic and Exotic Diseases, Royal Veterinary College, Hawkshead Lane, North Mymms, Herts, AL9 7TA, UK.  
(Email: jdrewe@rvc.ac.uk)

via food products – even where such conditions are unapparent in the animal itself [3]. The output of surveillance programmes assists in setting priorities and guiding effective prevention and control strategies. It also helps to monitor the progress and success of intervention programmes and, in the animal health field, to demonstrate the infection- and hazard-free status of animals and animal-derived products [4]. Ensuring that surveillance programmes are fit for purpose is therefore paramount.

The costs of obtaining surveillance information need to be balanced against the benefits derived. The importance of ensuring that public health systems are efficient and effective is increasingly being recognized [5–7] and this applies equally to animal health surveillance systems [6, 8]. Improving the efficiency of surveillance is a key goal of the UK's Veterinary Surveillance Strategy [3]. Evaluation of surveillance programmes is essential to ensure that limited resources are effectively used to provide the evidence required for protecting animal (and human) health. Such evaluations can lead to changes in surveillance methods, resulting in considerable financial savings [9]. Evaluation of surveillance can play an essential part in establishing and maintaining international trust [10]. Quality assurance is essential to maintain credibility, which is particularly important for inter-community and international trade with animals and animal-derived products.

Evaluation is defined as the systematic and objective assessment of the relevance, adequacy, progress, efficiency, effectiveness and impact of a course of actions, in relation to objectives and taking into account the resources and facilities that have been deployed [11]. Currently, there is no universally accepted standard for evaluation of animal health surveillance. Many different approaches have been applied [e.g. 2, 12–16] without consistency or apparent agreement on what is optimal. Evaluation of human health surveillance systems is more commonly practised and several generic guidelines exist for public health surveillance evaluation [17–20]. These typically include the assessment of a series of attributes such as flexibility, acceptability and timeliness, using a combination of qualitative and quantitative techniques. Probably the most well-established guidelines for evaluating public health surveillance systems are those published by the Centers for Disease Control and Prevention (CDC) in the USA [17]. The CDC guidelines suggest ten attributes that may be assessed as part of a balanced evaluation process: simplicity, flexibility, data

quality, acceptability, sensitivity, positive predictive value, representativeness, timeliness, stability and usefulness. However, the applicability of these human health guidelines to animal health surveillance is unclear given differences in emphasis and prioritization of surveillance objectives between the disciplines. For example, cost-effectiveness might be considered to be most important in animal health surveillance programmes whereas diagnostic accuracy may be more valued in public health schemes because of the consequences of classification errors for the individual case. An animal health-specific evaluation framework that is both comprehensive and generic is required but not currently available.

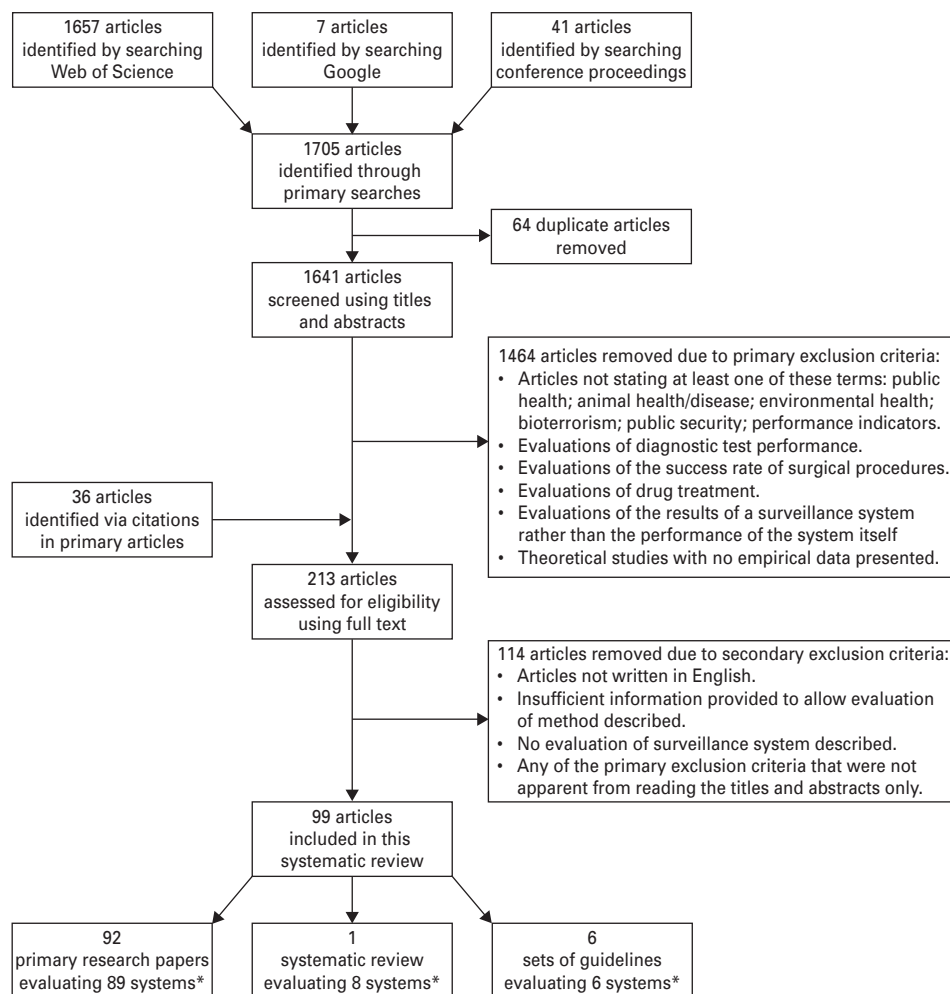
The aim of this systematic review is to identify and examine existing frameworks for surveillance evaluation in animal health, public health and allied disciplines, to discover which techniques are currently being used across the globe and to assess their strengths and weaknesses. This information will be used to inform the development of a generic evaluation framework for animal health surveillance systems in Great Britain.

## METHODS

We sought to identify published and unpublished reports of evaluations conducted on surveillance systems in the following areas: animal health/disease; public health; environmental health; bioterrorism; and public security.

### Literature sources and search strategies

Three sources were searched by one author (J.A.D.) for relevant reports: Web of Science databases of peer-reviewed articles; Google search of grey literature; and conference proceedings of the International Society for Veterinary Epidemiology and Economics, and the Society for Veterinary Epidemiology and Preventive Medicine. Additionally, we identified additional articles from the bibliographies of included articles. We searched for literature containing combinations of the terms 'surveillance', 'evaluation', 'analysis' and 'performance' using the Boolean query: Topic = surveillance AND Title = (surveillance AND (evaluat\* OR analy\* OR perform\*)) OR (evaluat\* AND perform\*). The use of wildcards (\*) ensured that articles containing any variation of each of the search terms were identified (e.g. evaluat\* would detect evaluate, evaluates, evaluation, evaluations,



**Fig. 1.** Flow chart documenting literature retrieval and criteria used to select articles for inclusion in the systematic review of frameworks for evaluating human and animal health surveillance systems. \* The total number of surveillance system evaluations was 101: one system included in the systematic review, and another that was included in one of the sets of guidelines, were both already identified and included in this review as primary research papers.

evaluating, evaluative and evaluator). We used identical search terms for all sources. We restricted the searches of Web of Science and Google to articles published in the last 15 years (i.e. 1995–2010) and that of the conference proceedings to articles published in the last 10 years (i.e. 2000–2010).

