Ministry for Primary Industries
Manatū Ahu Matua

Surveillance Evaluation Framework (SurF)

Appendix 1: SurF Methods Catalogue

17468 Deliverable: Draft Framework

MPI Technical Paper No: 2017/38

Prepared for Investigation and Diagnostic Centres and Response Directorate
Ministry for Primary Industries
Pastoral House, 25 The Terrace
Wellington, New Zealand
www.mpi.govt.nz

by Dr Petra Muellner (Epi-interactive Ltd)
Prof Katharina D.C. Stärk (SAFOSO AG)
Jonathan Watts (MPI)

ISBN No: 978-1-77665-566-3 (online)
ISSN No: 2253-3923 (online)

March 2016

Disclaimer

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04 894 0300

This publication is also available on the Ministry for Primary Industries website at http://www.mpi.govt.nz/news-and-resources/publications/

Acknowledgements

The authors would like to thank Paul Stevens, Mark Bullians, Daan Vink, Anjali Pande, Paul Bingham, Tim Riding and Brendan Gould (all MPI) for their guidance and input.

Surveillance Evaluation Framework (SurF)

Prepared for

Investigation and Diagnostic Centres
and Response Directorate
Ministry for Primary Industries
as 17468 Deliverable: Draft Framework

by

Dr Petra Muellner (Epi-interactive Ltd)
Prof Katharina D.C. Stärk (SAFOSO AG)
Jonathan Watts (MPI)

March 2016

Contents

Page

Appendix 1: SurF Methods Catalogue

This Appendix provides a catalogue of references describing methods of assessment for the different attributes used in SurF, where available. It completes the guidance notes. Users are encouraged to add to the current catalogue as SurF is being used and to document the use of the methods in different SurF evaluations.

ACCEPTABILITY AND ENGAGEMENT

Reference	Abstract	Date used	Name of evaluation	Comments
Buehler *et al.*, Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recommendations and Reports* 53: 1–11, 2004	The threat of terrorism and high-profile disease outbreaks has drawn attention to public health surveillance systems for early detection of outbreaks. State and local health departments are enhancing existing surveillance systems and developing new systems to better detect outbreaks through public health surveillance. However, information is limited about the usefulness of surveillance systems for outbreak detection or the best ways to support this function. This report supplements previous guidelines for evaluating public health surveillance systems. Use of this framework is intended to improve decision-making regarding the implementation of surveillance for outbreak detection. Use of a standardized evaluation methodology, including description of system design and operation, also will enhance the exchange of information regarding methods to improve early detection of outbreaks. The framework directs particular attention to the measurement of timeliness and validity for outbreak detection. The evaluation framework is designed to support assessment and description of all surveillance approaches to early detection, whether through traditional disease reporting, specialized analytic routines for aberration detection, or surveillance using early indicators of disease outbreaks, such as syndromic surveillance.			
Riera-Montes and Velicko. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008. We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the system was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			

BENEFIT

Reference	Abstract	Date used	Name of evaluation	Comments
Dijkhuizen *et al.* Economic analysis of animal diseases and their control. *Preventive Veterinary Medicine*, 25: 135–49, 1995	Animal Health Economics is a relatively new discipline, which is progressively developing a solid framework of concepts, procedures, and data to support the decision-making process in optimizing animal health management. Research in this field primarily deals with three interrelated aspects: (1) quantifying the financial effects of animal disease, (2) developing methods for optimizing decisions when individual animals, herds or populations are affected, and (3) determining the costs and benefits of disease control measures. In the paper the four most common economic modelling techniques (i.e. partial budgeting, cost-benefit analysis, decision analysis, and systems simulation) are described and applied on three levels of veterinary decision making: the animal, herd, and national level. Outcomes so far are summarized, and shortcomings indicated and discussed. The importance of a close link between economics and epidemiology is stressed for future development, as well as the need and possibilities for an international exchange of models and procedures.			
Dufour B. Technical and economic evaluation method for use in improving infectious animal disease surveillance networks. *Veterinary Research* 30: 27–37, 1999	With hope of improving the increasing number of epidemiological surveillance networks for animal diseases set up in recent years, a qualitative and quantitative technical and economic evaluation tool was developed and then applied to three epidemiological surveillance networks: RENESA (a French surveillance network for salmonella and mycoplasma contamination in poultry production units subject to official sanitary controls), the French Foot and Mouth Disease Epidemiovigilance Network and REPIMAT (the epidemiological surveillance network in Chad for major cattle diseases). We identified critical points in epidemiological surveillance networks using a modified version of the hazard analysis: critical control point (HACCP) method. An evaluation grid was then developed and validated by experts who were consulted in accordance with the Delphi method. A questionnaire to collect the information required for the evaluation and a scoring guide were then designed. Our evaluation procedure also included a calculation of the annual operating costs for two of the three networks studied. On the basis of the detailed results of the technical and economic evaluation, we formulated specific suggestions for improving the networks. The cost of implementing these proposals was calculated. We then simulated the effects of implementing each of the proposed improvements and a new global evaluation score was determined for each network. The 'cost per point' of each improvement was then calculated and discussed. This tool for the technical and economic evaluation of epidemiological surveillance networks for animal diseases is proposed so that it may be tested on a far wider scale and eventually be used in improving the functioning of such networks and for risk analysis in international trade.			
Häsler *et al*. Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for	This article originates from a research project to develop a conceptual framework and practical tool for the economic evaluation of surveillance. Exploring the technical relationship between mitigation as a source of economic value and surveillance and intervention as sources of economic cost is crucial. A framework linking the key technical relationships is proposed. Three			

economic analysis. *BMC Health Services Research* 11: 225, 2011	conceptually distinct stages of mitigation are identified. Avian influenza, salmonella, and foot and mouth disease are presented to illustrate the framework.			
Veterinary epidemiology and economics in Africa: a manual for use in the design and appraisal of animal health policy. International Livestock Research Institute, 1988	The value of epidemiological investigation as a basis for the treatment and control of animal disease has been recognised for many decades, but the need to apply economic techniques to the formulation and assessment of disease control activities only became apparent about 15 years ago. This arose in part from burgeoning veterinary expenditure demands associated with new, but costly, technology and in part from growing awareness of the significant influence of economic and social factors on patterns of ill-health and disease. FAO published a collation of disease losses in 1963, but it was concern in WHO over the zoonoses which led to the first international initiative, at Reading University in 1972, to develop new methods for the economic, as well as epidemiological, evaluation of animal health programmes. Since then many national and international agencies have become involved and research and training units have sprung up at several universities around the world. An international society and various national societies have also been formed to provide forums for discussion of the more profound understanding that is emerging of how to improve the health, welfare and productivity of animals. The team which has prepared this manual has demonstrated how representatives of a wide variety of disciplines can, and should, work together not only to control and avoid the major disease hazards which can still decimate animal populations, but also to define how genetics, management, nutrition and environmental adjustment can complement specific veterinary measures. Each member has contributed to a wide variety of research projects and field investigations over the past decade and in so doing, has crystallised a contribution to the training of disease control planners and animal health advisers in Britain and overseas. Recognising the need to provide such material for reference purposes and a wider range of training activities in Africa, ILCA and VEERU decided to join forces in publishing this manual. While Africa is the main focus, we feel sure that this manual will prove useful in other continents of the world and will further the long-term wellbeing of animals, in their many roles, as well as of people.			
McInerney *et al.* A framework for the economic analysis of disease in farm livestock. *Preventive Veterinary Medicine* 13: 137–54, 1992	Economic studies on disease in farm livestock have focused largely on cost/benefit analyses of control programs or gross estimates of the costs due to particular diseases. Neither offers an adequate basis to guide economic decisions. This paper develops a framework with a more rigorous underpinning in economic concepts and principles. First, the total economic cost of a disease is explained in terms of two distinct components: (i) output losses following disease occurrence; (ii) expenditures made to treat disease or prevent its occurrence. We then develop the general relationship between these two components, calling it the 'loss-expenditure frontier', and use it to define the economically optimal level of disease costs. This shows that the relevant measures of the economic cost of disease are not the total cost figures usually estimated. The paper explains how, in principle, the loss-expenditure frontier can provide the appropriate information for decisions on disease control, and demonstrates this using survey data on subclinical mastitis in the UK dairy national herd.			

Moran and Fofana. An economic evaluation of the control of three notifiable fish diseases in the United Kingdom. *Preventive Veterinary Medicine* 80: 193–208, 2007	We summarised the challenges faced in an ex ante cost–benefit appraisal of United Kingdom government spending on disease surveillance for three notifiable fish diseases: infectious salmon anaemia (ISA), viral haemorrhagic septicaemia (VHS) and infectious haemorrhagic necrosis (IHN). We used a social cost–benefit analysis and adopted a national perspective. We compared costs of current public and private surveillance effort with the benefits stated in terms of the avoided private and social costs of potential disease outbreaks. Spending on ISA and VHS were predicted to be efficient; the benefit–cost ratios were always ≥3.2 for ISA and ≥5.8 for VHS for all nine scenarios examined for each infection. However, the benefit–cost ratio for IHN was predicted never to exceed 1.6, and was <1.0 in five of the nine scenarios-so spending on IHN would be harder to justify.			
Morris *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* 18: 415–22, 1996	BACKGROUND: In England and Wales, surveillance of communicable disease is carried out and co-ordinated by the Public Health Laboratory Service (PHLS). The surveillance of HIV infection and AIDS is undertaken by the PHLS AIDS Centre at the Communicable Disease Surveillance Centre (CDSC). Epidemiological data derived from surveillance are not, however, a free good: they are a resource with an associated opportunity cost and should therefore be open to economic appraisal alongside other users of health care resources such as medical interventions. This paper assembles information on the current surveillance of HIV and AIDS in England and Wales, and explores methods for performing an economic evaluation of such activities. METHODS: An examination of the cost and effectiveness of the PHLS AIDS Centre's epidemiological surveillance mechanisms for HIV and AIDS in England and Wales was undertaken. The total costs of each component of surveillance of HIV and AIDS in England and Wales were calculated. Two categories of cost were estimated: peripheral costs incurred by reporters in reporting AIDS cases or HIV infections or by laboratories in collecting samples; and central costs incurred by the PHLS AIDS Centre in processing and analysing incoming data. Using these cost data and information from a cost-effectiveness register, the additional health gains that would have to be obtained from surveillance to make the programme broadly cost-effective in comparison with other accepted uses of health service resources were then estimated. RESULTS: In the financial year 1993-1994 the total costs of surveillance were estimated to be 1.4 million pounds. To avoid being considered relatively cost-ineffective at least 3.5 infections per annum need to be averted. To be considered favourably cost-effective, approximately 9.5 infections per annum need to be averted. CONCLUSIONS: In 1993-1994, expenditure on surveillance of HIV and AIDS accounted for less than 1 per cent of the total allocation of resources to the National Health Service for all HIV and AIDS activities. Given these cost estimates, the number of infections which surveillance would have to contribute towards preventing in order to be considered cost-effective is low.			
Rushton *et al*. Investment in preventing and preparing for biological emergencies and disasters: social and economic	Biological emergencies such as the appearance of an exotic transboundary or emerging disease can become disasters. The question that faces Veterinary Services in developing countries is how to balance resources dedicated to active insurance measures, such as border control, surveillance, working with the governments of developing countries, and investing in improving			

costs of disasters versus costs of surveillance and response preparedness. *Scientific and Technical Review* 25(1): 375–88, 2006.	veterinary knowledge and tools, with passive measures, such as contingency funds and vaccine banks. There is strong evidence that the animal health situation in developed countries has improved and is relatively stable. In addition, through trade with other countries, developing countries are becoming part of the international animal health system, the status of which is improving, though with occasional setbacks. However, despite these improvements, the risk of a possible biological disaster still remains, and has increased in recent times because of the threat of bioterrorism. This paper suggests that a model that combines decision tree analysis with epidemiology is required to identify critical points in food chains that should be strengthened to reduce the risk of emergencies and prevent emergencies from becoming disasters.			
Rushton *et al*. Methods of economic impact assessment. *Scientific and Technical Review* 18: 315–42, 1999	A broad framework for the assessment of the economic impact of animal diseases and control of these diseases is described. Various levels of analysis are discussed, from the herd, to the household, the sector, the nation and the region. For each of these levels, methods of economic impact assessment are outlined and discussed with respect to the objectives, data requirements and outputs of each type of assessment. In addition, the authors attempt to describe the conditions under which the various methods might be used. To conclude, a discussion of the state of the art is presented, and areas of further research are identified.			
Zinsstag *et al*. Human benefits of animal interventions for zoonosis control. *Emerging Infectious Diseases* 13(4): 527–31, 2007	Although industrialized countries have been able to contain recent outbreaks of zoonotic diseases, many resource-limited and transitioning countries have not been able to react adequately. The key for controlling zoonoses such as rabies, echinococcosis, and brucellosis is to focus on the animal reservoir. In this respect, ministries of health question whether the public health sector really benefits from interventions for livestock. Cross-sectoral assessments of interventions such as mass vaccination for brucellosis in Mongolia or vaccination of dogs for rabies in Chad consider human and animal health sectors from a societal economic perspective. Combining the total societal benefits, the intervention in the animal sector saves money and provides the economic argument, which opens new approaches for the control of zoonoses in resource-limited countries through contributions from multiple sectors.			

