Project no. 310806

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

KBBE
Collaborative Project
FP7-KBBE-2012-6

http://www.fp7-risksur.eu

Deliverable No. 3.10

Review of epidemiological methods used for surveillance systems certifying freedom from disease

WP 3 – WP Demonstration of freedom from disease

Authors: Schauer, Birgit (FLI)
Staubach, Christoph (FLI)
Vergne, Timothee (RVC)
Comin, Arianna (SVA)
Fröhlich, Andreas (FLI)
Höreth-Böntgen, Detlef (FLI)
Stärk, Katharina (Safoso)
Cameron, Angus (AusVet)
Conraths, Franz (FLI)

Lead participant: FLI

Submission date: October 2013
Type: Deliverable

TABLE OF CONTENTS

1 ABSTRACT

Demonstration of disease freedom is a prerequisite for safe trade of animals and animal-derived products with both domestic and international trade partners. Efficient surveillance designs have the potential to reduce the cost of surveillance, whilst preserving the power of confirmation of freedom. The aim of this review was to evaluate recent progress in the development of surveillance methods to demonstrate freedom from disease, in order to determine if and where conventional strategies can be replaced by novel statistical methods. Data related to demonstrating disease freedom in animal populations were extracted from 132 identified publications. This report includes a descriptive analysis and an overview of the development and interrelation of methods relating to demonstrating disease freedom including sample size calculations, scenario tree models and simulation models. The number of identified publications continuously increased since 1995 with more frequent application of risk-based methods after 2005. Forty-one, 29 and 28 publications were identified for the methodological groups 'sample size calculation', 'scenario tree model' and 'simulation model', respectively. Within each group, patterns existed regarding methodological issues, imperfection and uncertainty of test characteristics, software application and use of risk information. It is discussed, which methodologies are well established (e.g. two-stage sampling in clustered populations), require further assessment (e.g. approximate versus exact methods to calculate sample size; value of historical information) or would benefit from clearer guidelines (e.g. imperfect and uncertain test characteristics; expert opinion; epidemiological models). Risk-based methods appear to be particularly useful when demonstrating disease freedom as the confidence from previous surveys and targeted selection of higher risk strata are highly suitable to increase the efficiency of a surveillance programme. It is hypothesized that revision of international standards to guide decision makers, a well-established decision making framework, better analytical tools (i.e. exact methods to calculate sample size; estimation of value of historic info) and comprehensive software would promote the application of risk-based methodologies in the future and thus offer opportunities to decrease surveillance costs.

Keywords: Surveillance, disease freedom, sample size, scenario tree model, simulation model, sensitivity, animal health.

2 ABBREVIATIONS

AI	Avian influenza
ASF	African Swine Fever
BHV-1	Bovine Herpes Virus 1
BSE	Bovine Spongiform Encephalopathy
BT	Bluetongue disease
CSF	Classical Swine Fever
EM	Echinococcus multilocularis
EU	European Union
FLI	Friedrich-Loeffler-Institut
FMD	Foot and Mouth Disease
G1	1st generation of output-based standards: See Cameron (2012)
G2	2nd generation of output-based standards: See Cameron (2012)
MSR	Multistage random sampling
n	Number of publications
ND	Newcastle disease
OIE	World Organization for Animal Health
P*	Design prevalence
P^*_A	Animal or within herd design prevalence
P^*_H	Herd or between herd design prevalence
PRRS	Porcine Reproductive and Respiratory Syndrome
RVC	Royal Veterinary College
Se	Sensitivity
Sp	Specificity
SPS	Sanitary and Phytosanitary
SRS	Simple random sampling
STM	Scenario tree method
TB	Tuberculosis
VHS	Viral hemorrhagic septicemia
WP	Work package
WTO	World Trade Organization

3 INTRODUCTION

Freedom from infection is identified as 'the absence of a pathogenic agent in the country, zone or compartment' (OIE, 2011). The objective to demonstrate 'freedom from infection' can however be ambitious as the surveillance method, the diagnostic test or the vaccination status of the animals may not allow detecting infected animals. For example, passive clinical farmer reporting and serological testing would not detect infected animals during the incubation period and during the pre-seroconversion window period, respectively, and the use of non-marker vaccines would not allow to distinguish between infected and vaccinated animals (Kitching, 2003). Therefore, it was decided to use the more general term 'freedom from disease' throughout the report to cover both, freedom from disease and infection. In a wider sense, this term shall also cover non-infectious or non-pathogenic substances such as antibiotic residuals (hazard freedom).

Obtaining the status of freedom from disease can carry important implications not just for trade, but also for non-trade purposes such as to improve public health, to decide when to stop an eradication programme and to eliminate production losses and control costs due to endemic disease. Thus, it has relevance for a country, a zone as well as for individual herds (certification programmes). However, due to imperfect measurement methods (i.e. test specificity <100%) and the impracticality of testing every animal in the population, it is not possible to prove absence of disease with absolute certainty. Instead, demonstration of freedom from disease involves providing sufficient evidence to show that if a particular pathogen is present, it is present in less than a specified proportion of the population (design prevalence, P*) at a given level of statistical confidence. Hence, rather than aiming to document absolute freedom, the aim is to estimate the "probability of freedom from disease" and its opposite, the "probability of disease" despite having negative test results (Schuppers, et al., 2012). Applying this probabilistic approach allows considering accumulative evidence (Cameron, 2012), such as taking results from different surveillance activities (structured and non-structured) and from previous surveys into account. However, this also means that applied statistical methods, determination of risk-pathways and estimation of risk ratios need to be scientifically sound.

Traditionally, input-based standards have been used, which prescribe the surveillance activities to be carried out (i.e. sampling strategy, sample size, choice of test and frequency of testing), assuming that the population properties of herds are homogeneously distributed. However, herds vary in many factors: The most relevant ones are herd size, risk of infection and spread as well as the likelihood of pathogen detection. Therefore, this method has proven to be inefficient as it wastes resources in large and in low-risk herds. Moreover, it is ineffective as it provides only insufficient evidence in small and high-risk herds (De Massis, et al., 2005, Cameron, 2012). Although input-based standards are simple to implement and audit, there is an increasing tendency to set output-based standards, which prescribe what surveillance has to achieve. Output-based standards grant the flexibility required to find the most effective surveillance approach for the specific population under surveillance. However, there is a need to a) define clear guidelines how to achieve the expected outputs and to document the applied decisions and procedures, b) update / develop tools to allow application of more complex methodologies and c) create a framework for validating the surveillance system so as to ensure standardized and reliable implementation of output-based standards (Willeberg, et al., 2012).

Risk-based surveillance methods are particularly useful for demonstrating freedom from disease (Cameron, 2012, Oidtmann, B., et al., 2013). First, targeted selection of high risk strata is particularly efficient when a disease is

rare as it increases the likelihood of detection. Secondly, the value of historical information is higher when the aim is to demonstrate disease freedom than for endemic diseases, as the impact of disease spread or effectiveness of control measures plays no role if the disease is truly absent and no new introduction has occurred. The definition of risk in the context of disease freedom however is still under debate: The standard definition of risk in the field of risk analysis for animals is 'the chance of encountering some form of harm, loss or damage' (OIE, 2004), which implies that both likelihood (contracting disease) and consequences (spreading disease) should be considered. Cameron (2012) argued for demonstrating disease freedom the question is whether the population is infected, rather than whether it becomes infected. Based on this argument, the authors suggest that the only focus should be on the likelihood of infection when selecting high-risk strata. This recent discussion indicates the need to agree on standard terminologies related to risk-based surveillance, which may differ between surveillance objectives (early detection, endemic disease, disease freedom).

Due to trade globalization, it would in the long term also be desirable to reach agreements at the international level on how to define surveillance requirements to demonstrate disease freedom in a consistent and transparent way. For example, the handling of different hazards within the European Union (EU) does not follow a consistent strategy (Reist, et al., 2012), as some programmes still rely on input-based standards, while others have adopted output-based standards, sometimes also promoting risk-based surveillance methods. It would be useful to define requirements for demonstrating freedom from different hazards consistently and in a transparent way, outlining how the biology of the disease and other factors influence the decision on selected design criteria. Similarly, increased consistency between organizations / regions may contribute to a more uniform approach for trade regulation. For example, the World Organization for Animal Health (OIE) was designated in the Sanitary and Phytosanitary (SPS) Agreement of the World Trade Organization (WTO) as the standard-setting organization for animal health, so that any differences in regulations set by the EU should exceed the minimum requirements by the OIE. However, this may not always be the case as demonstrated by Schuppers, et al. (2011). The requirements set by the OIE and EU for detecting Trichinella spp. infections differ with respect to testing method (serology versus meat inspection), P* (0.02/0.01% and 0.2/0.5%1 versus 0.0001%) and the level at which disease freedom can be testified (country only versus country and herd), so that it is not straightforward to assess the comparability. Moreover, for comparability of surveillance systems, it would be desirable to provide clear guidelines based on peer review by experts which criteria to consider when defining fundamental parameters such as P* and propose standard values for each disease of international interest. The expectation of sound justifications that need to undergo independent review would oblige trading partners to base parameters on scientific rather than on political or economic grounds. For those diseases, for which no international standards exist, the general guidelines could help trading partners to select suitable criteria.

Standardization regarding both the implementation and validation of output-based standards as well as the definition of national and international requirements should account for new methodological developments, which could lead to a more efficient use of existing resources. Recent reviews have outlined new developments regarding the design of surveillance systems to demonstrate freedom from disease (Dufour, et al., 2001, Stärk, et al., 2006, Cameron, 2012, Oidtmann, B., et al., 2013). For example, new methods have been developed accounting for imperfect diagnostic tests and finite population sizes (Cameron, et al., 1998a) as well as clustering

[1] The OIE specified different design prevalences for the sow and finisher pig population during the first 5 years (0.02% and 0.01%, respectively) and thereafter (0.2% and 0.5%, respectively).

and heterogeneity in P* (Audigé, et al., 1999). Furthermore, multistage (Cameron, et al., 1998b) and risk-based (i.e. targeted) sampling (Dufour, et al., 2001), pooling of samples (Christensen, et al., 2000) as well as using information derived from multiple data sources and historical surveillance results (Martin, et al., 2007b) may allow reducing the number of samples or tests required. Such methodological approaches may have the potential to increase the accuracy and cost-effectiveness of surveys to demonstrate freedom from disease, but also increase the complexity of methods and therefore make it more difficult to evaluate the adequacy of the system.

The aim of this work was to review methodological developments for surveillance systems to demonstrate freedom from disease and identify issues that are well established or may require further assessment. We operate on the assumption that future methods need to preserve the power of confirmation of freedom and thus have the potential to fulfill the legal requirements to the same extent as the established conventional methods.

4 MATERIALS AND METHODS

The initial search for publications was done by colleagues from the Royal Veterinary College (RVC) to ensure a homogeneous search for reviews between the three work packages 2-4. Their approach was based on an automated search of peer-reviewed literature of the 'CabAbstract' and 'Scopus' databases. These databases have been chosen as they were thought to cover more than 91% of journals relevant for veterinary studies (Grindlay, et al., 2012). Title, abstract and keywords were included in the search. This search was limited to studies published in English in the last 20 years (1993 to 2012). The search was conducted in two stages:

First, a general theme was defined for all three WPs, targeting publications dealing with surveillance of animal health based on the following search terms:

surveillance OR monitor* OR survey* OR sampling
AND
animal* OR livestock OR veterinary* OR fish* OR wildlife OR "food system*" OR herd* OR farm* OR cattle OR cow* OR bovine OR ruminant* OR pig* OR porcine OR swine OR sheep OR goat* OR poultry OR bird* OR avian OR horse* OR equine OR equid* OR cat* OR dog*
AND
disease* OR health OR infection* OR outbreak

Next, specific searches were carried out for each work package (WP). For WP3, the search terms 'freedom' OR 'negative predictive value' OR 'sensitivity' OR 'absence of disease' were used. The sensitivity of these search terms was evaluated by cross-checking the resulting publications with a list of publications considered relevant by the FLI group. The search was subsequently performed without further modifications on 21st January 2013.

Titles and abstracts of resulting publications were screened for eligibility based on exclusion criteria 1-13 (Table 1). For all remaining publications, full texts were obtained and forwarded to the FLI group. One researcher at the FLI carried out all of the subsequent steps described. Full texts were screened to identify publications to be included based on exclusion criteria 1-15. Conference proceedings (e.g. ICAHS proceedings published in a special edition of the journal 'Épidémiologie et Santé Animale') were excluded as they were not considered to provide enough details on methodologies and results to allow a detailed comparison with full text publications. The search of 'CabAbstract' and 'Scopus' databases and subsequent steps were repeated on 8th July 2013 to update the database with articles published in the meantime.

A systematic search within the 'cited by lists' of each selected article was performed in the database in 'Web of Knowledge', using the words 'free*' OR 'absence' to identify additional relevant publications that may have been missed in the initial search. 'Sensitivity' was not included as a search term in this search anymore as it was considered too broad to be included at that stage of verification. Furthermore, relevant publications from the reference list of included publications as well as methodological key publications published before 1993 were included afterwards (non-systematic search). Full texts of these publications were also downloaded and screened for eligibility based on the same exclusion criteria.

It was assessed how many of the publications in the final publication list were identified when searching title/abstracts/key words by one search term at a time. The search term used for the initial search ('freedom'; 'negative predictive value'; 'sensitivity'; 'absence of disease') and the two modified search terms ('free'; 'absence') were assessed. The percentage of publications identified by the respective search terms indicates how sensitive this search term was in identifying relevant publications. The following variables were extracted from the publications, where applicable:

- General information: author, year, country, species group, disease, objective as stated in the article, level (herd, regional, national, international), type of data (real, simulated, both);
- Methodological categorizations (details see below): primary objective of paper, type of method, application of risk-based methods;
- Surveillance details: surveillance components, sampling method, time units, P*, confidence level, sensitivity, specificity, risk factors applied;
- Methodological details: outcome of interest, calculation method, uncertainty in test characteristics (and other parameters), category nodes (for scenario tree model), sensitivity analysis.