### Study selection and data abstraction

The literature retrieval process is illustrated in Figure 1. Articles were screened using the criteria detailed in Figure 1. Primary exclusion criteria were applied to the titles and abstracts of articles. The full texts of these articles were then obtained and the secondary exclusion criteria applied. Data extracted from the articles about each surveillance system included: aim of evaluation exercise, surveillance system

evaluated, location, species involved, disease/condition, data collected, collection method, analysis method performed, specific attributes assessed, use of performance indicators (PIs), and the perceived strengths and weaknesses of each evaluation approach. Articles were included in this review if they presented data from the estimation of at least one attribute of a surveillance system.

### RESULTS

A total of 1741 articles were screened for this review: 1705 primary articles identified through searching Web of Science, Google and epidemiology conference proceedings, and 36 additional articles identified by examining the citation lists of these primary articles (Fig. 1). Nineteen articles written in languages other

Table 1. *Species for which health surveillance systems were evaluated*

Species	No. of articles	References
Humans	73	[7, 17, 21, 24–37, 39, 40, 42, 43, 45, 46, 48–52, 54–57, 60, 62–72, 74, 76, 77, 79–82, 85–88, 90–93, 95–105, 107–109]
Cattle	13	[9, 13, 22, 23, 38, 41, 53, 58, 61, 75, 78, 84, 106]
Sheep	3	[12, 16, 73]
Animals (>1 species)	2	[14, 47]
Horses	2	[83, 89]
Pigs	2	[59, 94]
Wild birds	2	[8, 44]
Fish	1	[15]
Humans and animals	1	[6]
<b>Total</b>	<b>99</b>	

than English were excluded. After applying all exclusion criteria (Fig. 1), 99 articles remained. Of these, 92 were primary research papers [6–9, 12–16, 21–103], one was a systematic review [104], and six were sets of guidelines for evaluating surveillance systems [17, 105–109] (Fig. 1). Data from these 99 articles were extracted and included in this review.

#### Health conditions for which surveillance systems were evaluated

The 99 articles included evaluations of 101 different surveillance systems: some articles evaluated more than one system and some systems were evaluated (in different ways) by more than one article. Most (73/99) of the articles on surveillance system evaluations were for human diseases, with far fewer for animal diseases: cattle were the most frequent subject of the animal health surveillance systems that had been evaluated (13/99) (Table 1). Only one article integrated the evaluation of human and animal health surveillance, in a study of West Nile virus epidemiology [6].

Surveillance systems for 38 named diseases (27 infectious and 11 non-infectious) were evaluated within the 99 articles (Table 2). Influenza was the disease for which surveillance was most frequently evaluated (eight articles: seven in humans and one in wild birds). While the majority of surveillance system evaluations focused on a single disease, about one quarter

(27/101) did not specify a particular disease with the implication that the evaluation was applied generically across a range of conditions. Of those that could be considered generic, a variety of information sources were exploited. Two evaluations were internet-based, using the web as an automated system for real-time monitoring of online information about emerging diseases [6, 36]. One evaluated system logged and rapidly analysed telephone calls made by members of the public to the NHS Direct helpline [46]. Four evaluations focused on the ability of surveillance systems to detect disease outbreaks early, including two dedicated to illnesses among military personnel engaged in battle [7, 65, 66, 105]. Finally, three sets of generic guidelines were presented from which more specific frameworks could be developed [17, 91, 107]. Of these, one outlined a novel stepwise process by which a prioritized list of evaluation criteria may be generated for any chosen disease or health indicator [107]. Although this system was developed for the selection and evaluation of public health indicators from environmental threats, its high flexibility suggests it should be adaptable for the evaluation of animal health surveillance.

#### Locations of surveillance systems evaluated

The 99 evaluations included in this review covered 101 surveillance systems located unevenly across the globe and dominated by the USA (Table 3). Most evaluations related to surveillance systems implemented in North America (35/101), followed by Europe (26/101), Australasia (11/101), Asia (6/101), Africa (5/101) and South America (4/101). Thirteen articles described the evaluation of surveillance systems located in more than one country (Table 3).

#### Methods employed to evaluate surveillance systems

Quantitative approaches were applied far more commonly than qualitative approaches, and this was especially true for evaluations of animal health surveillance. A summary of the methods used for evaluating surveillance systems appears in Table 4. The commonest method (employed in 39/101 evaluations) was to apply simple quantitative approaches such as measuring the proportion of actual cases reported or the percentage of record cards completed [e.g. 39, 56, 75, 77, 92, 98, 101]. Comparison of one surveillance system with another to estimate the relative sensitivity of each (the proportion of cases in the population

Table 2. *Health conditions for which surveillance systems were evaluated*

Health condition	No. of articles	References
Infectious disease (generic or no specific disease stated)	27	[6, 7, 14, 15, 17, 31, 33, 36, 39, 46, 47, 52, 60, 65, 66, 68, 78, 80, 89, 91, 92, 98, 101, 103–105, 107]
Influenza	8	[8, 42, 48, 49, 67, 69, 71, 96]
Congenital illness	5	[50, 51, 57, 87, 95]
HIV/AIDS	5	[21, 25, 54, 81, 99]
Malaria	4	[30, 32, 35, 40]
Measles	3	[55, 56, 88]
Physical injury	3	[28, 74, 108]
Scrapie	3	[12, 16, 73]
Salmonellosis	3	[13, 58, 94]
Tuberculosis	3	[53, 79, 86]
Bacterial meningitis	2	[27, 45]
Brucellosis	2	[9, 22]
Infectious abortion	2	[38, 83]
Infectious bovine rhinotracheitis	2	[41, 84]
Infectious hepatitis	2	[77, 93]
Rinderpest (cattle plague)	2	[75, 106]
Anthrax	1	[64]
Bioterrorism-related illness	1	[109]
Chemical injury	1	[34]
Creutzfeldt–Jakob disease	1	[90]
Carbon monoxide poisoning	1	[63]
Dracunculiasis (guinea worm disease)	1	[26]
Enteric disease	1	[43]
Enzootic bovine leucosis	1	[61]
Foot-and-mouth disease	1	[59]
Haemolytic-uraemic syndrome	1	[70]
Legionellosis	1	[82]
Listeriosis	1	[102]
Polio	1	[100]
Respiratory infection	1	[37]
Rotavirus	1	[29]
Rubella	1	[85]
Sexually transmitted infection	1	[62]
Staphylococcal toxic-shock syndrome	1	[72]
Streptococcal mastitis	1	[23]
Syphilis	1	[24]
Tsutsugamushi disease	1	[76]
Varicella (chickenpox)	1	[97]
West Nile virus infection	1	[44]
<b>Total</b>	<b>99</b>	

under surveillance that are detected by the surveillance system) was also frequently done [e.g. 9, 32, 49, 61, 72, 78]. Several advanced statistical approaches were employed to evaluate surveillance systems. These included 19 articles that employed simulation modelling (used far more commonly in evaluations of animal than human health surveillance), seven articles on stochastic scenario tree modelling [8, 9, 14, 53, 59, 61, 100], five articles using the capture–recapture (CRC) technique [30, 40, 52, 82, 87] and one article

which quantified the effort applied in looking for infection (expressed as the number of negative measles test results per 100 000 population) which gave an indication of the confidence that could be associated with the conclusion of infection being absent if it were not found [55]. Qualitative approaches based on subjective scoring systems or expert opinion were less commonly used and mainly restricted to evaluations of human health surveillance [e.g. 57, 62, 66, 90, 94, 101]. Many articles used more than one approach to