COVERAGE

Reference	Abstract	Date used	Name of evaluation	Comments
Brooker *et al*. The use of schools for malaria surveillance and programme evaluation in Africa. *Malaria Journal* 8: 231, 2009	Effective malaria control requires information on both the geographical distribution of malaria risk and the effectiveness of malaria interventions. The current standard for estimating malaria infection and impact indicators are household cluster surveys, but their complexity and expense preclude frequent and decentralized monitoring. This paper reviews the historical experience and current rationale for the use of schools and school children as a complementary, inexpensive framework for planning, monitoring and evaluating malaria control in Africa. Consideration is given to (i) the selection of schools; (ii) diagnosis of infection in schools; (iii) the representativeness of schools as a proxy of the communities they serve; and (iv) the increasing need to evaluate interventions delivered through schools. Finally, areas requiring further investigation are highlighted.			
Del Rio Vilas and Pfeiffer. The evaluation of bias in scrapie surveillance: A review. *Veterinary Journal* 185: 259–64, 2010	Evaluation of surveillance systems is a common practice in the context of human health, but only recently has been applied in the veterinary field. Commonly, a series of attributes are monitored to assess the system. Suboptimal performance of the surveillance in relation to any of these attributes may lead to bias in the surveillance results. The intensity of scrapie surveillance has increased considerably in recent years as a result of public health concerns. In this paper, a number of approaches described in the literature for the evaluation of surveillance systems are reviewed, with a focus on the sensitivity and representativeness attributes of scrapie surveillance systems in the European Union. Many of the methods applied in other fields, such as ecology and public health, are exchangeable and relevant for scrapie surveillance.			
Del Rio Vilas and Böhning. Application of one-list capture-recapture models to scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 85: 253–66, 2008	In this paper, we apply one-list capture–recapture models to estimate the number of scrapie-affected holdings in Great Britain. We applied this technique to the Compulsory Scrapie Flocks Scheme dataset where cases from all the surveillance sources monitoring the presence of scrapie in Great Britain, the abattoir survey, the fallen stock survey and the statutory reporting of clinical cases, are gathered. Consequently, the estimates of prevalence obtained from this scheme should be comprehensive and cover all the different presentations of the disease captured individually by the surveillance sources. Two estimators were applied under the one-list approach: the Zelterman estimator and Chao's lower bound estimator. Our results could only inform with confidence the scrapie-affected holding population with clinical disease; this moved around the figure of 350 holdings in Great Britain for the period under study, April 2005–April 2006. Our models allowed the stratification by surveillance source and the input of covariate information, holding size and country of origin. None of the covariates appear to inform the model significantly.			
Del Rio Vilas *et al*. A case of capture- recapture methodology using scrapie surveillance data in	We applied capture–recapture methodology (CRC) to data from three surveillance sources (statutory notification, abattoir survey (AS) and fallen stock (FS) survey) to estimate the number of holdings infected with scrapie in Great Britain and to assess the sensitivity of the surveillance			

Great Britain. *Preventive Veterinary Medicine* 67: 303–17, 2005	network. Between January 1, 2002 and March 31, 2003, 144 holdings were identified by the three sources. Using CRC modelling techniques, we estimated a minimum lower bound for the total number of holdings infected as 642. A biologically plausible positive dependence between the statutory reporting and the fallen stock survey was found statistically significant. The sensitivity of the three sources combined was very low. The integration of the three overlapping sources provided a better understanding of the interactions within the surveillance network. However, the scarcity of the data and reduced overlapping among sources only allowed for very cautious inferences to be drawn about the true proportion of scrapie affected holdings in the national population. Future surveys and surveillance activities should be planned such that the resulting data can be used more effectively as part of CRC modelling approaches.			
Walker *et al*. Epidemiological analysis of the quality of HIV sero-surveillance in the world: how well do we track the epidemic? *AIDS* 15: 1545–54, 2001	OBJECTIVE: The objective of this paper was to analyse the quality of HIV/AIDS sentinel surveillance systems in countries and the resulting quality of the data used to make estimates of HIV/AIDS prevalence and mortality. METHODS: Available data on sero-surveillance of HIV/AIDS in countries were compiled in the process of making the end of 1999 estimates of HIV/AIDS. These data came primarily from the HIV/AIDS Surveillance Database developed by the United States Census Bureau, from a database maintained by the European Centre for the Epidemiological Monitoring of AIDS and all country reports on sentinel surveillance that had been provided to World Health Organization or UNAIDS. Procedures were developed to score quality of surveillance systems based on four dimensions of quality: timeliness and frequency; appropriateness of groups; consistency of sites over time; and coverage provided by the system. In total, the surveillance systems from 167 countries were analysed. RESULTS: Forty-seven of the 167 countries whose surveillance systems were rated were judged to have fully implemented sentinel surveillance systems; 51 were judged to have systems that had some or most aspects of a good HIV surveillance system in place and 69 were rated as having poorly functioning or non-existent surveillance systems. CONCLUSION: This analysis suggests that the quality of HIV surveillance varies considerably. The majority of countries most affected by HIV/AIDS have systems that are providing sufficient sero-prevalence data for tracking the epidemic and making reasonable estimates of HIV prevalence. However, many countries have poor systems and strengthening these is an urgent priority.			

DATA ANALYSIS

Reference	Abstract	Date used	Name of evaluation	Comments
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

DATA AND INFORMATION COLLECTION

Reference	Abstract	Date used	Name of evaluation	Comments
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

DATA COMPLETENESS AND CORRECTNESS

Reference	Abstract	Date used	Name of evaluation	Comments
Harpaz *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* 189: S196–S203, 2004	As part of a strategy to eliminate measles, 7 indicators were adopted in the United States in 1996 to ensure the quality of measles surveillance. This report summarizes the US experience with these indicators during 1996-1998. The indicators are compiled from data reported to the Centers for Disease Control and Prevention (CDC) during routine surveillance supplemented with information collected directly from states. Measles case investigations are generally thorough, and sufficient information is collected to control and monitor disease. A high proportion of measles cases are imported from other countries, suggesting that investigations are complete. For some states, the lag from disease onset to reporting is long, and the number of health department investigations of measles like illnesses is low. Most of these investigations include laboratory testing of clinical specimens. Collection of measles virus specimens from cases for genetic analysis needs improvement. The CDC and health departments need to continue efforts directed at health care professionals to ensure the recognition, proper diagnostic workup, and reporting of measles.			
Miller *et al*. Evaluation of Australia's National Notifiable Disease Surveillance System. *Communicable Diseases Intelligence* 28: 311–23, 2004	The Australian National Notifiable Diseases Surveillance System (NNDSS) is a passive surveillance system that collects information on communicable diseases. The Australian Government manages NNDSS under the auspices of the Communicable Diseases Network Australia (CDNA). Data collected by each state and territory are collated, analysed and disseminated by the Australian Government Department of Health and Ageing. We report the first evaluation of NNDSS since it was established in 1991. Three primary stakeholder groups were surveyed: (a) CDNA members, (b) the National Surveillance Committee and (c) the readership of Communicable Diseases Intelligence, the primary means of data dissemination from NNDSS. The evaluation revealed that the system was acceptable, structurally simple, and that the data collected were actively used by stakeholders. However, the lack of clearly documented aims and objectives for NNDSS, inflexibility to changing needs, lack of timeliness and complexity in processes were seen as problematic. The results of this evaluation, supported by recent federal funding to enhance national biosecurity, will provide the framework for enhancing NNDSS to meet national communicable disease surveillance requirements in Australia.			
Pipino *et al.* Data quality assessment. *Communications of the ACM* 45(4): 211, 2002	N/A			
Riera-Montes and Velicko The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008.			

system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the sys- tem was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			
Rumisha *et al*. Monitoring and evaluation of integrated disease surveillance and response in selected districts in Tanzania. *Tanzania Health Research Bulletin* 9: 1–11, 2007	Integrated Disease Surveillance and Response (IDSR) is a strategy developed by the World Health Organization Regional Office for Africa in 1998. The Ministry of Health, Tanzania has adopted this strategy for strengthening communicable diseases surveillance in the country. In order to improve the effectiveness of the implementation of IDSR monitoring and evaluating the performance of the surveillance system, identifying areas that require strengthening and taking action is important. This paper presents the findings of baseline data collection for the period October-December 2003 in 12 districts representing eight regions of Tanzania. The districts involved were Mbulu, Babati, Dodoma Rural, Mpwapwa, Igunga, Tabora Urban, Mwanza Urban, Muleba, Nkasi, Sumbawanga Rural, Tunduru and Masasi. Results are grouped into three key areas: surveillance reporting, use of surveillance data and management of the IDSR system. In general, reporting systems are weak, both in terms of receiving all reports from all facilities in a timely manner, and in managing those reports at the district level. Routine analysis of surveillance data is not being done at facility or district levels, and districts do not monitor the performance of their surveillance system. There was also good communication and coordination with other sectors in terms of sharing information and resources. It is important that districts' capacity on IDSR is strengthened to enable them monitor and evaluate their own performance using established indicators.			

DATA MANAGEMENT AND STORAGE

Reference	Abstract	Date used	Name of evaluation	Comments
M. Mosley (Ed). *The DAMA Guide to the Data Management Body of Knowledge*. DAMA-DMBOK, 2009	Until very recently, data management in most IT departments has been largely ignored as a formal discipline. This lack of understanding has brought us to a point today where the biggest problems in most major IT projects revolve around data integration and information management. DAMA has now addressed this problem by creating the DAMA Data Management Guide to the Body of Knowledge (DAMA-DMBOK) – the first authoritative resource for data management best practices in 40 years of IT practice.			
Office for National Statistics (UK). Implementation Guides.	These guides are for staff in public authorities who are responsible for records management and, in particular, for compliance with the UK records management code issued under section 46 of the Freedom of Information Act. They provide an introduction to records management concepts, explain the good practice recommendations in the Code and give guidance on how to apply them within an organisation. They are aimed at people with little or no experience or knowledge of records management.			
Pipino *et al.* Data quality assessment. *Communications of the ACM* 45(4): 211, 2002	N/A			
Woolhouse *et al*. *Guide to good practice for quantitative veterinary epidemiology.* Chapter 1. EERA, 2011	There are no generally accepted standards for 'good practice' in quantitative veterinary epidemiology. This document addresses a variety of issues to do with data collection, ways of analysing and modelling data and communication of the results of such analyses. These issues are relevant in a wide variety of contexts, from ongoing attempts to analyse the impact of badger culling on the incidence of bovine tuberculosis through to the use of mathematical models to inform policy during the epidemics of foot-and-mouth disease in 2001 or of 'mad cow' disease (bovine spongiform encephalopathy) in the 1990s. Further, the issues raised and the good practice proposed is relevant beyond the disciplines making up quantitative veterinary epidemiology and certainly apply more broadly to other branches of epidemiology and data driven science. An underlying aim here is to move away from the current position, where quantitative methods are all too often viewed by policy makers and other stakeholders as a mysterious 'black box', to their becoming familiar and valued tools.			

DECISION SUPPORT

Reference	Abstract	Date used	Name of evaluation	Comments
Staerk and Haesler. The value of information: Current challenges in surveillance implementation. *Preventive Veterinary Medicine*, doi: 10.1016/j.prevetmed.2015.05.002, 2015	Animal health surveillance is a complex activity that involves multiple stakeholders and provides decision support across sectors. Despite progress in the design of surveillance systems, some technical challenges remain, specifically for emerging hazards. Surveillance can also be impacted by political interests and costly consequences of case reporting, particularly in relation to international trade. Constraints on surveillance can therefore be of technical, economic and political nature. From an economic perspective, both surveillance and intervention are resource-using activities that are part of a mitigation strategy. Surveillance provides information for intervention decisions and thereby helps to offset negative effects of animal disease and to reduce the decision uncertainty associated with choices on disease control. It thus creates monetary and non- monetary benefits, both of which may be challenging to quantify. The technical relationships between surveillance, intervention and loss avoidance have not been established for most hazards despite being important consideration for investment decisions. Therefore, surveillance cannot just be maximised to minimise intervention costs. Economic appraisals of surveillance need to be done on a case by case basis for any hazard considering both surveillance and intervention performance, the losses avoided and the values attached to them. This can be achieved by using an evaluation approach which provides a systematic investigation of the worth or merit of surveillance activities. Evaluation is driven by a specific evaluation question which for surveillance systems commonly considers effectiveness, efficiency, implementation and/or compliance issues. More work is needed to provide guidance on the appropriate selection of evaluation attributes and general good practice in surveillance evaluation. Due to technical challenges, economic constraints and variable levels of capacity, the implementation of surveillance systems remains variable. Political and legal issues are also influential. A particular challenge exists during outbreaks when surveillance needs to be conducted under emergency conditions. Decision support systems can help make epidemiologically and economically sound choices amongst surveillance options. However, contingency planning is advisable so that pre-defined options allow for rapid decision making.			