Publications were grouped based on their primary objective ('application of a method', 'comparison of methods', 'new method', 'description of a programme', 'input parameters', 'evaluation of methods', 'simulation model', 'opinion paper', 'review paper). If more than one objective applied, the main objective was defined. Furthermore, publications were grouped according to the applied method(s) ('sample size calculation', 'scenario tree model', 'simulation model', 'other methods', 'not methodological'). Publications with the objectives 'description of a programme', 'evaluation of methods', 'opinion paper' and 'review' may not necessarily appear under one of the methodological groups. Lastly, publications were grouped as to whether they applied risk-based strategies. For that purpose, risk-based strategies were defined to be applied if a publication targeted subpopulations that have a higher risk of infection than the population as a whole (targeted surveillance) (Salman, et al., 2003) or took into account the level of confidence from previous surveys (risk-based sample size calculation) (Cannon, 2001, Stärk, et al., 2006). An article was not grouped as risk-based if it only targeted one risk group without further risk differentiation within this group.

Extracted data were entered into a customized Access database. The Fisher exact test was used to test for statistical differences in the application of risk-based methods between the species groups 'Cattle', 'Small ruminants', 'Pigs', 'Poultry' and 'Wildlife'.

Analyses and graphical representations of the data were performed using R version 3.0.1 (R Development Core Team: www.r-project.org).

5 RESULTS

5.1 Search results

For the review, 131 of the 162 screened publications were included, of which 19 dealt with new methodologies and 14 were review or opinion papers (Figure 1). The percentage of publications identified through the initial database search, the search of 'cited by lists' and the reference lists was 57.3%, 34.4% and 8.4%, respectively. The most sensitive single search terms used for this review were 'freedom' (47.2% of publications; compared to the modified search term 'free': 59.8%) and 'sensitivity' (37.0%). 'Absence of disease' (3.1%; compared to the modified search term 'absence': 9.4%) and 'negative predictive value' (0.8%) appeared in the title/abstract/keywords of only few publications, most of which were covered by other search terms as well. Most exclusions were based on the criteria 'Conference proceeding' (n = 11), followed by the criteria 'Control measures' (n = 6) and 'Endemic' (n = 4) (Table 1).

5.2 Descriptive results

Table 2 shows the grouping of all 131 publications based on primary purpose category and application of risk-based methods. Risk-based methods were applied in 66.4% of the publications.

Of the 105 publications that specified the area under investigation, 63.8% referred to a country/region in Europe, 16.2% in Australia, 15.2% in the Americas and 4.8% in Asia (Figure 2). The highest number of publications referred to Switzerland (n = 20), followed by Australia (n = 11) and Denmark (n = 11). More than 90% of the publications referring to Australia and Denmark included application of risk-based methods (Switzerland: 60%). Figure 3 illustrates a continuous increase in the number of publications on disease freedom from 1995 to 2012. After 2005, the annual median percentage of publications including risk-based methods was higher than 75%. Of the 115 publications referring to a specific level, disease freedom was mostly applied at the national (53.0%) or multinational level (3.5%), followed by the regional (20.9%) and herd level (22.6%) (Figure 4).

Figure 5 shows that most publications referred to ruminants (n = 57) and pigs (n = 21), followed by aquacultural species (fish, crustaceae; n = 9), poultry (n = 8) and wildlife (n = 8). There was no statistical difference in the application of risk-based methods between species. Common diseases covered included Paratuberculosis (n = 16), TB, FMD and AI (8 publication each), Trichinella (n = 7), PRRS (n = 5), brucellosis (n = 4), VHS, Scrapie, EM, BSE and BT (3 publications each), ND, CSF, BHV-1 and ASF (2 publication each) (data not shown). Apart from brucellosis, risk-based methods have been applied to every disease, for which the review has identified at least two publications.

5.3 Methodologies

Important methods for demonstrating disease freedom include the calculation of the sample size necessary to meet certain survey requirements (section 5.3.1) and the epidemiological models used to estimate the probability that a population is free from a specified disease (section 5.3.2).

5.3.1 Sample size calculation

The review identified 41 publications reporting the calculation of sample sizes (Table 3), of which nine applied simple random sampling, 11 multi-stage random sampling and 21 risk-based sampling (including five publications calculating sample size via a scenario tree model).

Sampling is usually performed without replacement, apart from special situations such as sampling of rodents in live traps, which are subsequently released and recapturing is possible (Seber, 1986), even if the probability may be low. When sampling without replacement, the hypergeometric distribution is mathematically the exact method to use (Feller, 1968). However, this formula is impossible to solve exactly for large population sizes. Therefore, different kinds of binomial approximations were proposed: Cannon (1982) illustrated the use of the simple binomial approximation and Cameron (1998a) developed the modified binomial approximation to the hypergeometric formula. Cameron et al. (1998a) compared these two approximations and showed that the modified binomial approximation formula is more suitable for small population sizes than the simple binomial formula. A generally accepted recommendation is that the binomial approximation is only appropriate to use when the sampling fraction is less than 10% of the population size (Dohoo, et al., 2003, Fosgate, 2009, Williams, et al., 2009a).

Cannon (1982) applied the simple binomial approximation for simple random sampling (SRS), which assumes a homogeneous population. Cameron (1998b) applied the modified binomial approximation formula also to multi-stage random (MSR) designs, as described by Levy (1991), in order to address the issue of disease clustering. By calculating sample sizes for separate sampling units (e.g. herds, animals), a different P* can be assigned (e.g. between herd design prevalence, P*H, for the selection of herds; animal or within herd design prevalence, P*A, for the selection of animals within the herd). Hence, sample sizes can be calculated for each stage accounting for imperfect test characteristics. The selection of individual units from within a risk stratum is usually done based on SRS to ensure that the sample from each stratum is representative (Cameron, 2012). As the herd has been increasingly recognized as the unit of interest in international legislation concerning domestic animals (Stärk, et al., 2000, OIE, 2011), MRS has become the standard approach in livestock populations that are clustered into herds, flocks or other groups although it results in larger sample sizes than SRS as demonstrated by Stärk (2000). All but one (Rodríguez, et al., 2012) of nine reviewed publications applying SRS actually dealt only with one level (i.e. herd as unit of interest, census of herds). Rodríguez (2012) justified the application of SRS given that, in their study, the domestic ruminant population was kept on an island.

Early sample size calculations were based on the assumption of a perfect diagnostic test (Cannon, 1982). Therefore, a major development has been the inclusion of test characteristics (i.e. sensitivity, Se and specificity, Sp) as a perfect test hardly exists (Fosgate, 2009). The modified binomial approximation formula by Cameron et al. (1998a) allows incorporating test characteristics. Cannon (2001) proposed two new approximations to calculate sample size based on the maximum number of diseased animals accounting for imperfect Se and Sp or imperfect Se at perfect Sp. Only one article has been identified in the review (Hadorn, et al., 2002) that applied this calculation method, making use of the advantage that the required level of confidence can be adjusted. Of those 35 publications that specified test characteristics, 16 accounted for imperfect Se and Sp, whilst six assumed perfect test Se and Sp and eleven assumed perfect Sp (allowing for imperfect Se).

If a perfect test is assumed, the formula of Cannon (1982) can be solved using a simple hand calculator or the freeware application WinEpiscope (www.clive.ed.ac.uk/winepiscope/) developed by Thrusfield et al. (2001).

WinEpiscope includes a sample size adjustment factor to correct for small population sizes (Fosgate, 2009). Three publications used WinEpiscope. The modified binomial approximation formula by Cameron et al. (1998a) including the aspect of imperfect test characteristics and the option to incorporate multi-stage random sampling is implemented in the freeware application FreeCalc (part of SurveyToolbox: available at http://epitools.ausvet.com.au/content.php?page=SurveyToolbox). The latter can also be used for the analysis of survey results (Audigé et al., 1999, Baldock, 1998, Deliberto et al., 2009, Gustafson et al., 2010, McFadden et al., 2009, Mur et al., 2012, Paton et al., 2006, Reber et al., 2012, Ryan et al., 2012, Van Schaik et al., 2003). FreeCalc has since been the most commonly used software when calculating sample size to demonstrate freedom from disease (20 out of 38 articles published since its development).

Johnson (2003) extended the work of previous authors (Cameron, et al., 1998b, a, Cannon, 2001), accounting for uncertainty in test Se and Sp using a Bayesian approach in a single cluster setting. They argued that test characteristics are rarely known with certainty. When comparing sample sizes with those estimated by Cameron (1998a), Johnson (2003) demonstrated that sample sizes were comparable if Se and Sp were known with high certainty, but were higher with increasing uncertainty. Branscum, et al. (2006) extended the approach of Johnson (2003) to a multiple cluster setting, also allowing for clusters with zero prevalence, variability in prevalence among clusters and uncertainty in the within cluster prevalence. This approach was subsequently applied by Kostoulas (2012).

Risk-based sampling

Risk-based sampling as defined in this report can involve targeting subpopulations that have a higher risk of infection than the population as a whole (targeted sampling) or accounting for the level of confidence from previous surveys (risk-based sampling). Four comprehensive reviews provide a good outline of the historical development of risk-based surveillance (Dufour, et al., 2001, Stärk, et al., 2006, Cameron, 2012, Oidtmann, B., et al., 2013).

Hadorn, et al. (2002) developed a method based on the work of Cannon (2001) and Audigé (1999) that allows adjusting the required level of confidence for repeated surveys and thus reducing the sample size. Aspects that may affect the value of historical information include the probability of introduction (Hadorn, et al., 2002) (e.g. through legal or illegal trade, wildlife, vectors) and the consequences of residual undetected infection (Knopf, et al., 2007). Schwermer (2009) developed this method further for non-highly contagious diseases by incorporating the time dependent reduction process of the value of historically gained testing information by multiplying the confidence acquired in the past with the proportion of surviving animals. Furthermore, this method of accounting for the ageing of historical information reduced the degree of undulation of the confidence level compared to the method by Hadorn (2002).

Several authors demonstrated reductions in sample sizes (Hadorn, et al., 2002, Knopf, et al., 2007, Schwermer, et al., 2009, Williams, et al., 2009a, b, Schuppers, et al., 2010a, Blickenstorfer, et al., 2011) and survey cost (Blickenstorfer, et al., 2011, Reist, et al., 2012) when risk-based methods were applied. The magnitude of reductions in sample sizes depend on various factors and can thus not be generalized. However, risk-based surveillance taking historic records into account was demonstrated to achieve reductions in sample sizes by up to 80% (Hadorn, et al., 2002, Knopf, et al., 2007) and reductions in survey costs by over 6 million Euro (1.13% of total cost) (Reist, et al., 2012). Examples for applying targeted sampling only (without taking historical

information into account) also showed reductions in sample sizes up to 80% (Williams, et al., 2009b) and a reduction in survey cost by 40% (Blickenstorfer, et al., 2011).

5.3.2 Epidemiological models

Epidemiological models are used to represent the logical or mathematical processes in the epidemiology of a disease and associated factors (Willeberg, et al., 2011). After specifying model input parameters obtained from real data, published literature review or expert opinion (Martin, et al., 2007b), the course of a past epidemic is assessed (deterministic model) or its future course projected (predictive model). Such models may be used for a wide range of purposes, such as to inform national, technical and administrative needs, scientific questions and international, political and trade-related decisions (Willeberg, et al., 2011). Vose (1997) was the earliest author identified in this review who described the two main modeling techniques based on the example of estimating risks associated with the import of animals and animal products, i.e. the scenario-pathway method and simulation modeling.

5.3.2.1 Scenario tree models (STM)

Martin (2007b) first applied scenario tree modeling (STM) for the aim to demonstrate disease freedom assuming a test with perfect specificity. The authors justified this assumption by the fact that commonly applied multi-stage ordered serial testing has perfect specificity and is therefore eligible to confirm potentially false-positive result. Based on this assumption, it is further assumed that all final results (i.e. after completion of any diagnostic follow-up) from the surveillance system are negative.

STM has been applied by 27 authors besides Vose (1997) and Martin (2007b) (Table 4). Most authors applied STM at the country level and only one author at the herd level. Of the 26 publications, for which details were specified, 21 accounted for historical information (ranging from 6 months to up to 16 years), 13 incorporated information from more than one surveillance component and 20 distinguished between different risk groups (see column 'Risk factors'). Different scenarios were compared by 22 publications (data not shown). All but six publications followed Martin, et al. (2007b) in their assumption of a perfect test or surveillance system specificity. @Risk, developed by Palisade Corporation, was the software used by all but two others (data not shown). One group used PopTools developed by Greg Hood, CSIRO, AUS (More, et al., 2009) and another one used ModelRisk developed by Vose Software (Welby, et al., 2012). These two alternative products and @Risk are all Microsoft Excel add-ins.

2007b) emphasized that their approach was quantitative, which should have a stochastic component, although a qualitative or semi-quantitative approach may be more feasible in data-poor environments. Twenty-one publications estimated quantitative input parameters partly by means of expert opinion, but only three provide a detailed description on how expert opinion was gathered. All but four publications allowed for uncertainty in test characteristics and often also other parameters such as risk ratios (stochastic model). Eighteen publications included a sensitivity analysis to identify influential parameters.