Table 3. Locations where the 101 surveillance systems whose evaluations were included in this review were implemented

Location	No. of surveillance systems	References
USA	26	[13, 24, 25, 30, 54, 56, 63, 64, 67, 71–73, 79, 83, 86, 89, 95, 103, 104, 107, 109]
More than one country	13	[6–8, 14, 17, 31, 36, 47, 55, 91, 99, 105]
UK	10	[9, 12, 15, 16, 22, 29, 34, 46, 62, 81, 87]
Australia	9	[42, 43, 50, 57, 59, 61, 80, 90, 100, 108]
Canada	8	[33, 37, 44, 51, 70, 74, 85, 102]
France	4	[47, 58, 82]
Denmark	3	[23, 38, 41]
Norway	3	[21, 84, 94]
Italy	2	[39, 52]
Japan	2	[48, 49, 76, 96]
New Zealand	2	[68, 78]
Pakistan	2	[27, 93]
Switzerland	2	[53, 88]
Chad	1	[47]
French Guiana	1	[66]
Ghana	1	[26]
Israel	1	[69]
Jamaica	1	[28]
Mexico	1	[32]
Netherlands	1	[98]
Niger	1	[45]
Peru	1	[60]
Poland	1	[77]
South Africa	1	[101]
Taiwan	1	[97]
Tanzania	1	[92]
Trinidad	1	[40]
Not specified	1	[65]
<b>Total</b>	<b>101</b>	

evaluate a surveillance system, for example combining quantitative measures of data completeness (an indication of quality) with qualitative impressions of acceptability of the system to users [e.g. 46, 57, 62, 66, 80, 101].

#### Attributes of surveillance systems

The range and number of attributes assessed by the different studies varied widely. In total, 23 different attributes of surveillance systems were assessed across the 99 articles (Fig. 2). These attributes are defined in Table 5. The most frequently assessed attributes were sensitivity, timeliness and data quality (Fig. 2). The frequency distribution of the number of attributes assessed per article was positively skewed, with approximately half the articles (48/99) assessing one or two attributes only and very few articles assessing

more than ten attributes (Fig. 3). Twenty-four articles reported on the assessment of a single attribute: sensitivity (13 articles); cost-effectiveness (four articles); representativeness (three articles); timeliness (three articles); and acceptability (one article). Attributes such as consistency of performance over time [99], system security [108], and surveillance feasibility [7] were assessed in a single article each (Fig. 2).

Almost a quarter (23/99) of the articles specifically stated as an objective to assess one or more of the ten attributes recommended in the CDC guidelines for evaluating public health surveillance systems [17]. Only five articles [15, 32, 38, 81, 86] did not assess any of the ten attributes recommended by CDC and all of these five articles assessed the cost-effectiveness of surveillance programmes. A variety of ways was used to assess cost-effectiveness. By determining the relative costs of several regimens to each generate a 1%

Table 4. Summary and distribution of methods for evaluating surveillance systems used in the 99 articles included in this review

Method of analysis	No. of articles		
	Animal	Human	Total*
<b>Quantitative approaches</b>	<b>25</b>	<b>64</b>	<b>89</b>
Calculation of percentage of complete records†	7	33	39
Comparison of one system with another	4	25	29
Simulation modelling or statistical algorithms	14	5	19
Scenario tree modelling	6	1	7
Cost–benefit analysis	2	4	6
Capture–recapture technique	0	5	5
Performance indicators	3	2	5
Odds ratios of disease detection probability	1	0	1
Measurement of effort applied	0	1	1
<b>Qualitative approaches</b>	<b>2</b>	<b>24</b>	<b>26</b>
Subjective scoring system or expert opinion	1	22	23
Spatial mapping	1	1	2
Logic model	0	1	1

\* Figures do not sum to 99 because several articles used more than one approach.  
 † One article using this approach evaluated health surveillance in both animals and humans.

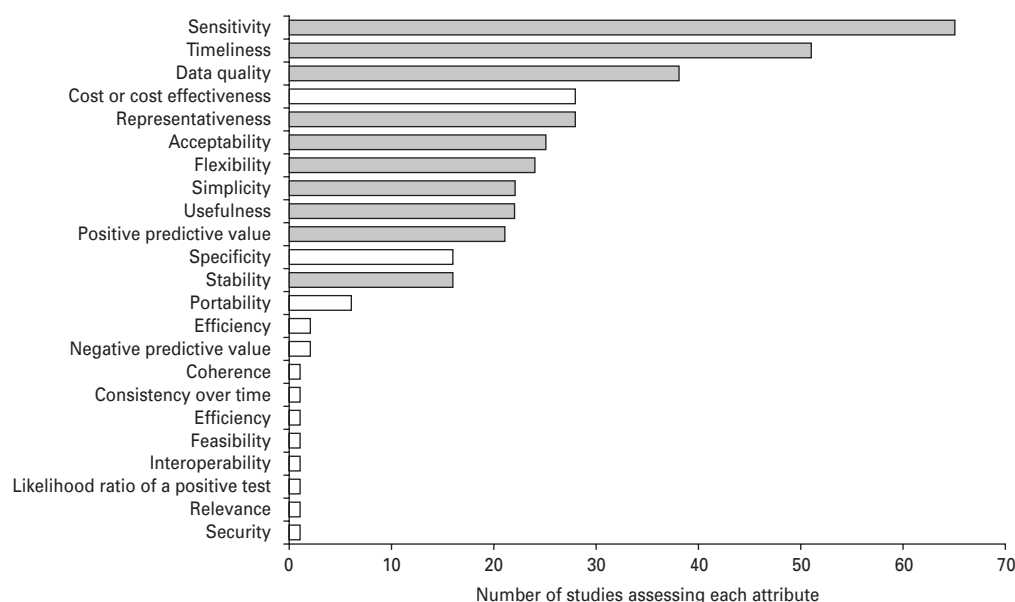


Fig. 2. Surveillance system attributes assessed by the 99 studies included in this review. Attributes recommended for evaluation in the Centers for Disease Control and Prevention guidelines [17] are shaded in grey.

increase in surveillance sensitivity, a comparison of efficacy was achieved which allowed the most cost-effective regimen to be recommended for surveillance of influenza in wild birds [8]. A slightly different approach was employed in an evaluation of human tuberculosis surveillance, where cost per 1% increase in PIs was used to determine the most efficient use of resources programmes [86]. While comparing the

cost-effectiveness of two or more surveillance regimens may be relatively straightforward, conducting a cost–benefit analysis may prove more difficult. Costs may be difficult to estimate [86] and social benefits may be difficult to quantify [81] which means precise cost–benefit evaluations may not always be possible.