EFFICIENCY

Reference	Abstract	Date used	Name of evaluation	Comments

EXTERNAL COMMUNICATION AND DISSEMINATION

Reference	Abstract	Date used	Name of evaluation	Comments
Buehler *et al*. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recommendations and Reports* 53: 1–11, 2004	The threat of terrorism and high-profile disease outbreaks has drawn attention to public health surveillance systems for early detection of outbreaks. State and local health departments are enhancing existing surveillance systems and developing new systems to better detect outbreaks through public health surveillance. However, information is limited about the usefulness of surveillance systems for outbreak detection or the best ways to support this function. This report supplements previous guidelines for evaluating public health surveillance systems. Use of this framework is intended to improve decision-making regarding the implementation of surveillance for outbreak detection. Use of a standardized evaluation methodology, including description of system design and operation, also will enhance the exchange of information regarding methods to improve early detection of outbreaks. The framework directs particular attention to the measurement of timeliness and validity for outbreak detection. The evaluation framework is designed to support assessment and description of all surveillance approaches to early detection, whether through traditional disease reporting, specialized analytic routines for aberration detection, or surveillance using early indicators of disease outbreaks, such as syndromic surveillance.			
Riera-Montes and Velicko. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008. We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the system was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			
Roberts *et al*. Implementing and evaluating a practice-based surveillance program for equine infectious disease in North Carolina. In: *Proceedings of the 11th Symposium of the International Society of Veterinary Epidemiology and Economics*, Cairns, Australia,	A novel software program developed with input from equine practitioners enabled case reporting by syndrome that could be further categorized, by clicking on the syndrome name, to a specific suspected disease entity. Each week participants could access a password-protected, on line reporting page maintained by the NCDA & CS to record the number of cases seen for each syndrome. Cumulative information was integrated and accessible on the web in real time. A seamless link to a GIS state map indicated case by syndrome totals for the practice. Laboratory confirmed diagnoses were requested monthly. The web-based surveillance pilot program was tested in 15 equine practices throughout NC for up to 8 months in 2004. All practices entered data; 9 contributed the entire period. Recording by syndrome was the preferred method. Most			

2006	frequently reported syndromes were as expected. Few confirmatory diagnoses were pursued. There were no reported disease outbreaks. Thirteen of 15 practices (87%) provided survey feedback to help refine the system. Program strengths included ease and rapidity of use, the syndrome diagnosis format, and straightforward layout. Limitations were absence of case definitions to obviate potential overlap and use of "other infectious agent" and "unknown" as default categories, and difficulties in continued data entry. Respondents were positive about working on line, participating in a timely applicable project, and expanding the program. Perceived benefits included raised awareness of infectious and emerging diseases, and evidence to support preventive health management recommendations. Active practitioner engagement and participation are essential to successfully implement an equine disease surveillance system.			

FIELD AND LABORATORY SERVICES

Reference	Abstract	Date used	Name of evaluation	Comments
International Accreditation Forum. http://www.iaf.nu/	N/A			
International Laboratory Accreditation Cooperation. http://www.ilac.org/	N/A			
International Organisation for Standardisation. http://www.iso.org/	N/A			

FLEXIBILITY

Reference	Abstract	Date used	Name of evaluation	Comments
Aavitsland *et al*. Anonymous reporting of HIV infection: An evaluation of the HIV/AIDS surveillance system in Norway 1983-2000. *European Journal of Epidemiology* 17: 479–89, 2001	Several European countries are considering implementing surveillance systems for HIV infection, but questions remain regarding patient confidentiality. The population-based HIV/AIDS surveillance system in Norway integrates anonymous HIV case reports from laboratories and clinicians and named AIDS case reports. This evaluation includes a description of the system, evidence of system attributes, estimation of resources for system operations, and documentation of the system's usefulness. HIV case reports provide a far better picture of the epidemic than AIDS reports. The median delay between positive HIV test and reporting was 30 days (interquartile range 18–49 days). Completeness of demographic and epidemiologic information in the surveillance database ranges from 60 to 100%. Information on pre-AIDS mortality and emigration is incomplete. The system cost euro 25,200 in 1999. Results are published every week and used for planning of health care and prevention. We conclude that the Norwegian surveillance system with anonymous reporting of HIV cases is simple, inexpensive and flexible, and capable of providing a representative and timely overview that guides prevention. The system fulfils its objectives while respecting confidentiality and adhering to ethical principles. A similar system may be considered in other countries.			
CDC Updated guidelines for evaluating public health surveillance systems. *Morbidity and Mortality Weekly Report*, 50, 2001	The purpose of evaluating public health surveillance systems is to ensure that problems of public health importance are being monitored efficiently and effectively. CDC's Guidelines for Evaluating Surveillance Systems are being updated to address the need for a) the integration of surveillance and health information systems, b) the establishment of data standards, c) the electronic exchange of health data, and d) changes in the objectives of public health surveillance to facilitate the response of public health to emerging health threats (e.g., new diseases). This report provides updated guidelines for evaluating surveillance systems based on CDC's Framework for Program Evaluation in Public Health, research and discussion of concerns related to public health surveillance systems, and comments received from the public health community. The guidelines in this report describe many tasks and related activities that can be applied to public health surveillance systems.			
Colby *et al*. Evaluation of two systems for managing emergency poultry diseases in intensive poultry production regions. *International Journal of Poultry Science* 2(3): 234–41, 2003	This paper provides an overview of two systems used for the management of infectious disease within the poultry industry on the Delmarva Peninsula. The first system is a paper-based Grid system that was developed in the 1970`s. This system divides the peninsula into a total of 3420 unique locations, each covering approximately 5 square miles. The second system is a Geographic Information System (GIS) database that is currently under construction. Each system is critiqued with respect to several of the criteria established by Klaucke *et al*., (1998) for the evaluation of a surveillance system: simplicity, flexibility, acceptability, representativeness and timeliness. In addition, the objectives of a disease management system in an intensive poultry production area are discussed. A grid-based system is most appropriately used in situations			

	involving a small, easily managed population or geographic location, especially when funding and geographic resources are limited. When multiple diseases or large geographic areas are the focus of a surveillance and monitoring system, or when several different risk factors are to be examined and funding and adequate resources are available, the use of a GIS-based system provides additional flexibility. The database management component of GIS allows for rapid updating of demographic and disease information, and the linkage of the database component with the spatial coordinates of a GIS provides the ability to examine the effects of several risk factors at the same time.			
Jefferson *et al*. Evaluation of a syndromic surveillance for the early detection of outbreaks among military personnel in a tropical country. *Journal of Public Health* 30: 375–83, 2008	BACKGROUND: To evaluate a new military syndromic surveillance system (2SE FAG) set up in French Guiana. METHODS: The evaluation was made using the current framework published by the Centers for Disease Control and Prevention, Atlanta, USA. Two groups of system stakeholders, for data input and data analysis, were interviewed using semi-structured questionnaires to assess timeliness, data quality, acceptability, usefulness, stability, portability and flexibility of the system. Validity was assessed by comparing the syndromic system with the routine traditional weekly surveillance system. RESULTS: Qualitative data showed a degree of poor acceptability among people who have to enter data. Timeliness analysis showed excellent case processing time, hindered by delays in case reporting. Analysis of stability indicated a high level of technical problems. System flexibility was found to be high. Quantitative data analysis of validity indicated better agreement between syndromic and traditional surveillance when reporting on dengue fever cases as opposed to other diseases. CONCLUSIONS: The sophisticated technical design of 2SE FAG has resulted in a system which is able to carry out its role as an early warning system. Efforts must be concentrated on increasing its acceptance and use by people who have to enter data and decreasing the occurrence of the frequency of technical problems.			
Riera-Montes and Velicko. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008. We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the sys- tem was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			

HISTORICAL DATA

Reference	Abstract	Date used	Name of evaluation	Comments
Gazarian *et al*. Evaluation of a national surveillance unit. *Archives of Disease in Childhood* 80: 21–7, 1999	AIM—The Australian Paediatric Surveillance Unit (APSU) facilitates national active surveillance of uncommon childhood conditions. This study assessed whether it fulfilled its objectives and satisfied criteria established by the Centers for Disease Control and Prevention (CDC) for evaluating surveillance systems. METHODS—Anonymous questionnaires were sent to users of the system, individual studies were reviewed, and data were collected from independent sources. RESULTS—Seven hundred and sixty-six clinicians, 48 investigators, and 15 public health professionals responded to the questionnaires. Clinicians reported that the APSU was useful, 33% saying information provided by the APSU informed or changed their clinical practice. Most (88%) reported that completing monthly report cards was not a burden. Impact on policy development was limited by suboptimal dissemination of information to public health professionals. Flexibility and timeliness were limited by design. Estimated sensitivity of APSU studies ranged from 92% (congenital rubella) to 31% (drowning/near drowning). Positive predictive value of notified cases was over 70% for most studies. CONCLUSION—The APSU fulfils most of its objectives and meets CDC criteria salient to these. Ways in which the APSU could be improved have been identified, as have methodological challenges and limitations in applying CDC guidelines to this type of unit.			

INTERNAL COMMUNICATION

Reference	Abstract	Date used	Name of evaluation	Comments
Buehler *et al*. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recommendations and Reports* 53: 1–11, 2004	The threat of terrorism and high-profile disease outbreaks has drawn attention to public health surveillance systems for early detection of outbreaks. State and local health departments are enhancing existing surveillance systems and developing new systems to better detect outbreaks through public health surveillance. However, information is limited about the usefulness of surveillance systems for outbreak detection or the best ways to support this function. This report supplements previous guidelines for evaluating public health surveillance systems. Use of this framework is intended to improve decision-making regarding the implementation of surveillance for outbreak detection. Use of a standardized evaluation methodology, including description of system design and operation, also will enhance the exchange of information regarding methods to improve early detection of outbreaks. The framework directs particular attention to the measurement of timeliness and validity for outbreak detection. The evaluation framework is designed to support assessment and description of all surveillance approaches to early detection, whether through traditional disease reporting, specialized analytic routines for aberration detection, or surveillance using early indicators of disease outbreaks, such as syndromic surveillance.			
Riera-Montes and Velicko. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008. We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the system was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			
Roberts *et al*. Implementing and evaluating a practice-based surveillance program for equine infectious disease in North Carolina. In: *Proceedings of the 11th Symposium of the International Society of Veterinary Epidemiology and Economics*, Cairns, Australia, 2006	A novel software program developed with input from equine practitioners enabled case reporting by syndrome that could be further categorized, by clicking on the syndrome name, to a specific suspected disease entity. Each week participants could access a password-protected, on line reporting page maintained by the NCDA & CS to record the number of cases seen for each syndrome. Cumulative information was integrated and accessible on the web in real time. A seamless link to a GIS state map indicated case by syndrome totals for the practice. Laboratory confirmed diagnoses were requested monthly. The web-based surveillance pilot program was tested in 15 equine practices throughout NC for up to 8 months in 2004. All practices entered data; 9 contributed the entire period. Recording by syndrome was the preferred method. Most			

	frequently reported syndromes were as expected. Few confirmatory diagnoses were pursued. There were no reported disease outbreaks. Thirteen of 15 practices (87%) provided survey feedback to help refine the system. Program strengths included ease and rapidity of use, the syndrome diagnosis format, and straightforward layout. Limitations were absence of case definitions to obviate potential overlap and use of "other infectious agent" and "unknown" as default categories, and difficulties in continued data entry. Respondents were positive about working on line, participating in a timely applicable project, and expanding the program. Perceived benefits included raised awareness of infectious and emerging diseases, and evidence to support preventive health management recommendations. Active practitioner engagement and participation are essential to successfully implement an equine disease surveillance system.			