The scenario trees can become fairly complex in that up to eight or nine nodes were assigned for the scenario tree structure displayed in the publication (Table 4), sometimes in combination with multiple trees being constructed for different surveillance components (Hadorn, et al., 2008b, Frössling, et al., 2009, Knight-Jones, et al., 2010, Wahlström, et al., 2010, Frössling, et al., 2013), species (Wahlström, et al., 2011) or production systems (Martin,

2008, Hadorn, et al., 2009, Alba, et al., 2010). Risk category nodes commonly related to potentially risky contacts (n = 14), followed by management factors (n = 12), animal factors (n = 12) and location (n = 6) (Table 5). This complexity was criticized by Hood, et al. (2009) who illustrated how a Bayesian belief network (BBN), a method already used for modeling ecological systems, could be applied to deal with complex systems in a better way. Compared to STM, BBN provide a compact diagram of the structure used, simplifies calculations and extends the range of software that can be used. This method has not been applied in any other published work for animal disease surveillance and only by one group in a publication on biosecurity surveillance (Whittle, et al., 2013).

5.3.2.2 Simulation models

Table 6 summarizes details of the 28 identified publications that used simulation models. A similar number of simulation models were applied at the country (n = 10), regional (n = 7) and herd level (n = 11). Models were most commonly used to calculate the surveillance system sensitivity (G1) (n = 13) or probability of freedom (G2) (n = 12), but also to assess the cost of surveillance (n = 7), calculate sample sizes (n = 2), estimate epidemiological parameters such as the intra-correlation coefficient or prevalence (n = 3), estimate the time needed to confirm disease freedom (n = 2) and simulate the spread of disease (n = 1). A longitudinal data set or multiple time periods (2S) was considered by 15 publications, with longitudinal time periods ranging from four months to 25 years. Five publications included multiple surveillance components, whilst the remainder only focused on one component, usually active surveillance (n = 21). All publications that specified test characteristics accounted for imperfect test sensitivity, whilst seven assumed perfect surveillance specificity.

Twenty publications applied a stochastic model allowing for uncertainty in input parameters, 17 of which considered uncertainty in test characteristics (Table 6). A variety of software was applied including simple spreadsheet models in Excel (n = 2), commercial modeling software (mainly @Risk, n = 10) freeware products (R, n = 5) and own models that were made publicly available (JohneSSim and BDFree). Only nine publications accounted for risk factors, most commonly previous disease status (n = 5).

6 DISCUSSION

In the last decade, a considerable shift has occurred in the design of surveillance systems through the increased application of output-based standards and propagation of risk-based methods. But it appears that there is a need for harmonization and standardized guidelines, which would make it easier to design, validate and compare different surveillance systems (Vanderstichel, et al., 2013). This is the first report in recent years reviewing methodological developments for surveillance systems to demonstrate freedom from disease. The aim of this report was to review methodological developments for surveillance systems to demonstrate freedom from disease and identify issues that are well established or may require further assessment.

The first part descriptively identified patterns of the number of identified publications and whether they applied risk-based methods. Risk-based methods were increasingly applied after 2005. This trend was possibly due to the relatively new evidence that past surveillance results can be used to strengthen confidence in disease freedom (Hadorn, et al., 2002), the publication of reviews (Dufour, et al., 2001, Stärk, et al., 2006) and the introduction of the scenario tree model as a relatively straightforward method, which allows incorporating risk category nodes (Martin, et al., 2007b). Since then, risk-based methodologies have become increasingly accepted, at least within the scientific community.

The number of publications referring to a particular country or region showed interesting publication patterns. Most publications referred to an application in Switzerland (n = 20), followed by Australia (n = 11) and Denmark (n = 11). This result might be explained by a mixture of factors such as strong exporting background (Denmark, Australia) and freedom from a comparatively large number of diseases, but also by the location of key institutes and key people for the development and application of epidemiological methods propagating risk-based approaches.

The predominance of publications referring to an application within a European country (63.5% and 43.8%, respectively, when including all of Europe or EU countries only) illustrates the importance of demonstrating disease freedom within this region, possibly combined with a strong financial support for research. Both, extra-EU exports and intra-community trade (measured in terms of dispatches) of live animals and animal products, play an important role in this economic and political union. Hence, the single market comprised of separate member states makes surveillance to demonstrate disease freedom very important. The predominance of publications relating to the national or multinational level indicates the international importance of disease freedom, as it is a prerequisite to maintain, open new or re-open (after an outbreak) export markets. However, also from a regional and herd level perspective, demonstrating disease freedom offers benefits as trade permits can be issued for a specific region or certified herds within a region.

Although risk–based methods have been described in the literature, they are often only science-based, while national or regional programmes need to follow international regulations to take advantage of international trade benefits. Legal frameworks are relatively slow in responding to new evidence as the validity of advancements needs to be well-established and the process in getting legal documents drafted and approved is rather slow. Therefore, it is important to develop sound scientific evidence, which can support policy makers in decision-making. Another potentially limiting factor is insufficient data/evidence to inform the choice of suitable risk factors and quantification of risk ratios (Oidtmann, B., et al., 2013). This is especially a problem when a disease has not been present in the country for a long time, so that estimates can only be derived from other, e.g. neighboring countries, from the literature or expert opinion. Furthermore, risk-based methods make it difficult to compare surveillance systems between countries (Stärk, et al., 2006) and to extrapolate the results to the general population if positive samples are detected (Williams, et al., 2009b). Finally, it is hypothesized that decision makers may be unsure of the correct implementation of such methods since officially approved guidelines, manuals or electronic tools are missing. In terms of the utility of risk-based surveillance, their economic benefits may be limited by the fact that substantial efforts are required to inform and update the designs. Formal economic evaluation would therefore be desirable, although it is anticipated that gains in the context of surveillance for freedom from disease will probably be sufficient to justify the investment.

The methodological part distinguished between three groups, which are a) methods to carry out sampling and calculate sample sizes, b) scenario tree models and c) simulation models. When calculating sample size, a major advancement over recent years has been the general acceptance that an infectious disease cannot be assumed to be homogeneously distributed throughout the population if animals are kept in groups (e.g. herd, flock). Although multi-stage random sampling results in a higher total number of samples (Stärk, et al., 2000), it addresses this cluster effect of herds. Consequently, over the last decade multi-stage random sampling has become a general legal requirement when animals are kept in herds.

Another major advancement has been the development of the modified binomial approximation to the hypergeometric formula (Cameron, et al., 1998a) instead of the simple binomial approximation (Cannon, 1982) as it is more suitable for small populations and allows incorporating test characteristics, even for multi-stage sampling. Other approximations may exist, but have not been applied by the identified publications. However, there is insufficient evidence how the approximation formulas compares to the exact calculation method. It should be explored whether complex issues such as the calculation of confidence intervals for the hypergeometric formula could be incorporated into a software, which has not yet been the case, at least not when Fosgate (2009) published his work.

Interestingly, uncertainty in test characteristics was rarely accounted for when calculating sample sizes (8/39), although Bayesian methods suitable for this purpose have been proposed at least a decade ago (Johnson, et al., 2003). In comparison, when scenario tree models or simulation models were applied, uncertainty in test characteristics were incorporated by 22/26 and 17/27 publications, respectively. This raises the question, whether uncertainty in test characteristics should be included or not (regardless of the technical limitations such as software) and to what extent it depends on epidemiological parameters (e.g. P*, confidence level), the disease of interest and the applied test.

Another issue that was handled differently between methods was how publications dealt with imperfect specificity. Perfect specificity is a common assumption when demonstrating disease freedom (Cannon, 2002, Martin, et al., 2007b, Cameron, 2012). The key reference (Martin, et al., 2007b) for scenario tree modeling worked on the assumption of perfect surveillance specificity, based on the hypothesis that disease freedom can be rejected as soon as a true positive sample is found. Therefore, serial testing is generally used to verify any positive results, so to increase specificity to the highest possible degree. Subsequently, more than 70% of publications applying STM assumed perfect specificity. In comparison, only 47.1% (n = 16) and 28.0% (n = 7) of publications assumed perfect specificity when calculating sample size and applying simulation modeling, respectively. Therefore, it is proposed to clearly illustrate the effect of serial testing and assuming perfect specificity on overall survey sensitivity for a range of scenarios and provide guidelines whether this practice is acceptable (or only under certain conditions).

Scenario tree modeling provides a relatively easy to use tool to determine the sensitivity of the surveillance system as a whole or of individual branches (e.g. components, species, production types etc.), calculate sample size and evaluate what-if scenarios. This method allows estimating surveillance sensitivity based on non-random data (not just from structured random surveys) and incorporating information from multiple data sources. By incorporating Bayesian methods, Martin (2007b) also demonstrated the possibility of accounting for historical information with this easy to use tool. Scenario tree modeling is valuable for demonstrative purposes as the target audience can easily follow the steps undertaken in the analysis as long as it is not too complex (Hood, et al., 2009).

Incorporating historic information to adjust the probability of freedom and thus reduce the required sample size of the upcoming survey (risk-based sample size calculation) has become a popular tool (e.g. applied in five out of 26 publications applying scenario tree models). New disease introductions (through legal or illegal trade, movement of wildlife, vectors, …) and residual infection may contribute to a loss in value of historical information. It would be beneficial to summarize relevant aspects to consider according to the country's specific situation in order to adequately estimate the loss in value of historical information.

Simulation models provide maximum flexibility. For example, whilst STM assumes two infection states (infected, not infected), simulation models allows accounting for different infection states (e.g. susceptible, non-susceptible, latent-infected, lowly infectious, highly infectious and clinical disease), various infection routes and different design prevalences. On the one hand, this flexibility can lead to an improved model, which allows assessing a wide range of scenarios. On the other hand, it may become difficult for the audience to assess the adequacy and validity of the model structure and input parameters. Guidance would be useful for policy makers when to use a STM or a simulation model and what standards need to be followed to document the decision making process when designing a model. Guidelines would also be valuable how to validate the model and related outputs so that also a less experienced target audience has the option to assess its adequacy.

One problem with both modeling methods is that they often require many input data, which are not known with certainty. Sensitivity analysis was commonly applied for scenario tree models (69.2%), but not for simulation modeling (28.6%) to assess whether the model was affected by the uncertainty and variation associated with input parameters. Given the difficulties in accurately defining input parameters, sensitivity analysis could be proposed as a general requirement to assess which parameters have a large impact on results and could thus potentially be targeted by future research (e.g. epidemiological studies, simulation studies, more intense literature review) (Fahrion, et al., 2011). Another weakness was that expert opinion was widely used to define input parameters, but often no details were provided on the process and number of experts included, thus not informing the reader about the validity of assumptions, although they have a high influence on model outputs. Hence, standard guidelines on how to obtain and validate expert opinion and other input parameters could be a useful advancement to ensure validity and thus acceptance of epidemiological models (Gustafson, et al., 2013).

There are some methodological limitations to the review article. First, the original database search identified 75 of the 131 included publications, whilst the search of the 'cited by lists' identified 46 publications. Hence, the latter search method may present an effective way of finding additional references and could possibly be recommended for future reviews if the related workload is manageable and the initial database search appears to have resulted in incomplete coverage. However, if many publications have been missed in the initial database search, bias may occur as certain streams of publications may be missed completely. The results presented here cannot be assumed to cover the complete range of potentially relevant publications. However, we are confident that relevant methodological publications were picked up through the search, as it included three different components, two systematic ones (initial database search and search of the 'cited by lists') and one non-systematic one (reference lists). Furthermore, scientists working in the field did not detect any important publications being missed. Secondly, at this stage only one researcher has screened full texts and extracted the information, so that this review cannot be termed a systematic review yet. However, prior to publication in a peer-reviewed journal, other FLI researchers will split the publications amongst themselves and extract the variables covered in this report. Inconsistencies in results shall be resolved based on the outcome of subsequent group discussions.

This review examined the three most common methodologies to demonstrate disease freedom and identified aspects that are well established, require further evaluation or would benefit from clearer guidelines. These aspects need to be further discussed with experts to verify and prioritize them according to their impact. Some additional methodologies referred to in publications included in the descriptive part were detected by this review, but were not used by any subsequent authors and did not clearly fit the context chosen for this report. But the usefulness of these and as well as potential other statistical methodologies, which may not have been documented yet in the context of demonstrating freedom from disease, will need to be further explored.

7 CONCLUSION

This review outlined major methodological developments and suggested areas that may benefit from further exploration. For example, the RISKSUR project could address the following gaps in knowledge identified in this review:

- Application of exact hypergeometric calculations to calculate sample size and confidence intervals;
- What aspects should be considered when estimating the value of historic information?
- Simulation of methodological aspects such as sample size calculation methods, serial testing and uncertainty of test characteristics over a continuous range of scenarios, which can inform the illustrate differences and thus inform the choice of appropriate methods;
- Development of guidelines and training materials on how to select appropriate methods depending on the surveillance objective and the country situation.

It is hypothesized that improved international standards to guide decision makers, a well-established decision making framework, better illustrated analytical tools and comprehensive software could promote the application of risk-based methodologies in the future. However, such guidelines and tools would need to be revised on a regular basis to take into account new developments.

8 ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7 / 2007-2013) under grant agreement n° 310806. We gratefully acknowledge funding provided through the Seventh Framework Programme of the European Union (F7-KBBE.2012.2.3-01; grant agreement no: 310806).

9 FIGURES AND TABLES

Table 1. Exclusion criteria for selection of publications in the review of methods to demonstrate freedom from disease.