Other non-CDC attributes that were commonly assessed included specificity (the proportion of true

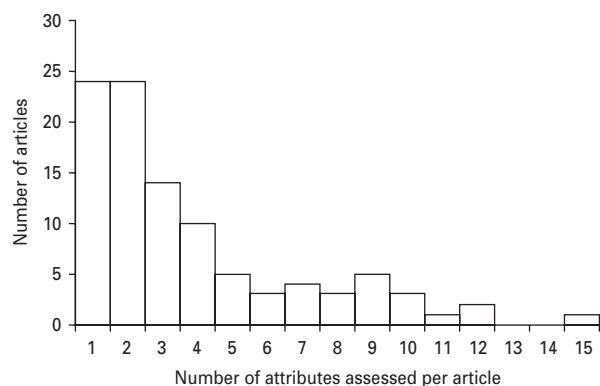
Table 5. *Definitions of surveillance system attributes assessed in the 99 articles included in this review (sources of definitions: [7, 17, 99, 105, 108])*

Attribute	Definition
Acceptability	Willingness of persons and organizations to participate in the surveillance system
Coherence	Link between the different components and the development stages of a surveillance system
Consistency over time	Deliberate repetition in sampling the same geographical sites over time to allow trends in epidemics to be measured
Cost-effectiveness	Relationship between the expected outcomes (such as the number of lives saved) and the costs of surveillance required to achieve this. May be expressed as a measure of efficiency, whereby the system operates at the least possible cost or makes the best use of available resources
Data quality	Completeness and validity of the data recorded
Efficacy	Extent to which the system objectives are achieved
Efficiency	Link between the resources implemented and the results obtained. An efficient system will accomplish a job with minimum expenditure of time, human effort and cost
Feasibility	Extent to which the available means meet the system's needs. A surveillance system may be unfeasible if there are not the means to run it
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
Interoperability	Ease with which one surveillance system can be integrated into another, in an appropriate format for people in other countries to easily use
Likelihood ratio of a positive test	Ratio of the probability of a surveillance system detecting an infected individual to the probability of the system incorrectly identifying them as infected when they are in fact not. Likelihood ratios do not vary with disease prevalence and so are stable expressions of system performance
Negative predictive value	Probability that infection is truly absent given that it is not detected
Portability	How well the system can be duplicated in another setting
Positive predictive value	Proportion of reported cases that actually have the infection of interest
Relevance	Assessment of the how closely the outputs of a surveillance system meet its objectives. Also referred to as pertinence
Representativeness	Extent to which features of the population of interest (e.g. herd size, age, location) are reflected in the surveillance data that are collected. A surveillance system that is representative accurately describes the distribution of infection in the population by place and person (or animal). Bias reduces representativeness
Security	Measures taken to assure authorized computer system access and to maintain confidentiality where needed
Sensitivity	The sensitivity of a surveillance system can be considered on two levels. For endemic infections, sensitivity refers to the proportion of cases of a disease detected by the surveillance system (this usually requires a gold standard test to indicate the actual number of cases). For non-endemic infections, sensitivity can refer to the ability of a surveillance system to detect disease outbreaks
Simplicity	Refers to the surveillance system structure, ease of operation and flow of data through the system. Surveillance systems should be as simple as possible while still meeting their objectives
Specificity	Proportion of true non-events correctly classified as such, the inverse being the false alarm rate
Stability	Reliability (function without failure) and availability (operational when needed)
Timeliness	Speed between steps in a surveillance system. For outbreak detection, timeliness refers to the time between exposure to the infectious agent and the initiation of interventions to control infection
Usefulness	Actions taken to protect health based on the information provided by the surveillance system. A measure of the impact of the surveillance system

non-events correctly classified as such, the inverse being the false alarm rate: assessed in 16/99 articles) and portability (how well the system can be duplicated in another setting: assessed in 6/99 articles). Several

articles described schemes which set out to assess more attributes than were subsequently possible due to lack of reliable data. This was particularly apparent for specificity [28, 36], sensitivity [50, 58] and





**Fig. 3.** Frequency distribution of the number of surveillance attributes assessed per article.

representativeness [46, 50], where accurate data on the presence/absence of infection and the demographic heterogeneity is needed if these attributes are to be calculated. Representativeness in particular is rarely assessed fully [73, 110].

Relationships between attributes were rarely investigated. An exception was a Bayesian network model for analysis of detection performance in surveillance systems, an approach which offered insights into the trade-offs between sensitivity, specificity and timeliness [65].

### Performance indicators

Five articles described the development and calculation of PIs to evaluate surveillance systems for tuberculosis in humans [79, 86], rinderpest (cattle plague) [75, 106] and bovine clinical salmonellosis [58]. PIs are time-delimited, denominator-based statistics [75] which can be used to monitor the implementation of surveillance systems rather than for the periodic evaluation of surveillance activities to determine whether these activities are meeting their objectives. They allow the progress of surveillance to be monitored by providing quantitative comparisons of elements of the activity over time [106]. An example of a PI would be the number of cases of the condition of interest properly reported within 7 days of diagnosis, per 100 000 population. In the case of the rinderpest eradication programme, PIs were further subdivided into diagnostic indicators [106]. Diagnostic indicators are more detailed than PIs and measure specific sub-steps in the surveillance process [75]. Examples of diagnostic indicators include the number of individuals sampled for which results were reported, or the percentage of districts with up-to-date report registries.

### Generic evaluation frameworks

Four articles described three generic evaluation frameworks which could be applied to a range of diseases and situations [7, 17, 105, 107]. The generic nature of these frameworks comes about from their common structure which allows priorities to be varied according to the specific objectives of each surveillance programme. A series of core elements (such as zoonotic importance or public concern) reflect the different purposes of surveillance and may be chosen accordingly. Each of these core elements contains a selection of criteria to be evaluated (such as strength of evidence). The criteria are judged through assessment of attributes of the surveillance system. By varying the priority of the core elements depending on the surveillance objectives and choosing a different selection of criteria to be evaluated using a range of attributes each time, these frameworks appear flexible and truly generic.

### DISCUSSION

A distinct lack of standardization exists regarding the best approach for evaluating surveillance systems in order to facilitate decision-making in the fields of animal or public health. The ten attributes recommended for evaluation by CDC [17] – simplicity, flexibility, data quality, acceptability, sensitivity, positive predictive value, representativeness, timeliness, stability and usefulness – were often assessed but usually singly or in pairs rather than all ten together. An evaluation based on only one or two attributes is not likely to provide a complete, unbiased evaluation of a surveillance system since multiple indicators are needed for tracking the implementation and effects of a programme [17]. Given that evaluation is defined as the *systematic* assessment of the quality of something, the large proportion of articles included in this review that assessed only one or two attributes cannot be considered complete evaluations. Indeed, it could be argued that only about one quarter of the articles in this review (27/99) performed a systematic assessment, by addressing five or more attributes (Fig. 3) to form a balanced evaluation of a surveillance system. While 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. Defining too few will not result in a systematic assessment and defining too many can detract

from the evaluation's goal by making it a huge task to gather data and making interpretation difficult. In some cases a complete evaluation may not be required to achieve the objectives of the evaluation process. For example, an evaluation of a limited number of parameters allowed the relative value of different surveillance strategies to be assessed resulting in recommendations that allowed substantial cost savings [9].

Focusing on the relative value and relationships between attributes may allow the identification of a limited number of 'core' characteristics which when all considered will allow for a holistic evaluation. For example, simplicity appears to be positively related to acceptability, with staff willingness to participate in surveillance being high if the system is simple and easy to use [105]. In the same way, a reliable system (one that functions without failure, which often means absence of complex technology) is likely to have higher acceptability to users than a system that frequently fails [66]. Thus, assessment of acceptability should capture much of the essence of simplicity and reliability too. Similarly, the three attributes sensitivity, specificity and positive predictive value all give related information and so assessment of one or two might be sufficient. Some authors have made suggestions for grouping related or comparable attributes [107, 109]. However, it may still be important to evaluate several related attributes individually. For example, a system could be extremely sensitive (detecting all cases of a disease) but if specificity was low, many of the apparently positive cases would in fact be false positives. This would dilute any benefit provided by the high sensitivity. Because sensitivity and specificity are related, but provide different information, they ought to be estimated simultaneously [5], taking into account the evaluation objectives. Theoretical work indicates it may be possible to incorporate sensitivity, specificity and timeliness into a single metric [111] although interpretation of the combined measure is not straightforward. In addition, some attributes may provide information that is more relevant to the assessment of the worth of a surveillance system than other related attributes. For example, it has been suggested that the number of lives saved could be used rather than timeliness to evaluate surveillance systems for outbreak detection [112].