INTEROPERABILITY

Reference	Abstract	Date used	Name of evaluation	Comments

MULTIPLE UTILITY

Reference	Abstract	Date used	Name of evaluation	Comments
Izadi M, *et al*. A Bayesian network model for analysis of detection performance in surveillance systems. AMIA Annual Symposium Proceedings 2009: 276–80, 2009	Worldwide developments concerning infectious diseases and bioterrorism are driving forces for improving aberrancy detection in public health surveillance. The performance of an aberrancy detection algorithm can be measured in terms of sensitivity, specificity and timeliness. However, these metrics are probabilistically dependent variables and there is always a trade-off between them. This situation raises the question of how to quantify this tradeoff. The answer to this question depends on the characteristics of the specific disease under surveillance, the characteristics of data used for surveillance, and the algorithmic properties of detection methods. In practice, the evidence describing the relative performance of different algorithms remains fragmented and mainly qualitative. In this paper, we consider the development and evaluation of a Bayesian network framework for analysis of performance measures of aberrancy detection algorithms. This framework enables principled comparison of algorithms and identification of suitable algorithms for use in specific public health surveillance settings.			
Malecki KC, *et al*. Effective Environmental Public Health Surveillance Programs: A Framework for Identifying and Evaluating Data Resources and Indicators. *Journal of Public Health Management Practice* 14: 543–51, 2008	The complexity and multidisciplinary nature of environmental public health (EPH) surveillance call for a systematic framework and a concrete set of criteria to guide development, selection, and evaluation of environmental public health indicators. Environmental public health indicators are the foundation of a comprehensive EPH surveillance system, providing quantitative summary measures and descriptive information about spatial and temporal trends of hazard, exposure, and health effects over person, place, and time. A case-synthesis review of environmental regulatory and public health indicator models was employed to develop a framework and outline a methodological approach to EPH surveillance system development, including the selection of content areas and the corresponding data and environmental public health indicators. The framework is organized around three assessment phases: (1) scientific basis and relevance, (2) analytic soundness, and (3) feasibility, interpretation and utility. By outlining a process and identifying important constructs and criteria, the framework provides practitioners with an effective and systematic tool for making scientifically valid programmatic decisions about EPH content development. Improved decision-making ensures more effective EPH surveillance systems and enhanced opportunities to understand and protect the public health from environmental threats.			

NEGATIVE PREDICTIVE VALUE

Reference	Abstract	Date used	Name of evaluation	Comments
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

ORGANISATION AND MANAGEMENT

Reference	Abstract	Date used	Name of evaluation	Comments

PERFORMANCE INDICATORS AND EVALUATION

Reference	Abstract	Date used	Name of evaluation	Comments

POSITIVE PREDICTIVE VALUE

Reference	Abstract	Date used	Name of evaluation	Comments
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

PRECISION

Reference	Abstract	Date used	Name of evaluation	Comments
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

QUALITY ASSURANCE

Reference	Abstract	Date used	Name of evaluation	Comments

RARR (RELIABILITY, AVAILABILITY, REPEATABILITY, AND ROBUSTNESS)

Reference	Abstract	Date used	Name of evaluation	Comments
Clothier *et al*. An evaluation of the Australian Sentinel Practice Research Network (ASPREN) surveillance for influenza-like illness. *Communicable Disease Intelligence* 29: 231–47, 2005	The Australian Sentinel Practice Research Network (ASPREN) is a national network of general practitioners (GPs) who collect and report data on selected conditions, including influenza-like illness (ILI). The Australian Government Department of Health and Ageing initiated an evaluation of ASPREN, aiming to assess its potential to contribute to surveillance of emerging infectious diseases including pandemic influenza. System attributes and utility for decision-making were elucidated from stakeholder surveys. ASPREN ILI data for 2002 to 2004 were compared with ILI data from South Australia and New South Wales. In 2004, 50 GPs participated in the ASPREN surveillance, with proportionately more in New South Wales (30%) and South Australia (30%) than in other states. The majority (78%) of GPs were in metropolitan practices. Compliance with the manual data collection system was not optimal, nor consistent by state. ASPREN ILI data compared favourably with that of other surveillance systems. No formal structures were in place by which to assess data trends, provide alerts or initiate public health action. To maximise the contribution to biosecurity surveillance, ASPREN would require targeted GP recruitment to achieve geographic representativeness; exploration of alternative technologies for data collection and reporting; provision of committed resources adequate for system operation; and negotiation with state-based public health reference laboratories to provide laboratory support. The main potential of ASPREN is to permit rapid dissemination of a syndromic case definition and acquisition of nationwide community level clinical presentation data. These evaluation findings will be used to inform redevelopment of ASPREN as part of the Biosecurity Surveillance System project.			
Hendrikx *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* 52: 465–75, 2005	Performance indicators are variables used to calculate on a continuous basis the operational level of a surveillance network's priority activities. A 10-step process was developed to enable network coordinators to identify specific performance indicators to help them monitor their networks. The methodology was based on a listing of surveillance activities, the definition and choice of network's global priority objectives, the construction of performance indicators and monitoring tables and the implementation of the system within the network. This process was implemented for the bovine clinical digestive salmonellosis surveillance network in France (RESSAB). The process produced a list of 26 activities synthesized into 15 global objectives, from which 12 were retained as priority objectives. This made it possible to develop 12 performance indicators. Indicators retrospectively calculated for the year 2003 indicated that RESSAB was operating according to the objectives set by the network's stakeholders and its financial supporter. The methodology was implemented successfully and was not very time consuming (12 person-days) or expensive. The decision makers and actors in the network quickly appropriated the system. The methodology can thus be considered validated through this example. Nevertheless, the risks inherent in the use of performance indicators must be addressed by ensuring the relevance of the selected indicators through external assessment and by prioritizing an internal and participatory			

	approach to avoid a misuse of the performance indicators. In addition, considering that performance indicators address only the operation of the network, relevance of surveillance procedure should be addressed by external evaluation.			
Walker *et al*. Epidemiological analysis of the quality of HIV sero-surveillance in the world: how well do we track the epidemic? *AIDS* 5: 1545–54, 2001	OBJECTIVE: The objective of this paper was to analyse the quality of HIV/AIDS sentinel surveillance systems in countries and the resulting quality of the data used to make estimates of HIV/AIDS prevalence and mortality. METHODS: Available data on sero-surveillance of HIV/AIDS in countries were compiled in the process of making the end of 1999 estimates of HIV/AIDS. These data came primarily from the HIV/AIDS Surveillance Database developed by the United States Census Bureau, from a database maintained by the European Centre for the Epidemiological Monitoring of AIDS and all country reports on sentinel surveillance that had been provided to World Health Organization or UNAIDS. Procedures were developed to score quality of surveillance systems based on four dimensions of quality: timeliness and frequency; appropriateness of groups; consistency of sites over time; and coverage provided by the system. In total, the surveillance systems from 167 countries were analysed. RESULTS: Forty-seven of the 167 countries whose surveillance systems were rated were judged to have fully implemented sentinel surveillance systems; 51 were judged to have systems that had some or most aspects of a good HIV surveillance system in place and 69 were rated as having poorly functioning or non-existent surveillance systems. CONCLUSION: This analysis suggests that the quality of HIV surveillance varies considerably. The majority of countries most affected by HIV/AIDS have systems that are providing sufficient sero-prevalence data for tracking the epidemic and making reasonable estimates of HIV prevalence. However, many countries have poor systems and strengthening these is an urgent priority.			

REPRESENTATIVENESS AND BIAS

Reference	Abstract	Date used	Name of evaluation	Comments
Del Rio Vilas and Pfeiffer. The evaluation of bias in scrapie surveillance: A review. *Veterinary Journal* 185: 259–64, 2010	Evaluation of surveillance systems is a common practice in the context of human health, but only recently has been applied in the veterinary field. Commonly, a series of attributes are monitored to assess the system. Suboptimal performance of the surveillance in relation to any of these attributes may lead to bias in the surveillance results. The intensity of scrapie surveillance has increased considerably in recent years as a result of public health concerns. In this paper, a number of approaches described in the literature for the evaluation of surveillance systems are reviewed, with a focus on the sensitivity and representativeness attributes of scrapie surveillance systems in the European Union. Many of the methods applied in other fields, such as ecology and public health, are exchangeable and relevant for scrapie surveillance.			
Del Rio Vilas and Böhning. Application of one-list capture-recapture models to scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 85: 253–66, 2008	In this paper, we apply one-list capture–recapture models to estimate the number of scrapie-affected holdings in Great Britain. We applied this technique to the Compulsory Scrapie Flocks Scheme dataset where cases from all the surveillance sources monitoring the presence of scrapie in Great Britain, the abattoir survey, the fallen stock survey and the statutory reporting of clinical cases, are gathered. Consequently, the estimates of prevalence obtained from this scheme should be comprehensive and cover all the different presentations of the disease captured individually by the surveillance sources. Two estimators were applied under the one-list approach: the Zelterman estimator and Chao's lower bound estimator. Our results could only inform with confidence the scrapie-affected holding population with clinical disease; this moved around the figure of 350 holdings in Great Britain for the period under study, April 2005–April 2006. Our models allowed the stratification by surveillance source and the input of covariate information, holding size and country of origin. None of the covariates appear to inform the model significantly.			
Dohoo I, Martin W, Stryhn H (Eds). *Veterinary Epidemiologic Research*. Second Edition. VER Inc., 2009	Veterinary Epidemiologic Research, Second Edition is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar). New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures. All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case–control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.			

Guasticchi *et al.* Syndromic surveillance: sensitivity and positive predictive value of the case definitions. *Epidemiology and Infection* 137: 662–71, 2009	The aim of the study was to measure the positive predictive value (PPV) and sensitivity of operational case definitions of 13 syndromes in a surveillance system based on the Emergency online database of the Lazio region. The PPVs were calculated using electronic emergency department (ED) medical records and subsequent hospitalizations to ascertain the cases. Sensitivity was calculated using a modified capture-recapture method. The number of cases that fulfilled the case definition criteria in the 2004 database ranged from 27 320 for gastroenteritis to three for haemorrhagic diarrhoea. The PPVs ranged from 99.3 to 20; sepsis, meningitis-like and coma were below 50%. The estimated sensitivity ranged from 90% for coma to 22% for haemorrhagic diarrhoea. Syndromes such as gastroenteritis, where the signs, symptoms, and exposure history provide immediate diagnostic implications fit this surveillance system better than others such as haemorrhagic diarrhoea, where symptoms are not evident and a more precise diagnosis is needed.			
Hendrikx *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* 52: 465–75, 2005	Performance indicators are variables used to calculate on a continuous basis the operational level of a surveillance network's priority activities. A 10-step process was developed to enable network coordinators to identify specific performance indicators to help them monitor their networks. The methodology was based on a listing of surveillance activities, the definition and choice of network's global priority objectives, the construction of performance indicators and monitoring tables and the implementation of the system within the network. This process was implemented for the bovine clinical digestive salmonellosis surveillance network in France (RESSAB). The process produced a list of 26 activities synthesized into 15 global objectives, from which 12 were retained as priority objectives. This made it possible to develop 12 performance indicators. Indicators retrospectively calculated for the year 2003 indicated that RESSAB was operating according to the objectives set by the network's stakeholders and its financial supporter. The methodology was implemented successfully and was not very time consuming (12 person-days) or expensive. The decision makers and actors in the network quickly appropriated the system. The methodology can thus be considered validated through this example. Nevertheless, the risks inherent in the use of performance indicators must be addressed by ensuring the relevance of the selected indicators through external assessment and by prioritizing an internal and participatory approach to avoid a misuse of the performance indicators. In addition, considering that performance indicators address only the operation of the network, relevance of surveillance procedure should be addressed by external evaluation.			
Lynn *et al*. An evaluation of scrapie surveillance in the United States. *Preventive Veterinary Medicine* 81: 70–9, 2007	Animal health surveillance systems should reflect national disease control priorities and promote the best use of public resources by maximizing effectiveness and efficiency. A surveillance system should be routinely evaluated to assess the degree to which the system accomplishes these goals, fulfils its stated objectives, and meets accepted surveillance standards. In the United States, there are a number of disparate endemic disease surveillance and eradication programs. The National Animal Health Surveillance System is a federal initiative designed to combine animal health surveillance and monitoring activities into a comprehensive and coordinated system. A protocol has been developed to facilitate the evaluation of animal health surveillance systems and investigate opportunities for coordination between the different surveillance and eradication programs. The evaluation protocol was based largely on protocols developed for public health but			