No.	Exclusion criteria (short)	Description	n
1	Case report	The paper is a case report	0
2	Control measures	The paper focuses only on control measures (i.e. vaccination, stamping out...)	6
3	Descriptive paper	The paper is only descriptive (e.g. historical trend of a disease)	3
4	No freedom from disease	The paper is not relevant for demonstrating freedom from disease	3
5	Disease review	The paper is a review of a particular animal disease	0
6	Experimental infection	The paper involves experimental infections	0
7	Field survey	The paper is a field survey/report (incl. outbreak investigation)	0
8	Molecular characterization	The paper focuses on molecular characterization of pathogen	0
9	No animal disease	The paper does not focus on animal diseases or animal health	0
10	No surveillance	The paper does not focus on disease surveillance as defined for this project	1
11	Endemic	The disease of interest is endemic/enzootic in the study area, without focus on certifying disease freedom at the herd/flock level	4
12	Test evaluation	The paper deals with test evaluation and/or implementation of new diagnostic methods/tests	1
13	Vaccine evaluation	The paper deals with evaluation of vaccine efficacy	0
14	Proceeding	The paper is a conference proceeding published in a special edition of a scientific journal	11
15	Evaluation	The paper focuses on surveillance system evaluation	2
	Total		**31**

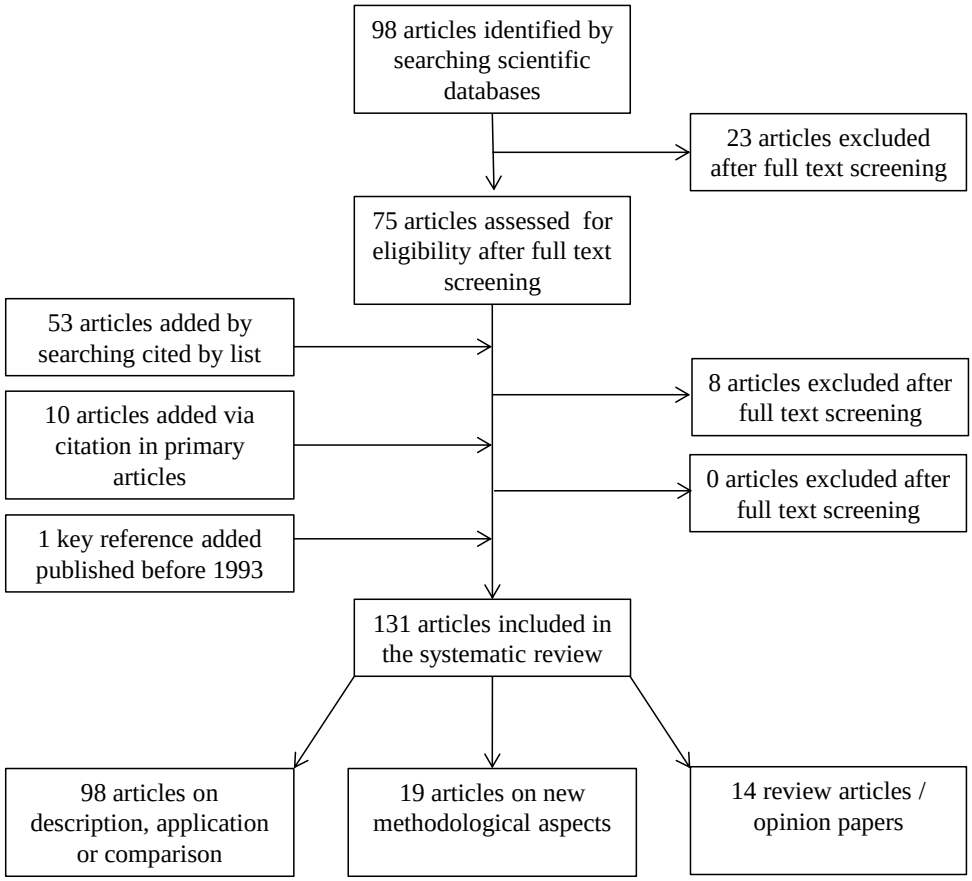

Figure 1. Flowchart illustrating the process of selecting peer reviewed publications (n = 131) included in the review of methods to demonstrate freedom from disease.

Table 2. List of publications included in the review of methods to demonstrate disease freedom grouped by primary objective and application of risk-based methods (n = 131).

Primary objective	Not risk-based		Risk-based		n
	n	References	n	References	
Application of a method	7	Thrusfield, et al., 2005, Knight-Jones, et al., 2010, Christensen, et al., 2011, Mur, et al., 2012, Murphy, et al., 2012, Ryan, et al., 2012, Trevennec, et al., 2012	13	Bradley, et al., 2005, Martin, et al., 2007a, Rolesu, et al., 2007, Martin, 2008, Deliberto, et al., 2009, Frössling, et al., 2009, Hadorn, et al., 2009, Wahlström, et al., 2010, Wahlström, et al., 2011, Goutard, et al., 2012, Rodríguez, et al., 2012, Calvo-Artavia, et al., 2013, Frössling, et al., 2013	20
Comparison of methods	8	Vose, 1997, Ziller, et al., 2002, Van Schaik, et al., 2003, Greiner, et al., 2005, Norby, et al., 2005, Su, et al., 2007, Hadorn, et al., 2008a, Häsler, et al., 2012	17	Good, et al., 2001, Hanson, et al., 2003, Weber, et al., 2004, Hadorn, et al., 2008b, Durand, et al., 2009, Schwermer, et al., 2009, Alba, et al., 2010, Efsa Panel on Biological Hazards, 2010, Schuppers, et al., 2010a, Welby, et al., 2010, Blickenstorfer, et al., 2011, Handel, et al., 2011, Willeberg, et al., 2011, Reber, et al., 2012, Reist, et al., 2012, Bessell, et al., 2013, Boklund, et al., 2013	25
New method	8	Cannon, 1982, Cameron, et al., 1998b, a, Audigé, et al., 1999, Cannon, 2001, Johnson, et al., 2003, Branscum, et al., 2006, Schuppers, et al., 2012	11	Schlosser, et al., 2001, Cannon, 2002, Hadorn, et al., 2002, Böhning, et al., 2006b, Martin, et al., 2007b, Hood, et al., 2009, Joly, et al., 2009, Williams, et al., 2009a, b, Martinez, et al., 2010, Kostoulas, et al., 2012	19
Description of a programme	7	Thornton, et al., 1995, Gohm, et al., 1999, Rautiainen, et al., 2001, Feliziani, et al., 2005, Cagienard, et al., 2006, Rawdon, et al., 2010, Plischuk, et al., 2011	21	McLaughlin, 1995, Garner, et al., 1997, Ellis, et al., 1998, Black, et al., 2001, Appleyard, et al., 2002, East, et al., 2004, Mintiens, et al., 2005, Corbellini, et al., 2006, Racloz, et al., 2006, Radunz, 2006, Alban, et al., 2008, O'Grady, et al., 2008, Paré, et al., 2008, Carlsson, et al., 2009, McDonald, et al., 2009, McFadden, et al., 2009, Schuppers, et al., 2010b, Sergeant, et al., 2011, Windsor, et al., 2011, Dukpa, et al., 2012, Learmount, et al., 2012	28
Input parameters	2	Peeler, et al., 2008, Sanchez-Vizcaino, et al., 2010a	5	Gustafson, et al., 2010, Sanchez-Vizcaino, et al., 2010b, VHSV Expert Panel Working Group, 2010, Oidtmann, B. C., et al., 2011, Willeberg, et al., 2012	7
Evaluation of methods	6	Thorburn, 1996, James, et al., 2002, Branscum, et al., 2005, De Massis, et al., 2005, Nérette, et al., 2008, Tavornpanich, et al., 2012	5	Ezzano, et al., 2005, Böhning, et al., 2006a, More, et al., 2009, Hernández-Jover, et al., 2011, Welby, et al., 2012	11
Simulation model[a]	3	Suess, et al., 2002, Heuer, et al., 2007, Ebel, et al., 2008	4	Sergeant, et al., 2002, Knopf, et al., 2007, Sergeant, et al., 2008, More, et al., 2013	7
Opinion paper	1	Baldock, 1998	3	Burr, 2007, Alban, et al., 2011, Gustafson, et al., 2013	4
Review paper	2	Christensen, 2001, Fosgate, 2009	8	Christensen, et al., 2000, Stärk, et al., 2000, Doherr, et al., 2001, Dufour, et al., 2001, Corbellini, et al., 2006, Paton, et al., 2006, Cameron, 2012, Peeler, 2012	10
Total	**44**		**89**		**131**

[a] Please note, that this grouping is based on the primary objective, so that the number here (n = 7) deviates from publications listed in Table 6 (n = 28).

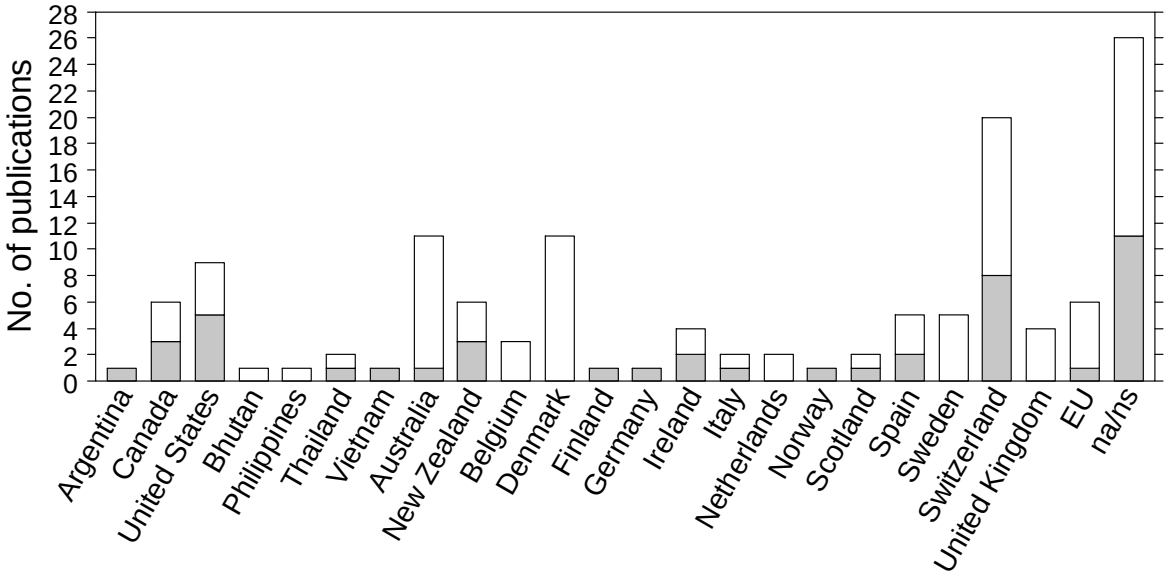

Figure 2. Number of publications identified by the review of methods to demonstrate disease freedom by country (*n* = 131) sorted by continent (Americas: n = 16; Asia: n = 5; Australia and Oceania: n = 18; Europe: n = 67) and stratified by application of risk-based methods (grey shading: Not risk-based; white shading: Risk-based). na/ns: not applicable/not specified.

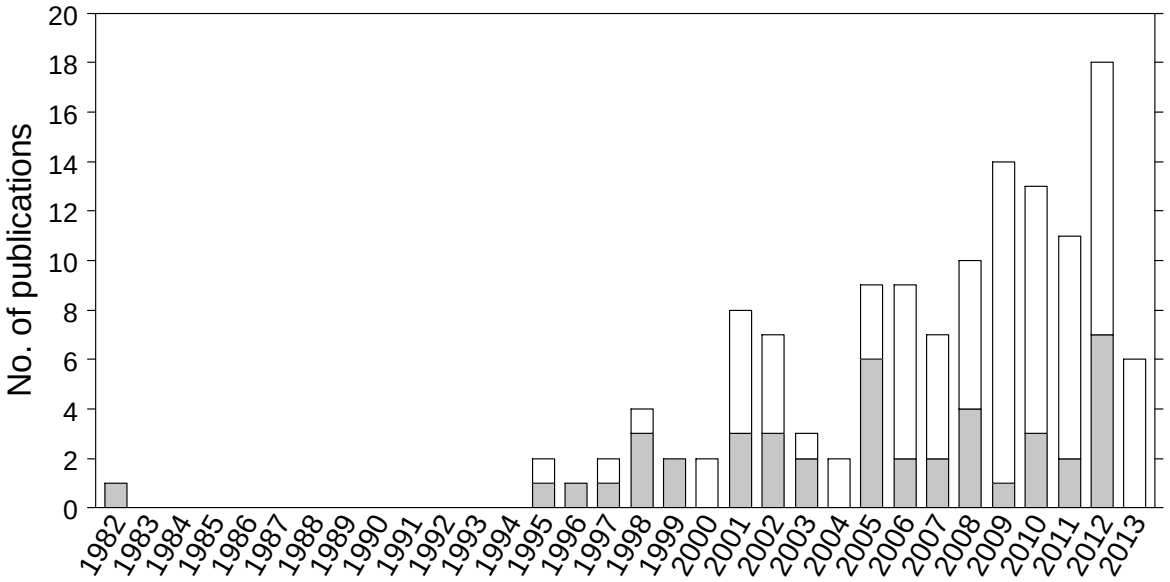

Figure 3. Number of publications identified by the review of methods to demonstrate disease freedom by year (*n* = 131), stratified by application of risk-based methods (grey shading: Not risk-based; white shading: Risk-based).