At best, only moderate agreement seems to exist concerning which attributes of a surveillance system should be assessed. This may be because the value of each attribute to decision-makers varies depending on

the surveillance objectives. A surveillance system designed to prove freedom from infection will require a higher sensitivity than a system which tracks the prevalence of a widespread endemic disease, for example. Although sensitivity and timeliness were each assessed in over half of the studies included in this review, this may be as much due to data availability and their ease of calculation rather than their particular usefulness to decision-makers in facilitating early detection of infection. Surveillance objectives were often not stated in the articles reviewed and so the reasons for choosing certain attributes for assessment were not always apparent. The objective of the evaluation process should be clearly stated and the evaluation designed accordingly, rather than being dictated by convenience. An assessment of the purpose of the surveillance activity should be included as part of the evaluation process.

Both quantitative and qualitative methods were used as part of the evaluation process. The commonest approach to evaluation used in the systems reviewed – comparing surveillance output (such as the number of reported cases of the disease under surveillance) with another dataset (e.g. official records of the incidence of the same disease) – should be applied with caution. In the absence of a reliable gold standard against which to compare, this approach has the potential to introduce significant bias [110]. In such situations, relative sensitivity rather than true sensitivity of the surveillance system is being determined [e.g. 73, 86, 97]. These figures may be very different, and artificially inflate the apparent sensitivity of the system. Sensitivity need not be high for a surveillance system to be useful (exactly how high it needs to be will vary with surveillance objectives) but the surveillance methodology must remain stable over time if trends in sensitivity are to be interpreted meaningfully [17, 113]. A single measurement of relative sensitivity is on its own arguably of little use.

A possible solution to under-ascertainment is the use of CRC methods to estimate the unobserved population with the event under study. This gives information on the size of the 'true' population against which surveillance output can be assessed. This approach, commonly used in ecological studies [114], is increasingly being applied to evaluate public health surveillance [30, 40, 52, 82, 87], but has rarely been applied in animal health surveillance evaluation [115]. The aim of CRC models is to estimate the number of individuals with the characteristics of interest that are not detected by any of the surveillance sources in

place. Once this estimate is obtained, it is possible to compute the total population with that characteristic, as well as the sensitivity of the surveillance system [110]. Greater application of CRC methods would enhance animal health surveillance evaluation by improving the accuracy of sensitivity estimates. It is likely to be beneficial to use a combination of quantitative (e.g. CRC) and qualitative (e.g. interviewing the surveillance actors) techniques to assess each attribute so that information is captured on the possible reasons for a certain measured level and perhaps indicate actions which may be undertaken to improve it.

It is surprising that economic evaluation is not an integral part of more surveillance evaluation programmes: only 28/99 articles included an assessment of cost or cost-effectiveness. The cost of obtaining surveillance information needs to be balanced against the benefits derived, so examining the outputs of surveillance is only half the process. The most commonly followed guidelines – those of CDC [17] – suggest that costs may be judged relative to benefits but do not explicitly advocate that this be an integral part of all surveillance evaluations nor indicate how this may be done. This is an area that could usefully be expanded, since the CDC guidelines were the source most frequently referred to by articles included in the present review. While ensuring cost-effectiveness might be expected to be of higher priority in animal rather than human health surveillance (where social factors might be valued higher), there appears to be no difference between human and animal surveillance schemes, with 29% (21/73) of human studies and 27% (7/26) of animal studies including cost-effectiveness as part of the surveillance evaluation process.

PIs were used in the evaluation of only four surveillance programmes covering three diseases (two of animals and one of humans). Given that PIs allow continuous monitoring – as opposed to periodic evaluation – of a surveillance system, it is perhaps surprising that they are not more widely documented. One reason may be a perception that they require the collection and analysis of lots of detailed data. However, this need not – in fact, should not – be the case [106]. PIs should reduce the wide range of quantitative information to one or two measures. Such measures can be even more informative if they include a measure of cost-effectiveness, such as cost per percentage point increase in PIs [86]. PIs should be set at a realistic target balancing minimal requirements with objectives for improvement [75]. Another possible

reason for the scarcity of PIs in the literature could be if they are used as part of an ongoing internal monitoring programme but not published as part of a formal evaluation. Proof that PIs are a valuable and effective tool comes from their pivotal role in the recent announcement of global rinderpest elimination, the first ever animal disease to be wiped out [116]. Increased use – and reporting – of PIs would enhance animal health surveillance evaluation by providing robust summary quantitative measures of changes in disease patterns over time.

Clear definitions and agreement on what each attribute, indicator or criterion actually measure is essential if surveillance evaluations are to be comparable and universally understood. The most accepted reference for definitions of attributes appears to be the CDC guidelines for evaluating public health surveillance systems [17], although not everyone appears to follow these. For example, completeness and accuracy are included under the definition of sensitivity in one paper [95] while they are taken to indicate quite different things elsewhere [17]. Use of non-standard terms can also lead to confusion. The distinction between ‘reporting efficiency’ and sensitivity is not clear in one article [29]. Similarly, the difference between generalizability and portability is not explicit in an evaluation framework [109]. Last, the use of the word ‘quality’ to describe surveillance schemes in a general sense [99] (as opposed to meaning the completeness and accuracy of data) has the potential to introduce misunderstanding. Definitions for the terms used in this review are included in Table 5 and discussions have been initiated within the veterinary surveillance community to clarify and where possible standardize terminology [117]; these discussions continued at a workshop prior to the International Conference on Animal Health Surveillance in May 2011 (<http://www.animalhealthsurveillance.org/>).

In conclusion, there is currently no universally accepted standardized process to evaluate surveillance systems for animal and human health. The most commonly cited guidelines for evaluating public health surveillance systems – those of CDC [17] – have been adapted for specific situations in the public health field [7] and could be adapted for animal health use as they were to produce an evaluation protocol that was applied to evaluate scrapie surveillance in the USA [73]. However, the CDC guidelines do not provide an ‘off-the-shelf’ framework; rather they include a broad selection of attributes whose use needs to be tailored to each surveillance evaluation, a process which may

be far from straightforward. Due to the wide range of system attributes that may be assessed, methods which collapse these down into a small number of grouped characteristics by focusing on the relationships between attributes and their links to the objectives of the surveillance system should be explored further. The application of methods such as CRC and scenario-tree analysis to improve sensitivity estimates is advised. A generic and comprehensive evaluation framework could then be developed consisting of a limited number of common attributes together with several sets of secondary attributes which could be selected depending on the disease or range of diseases under surveillance. If there is to be a benefit to decision-makers, and ultimately result in maximum impact, the outputs of the surveillance need to be interpreted correctly and communicated clearly to all who make use of the system. Economic evaluation should be an integral part of the surveillance evaluation process. This would provide a significant benefit to decision-makers who often need to make choices based on limited or diminishing resources.