	adapted for the specific needs and goals of animal health surveillance. The evaluation process was designed to identify program strengths and areas for improvement and facilitate the system's adaptability to changing situations. The evaluation protocol was applied to the scrapie surveillance system in the United States; scrapie surveillance was found to be an important part of surveillance for transmissible spongiform encephalopathies. Results from the evaluation of sensitivity, sampling methods and representativeness are presented.			
Macarthur and Pless. Evaluation of the quality of an injury surveillance system. *American Journal of Epidemiology* 149: 586–92, 1999	The sensitivity, positive predictive value, and representativeness of the Canadian Hospitals Injury Reporting and Prevention Program (CHIRPP) were assessed. Sensitivity was estimated at four centers in June through August 1992, by matching independently identified injuries with those in the CHIRPP database. The positive predictive value was determined by reviewing all "injuries" in the database (at Montreal Children's Hospital) that could not be matched. Representativeness was assessed by comparing missed with captured injuries (at Montreal Children's Hospital) on demographic, social, and clinical factors. Sensitivity ranged from 30% to 91%, and the positive predictive value was 99.9% (i.e., the frequency of false-positive capture was negligible). The representativeness study compared 277 missed injuries with 2,746 captured injuries. The groups were similar on age, sex, socioeconomic status, delay before presentation, month, and day of presentation. Injuries resulting in admissions, poisonings, and those presenting overnight were, however, more likely to be missed. The adjusted odds ratio of being missed by CHIRPP for admitted injuries (compared with those treated and released) was 13.07 (95% confidence interval 7.82-21.82); for poisonings (compared with all other injuries), it was 9.91 (95% confidence interval 5.39-18.20); and for injuries presenting overnight (compared with those presenting during the day or evening), it was 4.11 (95% confidence interval 3.11-5.44). These injuries were probably missed because of inadequate education of participants in the system. The authors conclude that CHIRPP data are of relatively high quality and may be used, with caution, for research and public health policy.			
Morignat *et al*. Estimates of the prevalence of spongiform encephalopathies in sheep and goats in France in 2002. *Veterinary Record* 158: 683–7, 2006	An active surveillance programme for transmissible spongiform encephalopathies (TSES) in sheep and goats was implemented in France in 2002 at abattoirs and rendering plants. The analysis of the results of this programme highlighted three biases: a potentially non-random sampling scheme in both rendering plants and abattoirs, a heterogeneous geographical sampling ratio, and the use of two diagnostic tests of unequal sensitivity. Simulations were run to estimate the prevalence of TSES by taking these biases into account. A comparison of the prevalence of TSES calculated from the raw data with the simulation results showed that the effects of non-random sampling were minor in comparison with the effects of the heterogeneous geographical sampling ratio and the use of two diagnostic tests.			

Van Benthem and van Vliet. Reflections on an evaluation of the Dutch Infectious Diseases Surveillance Information System. *European Surveillance* 13(11), 2008	The Netherlands' Infectious diseases Surveillance Information System (ISIS) was developed 12 years ago as an early warning system for the country. The initial objective was to establish a surveillance system that gathered the test results of all microorganisms from all medical microbiology laboratories (MMLs) in the Netherlands on a daily basis in order to create an early warning system. This paper analyses the most important results of a recent evaluation of the system. The evaluation was based on an analysis of early warning signals to detect outbreaks, number of visits to the ISIS website, and interviews with stakeholders, documentation on the ISIS system, and analyses of the ISIS MML database. While the daily collection of data on all micro-organisms for early warning has been achieved, the connection of all 85 MMLs in the Netherlands to the central ISIS MML database has not been achieved - only 18 MMLs have been connected. This has resulted in a low coverage and non-representative selection of MMLs for the Netherlands and therefore national outbreaks were missed. Data were used to determine trends in antimicrobial resistance over time. The ISIS system was not found suitable for early warning since outbreaks were detected via other systems. However, with some adaptations the ISIS system could be suitable for the surveillance of antimicrobial resistance. Furthermore, the discontinuation of this network would cause the loss of the most important data system for antimicrobial resistance in the Netherlands, since there is no other national system that gathers data on this topic. This evaluation resulted in a restart of the network.			
Wells *et al*. Use of epidemiologic information in targeted surveillance for population inference. *Preventive Veterinary Medicine* 89: 43–50, 2009	Epidemiologic information, including animal characteristics (e.g., observable risk factors or clinical signs) predisposing to animal disease, is frequently used for design of targeted surveillance systems, but this information is infrequently used for population inference. In this study, we report the evaluation of use of epidemiologic information for population inference in targeted surveillance in three animal disease scenarios. We adapted sampling theory using Monte Carlo methods to determine target population sample size to detect disease with 95% confidence, using information from the epidemiologic parameters risk ratio and fraction of the population with the characteristic. These parameters and their uncertainties were derived from a reference population. The next step was to use a second (sampled) population to evaluate effects of sampling the targeted population. The focus of the study was on estimation of prevalence. Our results showed that if one is less certain of the epidemiologic parameters, a rational decision is to model the input parameter distributions reflecting this uncertainty, thereby increasing the sample size above the minimum needed for the detection of the disease with a known confidence. Targeted surveillance is appropriate for prevalence estimation when one has representative and justifiable estimates of key epidemiologic parameters.			
Williams *et al*. Population inferences from targeted sampling with uncertain epidemiologic information. *Preventive Veterinary Medicine* 89: 25–33, 2009	Targeted sampling is an increasingly popular method of data collection in animal-based epidemiologic studies. This sampling approach allows the user to exclusively choose samples from subpopulations that have a higher likelihood of the disease of interest. This is achieved by selecting animals from a subpopulation that exhibits some characteristic that indicates a higher probability of the presence of the disease. Inferences drawn from a targeted sample require information regarding the epidemiology of the disease under surveillance, which is generally not known with certainty. This study describes estimators for both the detection of disease and the estimation of prevalence when targeted sampling is employed. Modifications of these estimators			

	are provided that account for the uncertainty in the parameters that describe the epidemiology of the disease. Results of a simulation study are provided to illustrate the effect of the uncertainty in these parameters.			

RESOURCE AVAILABILITY

Reference	Abstract	Date used	Name of evaluation	Comments
Moran and Fofana. An economic evaluation of the control of three notifiable fish diseases in the United Kingdom. *Preventive Veterinary Medicine* 80: 193–208, 2007	We summarised the challenges faced in an ex ante cost–benefit appraisal of United Kingdom government spending on disease surveillance for three notifiable fish diseases: infectious salmon anaemia (ISA), viral haemorrhagic septicaemia (VHS) and infectious haemorrhagic necrosis (IHN). We used a social cost–benefit analysis and adopted a national perspective. We compared costs of current public and private surveillance effort with the benefits stated in terms of the avoided private and social costs of potential disease outbreaks. Spending on ISA and VHS were predicted to be efficient; the benefit–cost ratios were always ≥3.2 for ISA and ≥5.8 for VHS for all nine scenarios examined for each infection. However, the benefit–cost ratio for IHN was predicted never to exceed 1.6, and was <1.0 in five of the nine scenarios-so spending on IHN would be harder to justify.			
Morris *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* 18: 415–22, 1996	BACKGROUND: In England and Wales, surveillance of communicable disease is carried out and co-ordinated by the Public Health Laboratory Service (PHLS). The surveillance of HIV infection and AIDS is undertaken by the PHLS AIDS Centre at the Communicable Disease Surveillance Centre (CDSC). Epidemiological data derived from surveillance are not, however, a free good: they are a resource with an associated opportunity cost and should therefore be open to economic appraisal alongside other users of health care resources such as medical interventions. This paper assembles information on the current surveillance of HIV and AIDS in England and Wales, and explores methods for performing an economic evaluation of such activities. METHODS: An examination of the cost and effectiveness of the PHLS AIDS Centre's epidemiological surveillance mechanisms for HIV and AIDS in England and Wales was undertaken. The total costs of each component of surveillance of HIV and AIDS in England and Wales were calculated. Two categories of cost were estimated: peripheral costs incurred by reporters in reporting AIDS cases or HIV infections or by laboratories in collecting samples; and central costs incurred by the PHLS AIDS Centre in processing and analysing incoming data. Using these cost data and information from a cost-effectiveness register, the additional health gains that would have to be obtained from surveillance to make the programme broadly cost-effective in comparison with other accepted uses of health service resources were then estimated. RESULTS: In the financial year 1993-1994 the total costs of surveillance were estimated to be 1.4 million pounds. To avoid being considered relatively cost-ineffective at least 3.5 infections per annum need to be averted. To be considered favourably cost-effective, approximately 9.5 infections per annum need to be averted. CONCLUSIONS: In 1993-1994, expenditure on surveillance of HIV and AIDS accounted for less than 1 per cent of the total allocation of resources to the National Health Service for all HIV and			

	AIDS activities. Given these cost estimates, the number of infections which surveillance would have to contribute towards preventing in order to be considered cost-effective is low.			
Phillips *et al*. Evaluation of program performance and expenditures in a report of performance measures via a case study of two Florida tuberculosis programs. *Evaluation and Program Planning* 33: 373–8, 2010	Health Department (HD) managers at both state and local levels are in desperate need of tools to assist in monitoring and evaluating programs. The purpose of this study is to assess the feasibility and utility of linking program performance scores and expenditures into a Report of Performance Measures (RPM). We analysed secondary data on performance indicators, selected by HD staff, and expenditures, related to six surveillance activities, from two, similar, high-incidence, tuberculosis (TB) programs in Florida from 2002 to 2003. We compared the findings between the county HDs as an illustration of basic cost-effectiveness benchmarking, based on the cost-effectiveness grid. Data included here provide examples of: (1) two instances in which one county was operating relatively inefficiently compared to the other; (2) two instances in which performance and expenditures were similar for the counties; and (3) two instances in which one county spent more for higher performance scores than the other. These data illustrate how the RPM can be used to facilitate benchmarking, a basic evaluation tool. They also demonstrate ways to identify potential operational inefficiencies in a single time period and ultimately over time. It is thus likely to be a feasible and useful management tool.			

SENSITIVITY

Reference	Abstract	Date used	Name of evaluation	Comments
Audigé and Becket. Quantitative assessment of the validity of animal-health surveys using stochastic modelling. *Preventive Veterinary Medicine* 38: 259–76, 1999	This paper presents a stochastic simulation model to evaluate the efficacy of regional or national surveys aimed at identifying infection in populations of animals. The process of evaluation involves specification or calculation of cluster-level test sensitivity and specificity, which are derived from two probability distributions of the number of individual-level positive tests expected from non-infected and infected clusters, respectively. Probability distributions for the number of positive clusters expected in a situation of freedom from infection and under various levels of cluster prevalence are specified and used to determine survey properties (the survey being considered a diagnostic system), and ROC curves are drawn. Likelihood ratios allow investigators to state the extent to which a survey result is more likely to be observed if the region or country is infected at a given prevalence than if it is free from infection. The result of a survey carried out to investigate the presence of porcine reproductive and respiratory syndrome (PRRS) in Switzerland is used to illustrate this approach. The model can be adapted to a wide range of survey designs.			
Branscum *et al*. Sample size calculation for disease freedom and prevalence estimation surveys. *Statistics in Medicine* 25: 2658–74, 2006	We developed a Bayesian approach to sample size calculations for studies designed to estimate disease prevalence that uses a hierarchical model for estimating the proportion of infected clusters (cluster-level prevalence) within a country or region. The clusters may, for instance, be villages within a region, cities within a state, or herds within a country. Our model allows for clusters with zero prevalence and for variability in prevalences among infected clusters. Moreover, uncertainty about diagnostic test accuracy and within-cluster prevalences is accounted for in the model. A predictive approach is used to address the issue of sample size selection in human and animal health surveys. We present sample size calculations for surveys designed to substantiate freedom of a region from an infectious agent (disease freedom surveys) and for surveys designed to estimate cluster-level prevalence of an endemic disease (prevalence estimation surveys). In disease freedom surveys, for instance, assuming the cluster-level prevalence for a particular infectious agent in the region is greater than a maximum acceptable threshold, a sample size combination consisting of the number of clusters sampled and number of subjects sampled per cluster can be determined for which authorities conducting the survey detect this excessive cluster-level prevalence with high predictive probability. The method is straightforward to implement using the Splus/R library emBedBUGS together with WinBUGS.			
Buckeridge. Outbreak detection through automated surveillance: a review of the determinants of detection. *Journal of Biomedical Informatics* 40: 370–9, 2007	Public health agencies and other groups have invested considerable resources in automated surveillance systems over the last decade. These systems generally follow syndromes in pre-diagnostic data drawn from sources such as emergency department visits. A main goal of syndromic surveillance systems is to detect outbreaks rapidly and the number of studies evaluating outbreak detection has increased recently. This paper reviews these studies with the goal of identifying the determinants of outbreak detection in automated syndromic surveillance systems. The review identified 35 studies with 22 studies (63%) relying on naturally occurring			