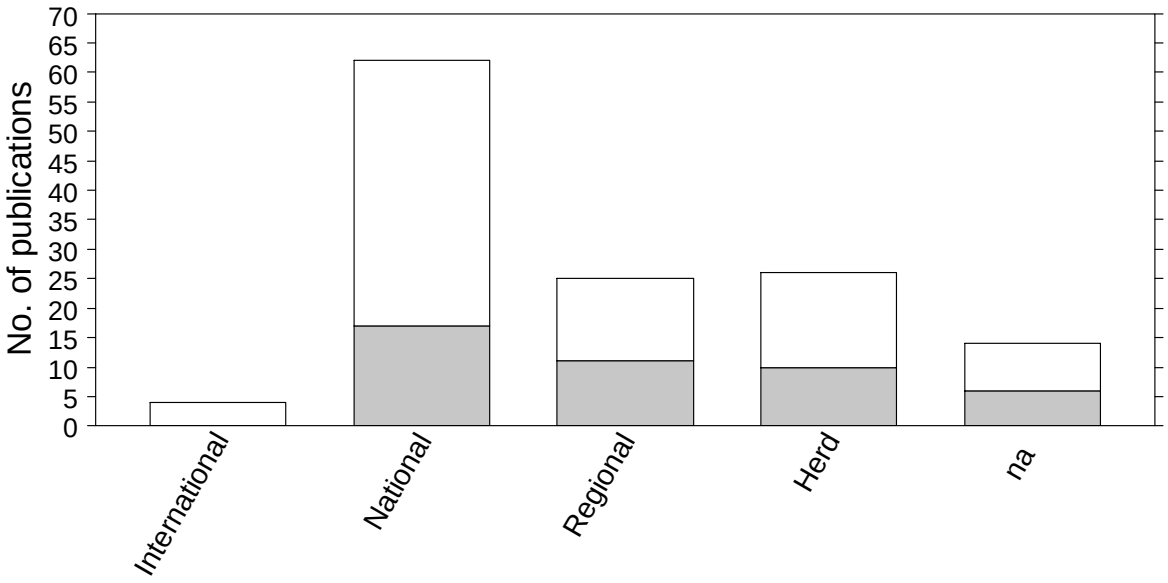

Figure 4. Number of publications identified by the review of methods to demonstrate disease freedom by unit level (n = 131), stratified by application of risk-based methods (grey shading: Not risk-based; white shading: Risk-based). Na: Not applicable.

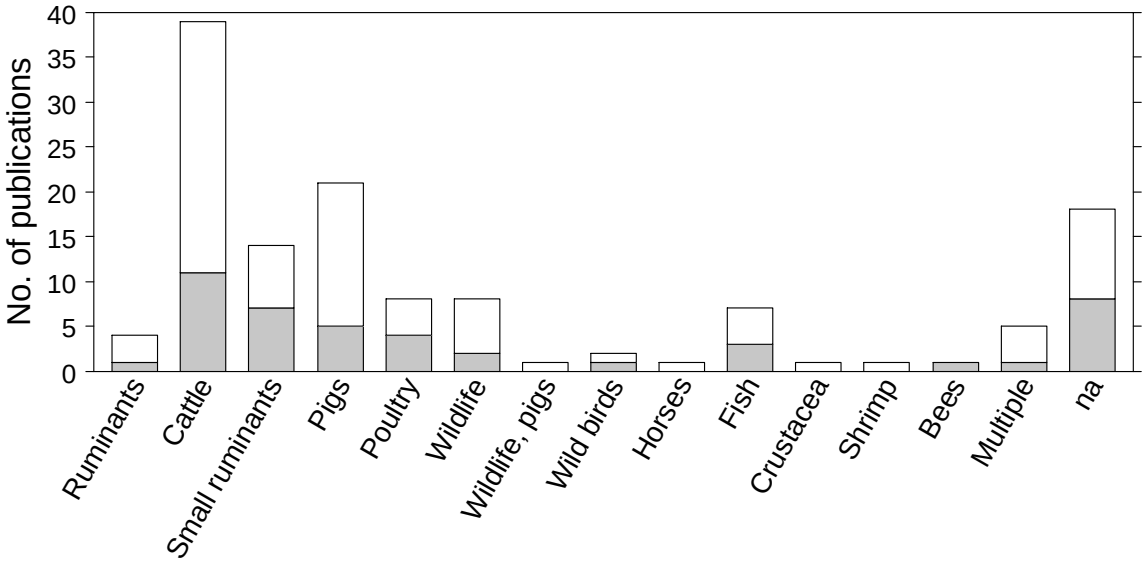

Figure 5. No. of publications identified by the review of methods to demonstrate disease freedom by species (n = 131), stratified by application of risk-based methods (grey shading: Not risk-based; white shading: Risk-based). Multiple: Inclusion of multiple domestic animals species. Na: Not applicable.

Table 3. Detailed list of publications identified by the review of methods to demonstrate disease freedom, in which sample sizes were calculated (n = 41).

No.	Reference	Type [a]	Level [b]	Methods$_S$ [c]	P_H / P_A (%) [d]	Software	Ref. no. [e]	Adjustments [f]			Method$_C$ [g]
								Se	Sp	Uncertainty	
1	Cannon (1982)	N	na	SRS	na/na	-	na	-	-	-	B
2	McLaughlin (1995)	A	R	SRS	na/7	na	1	ns	ns	-	ns
3	Garner, et al. (1997)	A	C	MSR	5/25		ns	✓	✓	-	ns
4	Cameron, et al. (1998a)	NC	H	SRS	na/30	FreeCalc	na	✓	✓	-	B, HG
5	Cameron, et al. (1998b)	N	na	MSR	na	FreeCalc	4	✓	✓	-	B, HG
6	Stärk, et al. (2000)	AC	C	RB	ns	FreeCalc	4	✓	✓	-	ns
7	Black, et al. (2001)	A	R	RB	5/40	FreeCalc	4,5	✓	-	-	ns
8	Cannon (2001)	N	na	MSR	ns	Excel	-	ns	ns	-	O
9	Rautiainen, et al. (2001)	A	R	SRS	na/10	Statistix	1	-	-	-	B
10	Cannon (2002)	N	na	RB	ns	ns	-	ns	ns	-	ns
11	Hadorn, et al. (2002)	NC	C	RB	0.2/1.5-13	FreeCalc	4,5,8	✓	✓	-	
12	Ziller, et al. (2002)	C	C	MSR	<0.2/1	na	various	✓	✓	-	HG
13	Johnson, et al. (2003)	NC	H	SRS	na/5-30	epi.ucdavis	other	✓	✓	✓	B, HG
14	East, et al. (2004)	A	C	MSR	5/10	FreeCalc	4,5	✓	✓	-	ns
15	Greiner, et al. (2005)	C	H	SRS	na/ns	Stata	ns	✓	✓	-	B, P
16	Thrusfield, et al. (2005)	A	A	MSR	2/5	WinEpiscope	1	-	-	-	HG
17	Branscum, et al. (2006)	N	H	MSR	ns	WinBugs, SPlus	13	✓	✓	✓	B
18	Cagienard, et al. (2006)	A	C	MSR	2/25;40;60	FreeCalc	4,5	-	✓	-	ns
19	Corbellini, et al. (2006)	ReC	na	RB	na	na	8,11	na	na	-	na
20	Knopf, et al. (2007)	AC	C	RB	Various/various	FreeCalc	ns	✓	✓	✓	HG
21	O'Grady, et al. (2008)	A	H	RB	na/10	Freecalc	4	✓	✓	-	ns
22	Peeler, et al. (2008)	A	H	MSR	various	Freecalc	4,5	✓	-		HG
23	Fosgate (2009)	ReC	R	SRS	0.1/-	Freecalc	4	-	-	-	B, HG
24	McDonald, et al. (2009)	A	C	RB	2/na	FreeCalc	4	-	-	-	HG
25	Schwermer, et al. (2009)	A	N	RB	0.2/na	@Risk, FreeCalc	11	✓	-	-	ns
26	Williams, et al. (2009a)	N	H	RB	na/8	ns	various	-	-	-	O; HG
27	Williams, et al. (2009b)	N	na	RB	ns/>1	R	10	ns	ns	✓	na
28	Rawdon, et al. (2010)	A	C	MSR	5/30	FreeCalc	4,5	✓	-	-	ns
29	Schuppers, et al. (2010b)	A	C	RB	na/0.1-0.04	WinEpiscope, FreeCalc	4,5	✓	-	-	ns
30	Welby, et al. (2010)	A	C	RB	5/30;LM	FreeCalc	4	✓	-	✓	HG
31	Blickenstorfer, et al. (2011)	C	C	STM	0.2/na	@Risk	11	✓	-	-	ns
32	Christensen, et al. (2011)	A	C	STM	1/30	CanNAISS	ns	✓	-	-	ns
33	Plischuk, et al. (2011)	A	R	SRS	1.1/-	FreeCalc	4	✓	-	-	HG
34	Windsor, et al. (2011)	A	R	RB	10/20	FreeCalc	4,5	-	✓	-	ns
35	Dukpa, et al. (2012)	A	R	RB	20; 25	FreeCalc	4,5	✓	✓	-	ns
36	Kostoulas, et al. (2012)	N	H	RB	various	WinBugs	17	✓	✓	✓	B
37	Reist, et al. (2012)	C	C	STM	0.1-0.2/na	@Risk	10,11, 29	ns	ns	ns	ns
38	Rodríguez, et al. (2012)	A	R	SRS	na/10;5	FreeCalc	3	✓	✓	-	HG
39	Schuppers, et al. (2012)	A	C	MSR	0.2/10;40	WinEpiscope	ns	✓	✓	✓	ns
40	Welby, et al. (2012)	C	C	STM	0.1/-	@Risk	other	✓	-	-	ns
41	Boklund, et al. (2013)	C	C	STM	1/5	@Risk	other	✓	-	✓	ns

[a] Type: New method (N), application (A), comparison (C), review (Re).
[b] Level: Country (C), region (R), herd (H).
[c] Methods$_S$: Simple random sampling (SRS), multi-stage random sampling (MSR), risk-based sampling (RB), scenario tree model (STM); NOTE: sampling method can be a mixture: E.g. SRS generally applied at a given level for MSR; MSR generally applied of RB; STM as a special application of RB.
[d] P^*_H / P^*_A (%): Herd / animal design prevalence; NOTE: this also indicates if the herd or animal was the unit of interest; LM: Limited sampling of a pre-fixed number of animals per herd.
[e] Ref. no.: No. of the publication in this table (see first column) used as a reference for calculation method.
[f] Adjustment for sensitivity (Se) and specificity (Sp) or uncertainty in test characteristics;
[g] Calculation method (Method$_C$) based on the binomial approximation (B), modified binomial approximation to the hypergeometric (HG) formula, Poisson (P) or other methods (O);
Na: not applicable; ns: not specified; ✓: Yes, -: No.

Table 4. Detailed list of publications identified by the review of methods to demonstrate disease freedom, in which a scenario tree model (STM) was applied (n = 29). The purpose of all publications was to estimate surveillance system sensitivity or probability of freedom, sometimes combined with sample size calculations (see Table 3).

Nr.	Author	Level [a]	Time units [b]	Compo-nents [c]	Nodes (R / I / D) [d]	Sensitivity analysis [e]	Adjustments [f]			Risk factors used [g]
							Se	Sp	Uncer-tainty	
1	Vose (1997)	----------Review using different examples----------								
2	Hadorn et al. (2002)	C	2S	1A	0/1/4	-	✓	✓	✓	-
3	Sergeant, et al. (2002)	C	12M	1A	-/1/2	-	✓	✓	✓	na (stratified by loc)
4	Martin (2007b)	----------New method using different examples----------								
5	Martin (2007a)	C	12M	1A	2/2/1	✓	✓	-	✓	Loc, age
6	Alban, et al. (2008)	C	16Y	1A	ns	✓	✓	ns	✓	Age, housing
7	Hadorn (2008b)	C	na	2A,1P	0/2/4	✓	✓	ns	✓	Wildlife, trade, human
8	Hadorn (2008a)	C	na	1P	(1)/1/4	✓	✓	-	-	-
9	Martin (2008)	R	11Y	1A,1P	1/2/5	✓	✓	-	✓	Contact, herd type, import, origin
10	Hood (2009)	----------Comparison of two methods using two examples----------								
11	Hadorn (2009)	C	12M	2A,2P	1/1/4	✓	✓	ns	✓	Loc
12	Frössling (2009)	C	6M	2A,1P	2/2/5	✓	✓	-	✓	Loc, contact
13	More (2009)	H	3Y	2A,1P	1/2/2	-	✓	-	✓	Biosec,import,loc
14	Alba (2010)	R	12Q	1A,1P	1/2/2	-	✓	ns	✓	Species, prod type
15	Knight-Jones, et al. (2010)	R	4Y	5A,1P	1/1/3	✓	✓	-	✓	Species
16	Schuppers, et al. (2010a)	C	15Y	1A	2/1/2	✓	✓	-	✓	Age, housing
17	Wahlström (2010)	C	13Y	3A	2/2/1	✓	✓	-	✓	Import, stream
18	Welby (2010)	C	na	1A	4/4/1	-	✓	-	✓	Outdoor, species
19	Blickenstorfer (2011)	C	na	1A	5;3/1/1	-	✓	-	✓	2xcontact, move, 2xloc, import
20	Christensen (2011)	C	6M	2A	0/2/5	✓	✓	-	✓	na
21	Hernández-Jover (2011)	C	na	1P	2/2/5	✓	✓	ns	✓	Feeding, herd size
22	Wahlström (2011)	C	10Y	Various A	0/1/3	✓	✓	-	✓	-
23	Willeberg, et al. (2011)	C	16Y	1A	0/1/1	-	✓	✓	-	Age, housing
24	Goutard (2012)	C	36M	1P,2A	2/2/4	✓	✓	✓	-	Loc, type
25	Murphy (2012)	C	11Y	1A	ns	-	✓	✓	-	-
26	Welby (2012)	C	5Y	1A	5/2/1	✓	✓	-	✓	Import, HF, DS
27	Boklund (2013)	C	24M	4A	3/1/1	✓	✓	-	✓	Biosec, type, age
28	Calvo-Artavia (2013)	C	7Y	1A	3/1/1	✓	✓	✓	✓	Gender,2HF
29	Frössling (2013)	C	4Y	2P,3A	4/2/3	✓	✓	-	✓	Import, DS, died, origin

[a] Level: Country (C), region (R), herd (H).
[b] Time units: Number and type of time units: Years (Y), quarters (Q), months (M).
[c] Components: Active (A), passive (P), enhanced passive (EP), risk-based (RB).
[d] Nodes: Number of risk category nodes (R), infection nodes (I) and detection nodes (D): If the no. of nodes differed between surveillance system components (SSCs) or different scenario trees (STs), then the SSC/ST with the maximum no. of nodes was chosen to indicate the complexity of a given SSC.
[e] Sensitivity analysis carried out: Yes (✓), no (-).
[f] Adjustment for sensitivity (Se) and specificity (Sp) or uncertainty in test characteristics;
[g] Risk factors used: Risk factors are either specifically listed or, if too detailed, were grouped as factors related to the animal (AF), herd (management) (HF), contact (CONT) or location (Loc). Details are provided in Table 5. The inclusion of historical data is covered under 'Time units'.
na: not applicable; ns: not specified; ✓: Yes, -: No.