#### ACKNOWLEDGEMENTS

Funding provided by the Department for Food, Environment and Rural Affairs (project code SE4302).

#### DECLARATION OF INTEREST

None.

#### REFERENCES

1. **Thacker SB, Berkelman RL.** Public health surveillance in the United States. *Epidemiologic Reviews* 1988; **10**: 164–190.
2. **Stärk KDC, et al.** Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches. *BMC Health Services Research* 2006; **6**: 1–8.
3. **Defra.** Partnership, priorities and professionalism: a strategy for enhancing veterinary surveillance in the UK. London: Department for Environment, Food and Rural Affairs; 2003.
4. **OIE.** Terrestrial Animal Health Code. Paris: Office international des épizooties, 2010.
5. **Bravata DM, et al.** Systematic review: surveillance systems for early detection of bioterrorism-related diseases. *Annals of Internal Medicine* 2004; **140**: 910–922.
6. **Hu PJH, et al.** A Web-based system for infectious disease data integration and sharing: Evaluating outcome, task performance efficiency, user information satisfaction, and usability. *Intelligence and Security Informatics: Biosurveillance, Proceedings* 2007; **4506**: 134–146.
7. **Meynard JB, et al.** Proposal of a framework for evaluating military surveillance systems for early detection of outbreaks on duty areas. *BMC Public Health* 2008; **8**: 146.
8. **Knight-Jones TJD, et al.** Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. *Veterinary Research* 2010; **41**: 50.
9. **Hesterberg U, Cook A, Stack JM, PAJ.** 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.
10. **Nabarro D.** Global disease surveillance and response: incentives and disincentives to timely disease reporting and response – lessons from the influenza campaign. In: *Relman DA, Choffnes ER, Mack A, eds. Infectious Disease Movement in a Borderless World*. Washington: National Academies Press, 2010, pp. 256–61.
11. **World Health Organisation.** Health systems strengthening – glossary (<https://www.healthintelligenceportal.org/landing/hssglossary.html>). Accessed 9 December 2010.
12. **Birch CPD, et al.** Spatial distribution of the active surveillance of sheep scrapie in Great Britain: an exploratory analysis. *BMC Veterinary Research* 2009; **5**: 23.
13. **Carpenter TE.** Evaluation and extension of the cusum technique with an application to Salmonella surveillance. *Journal of Veterinary Diagnostic Investigation* 2002; **14**: 211–218.
14. **Hadorn D, et al.** Using scenario tree modelling for the evaluation of surveillance strategies for zoonoses in disease-free and endemic situations. In: *Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine Annual Meeting*, 26–28 March 2008, Liverpool, UK, pp. 253–261, 2008.
15. **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.
16. **Vidal Diez A, Arnold ME, Del Rio Vilas VJ.** Evaluation and optimisation of scrapie surveillance in Great Britain: A Bayesian framework. In: *Proceedings of the 12th Meeting of the International Society of Veterinary Epidemiology and Economics*, 10–14 August 2009, Durban, South Africa, 2009.
17. **CDC.** Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *Morbidity and Mortality Weekly Report (Recommendations & Reports)* 2001; **50**: 1–35.
18. **WHO.** Communicable disease surveillance and response systems: guide to monitoring and evaluating. Geneva: World Health Organisation, 2006.
19. **Health Surveillance Coordinating Committee.** Framework and tools for evaluating health surveillance systems. Vancouver: Health Canada, 2004.

20. **WHO.** Protocol for the evaluation of epidemiological surveillance systems. Geneva: World Health Organisation, 1997.
21. **Aavitsland P, Nilsen O, Lystad A.** 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.
22. **Abernethy D, et al.** Evaluation of surveillance measures to detect *Brucella abortus* in a high density cattle population. In: *Proceedings of the 12th Meeting of the International Society of Veterinary Epidemiology and Economics*, 10–14 August 2009, Durban, South Africa, 2009.
23. **Andersen HJ, et al.** Evaluation of the surveillance program of *Streptococcus agalactiae* in Danish dairy herds. *Journal of Dairy Science* 2003; **86**: 1233–1239.
24. **CDC.** Evaluation of congenital syphilis surveillance system – New Jersey, 1993. *Morbidity and Mortality Weekly Report (Recommendations & Reports)* 1995; **44**: 225–227.
25. **CDC.** Evaluation of HIV case surveillance through the use of non-name unique identifiers – Maryland and Texas, 1994–1996. *Morbidity and Mortality Weekly Report* 1998; **46**: 1254–1258.
26. **WHO.** Evaluation of the dracunculiasis surveillance system in 4 districts in Ghana. *Weekly Epidemiological Record* 2005; **80**: 270–276.
27. **Ansari JA, et al.** Evaluation of the existing bacterial meningitis surveillance system in Islamabad, Pakistan. *International Journal of Infectious Diseases* 2008; **12**: E192.
28. **Arcscott-Mills S, Holder Y, Gordon G.** Comparative evaluation of different modes of a national accident and emergency department-based injury surveillance system: Jamaican experience. *International Journal of Injury Control and Safety Promotion* 2002; **9**: 235–239.
29. **Atchison CJ, et al.** Clinical laboratory practices for the detection of rotavirus in England and Wales: can surveillance based on routine laboratory testing data be used to evaluate the impact of vaccination? *Eurosurveillance* 2009; **14**: 1–6.
30. **Barat LM, et al.** Evaluation of malaria surveillance using retrospective, laboratory-based active case detection in four southwestern states, 1995. *American Journal of Tropical Medicine and Hygiene* 1999; **60**: 910–914.
31. **Betancourt JA, et al.** Evaluation of ICD-9 codes for syndromic surveillance in the electronic surveillance system for the early notification of community-based epidemics. *Military Medicine* 2007; **172**: 346–352.
32. **Betanzos-Reyes AF, et al.** Comparative analysis of two alternative models for epidemiological surveillance in the Mexican Malaria Control Program. *Health Policy* 2007; **80**: 465–482.
33. **Bingle CL, et al.** An evaluation of the Ontario rapid risk factor surveillance system. *Canadian Journal of Public Health – Revue Canadienne De Sante Publique* 2005; **96**: 145–150.
34. **Bowen HJ, et al.** Community exposures to chemical incidents: development and evaluation of the first environmental public health surveillance system in Europe. *Journal of Epidemiology and Community Health* 2000; **54**: 870–873.
35. **Brooker S, et al.** The use of schools for malaria surveillance and programme evaluation in Africa. *Malaria Journal* 2009; **8**: 231.
36. **Brownstein JS, Freifeld CC.** Evaluation of internet-based informal surveillance for global infectious disease intelligence. *International Journal of Infectious Diseases* 2008; **12**: E193–E194.
37. **Buckeridge DL, Verma A, Tamblin R.** Ambulatory e-prescribing: evaluating a novel surveillance data source. *Intelligence and Security Informatics: Biosurveillance, Proceedings* 2007; **4506**: 190–195.
38. **Carpenter TE, Chriel M, Greiner M.** An analysis of an early-warning system to reduce abortions in dairy cattle in Denmark incorporating both financial and epidemiologic aspects. *Preventive Veterinary Medicine* 2007; **78**: 1–11.
39. **Carrieri MP, et al.** Evaluation of the SIMI system, an experimental computerised network for the surveillance of communicable diseases in Italy. *European Journal of Epidemiology* 2000; **16**: 941–947.
40. **Chadee DD.** Evaluation of malaria surveillance in Trinidad (1988–1998). *Annals of Tropical Medicine and Parasitology* 2000; **94**: 403–406.
41. **Chriel M, Salman MD, Wagner BA.** Evaluation of surveillance and sample collection methods to document freedom from infectious bovine rhinotracheitis in cattle populations. *American Journal of Veterinary Research* 2005; **66**: 2149–2153.
42. **Clothier HJ, Fielding JE, Kelly HA.** An evaluation of the Australian Sentinel Practice Research Network (ASPREN) surveillance for influenza-like illness. *Communicable Diseases Intelligence* 2005; **29**: 231–47.
43. **Cretikos M, Telfer B, McAnulty J.** Evaluation of the system of surveillance for enteric disease outbreaks, New South Wales, Australia, 2000 to 2005. *NSW Public Health Bulletin* 2008; **19**: 8–14.
44. **David ST, et al.** A bird's eye view: using geographic analysis to evaluate the representativeness of corvid indicators for West Nile virus surveillance. *International Journal of Health Geographics* 2007; **6**: 3.
45. **de Chabalier F, Hassane A, Chippaux JP.** Evaluation of surveillance thresholds for prediction of meningitis epidemics using ongoing surveillance data at the district level, in Niger. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000; **94**: 251–252.
46. **Doroshenko A, et al.** Evaluation of syndromic surveillance based on National Health Service Direct derived data – England and Wales. *Morbidity and Mortality Weekly Report* 2005; **54** (Suppl.): 117–122.
47. **Dufour B.** Technical and economic evaluation method for use in improving infectious animal disease surveillance networks. *Veterinary Research* 1999; **30**: 27–37.
48. **Fujii H, et al.** Evaluation of the school health surveillance system for influenza, Tokyo, 1999–2000. *Japanese Journal of Infectious Diseases* 2002; **55**: 97–99.
49. **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.