	outbreaks and 13 studies (37%) relying on simulated outbreaks. In general, the results from these studies suggest that syndromic surveillance systems are capable of detecting some types of disease outbreaks rapidly with high sensitivity. The determinants of detection included characteristics of the system and of the outbreak. Influential system characteristics included representativeness, the outbreak detection algorithm, and the specificity of the algorithm. Important outbreak characteristics included the magnitude and shape of the signal and the timing of the outbreak. Future evaluations should aim to address inconsistencies in the evidence noted in this review and to identify the potential influence of other factors on outbreak detection.			
Cannon. Demonstrating disease freedom – combining confidence levels. *Preventive Veterinary Medicine* 52: 227–49, 2002	Part of the requirements for demonstrating disease freedom usually will be that sufficient testing be done to give a specified confidence of detecting the disease if it were present at a specified level. Often, this requirement is translated into a fixed testing regime that must be followed (an inflexible approach that might not be the most economic or practical solution). A more flexible approach is to specify the capabilities of the various tests that can be used to detect the disease, and let the party hoping to demonstrate disease freedom decide upon the testing regime. The question then arises as to how to combine information that can come from a variety of sources over a period of time to give an overall level of confidence. Two methods are given. The first, an exact method based on multiplying probabilities, would be more appropriate for a survey of an area in which no disease is thought to be present. The second method (more appropriate for a herd-assurance program within an infected area) is a point-based system that takes into account the different sensitivities of the methods used to detect disease and the change in prevalence over time. It allocates points for each test done proportional to the sensitivity of the test and the prevalence at the time of testing.			
del Rio Vilas *et al*. A case study of capture-recapture methodology using scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 67: 303–17, 2005	We applied capture–recapture methodology (CRC) to data from three surveillance sources (statutory notification, abattoir survey (AS) and fallen stock (FS) survey) to estimate the number of holdings infected with scrapie in Great Britain and to assess the sensitivity of the surveillance network. Between January 1, 2002 and March 31, 2003, 144 holdings were identified by the three sources. Using CRC modelling techniques, we estimated a minimum lower bound for the total number of holdings infected as 642. A biologically plausible positive dependence between the statutory reporting and the fallen stock survey was found statistically significant. The sensitivity of the three sources combined was very low. The integration of the three overlapping sources provided a better understanding of the interactions within the surveillance network. However, the scarcity of the data and reduced overlapping among sources only allowed for very cautious inferences to be drawn about the true proportion of scrapie affected holdings in the national population. Future surveys and surveillance activities should be planned such that the resulting data can be used more effectively as part of CRC modelling approaches.			
del Rio Vilas and Bohning D. Application of one-list capture-recapture models to scrapie surveillance data in Great Britain.	In this paper, we apply one-list capture–recapture models to estimate the number of scrapie-affected holdings in Great Britain. We applied this technique to the Compulsory Scrapie Flocks Scheme dataset where cases from all the surveillance sources monitoring the presence of scrapie in Great Britain, the abattoir survey, the fallen stock survey and the statutory reporting of clinical			

Preventive Veterinary Medicine 85: 253–66, 2008	cases, are gathered. Consequently, the estimates of prevalence obtained from this scheme should be comprehensive and cover all the different presentations of the disease captured individually by the surveillance sources. Two estimators were applied under the one-list approach: the Zelterman estimator and Chao's lower bound estimator. Our results could only inform with confidence the scrapie-affected holding population with clinical disease; this moved around the figure of 350 holdings in Great Britain for the period under study, April 2005–April 2006. Our models allowed the stratification by surveillance source and the input of covariate information, holding size and country of origin. None of the covariates appear to inform the model significantly.			
Fujii *et al*. Evaluation of a sentinel surveillance system for influenza, 1995-2000, Kyoto City, Japan. *Japanese Journal of Infectious Diseases* 55: 23–6, 2002	We compared a municipal sentinel surveillance system for influenza with the Microbial Isolation Surveillance System (MISS) in Kyoto City, Japan. Sensitivity, specificity, and predictive value positive (PVP) of the Influenza Sentinel Surveillance System (ISSS) compared to the MISS were calculated by influenza season and by month. Sensitivity ranged from 80 to 97%, specificity ranged from 55 to 77%, and PVP ranged from 29 to 52% by season ($P < 0.001$). On the other hand, sensitivity ranged from 86 to 100%, specificity ranged from 38 to 66%, and PVP ranged from 31 to 50% by month ($P < 0.001$). Specificity was calculated as 93% in November. The sensitivity of ISSS was found to be sufficient regardless of the magnitude of influenza activity. Specificity varied by season, indicating the difficulty of clinically diagnosing other respiratory illnesses. The PVP remained at less than 50% before and after the influenza seasons and it varied year by year. In general, the ISSS is a good surveillance system for monitoring influenza activity.			
Hadorn and Stärk. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Veterinary Research* 39(6): 57, 2008	Surveillance for rare and emerging infectious diseases poses a special challenge to veterinary services. Most emerging infectious diseases like bovine tuberculosis (bTB) are zoonoses, affecting both human and animal populations. Despite the low prevalence of such an emerging infectious disease at time of incursion, the surveillance system should be able to detect the presence of the disease as early as possible. Because passive surveillance is a relatively cost-effective and therefore commonly used process, it is the basic tool for infectious disease surveillance. Because of under-reporting in passive surveillance, cost-intensive active surveillance is often required to increase the sensitivity of the surveillance system. Using scenario tree modelling, the sensitivity of passive and active surveillance system components (SSC) can be quantified and an optimal, cost-effective surveillance system developed considering the contributions of each SSC. We illustrate this approach with the example of bTB surveillance in Switzerland where the surveillance system for bTB consists of meat inspection at the slaughterhouse (SLI), passive clinical surveillance on farm (CLIN) and human surveillance (HS). While the sensitivities for CLIN and HS were both negligible (<1%), SLI was assessed to be 55.6%. The scenario tree model showed that SLI is increasable up to 80.4% when the disease awareness of meat inspectors in Switzerland is enhanced. A hypothetical random survey (RS) was also compared with a targeted survey (TS) in high-risk strata of the cattle population, and the sensitivity of TS was 1.17-fold better than in RS but with 50% of the sample size.			
Hood *et al*. Alternative methods for	Stochastic scenario trees are a new and popular method by which surveillance systems can be			

computing the sensitivity of complex surveillance systems. *Risk Analysis* 29: 1686–98, 2009	analysed to demonstrate freedom from pests and disease. For multiple component systems-such as a combination of a serological survey and systematically collected observations-it can be difficult to represent the complete system in a tree because many branches are required to represent complex conditional relationships. Here we show that many of the branches of some scenario trees have identical outcomes and are therefore redundant. We demonstrate how to prune branches and derive compact representations of scenario trees using matrix algebra and Bayesian belief networks. The Bayesian network representation is particularly useful for calculation and exposition. It therefore provides a firm basis for arguing disease freedom in international forums.			
Jackson *et al*. A Simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Medical Informatics and Decision Making* 7: 6, 2007	BACKGROUND: The usefulness of syndromic surveillance for early outbreak detection depends in part on effective statistical aberration detection. However, few published studies have compared different detection algorithms on identical data. In the largest simulation study conducted to date, we compared the performance of six aberration detection algorithms on simulated outbreaks superimposed on authentic syndromic surveillance data. METHODS: We compared three control-chart-based statistics, two exponential weighted moving averages, and a generalized linear model. We simulated 310 unique outbreak signals, and added these to actual daily counts of four syndromes monitored by Public Health--Seattle and King County's syndromic surveillance system. We compared the sensitivity of the six algorithms at detecting these simulated outbreaks at a fixed alert rate of 0.01. RESULTS: Stratified by baseline or by outbreak distribution, duration, or size, the generalized linear model was more sensitive than the other algorithms and detected 54% (95% CI = 52%-56%) of the simulated epidemics when run at an alert rate of 0.01. However, all of the algorithms had poor sensitivity, particularly for outbreaks that did not begin with a surge of cases. CONCLUSION: When tested on county-level data aggregated across age groups, these algorithms often did not perform well in detecting signals other than large, rapid increases in case counts relative to baseline levels.			
Kleinman and Abrams. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 27: 4057–68, 2008	In modern surveillance of public health, data may be reported in a timely fashion and include spatial data on cases in addition to the time of their occurrence. This has led to many recent developments in statistical methods to detect events of public health importance. However, there has been relatively little work about how to compare such methods. One powerful rationale for performing surveillance is earlier detection of events of public health significance; previous evaluation tools have focused on metrics that include the timeliness of detection in addition to sensitivity and specificity. However, such metrics have not accounted for the number of persons affected by the events. We re-examine the rationale for this surveillance and conclude that earlier detection is preferred because it can prevent additional morbidity and mortality. On the basis this observation, we propose evaluating the number of cases prevented by each detection method, and include this information in assessing the value of different detection methods. Using this approach incorporates more information about the events and the detection and provides a sound basis for making decisions about which detection methods to employ.			

Knight-Jones *et al*. Evaluation of the effectiveness and efficiency of wild bird surveillance for avian influenza. *Veterinary Research* 41(4): 50, 2010	This study aimed to assess which method of wild waterbird surveillance had the greatest probability of detecting highly pathogenic avian influenza (HPAI) H5N1 during a period of surveillance activity, the cost of each method was also considered. Lake Constance is a major wintering centre for migratory waterbirds and in 2006 it was the site of an HPAI H5N1 epidemic in wild birds. Avian influenza surveillance was conducted using harmonised approaches in the three countries around the lake, Austria, Germany and Switzerland, from 2006-2009. The surveillance consisted of testing birds sampled by the following methods: live birds caught in traps, birds killed by hunters, birds caught in fishing nets, dead birds found by the public and catching live Mute Swans (*Cygnus olor*); sentinel flocks of Mallards (*Anas platyrhynchos*) were also used. Scenario tree analysis was performed including sensitivity analysis, followed by assessment of cost-effectiveness. Results indicated that if HPAI H5N1 was present at 1% prevalence and assuming HPAI resulted in bird mortality, sampling dead birds found by the public and sentinel surveillance were the most sensitive approaches despite residual uncertainty over some parameters. The uncertainty over the mortality of infected birds was an influential factor. Sampling birds found dead was most cost-effective, but strongly dependent on mortality and awareness of the public. Trapping live birds was least cost-effective. Based on our results, we recommend that future HPAI H5N1 surveillance around Lake Constance should prioritise sentinel surveillance and, if high mortality is expected, the testing of birds found dead.			
Lynn *et al*. An evaluation of scrapie surveillance in the United States. *Preventive Veterinary Medicine* 81: 70–9, 2007	Animal health surveillance systems should reflect national disease control priorities and promote the best use of public resources by maximizing effectiveness and efficiency. A surveillance system should be routinely evaluated to assess the degree to which the system accomplishes these goals, fulfils its stated objectives, and meets accepted surveillance standards. In the United States, there are a number of disparate endemic disease surveillance and eradication programs. The National Animal Health Surveillance System is a federal initiative designed to combine animal health surveillance and monitoring activities into a comprehensive and coordinated system. A protocol has been developed to facilitate the evaluation of animal health surveillance systems and investigate opportunities for coordination between the different surveillance and eradication programs. The evaluation protocol was based largely on protocols developed for public health but adapted for the specific needs and goals of animal health surveillance. The evaluation process was designed to identify program strengths and areas for improvement and facilitate the system's adaptability to changing situations. The evaluation protocol was applied to the scrapie surveillance system in the United States; scrapie surveillance was found to be an important part of surveillance for transmissible spongiform encephalopathies. Results from the evaluation of sensitivity, sampling methods and representativeness are presented.			

Mandl *et al*. Measuring outbreak-detection performance by using controlled feature set simulations. *Morbidity and Mortality Weekly Report* 53(Supp): 130–6, 2004	Introduction: The outbreak-detection performance of a syndromic surveillance system can be measured in terms of its ability to detect signal (i.e., disease outbreak) against background noise (i.e., normally varying baseline disease in the region). Such benchmarking requires training and the use of validation data sets. Because only a limited number of persons have been infected with agents of biologic terrorism, data are generally unavailable, and simulation is necessary. An approach for evaluation of outbreak-detection algorithms was developed that uses semisynthetic data sets to provide real background (which effectively becomes the noise in the signal-to-noise problem) with artificially injected signal. The injected signal is defined by a controlled feature set of variable parameters, including size, shape, and duration. Objectives: This report defines a flexible approach to evaluating public health surveillance systems for early detection of outbreaks and provides examples of its use. Methods: The stages of outbreak detection are described, followed by the procedure for creating data sets for benchmarking performance. Approaches to setting parameters for simulated outbreaks by using controlled feature sets are detailed, and metrics for detection performance are proposed. Finally, a series of experiments using semisynthetic data sets with artificially introduced outbreaks defined with controlled feature sets is reviewed. Results: These experiments indicate the flexibility of controlled feature set simulation for evaluating outbreak-detection sensitivity and specificity, optimizing attributes of detection algorithms (e.g., temporal windows), choosing approaches to syndrome groupings, and determining best strategies for integrating data from multiple sources. Conclusions: The use of semisynthetic data sets containing authentic baseline and simulated outbreaks defined by a controlled feature set provides a valuable means for benchmarking the detection performance of syndromic surveillance systems.			
Martin *et al*. Demonstrating freedom from disease using multiple complex data sources: A methodology based on scenario trees. *Preventive Veterinary Medicine* 79: 71–97, 2007	Current methods to demonstrate zone or country freedom from disease are based on either quantitative analysis of the results of structured representative surveys, or qualitative assessments of multiple sources of evidence (including complex non-representative sources). This paper presents a methodology for objective quantitative analysis of multiple complex data sources to support claims of freedom from disease. Stochastic scenario tree models are used to describe each component of a surveillance system (SSC), and used to estimate the sensitivity of each SSC. The process of building and analysing the models is described, as well as techniques to take into account any lack of independence between units at different levels within a SSC. The combination of sensitivity estimates from multiple SSCs into a single estimate for the entire surveillance system is also considered, again taking into account lack of independence between components. A sensitivity ratio is used to compare different components of a surveillance system. Finally, calculation of the probability of country freedom from the estimated sensitivity of the surveillance system is illustrated, incorporating the use and valuation of historical surveillance evidence.			