Table 5. Risk factors considered in publications identified in a review of methods to demonstrate disease freedom, in which scenario tree models were applied (n = 20).

Group	Risk factor	n	References
Animal factors (AF)		**12**	
	Age	5	Martin, et al., 2007a, Alban, et al., 2008, Schuppers, et al., 2010a, Willeberg, et al., 2011, Boklund, et al., 2013
	Gender	1	Calvo-Artavia, et al., 2013
	Origin of animal	2	Martin, 2008, Frössling, et al., 2013
	Previous disease status	2	Welby, et al., 2012, Frössling, et al., 2013
	Species	3	Alba, et al., 2010, Knight-Jones, et al., 2010, Welby, et al., 2010
Contact (CONT)		**18**	
	Animal contacts	3	Martin, 2008, Frössling, et al., 2009, Blickenstorfer, et al., 2011
	Animal movements	2	Blickenstorfer, et al., 2011, Welby, et al., 2012
	Grazing	2	Blickenstorfer, et al., 2011, Calvo-Artavia, et al., 2013
	High density of herds	1	Blickenstorfer, et al., 2011
	Human	1	Hadorn, et al., 2008b
	Importation	6	Martin, 2008, More, et al., 2009, Wahlström, et al., 2010, Welby, et al., 2010, Blickenstorfer, et al., 2011, Frössling, et al., 2013
	Outdoor	1	Welby, et al., 2010
	Trade	1	Hadorn, et al., 2008b
	Wildlife	1	Hadorn, et al., 2008b
Location (LOC)		**6**	
	Farm close to border	2	Martin, et al., 2007a, Blickenstorfer, et al., 2011
	Location	4	Frössling, et al., 2009, Hadorn, et al., 2009, More, et al., 2009, Goutard, et al., 2012
Management factors (MF)		**12**	
	Biosecurity	2	More, et al., 2009, Boklund, et al., 2013
	Farm type	4	Martin, 2008, Alba, et al., 2010, Goutard, et al., 2012, Boklund, et al., 2013
	Feeding	1	Hernández-Jover, et al., 2011
	Herd size	2	Hernández-Jover, et al., 2011, Welby, et al., 2012
	Housing	3	Alban, et al., 2008, Schuppers, et al., 2010a, Willeberg, et al., 2011
	Water	1	Calvo-Artavia, et al., 2013

Table 6. Detailed list of publications identified by the review of methods to demonstrate disease freedom, in which a simulation model was applied (n = 28).

No.	Reference	Level [a]	Outcome [b]	Time units [c]	Compo-nents [d]	Sensitivity analysis [e]	Adjustments [f]			Software	Risk factors [g]
							Se	Sp	Uncer-tainty		
1	Ellis, et al. (1998)	R	G2	2S	1A	-	✓	✓	-	@Risk	PD, CONT
2	Audigé, et al. (1999)	C	G1, G2	-	1A	-	✓	✓	✓	@Risk	
3	Gohm, et al. (1999)	C	G1	1S	1A	-	✓	✓	✓	@Risk	-
4	Good, et al. (2001)	H	G2	14Y	1A, 1P	-	✓	✓	✓	@Risk	-
5	Schlosser, et al. (2001)	H	G2	✓	1A	-	✓	-	✓	Visual Basics	-
6	Suess, et al. (2002)	C	G2	-	1A	-	✓	✓	-	Fortran, R	
7	Hanson, et al. (2003)	H	G1	-	1A	-	✓	✓	✓	ns	PD
8	Van Schaik, et al. (2003)	H	G1, $	-	1A	-	✓	✓	✓	@Risk	-
9	Weber, et al. (2004)	H	G1, P, $	20Y	1A	✓	✓	✓	(✓)	JohneSSim	HFs
10	Branscum, et al. (2005)	H	ICC	-	1A	-	✓	✓	✓	WINBUGS	-
11	De Massis, et al. (2005)	C	G1	12M	1A	-	ns	ns	-	@Risk	-
12	Ezzano, et al. (2005)	H	G2	25Y	1A	✓	✓	-	✓	Excel	-
13	Feliziani, et al. (2005)	C	G1	9Y	1A	-	ns	ns	ns	ns	PD
14	Greiner, et al. (2005)	H	G1	na	na	✓	✓	-	✓	Stata	-
15	Mintiens, et al. (2005)	R	G2	27M	1A,1P	✓	✓	✓	✓	Own (BDFree)	Loc
16	Corbellini, et al. (2006)	C	G2	2S	1A	-	✓	✓	✓	@Risk	-
17	Heuer, et al. (2007)	R	Time	na	1EP	✓	✓	-	(✓)	Berkeley Madonna	-
18	Knopf, et al. (2007)	C	G2, SS	12Y	1A	✓	✓	✓	✓	@Risk	
19	Ebel, et al. (2008)	R	P	5Y	1A	✓	✓	✓	✓	Winbugs	-
20	Sergeant, et al. (2008)	H	G1, $	na	1-2A	-	✓	✓	✓	R	(age)
21	Joly, et al. (2009)	R	G1	2Y	1A	-	ns	ns	-	ns	Loc
22	Handel, et al. (2011)	R	G1, $, time, spread	na	1A, 1P	-	ns	-	-	R	Inf.risk
23	Sergeant, et al. (2011)	R	G2	4M	As	-	✓	-	✓	R	HSize, PD
24	Häsler, et al. (2012)	C	$	10Y	1A	✓	✓	✓	(✓)	@Risk	-
25	Reber, et al. (2012)	C	G1, $	-	1A	-	✓	-	-	Excel, FreeCalc	ns
26	Schuppers, et al. (2012)	C	G2, SS	na	1A	-	✓	✓	✓	@Risk	-
27	Tavornpanich, et al. (2012)	H	G1	na	1A	-	✓	✓	-	Matlab	-
28	More, et al. (2013)	H	G2, $	5Y	1A	-	✓	✓	✓	R	PD

[a] Level: Country (C), region (R), herd (H).

[b] Outcome: Surveillance sensitivity (G1); probability of freedom (G2), cost ($), prevalence (P), sample size (SS), intracluster correlation coefficient (ICC), duration of surveillance being required (time).

[c] Time units: Number and type of time units: Years (Y), quarters (Q), months (M), surveys (S2).

[d] Components: Active (A), passive (P), enhanced passive (EP).

[e] Sensitivity analysis carried out: Yes (✓), no (-).

[f] Adjustment for sensitivity (Se) and specificity (Sp) or uncertainty in test characteristics;

[g] Risk factors used: Risk factors were grouped as factors related to contact (CONT) or location (Loc). Other risk factors listed were previous disease status (PD), herd size (HSize) and infection risk (Inf.risk). The inclusion of historical data is covered under 'Time units'.

na: not applicable; ns: not specified.

10 REFERENCES

Alba, A., Casal, J., Napp, S. and Martin, P. a. J. (2010). "Assessment of different surveillance systems for avian influenza in commercial poultry in Catalonia (North-Eastern Spain)." Preventive veterinary medicine **97**(2): 107-118.

Alban, L., Boes, J., Kreiner, H., Petersen, J. V. and Willeberg, P. (2008). "Towards a risk-based surveillance for Trichinella spp. in Danish pig production." Preventive veterinary medicine **87**(3-4): 340-357.

Alban, L., Boes, J., Sandberg, M., Dahl, J. and Kristensen, E. (2011). "The need for surveillance - from the livestock and meat industry's perspective." Épidémiologie et Santé Animale **59-60**: 342-344.

Appleyard, G. D., Forbes, L. B. and Gajadhar, A. A. (2002). "National serologic survey for trichinellosis in sows in Canada 1996-1997." The Canadian veterinary journal. La revue vétérinaire canadienne **43**(4): 271-273.

Audigé, L. and Beckett, S. (1999). "A quantitative assessment of the validity of animal-health surveys using stochastic modelling." Preventive veterinary medicine **38**(4): 259-276.

Baldock, F. C. (1998). "What constitutes freedom from disease in livestock?" Australian veterinary journal **76**(8): 544-545.

Bessell, P. R., Orton, R., O'hare, A., Mellor, D. J., Logue, D. and Kao, R. R. (2013). "Developing a framework for risk-based surveillance of tuberculosis in cattle: a case study of its application in Scotland." Epidemiology and infection **141**(2): 1-10.

Black, P. F., Cronin, J. P., Morrissy, C. J. and Westbury, H. A. (2001). "Serological examination for evidence of infection with Hendra and Nipah viruses in Queensland piggeries." Australian veterinary journal **79**(6): 424-426.

Blickenstorfer, S., Schwermer, H., Engels, M., Reist, M., Doherr, M. G. and Hadorn, D. C. (2011). "Using scenario tree modelling for targeted herd sampling to substantiate freedom from disease." BMC veterinary research **7**(1): 49-49.

Böhning, D. and Greiner, M. (2006a). "Evaluation of the cumulative evidence for freedom from BSE in birth cohorts." European journal of epidemiology **21**(1): 47-54.

Böhning, D. and Greiner, M. (2006b). "Modeling cumulative evidence for freedom from disease with applications to BSE surveillance trials." Journal of agricultural, biological, and environmental statistics **11**(3): 280-295.

Boklund, A., Dahl, J. and Alban, L. (2013). "Assessment of confidence in freedom from Aujeszky's disease and classical swine fever in Danish pigs based on serological sampling-Effect of reducing the number of samples." Preventive veterinary medicine **110**(2): 214-222.

Bradley, T. L. and Cannon, R. M. (2005). "Determining the sensitivity of abattoir surveillance for ovine Johne's disease." Australian veterinary journal **83**(10): 633-636.

Branscum, A. J., Gardner, I. A., Wagner, B. A., Mcinturff, P. S. and Salman, M. D. (2005). "Effect of diagnostic testing error on intracluster correlation coefficient estimation." Preventive veterinary medicine **69**(1-2): 63-75.

Branscum, A. J., Johnson, W. O. and Gardner, I. A. (2006). "Sample size calculations for disease freedom and prevalence estimation surveys." Statistics in medicine **25**(December 2004): 2658-2674.

Burr, P. (2007). "Does the herd have Johne's Disease? Simple steps to assess status." Cattle Practice **15**(Part 2): 172-174.

Cagienard, A., Thür, B., Griot, C., Hamblin, C. and Stärk, K. D. C. (2006). "No evidence of bluetongue virus in Switzerland." Veterinary microbiology **116**(1-3): 13-20.

Calvo-Artavia, F. F., Nielsen, L. R. and Alban, L. (2013). "Epidemiologic and economic evaluation of risk-based meat inspection for bovine cysticercosis in Danish cattle." Preventive veterinary medicine **108**(4): 253-261.

Cameron, A. R. (2012). "The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease." Preventive veterinary medicine **105**(4): 280-286.

Cameron, A. R. and Baldock, F. C. (1998a). "A new probability formula for surveys to substantiate freedom from disease." Preventive veterinary medicine **34**(1): 1-17.

Cameron, A. R. and Baldock, F. C. (1998b). "Two-stage sampling in surveys to substantiate freedom from disease." Preventive veterinary medicine **34**(1): 19-30.

Cannon, R. M. (1982). Livestock disease surveys: A field manual for veterinarians. Canberra, Australian Government Publishing Service.

Cannon, R. M. (2001). "Sense and sensitivity - designing surveys based on an imperfect test." Preventive veterinary medicine **49**(3-4): 141-163.

Cannon, R. M. (2002). "Demonstrating disease freedom - combining confidence levels." Preventive veterinary medicine **52**(3-4): 227-249.

Carlsson, U., Wallgren, P., Renström, L. H. M., Lindberg, A., Eriksson, H., Thorén, P., . . . Elvander, M. (2009). "Emergence of porcine reproductive and respiratory syndrome in Sweden: detection, response and eradication." Transboundary and emerging diseases **56**(4): 121-131.

Christensen, J. (2001). "Epidemiological concepts regarding disease monitoring and surveillance." Acta veterinaria Scandinavica. Supplementum **94**: 11-16.

Christensen, J. and Gardner, I. A. (2000). "Herd-level interpretation of test results for epidemiologic studies of animal diseases." Preventive veterinary medicine **45**(1-2): 83-106.

Christensen, J., Stryhn, H., Vallières, A. and El Allaki, F. (2011). "A scenario tree model for the Canadian Notifiable Avian Influenza Surveillance System and its application to estimation of probability of freedom and sample size determination." Preventive veterinary medicine **99**(2-4): 161-175.

Corbellini, L. G., Schwermer, H., Presi, P., Thür, B., Stärk, K. D. C. and Reist, M. (2006). "Analysis of national serological surveys for the documentation of freedom from porcine reproductive and respiratory syndrome in Switzerland." Veterinary microbiology **118**(3-4): 267-273.

De Massis, F., Petrini, A. and Giovannini, A. (2005). "Reliability evaluation of sampling plan fixed by Council Directive 91/68/EEC for the maintenance of officially brucellosis-free flock status." Journal of veterinary medicine. B, Infectious diseases and veterinary public health **52**(6): 284-290.

Deliberto, T. J., Swafford, S. R., Nolte, D. L., Pedersen, K., Lutman, M. W., Schmit, B. B., . . . Franklin, A. (2009). "Surveillance for highly pathogenic avian influenza in wild birds in the USA." Integrative zoology **4**(4): 426-439.

Doherr, M. G. and Audigé, L. (2001). "Monitoring and surveillance for rare health-related events: a review from the veterinary perspective." Philosophical transactions of the Royal Society of London. Series B, Biological sciences **356**(1411): 1097-1106.