50. **Gazarian M, et al.** Evaluation of a national surveillance unit. *Archives of Disease in Childhood* 1999; **80**: 21–27.
51. **Grenier D, et al.** Canadian Paediatric Surveillance Program evaluation: an excellent report card. *Paediatrics & Child Health* 2004; **9**: 379–384.
52. **Guasticchi G, et al.** Syndromic surveillance: sensitivity and positive predictive value of the case definitions. *Epidemiology and Infection* 2009; **137**: 662–671.
53. **Hadorn DC, Stark KDC.** Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Veterinary Research* 2008; **39**: 57.
54. **Hall HI, Mokotoff ED, Advisory G.** Setting standards and an evaluation framework for human immunodeficiency virus/acquired immunodeficiency syndrome surveillance. *Journal of Public Health Management and Practice* 2007; **13**: 519–523.
55. **Harpaz R, Papania MJ.** Can a minimum rate of investigation of measleslike illnesses serve as a standard for evaluating measles surveillance? *Journal of Infectious Diseases* 2004; **189**: S204–S209.
56. **Harpaz R, et al.** Lessons learned from establishing and evaluating indicators of the quality of measles surveillance in the United States, 1996–1998. *Journal of Infectious Diseases* 2004; **189**: S196–S203.
57. **He SW, Zurynski YA, Elliott EJ.** Evaluation of a national resource to identify and study rare diseases: The Australian Paediatric Surveillance Unit. *Journal of Paediatrics and Child Health* 2009; **45**: 498–504.
58. **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.
59. **Herndandez-Jover M, et al.** Evaluating post-farm-gate disease surveillance for the pig industry in Australia: livestock markets and abattoirs. In: *Proceedings of the 12th Meeting of the International Society of Veterinary Epidemiology and Economics*, 10–14 August 2009, Durban, South Africa, 2009.
60. **Huaman MA, et al.** Impact of two interventions on timeliness and data quality of an electronic disease surveillance system in a resource limited setting (Peru): a prospective evaluation. *BMC Medical Informatics and Decision Making* 2009; **9**: 16.
61. **Hutchison J, Martin T.** Evaluation of surveillance for enzootic bovine leucosis in the Australian dairy industry. In: *Proceedings of the 11th Meeting of the International Society of Veterinary Epidemiology and Economics*, 6–11 August 2006, Cairns, Australia, 2006.
62. **Ihekweazu C, et al.** Is STI surveillance in England meeting the requirements of the 21st century? An evaluation of data from the South West Region. *Eurosurveillance* 2007; **12**: 111–116.
63. **Iqbal S, et al.** Carbon monoxide-related hospitalizations in the US: evaluation of a web-based query system for public health surveillance. *Public Health Reports* 2010; **125**: 423–432.
64. **Irvin CB, Nouhan PP, Rice K.** Syndromic analysis of computerized emergency department patients' chief complaints: an opportunity for bioterrorism and influenza surveillance. *Annals of Emergency Medicine* 2003; **41**: 447–452.
65. **Izadi M, et al.** A Bayesian network model for analysis of detection performance in surveillance systems. *AMIA Annual Symposium Proceedings* 2009; **2009**: 276–280.
66. **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**: 375–83.
67. **Johnson HA, et al.** Analysis of web access logs for surveillance of influenza. *Medinfo 2004: Proceedings of the 11th World Congress on Medical Informatics, Pt 1 and 2* 2004; **107**: 1202–1206.
68. **Jones NF, Marshall R.** Evaluation of an electronic general-practitioner-based syndromic surveillance system – Auckland, New Zealand, 2000–2001. *Morbidity and Mortality Weekly Report* 2004; **53** (Suppl.): 173–178.
69. **Kaufman Z, et al.** Evaluation of a syndromic surveillance system using the WSARE algorithm for early detection of an unusual, localized summer outbreak of influenza B: Implications for bioterrorism surveillance. *Israel Medical Association Journal* 2007; **9**: 3–7.
70. **Laberge K, Galanis E.** Evaluation of the surveillance of hemolytic uremic syndrome in British Columbia – should it remain reportable? *Canadian Journal of Public Health – Revue Canadienne De Sante Publique* 2008; **99**: 286–289.
71. **Lenaway DD, Ambler A.** Evaluation of a school-based influenza surveillance system. *Public Health Reports* 1995; **110**: 333–337.
72. **Leshner L, et al.** Evaluation of surveillance methods for staphylococcal toxic shock syndrome. *Emerging Infectious Diseases* 2009; **15**: 770–773.
73. **Lynn T, et al.** An evaluation of scrapie surveillance in the United States. *Preventive Veterinary Medicine* 2007; **81**: 70–79.
74. **Macarthur C, Pless IB.** Evaluation of the quality of an injury surveillance system. *American Journal of Epidemiology* 1999; **149**: 586–592.
75. **Mariner JC, et al.** Rinderpest surveillance performance monitoring using quantifiable indicators. *Revue Scientifique et Technique de l'Office International des Epizooties* 2003; **22**: 837–847.
76. **Matsui T, et al.** Evaluation of national Tsutsugamushi disease surveillance – Japan, 2000. *Japanese Journal of Infectious Diseases* 2002; **55**: 197–203.
77. **Mazurek J, et al.** Evaluation of hepatitis C surveillance in Poland in 1998. *Epidemiology and Infection* 2002; **129**: 119–125.
78. **McIntyre L, et al.** The relative value of farmer, veterinary practitioner and diagnostic laboratory records in providing epidemiologically sound endemic disease surveillance data. In: *Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine Annual meeting*, 3–5 April 2002, Cambridge, UK, pp. 128–136.
79. **McNabb SJN, et al.** Applying a new conceptual framework to evaluate tuberculosis a surveillance and action performance and measure the costs, Hillsborough