Smith. Survey design for detecting rare freshwater mussels. *Journal of the North American Benthological Society* 25(3): 701–11, 2006	A common objective when surveying freshwater mussels is to detect the presence of rare populations. In certain situations, such as when endangered or threatened species are potentially in the area of a proposed impact, the survey should be designed to ensure a high probability of detecting species presence. Linking survey design to probability of detecting species presence has been done for quantitative surveys, but commonly applied designs that are based on timed searches have not made that connection. I propose a semiquantitative survey design that links search area and search efficiency to probability of detecting species presence. The survey can be designed to protect against failing to detect populations above a threshold abundance (or density). I illustrate the design for surveys to detect clubshell (*Pluerobema clava*) and northern riffleshell (*Epioblasma torulosa rangiana*) in the Allegheny River. Monte Carlo simulation indicated that the proposed survey design performs well under a range of spatial distributions and low densities (< 0.05 m(2)) where search area is sufficient to ensure that the probability of detecting species presence is predicted to be >= 0.85.			
Watkins *et al*. Approaches to the evaluation of outbreak detection methods. *BMC Public Health* 6: 263, 2006	BACKGROUND: An increasing number of methods are being developed for the early detection of infectious disease outbreaks which could be naturally occurring or as a result of bioterrorism; however, no standardised framework for examining the usefulness of various outbreak detection methods exists. To promote comparability between studies, it is essential that standardised methods are developed for the evaluation of outbreak detection methods. METHODS: This analysis aims to review approaches used to evaluate outbreak detection methods and provide a conceptual framework upon which recommendations for standardised evaluation methods can be based. We reviewed the recently published literature for reports which evaluated methods for the detection of infectious disease outbreaks in public health surveillance data. Evaluation methods identified in the recent literature were categorised according to the presence of common features to provide a conceptual basis within which to understand current approaches to evaluation. RESULTS: There was considerable variation in the approaches used for the evaluation of methods for the detection of outbreaks in public health surveillance data, and appeared to be no single approach of choice. Four main approaches were used to evaluate performance, and these were labelled the Descriptive, Derived, Epidemiological and Simulation approaches. Based on the approaches identified, we propose a basic framework for evaluation and recommend the use of multiple approaches to evaluation to enable a comprehensive and contextualised description of outbreak detection performance. CONCLUSION: The varied nature of performance evaluation demonstrated in this review supports the need for further development of evaluation methods to improve comparability between studies. Our findings indicate that no single approach can fulfil all evaluation requirements. We propose that the cornerstone approaches to evaluation identified provide key contributions to support internal and external validity and comparability of study findings, and suggest these be incorporated into future recommendations for performance assessment.			

Willeberg *et al*. Epidemiological models to support animal disease surveillance activities. *Scientific and Technical Review* 30(2): 603–14, 2011	Epidemiological models have been used extensively as a tool in improving animal disease surveillance activities. A review of published papers identified three main groups of model applications: models for planning surveillance, models for evaluating the performance of surveillance systems and models for interpreting surveillance data as part of ongoing control or eradication programmes. Two Danish examples are outlined. The first illustrates how models were used in documenting country freedom from disease (trichinellosis) and the second demonstrates how models were of assistance in predicting the risk of future cases, detected and undetected, of a waning infection of bovine spongiform encephalopathy. Both studies were successful in advancing European policy changes to reduce the cost of surveillance to appropriate levels given the magnitude of the respective hazards.			

SPECIFICITY

Reference	Abstract	Date used	Name of evaluation	Comments
Del Rio Vilas *et al*. A case of capture- recapture methodology using scrapie surveillance data in Great Britain. *Preventive Veterinary Medicine* 67: 303–17, 2005	We applied capture–recapture methodology (CRC) to data from three surveillance sources (statutory notification, abattoir survey (AS) and fallen stock (FS) survey) to estimate the number of holdings infected with scrapie in Great Britain and to assess the sensitivity of the surveillance network. Between January 1, 2002 and March 31, 2003, 144 holdings were identified by the three sources. Using CRC modelling techniques, we estimated a minimum lower bound for the total number of holdings infected as 642. A biologically plausible positive dependence between the statutory reporting and the fallen stock survey was found statistically significant. The sensitivity of the three sources combined was very low. The integration of the three overlapping sources provided a better understanding of the interactions within the surveillance network. However, the scarcity of the data and reduced overlapping among sources only allowed for very cautious inferences to be drawn about the true proportion of scrapie affected holdings in the national population. Future surveys and surveillance activities should be planned such that the resulting data can be used more effectively as part of CRC modelling approaches.			
Kleinman and Abrams. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 27: 4057–68, 2008	In modern surveillance of public health, data may be reported in a timely fashion and include spatial data on cases in addition to the time of their occurrence. This has led to many recent developments in statistical methods to detect events of public health importance. However, there has been relatively little work about how to compare such methods. One powerful rationale for performing surveillance is earlier detection of events of public health significance; previous evaluation tools have focused on metrics that include the timeliness of detection in addition to sensitivity and specificity. However, such metrics have not accounted for the number of persons affected by the events. We re-examine the rationale for this surveillance and conclude that earlier detection is preferred because it can prevent additional morbidity and mortality. On the basis this observation, we propose evaluating the number of cases prevented by each detection method, and include this information in assessing the value of different detection methods. Using this approach incorporates more information about the events and the detection and provides a sound basis for making decisions about which detection methods to employ.			

TECHNICAL COMPETENCE AND TRAINING

Reference	Abstract	Date used	Name of evaluation	Comments

TIMELINESS

Reference	Abstract	Date used	Name of evaluation	Comments
Carpenter. Evaluation and extension of the cusum technique with an application to Salmonella surveillance. *Journal of Veterinary Diagnostic Investigation* 14: 211–8, 2002	The cumulative sum (cusum) technique was examined for its use in a disease surveillance system for detecting temporal clusters of events. Optimal technique parameters were derived for scenarios not previously considered. Simulation modelling produced results that evaluated deviations from predefined rate increases. The cusum technique was less prone to false alarms and more efficient at detecting large rate increases than previously reported. As demonstrated using data obtained from a Salmonella surveillance system operated by a state animal diagnostic laboratory system, the cusum technique could provide early warning of an epidemic problem.			
CDC. Updated guidelines for evaluating public health surveillance systems. *Morbidity and Mortality Weekly Report* 50: 1–35, 2001	The purpose of evaluating public health surveillance systems is to ensure that problems of public health importance are being monitored efficiently and effectively. CDC's Guidelines for Evaluating Surveillance Systems are being updated to address the need for a) the integration of surveillance and health information systems, b) the establishment of data standards, c) the electronic exchange of health data, and d) changes in the objectives of public health surveillance to facilitate the response of public health to emerging health threats (e.g., new diseases). This report provides updated guidelines for evaluating surveillance systems based on CDC's Framework for Program Evaluation in Public Health, research and discussion of concerns related to public health surveillance systems, and comments received from the public health community. The guidelines in this report describe many tasks and related activities that can be applied to public health surveillance systems.			
CDC. Framework for evaluating public health surveillance systems for early detection of outbreaks. *Morbidity and Mortality Weekly Report* 53: 1–11, 2004	The threat of terrorism and high-profile disease outbreaks has drawn attention to public health surveillance systems for early detection of outbreaks. State and local health departments are enhancing existing surveillance systems and developing new systems to better detect outbreaks through public health surveillance. However, information is limited about the usefulness of surveillance systems for outbreak detection or the best ways to support this function. This report supplements previous guidelines for evaluating public health surveillance systems. Use of this framework is intended to improve decision-making regarding the implementation of surveillance for outbreak detection. Use of a standardized evaluation methodology, including description of system design and operation, also will enhance the exchange of information regarding methods to improve early detection of outbreaks. The framework directs particular attention to the measurement of timeliness and validity for outbreak detection. The evaluation framework is designed to support assessment and description of all surveillance approaches to early detection, whether through traditional disease reporting, specialized analytic routines for aberration detection, or surveillance using early indicators of disease outbreaks, such as syndromic surveillance.			

Colby *et al*. Evaluation of two systems for managing emergency poultry diseases in intensive poultry productions regions. *International Journal of Poultry Science* 2(3): 234–41, 2003	This paper provides an overview of two systems used for the management of infectious disease within the poultry industry on the Delmarva Peninsula. The first system is a paper-based Grid system that was developed in the 1970`s. This system divides the peninsula into a total of 3420 unique locations, each covering approximately 5 square miles. The second system is a Geographic Information System (GIS) database that is currently under construction. Each system is critiqued with respect to several of the criteria established by Klaucke *et al*., (1998) for the evaluation of a surveillance system: simplicity, flexibility, acceptability, representativeness and timeliness. In addition, the objectives of a disease management system in an intensive poultry production area are discussed. A grid-based system is most appropriately used in situations involving a small, easily managed population or geographic location, especially when funding and geographic resources are limited. When multiple diseases or large geographic areas are the focus of a surveillance and monitoring system, or when several different risk factors are to be examined and funding and adequate resources are available, the use of a GIS-based system provides additional flexibility. The database management component of GIS allows for rapid updating of demographic and disease information, and the linkage of the database component with the spatial coordinates of a GIS provides the ability to examine the effects of several risk factors at the same time.			
Del Rocio Amezcua *et al*. Evaluation of a veterinary-based syndromic surveillance system implemented for swine. *Canadian Journal of Veterinary Research* 74(4): 241–51, 2010	Practicing veterinarians play an important role in detecting the initial outbreak of disease in animal populations. A pilot study was conducted to determine the feasibility of a veterinary-based surveillance system for the Ontario swine industry. A total of 7 practitioners from 5 clinics agreed to submit information from July 1, 2007 to June 30, 2008. The surveillance program was evaluated in terms of timeliness, compliance, geographic coverage, and data quality. Our study showed that the veterinary-based surveillance system was acceptable to practitioners and produced useful data. The program obtained information from 25% of pig farms in Ontario during this time period. However, better communication with practitioners, more user-friendly recording systems that can be adapted to each clinic's management system, active involvement of the clinics' technical personnel, and the use of financial incentives may help to improve compliance and timeliness.			