Dohoo, I. R., Martin, W. and Stryhn, H. (2003). Veterinary epidemiologic research. Charlottetown, PE, Canada.

Dufour, B., Pouillot, R. and Toma, B. (2001). "Proposed criteria to determine whether a territory is free of a given disease." Veterinary research **32**(6): 545-563.

Dukpa, K., Robertson, I. D. and Ellis, T. M. (2012). "Serological and clinical surveillance studies to validate reported foot-and-mouth disease free status in Tsirang district of Bhutan." Preventive veterinary medicine **104**(1-2): 23-33.

Durand, B., Martinez, M.-J., Calavas, D. and Ducrot, C. (2009). "Comparison of strategies for substantiating freedom from scrapie in a sheep flock." BMC veterinary research **5**: 16-16.

East, I. J., Black, P. F., Mccoll, K. A., Hodgson, R. a. J. and Bernoth, E. M. (2004). "Survey for the presence of White Spot Syndrome virus in Australian crustaceans." Australian veterinary journal **82**(4): 236-240.

Ebel, E. D., Williams, M. S. and Tomlinson, S. M. (2008). "Estimating herd prevalence of bovine brucellosis in 46 USA states using slaughter surveillance." Preventive veterinary medicine **85**(3-4): 295-316.

Efsa Panel on Biological Hazards (2010). "Scientific Opinion on a second update on the risk for human and animal health related to the revision of the BSE monitoring regime in some Member States." EFSA Journal **8**(12): 1-74.

Ellis, T. M., Norris, R. T., Martin, P. a. J., Casey, R. H. and Hawkins, C. D. (1998). "Evidence for freedom from Johne's disease in cattle and goats in Western Australia." Australian journal of veterinary research **76**(9): 630-633.

Ezzano, P., Van Schaik, G., Weber, M. F. and Heesterbeek, J. a. P. (2005). "A modeling study on the sustainability of a certification-and-monitoring program for paratuberculosis in cattle." Veterinary research **36**(5-6): 811-826.

Fahrion, A. S., Schüpbach, G., Schwermer, H., Hauser, R. and Isoda, N. (2011). "Sensitivity analysis of modeling probability for disease freedom as a basis for risk-based sample size calculation." Épidémiologie et Santé Animale **59-60**: 103-104.

Feliziani, F., Maresca, C., Giovannini, A., Mammoli, R. and Rutili, D. (2005). "Statistical Evaluation of Classical Swine Fever Surveillance Plans in Italy." Journal of veterinary medicine **52**: 199-200.

Feller, W., Ed. (1968). An introduction to probability theory and its applications. New York, Wiley.

Fosgate, G. T. (2009). "Practical Sample Size Calculations for Surveillance and Diagnostic Investigations." Journal of veterinary diagnostic investigation **21**(1): 3-14.

Frössling, J., Agren, E. C. C., Eliasson-Selling, L. and Lewerin, S. S. (2009). "Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system." Preventive veterinary medicine **91**(2-4): 137-145.

Frössling, J., Wahlström, H., Agren, E. C. C., Cameron, A., Lindberg, A. and Sternberg Lewerin, S. (2013). "Surveillance system sensitivities and probability of freedom from Mycobacterium avium subsp. paratuberculosis infection in Swedish cattle." Preventive veterinary medicine **108**(1): 47-62.

Garner, M. G., Gleeson, L. J., Holyoake, P. K., Cannon, R. M. and Doughty, W. J. (1997). "A national serological survey to verify Australia's freedom from porcine reproductive and respiratory syndrome." Australian veterinary journal **75**(8): 596-600.

Gohm, D. S., Thür, B., Audigé, L. and Hofmann, M. A. (1999). "A survey of Newcastle disease in Swiss laying-hen flocks using serological testing and simulation modelling." Preventive veterinary medicine **38**(4): 277-288.

Good, C. M., Thorburn, M. A. and Mcewen, S. A. (2001). "Simulating the effect of eliminating routine bacterial screening on the negative predictive value of the Ontario fish hatchery disease monitoring program." Preventive veterinary medicine **50**(1-2): 101-108.

Goutard, F. L., Paul, M., Tavornpanich, S., Houisse, I., Chanachai, K., Thanapongtharm, W., . . . Roger, F. (2012). "Optimizing early detection of avian influenza H5N1 in backyard and free-range poultry production systems in Thailand." Preventive veterinary medicine **105**(3): 223-234.

Greiner, M. and Dekker, A. (2005). "On the surveillance for animal diseases in small herds." Preventive veterinary medicine **70**(3-4): 223-234.

Grindlay, D. J. C., Brennan, M. L. and Dean, R. S. (2012). "Searching the Veterinary Literature: A Comparison of the Coverage of Veterinary Journals by Nine Bibliographic Databases." Journal of veterinary medical education **39**(4): 404-412.

Gustafson, L. L., Gustafson, D. H., Antognoli, M. C. and Remmenga, M. D. (2013). "Integrating expert judgment in veterinary epidemiology: Example guidance for disease freedom surveillance." Preventive veterinary medicine **109**(1-2): 1-9.

Gustafson, L. L., Klotins, K., Tomlinson, S., Karreman, G., Cameron, A., Wagner, B., . . . Scott, A. (2010). "Combining surveillance and expert evidence of viral hemorrhagic septicemia freedom: a decision science approach." Preventive veterinary medicine **94**(1-2): 140-153.

Hadorn, D. C., Haracic, S. S. and Stärk, K. D. C. (2008a). "Comparative assessment of passive surveillance in disease-free and endemic situation: example of Brucella melitensis surveillance in Switzerland and in Bosnia and Herzegovina." BMC veterinary research **4**: 1-9 (article no. 52).

Hadorn, D. C., Racloz, V., Schwermer, H. and Staerk, K. D. C. (2009). "Establishing a cost-effective national surveillance system for Bluetongue using scenario tree modelling." Veterinary research **40**(6).

Hadorn, D. C., Rüfenacht, J., Hauser, R. and Stärk, K. D. C. (2002). "Risk-based design of repeated surveys for the documentation of freedom from non-highly contagious diseases." Preventive veterinary medicine **56**(3): 179-192.

Hadorn, D. C. and Stärk, K. D. C. (2008b). "Evaluation and optimization of surveillance systems for rare and emerging infectious diseases." Veterinary research **39**(6): 1-12.

Handel, I. G., De C Bronsvoort, B. M., Forbes, J. F. and Woolhouse, M. E. J. (2011). "Risk-targeted selection of agricultural holdings for post-epidemic surveillance: estimation of efficiency gains." PLOS ONE **6**(5): e20064-e20064.

Hanson, T. E., Johnson, W. O., Gardner, I. A. and Georgiadis, M. P. (2003). "Determining the infection status of a herd." Journal of agricultural biological and environmental statistics **8**(4): 469-485.

Häsler, B., Howe, K. S., Presi, P. and Stärk, K. D. C. (2012). "An economic model to evaluate the mitigation programme for bovine viral diarrhoea in Switzerland." Preventive veterinary medicine **106**(2): 162-173.

Hernández-Jover, M., Cogger, N., Martin, P. a. J., Schembri, N., Holyoake, P. K. and Toribio, J. a. L. M. L. (2011). "Evaluation of post-farm-gate passive surveillance in swine for the detection of foot and mouth disease in Australia." Preventive veterinary medicine **100**(3-4): 171-186.

Heuer, C., French, N. P., Jackson, R. and Mackereth, G. F. (2007). "Application of modelling to determine the absence of foot-and-mouth disease in the face of a suspected incursion." New Zealand Veterinary Journal **55**(6): 289-296.

Hood, G. M., Barry, S. C. and Martin, P. a. J. (2009). "Alternative methods for computing the sensitivity of complex surveillance systems." Risk analysis **29**(12): 1686-1698.

James, P. J., Garrett, J. A. and Moon, R. D. (2002). "Sensitivity of two-stage sampling to detect sheep biting lice (Bovicola ovis) in infested flocks." Veterinary parasitology **103**(1-2): 157-166.

Johnson, W. O., Su, C.-L., Gardner, I. A. and Christensen, R. (2003). "Sample Size Calculations for Surveys to Substantiate Freedom of Populations from Infectious Agents." Biometrics **60**(March): 165-171.

Joly, D. O., Samuel, M. D., Langenberg, J. A., Rolley, R. E. and Keane, D. P. (2009). "Surveillance to detect chronic wasting disease in white-tailed deer in Wisonsin." Journal of wildlife diseases **45**(4): 989-997.

Kitching, R. P. (2003). Foot and mouth disease diagnostics: requirements for demonstration of freedom from infection.

Knight-Jones, T. J. D., Hauser, R., Matthes, D. and Staerk, K. D. C. (2010). "Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza." Veterinary research **41**(4).

Knopf, L., Schwermer, H. and Stärk, K. D. C. (2007). "A stochastic simulation model to determine the sample size of repeated national surveys to document freedom from bovine herpesvirus 1 (BoHV-1) infection." BMC veterinary research **3**(10): 1-9.

Kostoulas, P., Nielsen, S. S., Browne, W. J. and Leontides, L. (2012). "Sample size estimation to substantiate freedom from disease for clustered binary data with a specific risk profile." Epidemiology and infection **141**(6): 1318-1327.

Learmount, J., Zimmer, I. A., Conyers, C., Boughtflower, V. D., Morgan, C. P. and Smith, G. C. (2012). "A diagnostic study of Echinococcus multilocularis in red foxes (Vulpes vulpes) from Great Britain." Veterinary parasitology **190**(3-4): 447-453.

Levy, P. S. and Lemeshow, S. (1991). Sampling of populations: Methods and applications. New York, John Wiley and Sons Inc.

Martin, P. a. J. (2008). "Current value of historical and ongoing surveillance for disease freedom: surveillance for bovine Johne's disease in Western Australia." Preventive veterinary medicine **84**(3-4): 291-309.

Martin, P. a. J., Cameron, A. R., Barfod, K., Sergeant, E. S. G. and Greiner, M. (2007a). "Demonstrating freedom from disease using multiple complex data sources 2: Case study - Classical swine fever in Denmark." Preventive veterinary medicine **79**(2-4): 98-115.

Martin, P. a. J., Cameron, A. R. and Greiner, M. (2007b). "Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees." Preventive veterinary medicine **79**(2-4): 71-97.

Martinez, M.-J., Durand, B., Calavas, D. and Ducrot, C. (2010). "Methodological approach for substantiating disease freedom in a heterogeneous small population. Application to ovine scrapie, a disease with a strong genetic susceptibility." Preventive veterinary medicine **95**(1-2): 108-114.

Mcdonald, W. L., Rawdon, T. G., Fitzmaurice, J., Bolotovski, I., Voges, H., Humphrey, S., . . . Mcintyre, L. (2009). "Survey of bulk tank milk in New Zealand for Mycoplasma bovis, using species-specific nested PCR and culture." New Zealand veterinary journal **57**(1): 44-49.

Mcfadden, A. M. J., Mcfadden, B. D., Mackereth, G. F., Clough, R. R., Hueston, L., Gradweradwell, B. and Dymond, M. (2009). "A serological survey of cattle in the Thames-Coromandel district of New Zealand for antibodies to Ross River virus." New Zealand veterinary journal **57**(2): 116-120.

Mclaughlin (1995). "Reconfirmation of the South Island's freedom from Aujeszky's disease." Surveillance **22**(2): 17-18.

Mintiens, K., Verloo, D., Venot, E., Laevens, H., Dufey, J., Dewulf, J., . . . Koenen, F. (2005). "Estimating the probability of freedom of classical swine fever virus of the East-Belgium wild-boar population." Preventive veterinary medicine **70**(3-4): 211-222.

More, S. J., Cameron, A. R., Greiner, M., Clifton-Hadley, R. S., Rodeia, S. C., Bakker, D., . . . Wierup, M. (2009). "Defining output-based standards to achieve and maintain tuberculosis freedom in farmed deer, with reference to member states of the European Union." Preventive veterinary medicine **90**(3-4): 254-267.

More, S. J., Sergeant, E. S. G., Strain, S., Cashman, W., Kenny, K. and Graham, D. (2013). "The effect of alternative testing strategies and bio-exclusion practices on Johne's disease risk in test-negative herds." Journal of dairy science **96**(3): 1581-1590.

Mur, L., Boadella, M., Martínez-López, B., Gallardo, C., Gortazar, C. and Sánchez-Vizcaíno, J. M. (2012). "Monitoring of African swine fever in the wild boar population of the most recent endemic area of Spain." Transboundary and emerging diseases **59**(6): 526-531.

Murphy, T. M., Wahlström, H., Dold, C., Keegan, J. D., Mccann, A., Melville, J., . . . Mcateer, W. (2012). "Freedom from Echinococcus multilocularis: an Irish perspective." Veterinary parasitology **190**(1-2): 196-203.

Nérette, P., Hammell, L., Dohoo, I. and Gardner, I. (2008). "Evaluation of testing strategies for infectious salmon anaemia and implications for surveillance and control programs." Aquaculture **280**(1-4): 53-59.

Norby, B., Bartlett, P. C., Grooms, D. L., Kaneene, J. B. and Bruning-Fann, C. S. (2005). "Use of simulation modeling to estimate herd-level sensitivity, specificity, and predictive values of diagnostic tests for detection of tuberculosis in cattle." American journal of veterinary research **66**(7): 1285-1291.

O'grady, L., O'neill, R., Collins, D. M., Clegg, T. A. and More, S. J. (2008). "Herd and within-herd BoHV-1 prevalence among Irish beef herds submitting bulls for entry to a performance testing station." Irish veterinary journal **61**(12): 809-815.