- county, Florida 2002. *Annals of Epidemiology* 2004; **14**: 640–645.
80. **Miller M, et al.** Evaluation of Australia's National Notifiable Disease Surveillance System. *Communicable Diseases Intelligence* 2004; **28**: 311–323.
  81. **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.
  82. **Nardone A, et al.** Repeat capture-recapture studies as part of the evaluation of the surveillance of Legionnaires' disease in France. *Epidemiology and Infection* 2003; **131**: 647–654.
  83. **Odoi A, et al.** Application of an automated surveillance-data-analysis system in a laboratory-based early-warning system for detection of an abortion outbreak in mares. *American Journal of Veterinary Research* 2009; **70**: 247–256.
  84. **Paisley L, Tharaldsen J, Harp J.** Evaluation of the Norwegian BHV-1 surveillance program with Monte Carlo simulation models. In: *Proceedings of the 9th Meeting of the International Society of Veterinary Epidemiology and Economics*, 6–12 August 2000, Breckenridge, Colorado, USA, 2000.
  85. **Perron L, De Wals P, Milord F.** Evaluation of the rubella surveillance system in Quebec. *Canadian Journal of Infectious Diseases* 2000; **11**: 313–316.
  86. **Phillips VL, et al.** Evaluation of program performance and expenditures in a report of performance measures (RPM) via a case study of two Florida county tuberculosis programs. *Evaluation and Program Planning* 2010; **33**: 373–378.
  87. **Rahi JS, Dezateux C, Brit Congenital Cataract Interest G.** Capture-recapture analysis of ascertainment by active surveillance in the British congenital cataract study. *Investigative Ophthalmology & Visual Science* 1999; **40**: 236–239.
  88. **Richard JL, Vidondo B, Mausezahl M.** A 5-year comparison of performance of sentinel and mandatory notification surveillance systems for measles in Switzerland. *European Journal of Epidemiology* 2008; **23**: 55–65.
  89. **Roberts M, et al.** Implementing and evaluating a practice-based surveillance program for equine infectious disease in North Carolina. In: *Proceedings of the 11th Meeting of the International Society of Veterinary Epidemiology and Economics*, 6–11 August 2006, Cairns, Australia, 2006.
  90. **Robotin M.** Evaluation of the Australian CJD surveillance system. *Communicable Diseases Intelligence* 2002; **26**: 265–72.
  91. **Romaguera P, German R, Klaucke D.** Evaluating public health surveillance. In: Teutsch S, Churchill R, eds. *Principles and Practice of Public Health Surveillance*, 2nd edn. New York: Oxford University Press, 2000, pp. 176–93.
  92. **Rumisha SF, et al.** Monitoring and evaluation of integrated disease surveillance and response in selected districts in Tanzania. *Tanzania Health Research Bulletin* 2007; **9**: 1–11.
  93. **Safdar RM, Khan SA, Asghar RJ.** Evaluation of hepatitis surveillance systems in Pakistan. *International Journal of Infectious Diseases* 2008; **12**: E192.
  94. **Sandberg M, et al.** An evaluation of the Norwegian Salmonella surveillance and control program in live pig and pork. *International Journal of Food Microbiology* 2002; **72**: 1–11.
  95. **Sekhobo JP, Druschel CM.** An evaluation of congenital malformations surveillance in New York State: An application of Centers for Disease Control and Prevention (CDC) guidelines for evaluating surveillance systems. *Public Health Reports* 2001; **116**: 296–305.
  96. **Takahashi T, et al.** Evaluation of the Japanese school health surveillance system for influenza. *Japanese Journal of Infectious Diseases* 2001; **54**: 27–30.
  97. **Tan HF, et al.** Evaluation of the national notifiable disease surveillance system in Taiwan: An example of varicella reporting. *Vaccine* 2007; **25**: 2630–2633.
  98. **van Benthem BHB, van Vliet JA.** Reflections on an evaluation of the Dutch Infectious diseases Surveillance Information System. *Eurosurveillance* 2008; **13**(11).
  99. **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.
  100. **Watkins RE, et al.** An evaluation of the sensitivity of acute flaccid paralysis surveillance for poliovirus infection in Australia. *BMC Infectious Diseases* 2009; **9**: 162.
  101. **Weber IB.** Evaluation of the notifiable disease surveillance system in Gauteng province, South Africa (M.Sc. thesis). University of Pretoria, 2007.
  102. **Wilson K, Brownstein JS.** Early detection of disease outbreaks using the Internet. *Canadian Medical Association Journal* 2009; **180**: 829–831.
  103. **Yih WK, et al.** Evaluating real-time syndromic surveillance signals from ambulatory care data in four states. *Public Health Reports* 2010; **125**: 111–20.
  104. **Jajosky RA, Groseclose SL.** Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* 2004; **4**: 29.
  105. **Buehler JW, et al.** Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *Morbidity and Mortality Weekly Report (Recommendations & Reports)* 2004; **53**: 1–11.
  106. **IAEA.** Performance indicators for rinderpest surveillance. IAEA-TECDOC-1261. Vienna, Austria: International Atomic Energy Agency, 2001.
  107. **Malecki KC, Resnick B, Burke TA.** Effective environmental public health surveillance programs: a framework for identifying and evaluating data resources and indicators. *Journal of Public Health Management and Practice* 2008; **14**: 543–551.
  108. **Mitchell RJ, Williamson AM, O'Connor R.** The development of an evaluation framework for injury surveillance systems. *BMC Public Health* 2009; **9**: 260.
  109. **Sosin DM.** Draft framework for evaluating syndromic surveillance systems. *Journal of Urban*

- Health – Bulletin of the New York Academy of Medicine* 2003; **80**: i8–i13.
110. **Del Rio Vilas VJ, Pfeiffer DU.** The evaluation of bias in scrapie surveillance: a review. *Veterinary Journal* 2010; **185**: 259–264.
  111. **Kleinman KP, Abrams AM.** Assessing surveillance using sensitivity, specificity and timeliness. *Statistical Methods in Medical Research* 2006; **15**: 445–464.
  112. **Kleinman KP, Abrams AM.** Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 2008; **27**: 4057–4068.
  113. **Magnani R, et al.** Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005; **19** (Suppl. 2): S67–S72.
  114. **Jolly GM.** Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* 1965; **52**: 225–247.
  115. **Del Rio Vilas VJ, et al.** A case study of capture-recapture methodology using scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 2005; **67**: 303–317.
  116. **Food and Agriculture Organisation.** Campaign against deadly cattle plague ending: End of field activities sets stage for rinderpest eradication (<http://www.fao.org/news/story/en/item/46383/icode/>). Accessed 4 November 2010.
  117. **Hoinville L, et al.** Appendix 2b: Summary of surveillance definitions and criteria for describing and evaluating surveillance activities developed after the workshop. In: *Discussing the Development and Application of Methods for Effective Surveillance in Livestock Populations*. Report of a workshop held prior to the ISVEE conference, Durban, South Africa, August 2009 ([http://www.defra.gov.uk/vla/reports/docs/rep\\_pres\\_isvee.pdf](http://www.defra.gov.uk/vla/reports/docs/rep_pres_isvee.pdf)). Accessed 7 March 2011.