Jackson *et al*. A Simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Medical Informatics and Decision Making* 7: 6, 2007	BACKGROUND: The usefulness of syndromic surveillance for early outbreak detection depends in part on effective statistical aberration detection. However, few published studies have compared different detection algorithms on identical data. In the largest simulation study conducted to date, we compared the performance of six aberration detection algorithms on simulated outbreaks superimposed on authentic syndromic surveillance data. METHODS: We compared three control-chart-based statistics, two exponential weighted moving averages, and a generalized linear model. We simulated 310 unique outbreak signals, and added these to actual daily counts of four syndromes monitored by Public Health--Seattle and King County's syndromic surveillance system. We compared the sensitivity of the six algorithms at detecting these simulated outbreaks at a fixed alert rate of 0.01. RESULTS: Stratified by baseline or by outbreak distribution, duration, or size, the generalized linear model was more sensitive than the other algorithms and detected 54% (95% CI = 52%-56%) of the simulated epidemics when run at an alert rate of 0.01. However, all of the algorithms had poor sensitivity, particularly for outbreaks that did not begin with a surge of cases. CONCLUSION: When tested on county-level data aggregated across age groups, these algorithms often did not perform well in detecting signals other than large, rapid increases in case counts relative to baseline levels.			
Jajosky and Groseclose. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. *BMC Public Health* 4: 29, 2004	Timeliness is a key performance measure of public health surveillance systems. The National Notifiable Diseases Surveillance System (NNDSS) data was evaluated to determine if it could support timely notification and state response to multistate outbreaks. When timeliness of NNDSS data was evaluated, the median national reporting delay, based on date of disease onset, ranged from 12 days for meningococcal disease to 40 days for pertussis. Diseases with the longer incubation periods tended to have a higher percentage of cases reported within its incubation period. Our analysis of NNDSS reporting timeliness indicated that among the conditions evaluated (except for acute hepatitis A infection), the long reporting lag and the variability across states limits the usefulness of NNDSS data and aberration detection analysis of those data for identification of and timely response to multistate outbreaks. A more standardized approach for evaluating and describing surveillance system timeliness should be considered. Further evaluation of the factors that contribute to NNDSS reporting timeliness is warranted.			
Kleinman and Abrams. Assessing the utility of public health surveillance using specificity, sensitivity, and lives saved. *Statistics in Medicine* 27: 4057-68, 2008	In modern surveillance of public health, data may be reported in a timely fashion and include spatial data on cases in addition to the time of their occurrence. This has led to many recent developments in statistical methods to detect events of public health importance. However, there has been relatively little work about how to compare such methods. One powerful rationale for performing surveillance is earlier detection of events of public health significance; previous evaluation tools have focused on metrics that include the timeliness of detection in addition to sensitivity and specificity. However, such metrics have not accounted for the number of persons affected by the events. We re-examine the rationale for this surveillance and conclude that earlier detection is preferred because it can prevent additional morbidity and mortality. On the basis this observation, we propose evaluating the number of cases prevented by each detection method,			

	and include this information in assessing the value of different detection methods. Using this approach incorporates more information about the events and the detection and provides a sound basis for making decisions about which detection methods to employ.			
Mandl *et al*. Measuring outbreak-detection performance by using controlled feature set simulations. *Morbidity and Mortality Weekly Report* 53(Supp): 130–43, 2004	INTRODUCTION: The outbreak-detection performance of a syndromic surveillance system can be measured in terms of its ability to detect signal (i.e., disease outbreak) against background noise (i.e., normally varying baseline disease in the region). Such benchmarking requires training and the use of validation data sets. Because only a limited number of persons have been infected with agents of biologic terrorism, data are generally unavailable, and simulation is necessary. An approach for evaluation of outbreak-detection algorithms was developed that uses semisynthetic data sets to provide real background (which effectively becomes the noise in the signal-to-noise problem) with artificially injected signal. The injected signal is defined by a controlled feature set of variable parameters, including size, shape, and duration. OBJECTIVES: This report defines a flexible approach to evaluating public health surveillance systems for early detection of outbreaks and provides examples of its use. Methods: The stages of outbreak detection are described, followed by the procedure for creating data sets for benchmarking performance. Approaches to setting parameters for simulated outbreaks by using controlled feature sets are detailed, and metrics for detection performance are proposed. Finally, a series of experiments using semisynthetic data sets with artificially introduced outbreaks defined with controlled feature sets is reviewed. RESULTS: These experiments indicate the flexibility of controlled feature set simulation for evaluating outbreak-detection sensitivity and specificity, optimizing attributes of detection algorithms (e.g., temporal windows), choosing approaches to syndrome groupings, and determining best strategies for integrating data from multiple sources. CONCLUSIONS: The use of semisynthetic data sets containing authentic baseline and simulated outbreaks defined by a controlled feature set provides a valuable means for benchmarking the detection performance of syndromic surveillance systems.			
Riera-Montes and Velicko. The Chlamydia surveillance system in Sweden delivers relevant and accurate data: Results from the system evaluation 1997-2008. *Eurosurveillance* 16(27): 2, 2011	This study evaluates the ability of the Chlamydia surveillance system to provide relevant information to inform prevention and control activities in Sweden. The system was evaluated, according to the Guidelines for Evaluating Public Health Surveillance Systems from the United States Centers for Disease Prevention and Control, using surveillance data from 1997 to 2008. We interviewed staff from the Swedish Institute for Communicable Disease Control, the National Board of Health and Welfare and one county medical officer (CMO). We conducted a survey among laboratories, CMOs and a sample of clinics. Satisfaction with the system was good for 86% of CMOs, all laboratories, and 99% of clinics. The interviewed stakeholders considered the system to deliver relevant and accurate information that is useful for health policy decision making. However, the objectives for Chlamydia surveillance should be clearly defined in order to adapt the system requirements, simplify data collection and improve timeliness.			

Siegrist *et al*. Bio-ALIRT biosurveillance detection algorithm evaluation. *Morbidity and Mortality Weekly Report* 53(Supp): 152–8, 2004	Early detection of disease outbreaks by a medical biosurveillance system relies on two major components: 1) the contribution of early and reliable data sources and 2) the sensitivity, specificity, and timeliness of biosurveillance detection algorithms. This paper describes an effort to assess leading detection algorithms by arranging a common challenge problem and providing a common data set. The objectives of this study were to determine whether automated detection algorithms can reliably and quickly identify the onset of natural disease outbreaks that are surrogates for possible terrorist pathogen releases, and do so at acceptable false-alert rates (e.g., once every 2-6 weeks). Historic de-identified data were obtained from five metropolitan areas over 23 months; these data included International Classification of Diseases, Ninth Revision (ICD-9) codes related to respiratory and gastrointestinal illness syndromes. An outbreak detection group identified and labelled two natural disease outbreaks in these data and provided them to analysts for training of detection algorithms. All outbreaks in the remaining test data were identified but not revealed to the detection groups until after their analyses. The algorithms established a probability of outbreak for each day's counts. The probability of outbreak was assessed as an "actual" alert for different false-alert rates. The best algorithms were able to detect all of the outbreaks at false-alert rates of one every 2-6 weeks. They were often able to detect for the same day human investigators had identified as the true start of the outbreak. Because minimal data exists for an actual biologic attack, determining how quickly an algorithm might detect such an attack is difficult. However, application of these algorithms in combination with other data-analysis methods to historic outbreak data indicates that biosurveillance techniques for analysing syndrome counts can rapidly detect seasonal respiratory and gastrointestinal illness outbreaks. Further research is needed to assess the value of electronic data sources for predictive detection. In addition, simulations need to be developed and implemented to better characterize the size and type of biologic attack that can be detected by current methods by challenging them under different projected operational conditions.			

Takahashi *et al*. Evaluation of a public health Salmonella surveillance system in King County, Washington. *American Journal of Infection Control* 32: 7–11, 2004	OBJECTIVES: Salmonella infection is the most common cause of foodborne bacterial outbreaks and deaths in the United States. The effectiveness of Salmonella surveillance for detecting outbreaks depends on timeliness of reporting. We evaluated the public health Salmonella surveillance system in King County, Washington, during an outbreak and at baseline. METHODS: We assessed the timeliness of the Salmonella surveillance system in King County using the Public Health-Seattle & King County Laboratory (County PHL) database from 1998 to 1999. We determined median days for key steps involved in the Salmonella identification and reporting process and the percentage of suspected Salmonella isolates confirmed. Time intervals for key steps during a Salmonella outbreak were compared to baseline surveillance. RESULTS: Of the 652 suspected Salmonella isolates sent to the County PHL from 22 clinical laboratories, 617 (94.6%) were confirmed as Salmonella. Salmonella confirmation rates improved from 1998 to 1999, and 41% of the submitting laboratories, contributing 32.4% of the isolates, had 100% confirmation rates for both years. The median total identification time during the outbreak did not differ significantly from baseline (13 days *vs* 17 days). The time interval for serotyping contributed most to the total identification time. CONCLUSIONS: King County's Salmonella surveillance system requires more than 2 weeks to confirm and report serotype results for Salmonella isolates. Variation in total identification time depends on serotyping time. A more detailed study of other surveillance systems may identify approaches to decrease total identification time.			
Yamamoto *et al*. Evaluation of surveillance strategies for bovine brucellosis in Japan using a simulation model. *Preventive Veterinary Medicine* 86: 57–74, 2008	Bovine brucellosis is caused by *Brucella abortus* and induces abortions in female cattle, with other cattle at risk of infection from the aborted fetus or contaminated placenta. In Japan, the number of cases has dramatically reduced due to national surveillance and eradication strategies. Bovine brucellosis is now believed to be eradicated in Japan. Here, we examine the surveillance strategies currently in place for early detection of infected cattle in the event of a future reintroduction of the disease. We compared current serological surveillance for the dairy population with bulk-milk surveillance and abortion surveillance, and used time to detection as the main criterion of surveillance efficacy. A stochastic individual-based model (IBM) was developed to simulate disease transmission within and between farms. Using outputs from the transmission model, a comparison of surveillance strategies was simulated. For evaluation of the robustness of the parameter values used in the transmission model, a sensitivity analysis was conducted. For the purpose of evaluating the direct costs of each surveillance strategy, the annual number of samples to be tested and the annual number of farms to be visited were estimated. Our results indicated that current serological surveillance with 60-month test intervals is not effective enough for rapid detection of a brucellosis outbreak. Bulk-milk surveillance appeared the most effective method based on the early detection of infected cows and a reduced number of samples required. The time to detection for abortion surveillance was greater than that of bulk-milk surveillance but varied widely depending on the reported ratio of abortions. Results from the surveillance model were consistent when alternative scenarios were applied to the transmission model. Although our model cannot exactly replicate an actual brucellosis outbreak, or the results of surveillance, our results may help decision-makers to choose the most effective surveillance strategy.			

UTILITY

Reference	Abstract	Date used	Name of evaluation	Comments
Carrieri *et al*. Evaluation of the SIMI system, an experimental computerised network for the surveillance of communicable diseases in Italy. *European Journal of Epidemiology* 16: 941–7, 2000	In Italy, the current communicable disease notification system is organised as follows: in each region, Local Health Units (LHU) fill in and forward case report forms (CRF) to the Regional Health Authority, which send aggregated and individual notifications to several central-level institutions. In most regions, all data are recorded manually on hardcopy. Although most relevant data from CRFs are eventually entered into a computerised database at the National Institute of Statistics (ISTAT), the national database is only available 3-4 years later and no data-quality control is performed at that time. To improve the quality and timeliness of notification, in 1994, the Istituto Superiore di Sanità (the National Institute of Health) began to develop an experimental computerised surveillance network for communicable diseases (referred to as 'SIMI'). Specifically, a software was created and distributed to the LHUs and the Regional Health Authorities; staff training was performed; and feedback and analyses of collected data was promoted. SIMI was evaluated in the 13 regions that were participating in 1997 (out of a total of 20 regions in Italy), using criteria commonly used for surveillance systems (i.e., completeness and coherence of data, case definitions, costs, timeliness, and feedback). SIMI was implemented at a limited cost and the data collected were observed to have had a high degree of completeness and internal consistency. The SIMI system has since been adopted for the routine notification of communicable diseases in nearly all regions. Similar evaluations will be necessary for assessing the performance of the various notification systems used across Europe and to include them in a European network.			
Hesterberg *et al*. Evaluation of the sensitivity of the British brucellosis surveillance system using stochastic scenario tree modelling. In: *Proceedings of the 12th Symposium of the International Society of Veterinary Epidemiology and Economics*, Durban, South Africa, 2009	N/A			
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* 11: 225, 2011	BACKGROUND: Surveillance and intervention are resource-using activities of strategies to mitigate the unwanted effects of disease. Resources are scarce, and allocating them to disease mitigation instead of other uses necessarily involves the loss of alternative sources of benefit to people. For society to obtain the maximum benefits from using resources, the gains from disease mitigation must be compared to the resource costs, guiding decisions made with the objective of achieving the optimal net outcome.			

	DISCUSSION: Economics provides criteria to guide decisions aimed at optimising the net benefits from the use of scarce resources. Assessing the benefits of disease mitigation is no exception. However, the technical complexity of mitigation means that economic evaluation is not straightforward because of the technical relationship of surveillance to intervention. We argue that analysis of the magnitudes and distribution of benefits and costs for any given strategy, and hence the outcome in net terms, requires that mitigation is considered in three conceptually distinct stages. In Stage I, 'sustainment', the mitigation objective is to sustain a free or acceptable status by preventing an increase of a pathogen or eliminating it when it occurs. The role of surveillance is to document that the pathogen remains below a defined threshold, giving early warning of an increase in incidence or other significant changes in risk, and enabling early response. If a pathogen is not contained, the situation needs to be assessed as Stage II, 'investigation'. Here, surveillance obtains critical epidemiological information to decide on the appropriate intervention strategy to reduce or eradicate a disease in Stage III, 'implementation'. Stage III surveillance informs the choice, timing, and scale of interventions and documents the progress of interventions directed at prevalence reduction in the population. SUMMARY: This article originates from a research project to develop a conceptual framework and practical tool for the economic evaluation of surveillance. Exploring the technical relationship between mitigation as a source of economic value and surveillance and intervention as sources of economic cost is crucial. A framework linking the key technical relationships is proposed. Three conceptually distinct stages of mitigation are identified. Avian influenza, salmonella, and foot and mouth disease are presented to illustrate the framework.			