Oidtmann, B., Peeler, E., Lyngstad, T., Brun, E., Jensen, B. B. and Staerk, K. D. C. (2013). "Risk-based methods for fish and terrestrial animal disease surveillance." Preventive Veterinary Medicine **112**(1-2): 13-26.

Oidtmann, B. C., Crane, C. N., Thrush, M. A., Hill, B. J. and Peeler, E. J. (2011). "Ranking freshwater fish farms for the risk of pathogen introduction and spread." Preventive veterinary medicine **102**(4): 329-340.

Oie (2004). Handbook on import risk analysis for animals and animal products. Volume 1: Introduction and qualitative risk analysis.

Oie. (2011). "Terrestrial animal health code." Chapter 1.4. Animal health surveillance Retrieved 13 August 2013, from http://www.oie.int/fileadmin/home/eng/health_standards/tahc/2010/en_chapitre_1.1.4.htm.

Paré, J., Geale, D. W., Koller-Jones, M., Hooper-Mcgrevy, K., Golsteyn-Thomas, E. J. and Power, C. A. (2008). "Serological status of Canadian cattle for brucellosis, anaplasmosis, and bluetongue in 2007-2008." The Canadian veterinary journal. La revue vétérinaire canadienne **53**(9): 949-956.

Paton, D. J., De Clercq, K., Greiner, M., Dekker, A., Brocchi, E., Bergmann, I., . . . Parida, S. (2006). "Application of non-structural protein antibody tests in substantiating freedom from foot-and-mouth disease virus infection after emergency vaccination of cattle." Vaccine **24**(42-43): 6503-6512.

Peeler, E. J. (2012). "Costs and benefits of freedom from shrimp diseases in the European Union." Journal of invertebrate pathology **110**(2): 188-195.

Peeler, E. J. and Oidtmann, B. C. (2008). "Demonstrating freedom from Gyrodactylus salaris (Monogenea: Gyrodactylidae) in farmed rainbow trout Oncorhynchus mykiss." Diseases of aquatic organisms **79**(1): 47-56.

Plischuk, S., Meeus, I., Smagghe, G. and Lange, C. E. (2011). "Apicystis bombi (Apicomplexa: Neogregarinorida) parasitizing Apis mellifera and Bombus terrestris (Hymenoptera: Apidae) in Argentina." Environmental microbiology reports **3**(5): 565-568.

Racloz, V., Straver, R., Kuhn, M., Thur, B., Vanzetti, T., Stärk, K. D. C., . . . Cagienard, A. (2006). "Establishment of an early warning system against Bluetongue virus in Switzerland." Schweizer Archiv für Tierheilkunde **148**(November): 593-598.

Radunz, B. (2006). "Surveillance and risk management during the latter stages of eradication: experiences from Australia." Veterinary microbiology **112**(2-4): 283-290.

Rautiainen, E., Oravainen, J., Virolainen, J. V. and Tuovinen, V. (2001). "Regional eradication of Mycoplasma hyopneumoniae from pig herds and documentation of freedom of the disease." Acta veterinaria Scandinavica **42**(3): 355-364.

Rawdon, T., Tana, T., Thornton, R. N., Mckenzie, J. S., Stanislawek, W. L., Kittelberger, R., . . . Cork, S. C. (2010). "Surveillance for avian influenza virus subtypes H5 and H7 in chickens and turkeys farmed commercially in New Zealand." New Zealand veterinary journal **58**(6): 292-298.

Reber, A., Reist, M. and Schwermer, H. (2012). "Cost-effectiveness of bulk-tank milk testing for surveys to demonstrate freedom from infectious bovine rhinotracheitis and bovine enzootic leucosis in Switzerland." Schweizer Archiv für Tierheilkunde **154**(5): 189-197.

Reist, M., Jemmi, T. and Stärk, K. D. C. (2012). "Policy-driven development of cost-effective, risk-based surveillance strategies." Preventive veterinary medicine **105**(3): 176-184.

Rodríguez, N. F., Tejedor-Junco, M. T., González-Martín, M., Santana Del Pino, A. and Gutiérrez, C. (2012). "Cross-sectional study on prevalence of Trypanosoma evansi infection in domestic ruminants in an endemic area of the Canary Islands (Spain)." Preventive veterinary medicine **105**(1-2): 144-148.

Rolesu, S., Aloi, D., Ghironi, A., Oggiano, N., Oggiano, A., Puggioni, G., . . . Montinaro, S. (2007). "Geographic information systems: a useful tool to approach African swine fever surveillance management of wild pig populations." Veterinaria italiana **43**(3): 463-467.

Ryan, E. G., Leonard, N., O'grady, L., Doherty, M. L. and More, S. J. (2012). "Herd-level risk factors associated with Leptospira Hardjo seroprevalence in Beef/Suckler herds in the Republic of Ireland." Irish veterinary journal **65**(1): 6-6.

Salman, M. D., Stärk, K. D. C. and Zepeda, C. (2003). "Quality assurance applied to animal disease." Revue scientifique et technique (International Office of Epizootics) **22**(2): 689-696.

Sanchez-Vizcaino, F., Perez, A., Lainez, M. and Manuel Sanchez-Vizcaino, J. (2010a). "A quantitative assessment of the risk for highly pathogenic avian influenza introduction into Spain via legal trade of live poultry." Risk analysis **30**(5): 798-807.

Sanchez-Vizcaino, F., Perez, A., Lainez, M. and Sanchez-Vizcaino, J. M. (2010b). "Quantification of the risk for introduction of virulent Newcastle disease virus into Spain through legal trade of live poultry from European Union countries." Avian pathology **39**(6): 459-465.

Schlosser, W. and Ebel, E. (2001). "Use of a Markov-chain Monte Carlo model to evaluate the time value of historical testing information in animal populations." Preventive veterinary medicine **48**(3): 167-175.

Schuppers, M. E., Frey, C. F., Gottstein, B., Stärk, K. D. C., Kihm, U. and Regula, G. (2010a). "Comparing the demonstration of freedom from Trichinella infection of domestic pigs by traditional and risk-based surveillance." Epidemiology and infection **138**(9): 1242-1251.

Schuppers, M. E., Pyburn, D. G. and Alban, L. (2011). "Alternative monitoring and surveillance scenarios for Trichinella spp . and their impact on public health." Épidémiologie et Santé Animale **59/60**: 120-122.

Schuppers, M. E., Rosenberg, G., Graf, R., Eidam, V., Wittwer, C., Zimmermann, W., . . . Frey, C. F. (2010b). "A study to demonstrate freedom from Trichinella spp. in domestic pigs in Switzerland." Zoonoses and public health **57**(7-8): e130-135.

Schuppers, M. E., Stegeman, J. A., Kramps, J. A. and Stärk, K. D. C. (2012). "Implementing a probabilistic definition of freedom from infection to facilitate trade of livestock: putting theory into praxis for the example of bovine herpes virus-1." Preventive veterinary medicine **105**(3): 195-201.

Schwermer, H., Reding, I. and Hadorn, D. C. (2009). "Risk-based sample size calculation for consecutive surveys to document freedom from animal diseases." Preventive veterinary medicine **92**(4): 366-372.

Seber, G. a. F. (1986). "A REVIEW OF ESTIMATING ANIMAL ABUNDANCE." Biometrics **42**(2): 267-292.

Sergeant, E. S. G. and Baldock, F. C. (2002). "The estimated prevalence of Johne's disease infected sheep flocks in Australia." Australian veterinary journal **80**(12): 762-768.

Sergeant, E. S. G., Nielsen, S. S. and Toft, N. (2008). "Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds." Preventive veterinary medicine **85**(1-2): 92-106.

Sergeant, E. S. G. and Wilson, G. (2011). "Demonstrating freedom from equine influenza in New South Wales, Australia, following the 2007 outbreak." Australian veterinary journal **89 Suppl 1**(July): 164-169.

Stärk, K. D. C., Mortensen, S., Olsen, A. M., Barfod, K., Botner, A., Lavritsen, D. T. and Strandbygard, B. (2000). "Designing serological surveillance programmes to document freedom from disease with special reference to exotic viral diseases of pigs in Denmark system." Revue scientifique et technique (International Office of Epizootics) **19**(3): 715-724.

Stärk, K. D. C., Regula, G., Hernandez, J., Knopf, L., Fuchs, K., Morris, R. S. and Davies, P. (2006). "Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches." BMC health services research **6**: 20-20.

Su, C. L., Gardner, I. A. and Johnson, W. O. (2007). "Bayesian estimation of cluster-level test accuracy based on different sampling schemes." Journal of agricultural biological and environmental statistics **12**(2): 250-271.

Suess, E. A., Gardner, I. A. and Johnson, W. O. (2002). "Hierarchical Bayesian model for prevalence inferences and determination of a country's status for an animal pathogen." Preventive veterinary medicine **55**(3): 155-171.

Tavornpanich, S., Wells, S. J., Fossler, C. P., Roussel, A. J. and Gardner, I. A. (2012). "Evaluation of an alternative method of herd classification for infection with paratuberculosis in cattle herds in the United States." Australian journal of veterinary research **73**(2): 248-256.

Thorburn, M. A. (1996). "Apparent prevalence of fish pathogens in asymptomatic salmonid populations and its effect on misclassifying population infection status." Journal of aquatic animal health **8**(4): 271-277.

Thornton, R. and Motha, J. (1995). "A serological survey to confirm New Zealand's freedom from maedi-visna." Surveillance **22**(2): 22-24.

Thrusfield, M., Mansley, L., Dunlop, P., Pawson, A. and Taylor, J. (2005). "The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 2: Serosurveillance, and efficiency and effectiveness of control procedures after the national ban on animal movements." Veterinary record **156**(9): 269-278.

Trevennec, K., Grosbois, V., Roger, F., Ho, T. H., Berthouly-Salazar, C. and Chevalier, V. (2012). "Evidence for freedom from swine influenza in a remote area of Northern Vietnam." Acta tropica **122**(1): 160-163.

Van Schaik, G., Stehman, S. M., Schukken, Y. H., Rossiter, C. R. and Shin, S. J. (2003). "Pooled fecal culture sampling for Mycobacterium avium subsp. paratuberculosis at different herd sizes and prevalence." Journal of veterinary diagnostic investigation **15**(3): 233-241.

Vanderstichel, R., Christensen, J., Stryhn, H. and Hurnik, D. (2013). "Standards for reporting surveillance information in freedom from infection models by example of Trichinella in Canadian market hogs." Preventive Veterinary Medicine **111**(1-2): 176-180.

Vhsv Expert Panel Working Group (2010). "Viral hemorrhagic septicemia virus (VHSV IVb) risk factors and association measures derived by expert panel." Preventive veterinary medicine **94**(1-2): 128-139.

Vose, D. J. (1997). "Risk analysis in relation to the importation and exportation of animal products." Revue scientifique et technique (International Office of Epizootics) **16**(1): 17-19.

Wahlström, H., Frössling, J., Lewerin, S. S., Ljung, A., Cedersmyg, M. and Cameron, A. (2010). "Demonstrating freedom from Mycobacterium bovis infection in Swedish farmed deer using non-survey data sources." Preventive veterinary medicine **94**(1-2): 108-118.

Wahlström, H., Isomursu, M., Hallgren, G., Christensson, D., Cedersmyg, M., Wallensten, A., . . . Hopp, P. (2011). "Combining information from surveys of several species to estimate the probability of freedom from Echinococcus multilocularis in Sweden, Finland and mainland Norway." Acta veterinaria Scandinavica **53**(1): 9-9.

Weber, M. F., Groenendaal, H., Van Roermund, H. J. W. and Nielen, M. (2004). "Simulation of alternatives for the Dutch Johne's disease certification-and-monitoring program." Preventive veterinary medicine **62**(1): 1-17.

Welby, S., Govaerts, M., Vanholme, L., Hooyberghs, J., Mennens, K., Maes, L. and Van Der Stede, Y. (2012). "Bovine tuberculosis surveillance alternatives in Belgium." Preventive veterinary medicine **106**(2): 152-161.

Welby, S., Van Den Berg, T., Marché, S., Houdart, P., Hooyberghs, J. and Mintiens, K. (2010). "Redesigning the serological surveillance program for notifiable avian influenza in Belgian professional poultry holdings." Avian diseases **54**: 597-605.

Whittle, P. J. L., Stoklosa, R., Barrett, S., Jarrad, F. C., Majer, J. D., Martin, P. a. J. and Mengersen, K. (2013). "A method for designing complex biosecurity surveillance systems: detecting non-indigenous species of invertebrates on Barrow Island." Diversity and distributions **19**(5-6): 629-639.

Willeberg, P., Nielsen, L. R. and Salman, M. (2012). "Designing and evaluating risk-based surveillance systems: potential unwarranted effects of applying adjusted risk estimates." Preventive veterinary medicine **105**(3): 185-194.

Willeberg, P., Paisley, L. G. and Lind, P. (2011). "Epidemiological models to support animal disease surveillance activities." Revue scientifique et technique (International Office of Epizootics) **30**(2): 603-614.

Williams, M. S., Ebel, E. D. and Wells, S. J. (2009a). "Poisson sampling: a sampling strategy for concurrently establishing freedom from disease and estimating population characteristics." Preventive veterinary medicine **89**(1-2): 34-42.

Williams, M. S., Ebel, E. D. and Wells, S. J. (2009b). "Population inferences from targeted sampling with uncertain epidemiologic information." Preventive veterinary medicine **89**(1-2): 25-33.

Windsor, P. A., Freeman, P. G., Abila, R., Benigno, C., Verin, B., Nim, V. and Cameron, A. (2011). "Foot-and-mouth disease control and eradication in the Bicol Surveillance Buffer Zone of the Philippines." Transboundary and emerging diseases **58**(5): 421-433.

Ziller, M., Selhorst, T., Teuffert, J., Kramer, M. and Schlüter, H. (2002). "Analysis of sampling strategies to substantiate freedom from disease in large areas." Preventive veterinary medicine **52**(3-4): 333-343.